UC San Diego

UC San Diego Previously Published Works

Title

NDF, a nucleosome-destabilizing factor that facilitates transcription through nucleosomes

Permalink

https://escholarship.org/uc/item/4nb950dp

Journal

Genes & Development, 32(9-10)

ISSN

0890-9369

Authors

Fei, Jia
Ishii, Haruhiko
Hoeksema, Marten A
et al.

Publication Date

2018-05-01

DOI

10.1101/gad.313973.118

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

NDF, a Nucleosome Destabilizing Factor that Facilitates Transcription through Nucleosomes

Jia Fei[1], Haruhiko Ishii[2], Marten A. Hoeksema[3,4], Franz Meitinger[2], George A. Kassavetis[1], Christopher K. Glass[3,4], Bing Ren[2,3,5], and James T. Kadonaga[1]*

[Key words: Nucleosome, Transcription, Gene Expression, RNA Polymerase II, Chromatin Dynamics]

Running Head: Pol II Transcription through Nucleosomes

[1]Section of Molecular Biology, University of California, San Diego, La Jolla, CA 92093, USA.

[2]Ludwig Institute for Cancer Research, San Diego Branch, La Jolla, CA 92093, USA.

[3]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA

[4]Department of Medicine, University of California, San Diego, La Jolla, CA 92093, USA

[5]Center for Epigenomics, Institute of Genome Medicine, Moores Cancer Center, University of California, San Diego, La Jolla, CA 92093, USA

*Corresponding author. Email: jkadonaga@ucsd.edu

Our understanding of transcription by RNA polymerase II (Pol II) is limited by our knowledge of the factors that mediate this critically important process. Here we describe the identification of NDF, a nucleosome destabilizing factor that facilitates Pol II transcription in chromatin. NDF has a PWWP motif, interacts with nucleosomes near the dyad, destabilizes nucleosomes in an ATP-independent manner, and facilitates transcription by Pol II through nucleosomes in a purified and defined transcription system as well as in cell nuclei. Upon transcriptional induction, NDF is recruited to the transcribed regions of thousands of genes and co-localizes with a subset of H3K36me3-enriched regions. Notably, the recruitment of NDF to gene bodies is accompanied by an increase in the transcript levels of many of the NDF-enriched genes. In addition, the global loss of NDF results in a decrease in the RNA levels of many genes. In humans, NDF is present at high levels in all tested tissue types, is essential in stem cells, and is frequently overexpressed in breast cancer. These findings indicate that NDF is a nucleosome destabilizing factor that is recruited to gene bodies during transcriptional activation and facilitates Pol II transcription through nucleosomes.

In the eukaryotic nucleus, transcription of protein-coding genes is carried out by RNA polymerase II (Pol II). Our understanding of Pol II transcription is limited, however, by our knowledge of the factors that mediate this essential process. In spite of considerable progress, many important elements of Pol II transcription remain unsolved.

One particularly intriguing aspect of transcription by Pol II is the ability of the enzyme to traverse the nucleosomal template. A key protein that promotes transcription and DNA replication in chromatin is FACT (facilitates chromatin transcription/transactions; see, for example: Reinberg and Sims 2006; Formosa 2008). FACT binds to nucleosomes and can remove an H2A-H2B dimer or reorganize the nucleosome into an altered structure. In humans, however, the SSRP1 subunit of FACT is not detectable in many tissue types (Garcia et al. 2011; Uhlén et al. 2015) (Supplemental Fig. S1A). It thus appears that critical components of the human transcription elongation machinery have yet to be identified.

In our studies of chromatin dynamics, we became drawn to the study of Pol II elongation in chromatin through our analysis of the prenucleosome, a conformational isomer of the nucleosome that associates with ~80 bp DNA rather than ~147 bp DNA as in a canonical nucleosome (Torigoe et al. 2011, 2013; Fei et al. 2015; Khuong et al. 2015). Notably, H3K56, which is located at the DNA entry and exit points of the nucleosome, can be acetylated by p300 in prenucleosomes but not in canonical nucleosomes (Fei et al. 2015). This observation is consistent with the partial unwrapping of the DNA in a prenucleosome relative to a nucleosome.

The properties of the prenucleosome led us to consider whether there might be factors that could partially disassemble a nucleosome and facilitate the acetylation of nucleosomal H3K56. These studies resulted in the identification of a protein that we term NDF (nucleosome destabilizing factor), which destabilizes nucleosomes, is enriched in the transcribed regions of many active genes, facilitates transcription by Pol II through nucleosomes, and is present in all tested human tissue types. The identification of NDF reveals a new Pol II elongation factor that is structurally and functionally distinct from FACT.

Results

Identification and purification of NDF, a novel nucleosome destabilizing factor

To identify factors that destabilize or disassemble nucleosomes, we employed an assay that was based on our observations of H3K56 accessibility in prenucleosomes versus nucleosomes (Fei et al. 2015). Specifically, we tested for activities that would enable nucleosomal H3K56 to be acetylated by purified p300. In principle, this assay would reveal factors that either destabilize nucleosomes or unwrap DNA from the ends of nucleosomes. As shown in Fig. 1A, we identified a heat-sensitive activity in a nuclear extract from *Drosophila* embryos that stimulates the ability of p300 to acetylate H3K56 with a chromatin substrate. This activity was purified by six column chromatography steps (Fig. 1B). The final Source 15Q column yielded a major ~75 kDa polypeptide as well as *Drosophila* yolk proteins (Fig. 1C), which are among the most abundant proteins in embryos. Mass spectrometry revealed that the ~75 kDa polypeptide is the CG4747 protein. To test whether the presence of CG4747 correlates with the biochemical activity, we synthesized full-length *Drosophila* CG4747 in bacteria and generated polyclonal antisera in rabbits. Western blot analysis with the antibodies revealed that CG4747 mostly correlates with the stimulation of nucleosomal H3K56 acetylation (Fig. 1C).

CG4747 was previously identified by ChIP-mass spectrometry as a factor that associates with the *Drosophila* MSL (male-specific lethal) complex, and it was found to facilitate the targeting of MSL to active genes in a manner that is dependent upon H3K36 trimethylation by Set2 (Wang et al. 2013). The human homolog of the *CG4747* gene is *GLYR1* (glyoxylate reductase 1 homolog; also known as N-PAC and NP60). The GLYR1 protein was identified in a screen for factors that preferentially bind to H3K36me3 relative to unmethylated H3K36 Vermeulen et al. 2010), and is present at high levels in the nuclei of all tested tissue types (Supplemental Fig. S1A). The *GLYR1* gene appears to be essential in human HUES62 stem cells but not in human A375 melanoma cells (Shalem et al. 2014). GLYR1 was observed to stimulate H3K4me1 and H3K4me2 demethylation by LSD2 (KDM1b) (Fang et al. 2013);

however, the region of human GLYR1 that interacts with LSD2 is not present in *Drosophila* CG4747 (Supplemental Fig. S1B). Both CG4747 and GLYR1 proteins contain a PWWP domain, which binds to methylated lysines (Qin and Min 2014), as well as a dehydrogenase domain (Fig. 2A, Supplemental Fig. S1B) that does not appear to be enzymatically active in human GLYR1 (Fang et al. 2013). In addition, human GLYR1 has a putative DNA-binding AT-hook motif, but *Drosophila* CG4747 does not (Supplemental Fig. S1B). Based on the shared biochemical activities of CG4747 and GLYR1 (as described in this work), we have designated these factors as NDF, for nucleosome destabilizing factor.

NDF destabilizes nucleosomes in an ATP-independent manner

To test the biochemical activities of dNDF and hNDF, we synthesized and purified the full-length recombinant proteins (Fig. 2B). We observed that dNDF and hNDF are each able to stimulate the acetylation of nucleosomal H3K56 by p300 in a dose-dependent and ATP-independent manner (Fig. 2C, Supplemental Figs. S2A and S2B). Purified dNDF and hNDF did not exhibit intrinsic histone acetyltransferase activity (Fig. 2C) and also did not stimulate the acetylation of free histones by p300 (Supplemental Fig. S2C). Thus, purified recombinant dNDF and hNDF can each mediate the stimulation of nucleosomal H3K56 acetylation, as seen in our original assay with the native *Drosophila* factor (Fig. 1).

We then examined whether NDF has the ability to disassemble nucleosomes. To this end, we investigated the ability of NDF to disrupt nucleosome-mediated supercoiling of covalently closed circular plasmid DNA. The wrapping of DNA around the core histone octamer in a nucleosome causes a change in the linking number of approximately −1 (Germond et al. 1975). Therefore, nucleosome disassembly can be monitored by the loss of negative supercoiling in chromatin in circular plasmid DNA. We reconstituted nucleosomes onto plasmid DNA by using salt dialysis methodology, and then treated the chromatin with NDF in the presence of topoisomerase I. The resulting DNA species were analyzed by agarose gel electrophoresis in the absence or the presence of chloroquine, which introduces positive

supercoils into the DNA and enables the differentiation between relaxed circular DNA and nicked circular DNA (Fig. 2D, Supplemental Figs. S2D and S2E). These experiments revealed that hNDF and dNDF are each able to mediate the partial disassembly of nucleosomes (as seen by the loss of negative DNA supercoiling) in a dose- and time-dependent manner. Hence, the results from the nucleosomal H3K56 acetylation and the chromatin disruption experiments indicate that hNDF and dNDF can each partially disassemble nucleosomes. In addition, because NDF functions in an ATP-independent manner, it appears that the binding of NDF to chromatin provides the energy for the destabilization of nucleosomes.

The recruitment of NDF to gene bodies upon transcriptional induction is accompanied by an increase in transcript levels

To analyze the properties of hNDF, we generated four hNDF knockout (KO) lines in HeLa cells as well as rabbit polyclonal antisera against the full-length protein (Fig. 3A, Supplemental Fig. S2F). hNDF appears to be essential in human HUES62 stem cells but not in human A375 melanoma cells (Shalem et al. 2014). The hNDF KO HeLa cells are viable but exhibit reduced growth rates relative to wild-type cells (Fig. 3B). The hNDF KO cells were found to have a higher percentage of cells in G1 phase than wild-type cells (Fig. 3C). Single-cell live imaging experiments showed that the hNDF KO cells have a longer duration of the cell cycle than wild-type cells and no apparent mitotic alterations (Supplemental Fig. S3).

To determine the localization of NDF, we carried out ChIP-seq analyses in wild-type and KO (as background control) cells. We observed that endogenous hNDF is often enriched over the transcribed regions of active genes and co-localizes with H3K36me3 over those genes (Figs. 3E and 3F, Supplemental Fig. S4A). This distribution of hNDF is similar to that seen with epitope-tagged dNDF in transgenic *Drosophila* larvae (Wang et al. 2013), and is likely to be due, at least in part, to the binding of the PWWP domain to H3K36me3 (Wang et al. 2013; Qin and Min 2014). Consistent with the presence of NDF over transcribed genes, hNDF is found

predominantly in introns and exons (Supplemental Fig. S4B). Strikingly, hNDF exhibits a distinct preference for longer genes relative to shorter genes (Fig. 3F).

To investigate the relationship between hNDF (ChIP), H3K36me3 (ChIP), and gene activity (RNA-seq), we subjected these three variables in 15,407 genes to k-means clustering and generated four clusters of genes (Figs. 4A and 4B, Supplemental Fig. S4C). Cluster 1 contains active genes that are enriched in NDF and H3K36me3 along the gene body. Cluster 2 has active genes with lower levels of NDF and intermediate to high levels of H3K36me3. Clusters 3 and 4 contain less active or inactive genes with low levels of NDF and H3K36me3. This analysis shows that active genes can be subdivided into a cluster that is enriched in NDF (cluster 1) and a cluster that is not enriched in NDF (cluster 2). It can also be seen that length distribution of genes in cluster 1 is distinct from that of genes in cluster 2 (Supplemental Fig. S4D).

To determine whether NDF affects gene activity, we performed RNA-seq analyses with wild-type and hNDF KO cells. These experiments revealed that the loss of hNDF results in a decrease in the steady-state RNA levels of many genes (Fig. 4C, Supplemental Fig. S4E). This effect was observed in genes with higher hNDF levels (cluster 1) as well as in genes with lower NDF levels (clusters 2-4). It is possible, for instance, that NDF is important for transcription of genes with low as well as high NDF occupancy. In addition, some of the effects could be indirect, such as through the downregulation of transcriptional activators or repressors. Nevertheless, it appears that NDF contributes to the expression of a broad range of genes.

We also carried out global run-on sequencing (GRO-seq) analysis of wild-type and hNDF KO cells. These experiments revealed differences in the enrichment of transcriptionally engaged polymerases over gene bodies in wild-type relative to KO cells (Supplemental Fig. S5). These findings further support a role of NDF in the transcription process.

We then examined the dynamics of mouse NDF (mNDF) in primary bone marrow-derived macrophages (BMDM). Because mNDF is nearly identical to hNDF (Supplemental Fig. S1B), our hNDF antibodies recognize mNDF. In these studies, we used Kdo2-Lipid A (KLA), a

highly specific TLR4 agonist (Raetz et al. 2006), to activate TLR4-inducible genes in BMDM (Kaikkonen et al. 2013; Oishi et al. 2017). The gene encoding mNDF is not induced by KLA (Supplemental Fig. S6A). In these experiments, we observed that mNDF is recruited to the gene bodies of over 2000 genes after 6 h of KLA activation (Figs. 5A and 5B, Supplemental Fig. S6B). This increase in the occupancy of NDF correlates with the increase in H3K36me3 as well as with transcript levels (Figs. 5C and 5D). These results thus reveal a dynamic role of NDF during gene activation.

These experiments show that NDF is recruited to the transcribed regions of over 2000 genes upon transcriptional induction (Figs. 5A and 5B). Moreover, this recruitment of NDF is accompanied by an increase in the transcript levels of many NDF-enriched genes (Fig. 5D). In addition, the global loss of hNDF results in a decrease in the RNA levels of many genes (Fig. 4C, Supplemental Fig. S4E). These findings, combined with the ability of NDF to destabilize nucleosomes, suggested that NDF may affect transcriptional elongation in chromatin.

NDF facilitates Pol II elongation in chromatin in vitro and in cell nuclei

To test whether NDF affects transcription through nucleosomes, we performed elongation assays with purified yeast RNA polymerase II (Pol II) on templates with a positioned downstream nucleosome. To this end, we employed a strategy that is similar to the method of Kireeva et al. (2002). As depicted in Fig. 6A, these experiments involved the assembly of a Pol II elongation complex with a 5'-labeled RNA primer and the subsequent ligation of a downstream nucleosome. Transcript elongation was initiated by the addition of ribonucleoside 5'-triphosphates, and the progression of Pol II was followed by analyzing the length of the 5' end-labeled transcripts.

We reconstituted mononucleosomes by salt dialysis methodology with purified bacterially-synthesized core histones and DNA containing the nucleosome positioning sequence from a 5S rRNA gene of *Xenopus borealis* (Rhodes 1985; Hayes et al. 1990). To analyze the positioning of the nucleosomes, we used a modified version of the histone-directed DNA

cleavage method of Flaus et al. (1996). These experiments showed that the reconstituted mononucleosomes were located at a single predominant position in the 5S rDNA (Supplemental Fig. S7).

We then purified yeast Pol II (Supplemental Fig. S8A) and performed the Pol II elongation experiments with the downstream 5S rDNA sequence as a mononucleosome or naked DNA (Fig. 6B). Consistent with well-established data (see, for example: Izban and Luse 1991; Kireeva et al. 2002), the downstream nucleosome was observed to inhibit Pol II elongation, and, as a control, the nucleosome-mediated inhibition was lost upon dissociation of the histones at 1000 mM KCl (Supplemental Fig. S8B). Nucleosome-specific pause sites were seen at −80, −73, and −20 nt relative to the dyad at position 0. When it is considered that the distance from the leading edge of Pol II to the active site is about 15 nt (Samkurashvili and Luse 1996; Kettenberger et al. 2004), the −80 pause site roughly corresponds to the site of Pol II entry into the nucleosome, and the −20 pause site corresponds to the approach of Pol II to the dyad. These pause sites have also been observed in *Drosophila* cells (Weber et al. 2014).

Upon addition of hNDF, there was a significant reduction in the pausing of Pol II in the nucleosome. The effects were particularly strong at the nucleosome entry (−80) and dyad (−20) pause sites, and could be seen at different reaction times (Fig. 6B) and hNDF concentrations (Supplemental Fig. S8C) as well as in the scanning profiles of the elongation products (Supplemental Fig. S8D). Moreover, in the presence of hNDF, there is a substantial amount of transcription through the nucleosome at the 8 and 16 min reaction times (Fig. 6B, Supplemental Fig. S8D). In addition, NDF reduces Pol II pausing through the naked DNA template (Fig. 6B); hence, it is possible that hNDF has a functional interaction with Pol II that is distinct from its ability to destabilize nucleosomes. We also tested dNDF in the Pol II elongation assay and observed that dNDF, like hNDF, is able to facilitate Pol II transcription through a nucleosome (Supplemental Fig. S8E). These data reveal that both hNDF and dNDF can facilitate Pol II elongation through a nucleosome.

We next examined whether NDF affects Pol II elongation in cells. To this end, we carried out nuclear run-on experiments with wild-type and NDF knockout cells. These experiments showed that the loss of NDF results in a decrease in Pol II elongation in nuclei (Supplemental Fig. S9A). Importantly, when the histones were stripped from the DNA with 1% (w/v) Sarkosyl (Green et al. 1975; Scheer 1978), we observed roughly the same amount of Pol II elongation in the wild-type cell nuclei as in the knockout cell nuclei (Supplemental Fig. S9B). Hence, the reduced levels of transcription in the NDF knockout cell nuclei appear to be due to blockage by nucleosomes. In addition, the transcription was sensitive to 4 μg/mL α-amanitin and was thus mediated by Pol II (Supplemental Fig. S9). These findings therefore indicate that NDF facilitates Pol II transcription through nucleosomes *in vitro* as well as in cell nuclei.

NDF interacts with nucleosomes in the vicinity of the dyad

To gain a better understanding of the function of NDF in nucleosome destabilization and transcription elongation in chromatin, we examined the interaction of NDF with naked DNA and nucleosomes. First, we performed gel mobility shift analyses. [In these studies, we used dNDF instead of hNDF because the hNDF-DNA and hNDF-mononucleosome species were not stable under the gel electrophoresis conditions.] These experiments revealed that NDF binds to 147 bp 5S rDNA sequences as either naked DNA or mononucleosomes (Fig. 7A). To test for any potential sequence-specificity in binding, we carried out identical experiments with a sequence from a *Drosophila hsp70* promoter and obtained essentially the same results (Supplemental Fig. S10A). Thus, NDF binds to naked DNA as well as to nucleosomes. In addition, the binding of NDF to 147 bp DNA-containing mononucleosomes suggests that the nucleosomal binding does not require linker DNA.

Then, to investigate the binding of NDF to DNA and nucleosomes at higher resolution, we performed DNase I footprinting experiments. With naked DNA, neither hNDF nor dNDF exhibited sequence specific interactions (Fig. 7B, Supplemental Figs. S10B, S11A, S11B). With mononucleosomes, DNase I digestion yielded the characteristic 10 bp digestion ladder (Noll

1974), and both hNDF and dNDF were observed to protect DNA in the vicinity of the dyad (Fig. 7B, Supplemental Figs. S10B and S11). Moreover, in the forward strand (*i.e.*, the non-template strand in the transcription experiments), we observed DNase I hypersensitivity in the flanking ends of the nucleosomal DNA (Fig. 7B, Supplemental Figs. S11A and S11C). This hypersensitivity may reflect the destabilization of the nucleosomes by NDF, and is consistent with the ability of NDF to enhance p300-mediated acetylation of H3K56, which is normally protected by nucleosomal DNA. Thus, the DNase I footprinting experiments revealed an interaction of NDF at the dyad and an alteration in histone-DNA contacts at the flanking nucleosomal DNA.

The interaction of NDF at the nucleosome dyad led us to ask whether the effects of NDF on nucleosomal transcription are similar to those of Sin (SWI-INdependent) mutations in histone H4. The *SIN* genes were identified in yeast on the basis of their ability to provide partial relief of the requirement of *SWI* genes for *HO* transcription (Sternberg et al. 1987). It was later found that some mutant alleles of the histone H4 gene have a Sin^- phenotype (Kruger et al. 1995). We were specifically interested in the histone H4R45A Sin mutation. H4R45 is located near the nucleosome dyad. The H4R45A mutation results in a Sin^- phenotype (Hodges et al. 2015), but has only a minimal effect on nucleosome structure (Muthurajan et al. 2004). H4R45A and the related H4R45C mutation reduce Pol II pausing on nucleosomal templates in vitro (Hsieh et al. 2010; Bintu et al. 2012). In addition, the H4R45H Sin mutation causes an increase in nuclease sensitivity as well as a reduction in nucleosome-mediated DNA supercoiling (Wechser et al. 1997). Somewhat strikingly, the effects of the H4R45 Sin mutations (*i.e.*, reduction of Pol II pausing on nucleosomes, loss of DNA supercoiling, and increased nuclease sensitivity) are similar to the effects of NDF on chromatin.

We therefore generated chromatin with H4R45A and tested the effect of the R45A mutation on Pol II elongation (Fig. 7C). In this manner, we could compare the effects of NDF-mediated nucleosome destabilization with the effects of the H4R45A mutation. These experiments revealed that the H4R45A mutation resulted in the loss of the −20 dyad pause site

and an accumulation of a downstream pause site, but did not affect the −80 (nucleosome entry) or −73 pause sites. The addition of NDF facilitated Pol II transcription through the −80, −73, and other pause sites. These findings indicate that the H4R45A mutation acts specifically at the nucleosomal dyad, whereas NDF functions throughout the entire nucleosome.

Discussion

In this study, we describe the properties of NDF, a nucleosome destabilizing factor that facilitates Pol II transcription through nucleosomes. We initiailly identified and purified NDF on the basis of its ability to facilitate the acetylation of nucleosomal histone H3K56 by the p300 acetyltransferase (Fig. 1). Both hNDF and dNDF can mediate the partial disassembly of nucleosomes, as detected by the loss of DNA supercoiling in circular chromatin templates (Fig. 2D, Supplemental Figs. S2D and S2E). hNDF is localized over gene bodies (Figs. 3D, 3E, 4A, 4B, Supplemental Fig. S4A). These results are consistent with the properties of epitope-tagged dNDF (CG4747) (Wang et al. 2013). Clustering analysis revealed that active genes can be subdivided into a cluster that is enriched in NDF and a cluster that is not enriched in NDF (Figs. 4A and 4B, Supplemental Fig. S4C). Thus, NDF is a nucleosome destabilizing factor that is present in the transcribed regions of a subset of active genes.

Importantly, upon induction of transcription by KLA in mouse primary BMDM, we found that NDF is recruited to gene bodies and that this recruitment is accompanied by an increase in the transcript levels of many of the NDF-enriched genes (Fig. 5). We also observed that the global loss of hNDF results in a decrease in the RNA levels of many genes (Fig. 4C, Supplemental Fig. S4E). NDF may thus have a direct role in gene activation. Intriguingly, NDF is preferentially associated with longer genes than shorter genes (Fig. 3F, Supplemental Fig. S4D). In this regard, it may be more advantageous for NDF to be located at longer genes than at shorter genes because more nucleosomes would have to be disassembled during transcription of longer genes than during transcription of shorter genes. Moreover, if there is some three dimensional clustering of actively transcribed genes, then there may be a higher local concentration of NDF at longer transcription units than at shorter transcription units.

We directly tested whether NDF affects Pol II elongation through nucleosomes, and found that hNDF and dNDF can relieve nucleosome-mediated Pol II pausing, such as at the nucleosome entry site and at the dyad (Figs. 6A and 6B, Supplemental Fig. S8). Moreover, in

the presence of NDF, there is efficient Pol II elongation through the nucleosome (Fig. 6B, Supplemental Fig. S8D). We then performed nuclear run-on experiments to test whether NDF contributes to Pol II elongation in cells, and found that NDF is important for Pol II elongation through chromatin in cell nuclei (Supplemental Fig. S9).

To gain further insight into this activity, we examined the interaction of NDF with chromatin, and found that both hNDF and dNDF interact with nucleosomes in the vicinity of the dyad (Figs. 7A and 7B, Supplemental Fig. S10 and S11). Comparison of the effects of NDF with those of the H4R45A Sin mutation revealed that H4R45A affects transcription at the dyad, whereas NDF affects transcription throughout the nucleosome (Fig. 7C). Altogether, these data indicate that NDF is a nucleosome binding and destabilizing factor that is recruited to gene bodies upon induction of transcription and facilitates Pol II elongation in chromatin (Fig. 7D).

NDF is present in most animals, and perhaps in plants, but it does not appear to have a homolog in *Saccharomyces cerevisiae*. NDF contains a conserved N-terminal PWWP motif and a conserved C-terminal dehydrogenase domain (Fig. 2A, Supplemental Fig. S1B), but hNDF does not exhibit dehydrogenase activity (Vermeulen et al. 2010). The PWWP methyl lysine binding domain (Qin and Min 2014) is likely to be involved in the binding of NDF to H3K36me3. It appears, however, that H3K36me3 is necessary but not sufficient for the recruitment of NDF to chromatin. The proper localization of dNDF in chromatin is dependent upon the Set2 histone H3K36 methyltransferase (Wang et al. 2013); however, many genes with moderate to high levels of H3K36me3 are not enriched in NDF (Figs. 4A and 4B, Supplemental Fig. S4C), and the sharp increase in NDF occupancy at gene bodies upon transcriptional induction is not mirrored by a similarly sharp increase in H3K36 trimethylation (Figs. 5A and 5B). Thus, H3K36 trimethylation does not appear to be sufficient for the recruitment of NDF. It is also notable that the biochemical activities of NDF, such as the destabilization of nucleosomes and relief of nucleosome-mediated Pol II pausing, do not absolutely require H3K36 methylation, as our experiments were performed with nucleosomes that were reconstituted with unmodified bacterially synthesized histones. These findings suggest that NDF, once recruited to its site of

action, can destabilize nucleosomes and facilitate transcription by a mechanism that is not strictly dependent upon H3K36 trimethylation.

Because NDF relieves nucleosome-mediated inhibition of Pol II elongation, it is essential to compare NDF with FACT, the well-studied factor that facilitates transcription and DNA replication in chromatin (see, for example: Reinberg and Sims 2006; Formosa 2008). FACT is present in many eukaryotes, including *S. cerevisiae*, *Drosophila*, and humans. FACT and NDF are not related in terms of their primary amino acid sequences or protein domains. Moreover, FACT binds to the histones and exhibits core histone chaperone activity, whereas NDF does not appear to be a core histone chaperone (data not shown). In yeast, the Spt16 subunit of FACT is enriched over gene bodies and depleted over promoter regions of active genes (True et al. 2016). In contrast, both subunits of mouse FACT were found to be strongly enriched in the vicinity of transcription start sites but also present over gene bodies (Mylonas and Tessarz 2018). These findings suggest that there are substantial differences in the biological functions of FACT in yeast versus mice. The localization of mouse and human NDF (Figs. 3E and 5B) is more similar to that of yeast FACT than mouse FACT. It is also notable that human NDF was observed to be present at high levels in the nuclei of 44/44 tissue types tested, whereas the levels of the SSRP1 subunit of FACT were found to be low (16/43) or not detectable (9/43) in the same tissue types (Garcia et al. 2011; Uhlén et al. 2015) (Supplemental Fig. S1A). In addition, it appears that NDF is essential in human HUES62 stem cells, whereas subunits of FACT are not (Shalem et al. 2014). Thus, in mammals, NDF and FACT are both able to destabilize nucleosomes and to facilitate transcription in chromatin, but they function by distinct mechanisms and in different tissues.

It is possible and perhaps likely that there is at least one other factor that is related to NDF. The existence of an NDF-related factor is suggested by the observation that NDF is not essential for the viability of HeLa cells or A375 cells. In addition, the absence of NDF does not result in a general loss of transcription of genes that are enriched in NDF (Fig. 4). These findings could be explained by the presence of an NDF-like factor that is recruited to mostly the

same genes as NDF (Cluster 1 genes, Fig. 4), such as via a PWWP motif. This NDF-like factor would then largely compensate for the absence of NDF, especially at genes that are enriched for both factors. The identification and characterization of an NDF-related factor would lead to significant new insights into the transcription process.

Based on the effects of NDF on transcription and gene expression, it is also important to consider whether alterations in the protein sequence or levels of NDF might be relevant to human diseases. In this regard, the *NDF/GLYR1* gene is overexpressed in about 21% of breast cancers (Supplemental Figs. S12A and S12B) (Forbes et al. 2017) and mutated in about 51% of microsatellite-unstable colorectal cancers (Alhopuro et al. 2012). The upregulation of *NDF/GLYR1* also correlates with a reduced survival probability in breast cancer patients (Supplemental Figs. S12C) (Uhlen et al. 2017). The knowledge of the biochemical activities of NDF would be useful in assessing the potential roles that this factor may have in breast cancer or other diseases.

In conclusion, the properties of NDF indicate that it has a key role in the transcription of chromatin by Pol II in animals. The further analysis of NDF should reveal additional factors and mechanisms that orchestrate the procession of Pol II through active genes.

Materials and methods

To ensure the reproducibility of the results, each experimental condition was performed independently at least two times

Antibodies

Rabbit polyclonal antisera against dNDF and hNDF were generated with purified full-length *Escherichia coli*-synthesized proteins at Pocono Rabbit Farm & Laboratory (Canadensis, PA). Commercial antibodies that were used in this study are as follows: anti-H3K56ac (Millipore Sigma, 07-677I, 1:450 dilution), anti-H3K56ac (Active Motif, 39281, 1:1000 dilution), anti-H3K56ac (Cell Signaling Technology, 4243, 1:1000 dilution), anti-H3 (Cell Signaling Technology, 4499, 1:4000 dilution), anti-H3K9ac (Abcam, ab8898, 1:1000 dilution), anti-H3K36me3 (Abcam, ab9050, 1:1000 dilution), anti-his (Santa Cruz Biotechnology, sc-803, 1:1000 dilution), and anti-pol II (Santa Cruz Biotechnology, sc-9001, 1:1000 dilution).

Purification of native NDF from Drosophila embryos

Native dNDF was purified from the 0.1 M KCl version of the soluble nuclear fraction (Kamakaka et al. 1991) from *Drosophila* embryos by using the scheme shown in Fig. 1B of the main text. Buffer N [20 mM Hepes, K^+, pH 7.5, 10% (v/v) glycerol, 0.5 mM EDTA, 0.5 mM EGTA, 1.5 mM $MgCl_2$, 0.1 M NaCl, 10 mM ß-glycerophosphate, 1 mM DTT and 0.2 mM PMSF] was used throughout the purification. Proteins in the peak activity fractions of the final Source 15Q chromatography step were identified by mass spectroscopy (Biomolecular/Proteomics Mass Spectrometry Facility, UCSD).

Purification of recombinant proteins

The cDNAs for dNDF (CG4747; *Drosophila* Genomics Resource Center, Bloomington, IN) and hNDF (*GLYR1*; Dharmacon, Lafayette, CO; note that three mismatches in the cDNA were

corrected) were subcloned into pET21b with N-terminal His6 tags to give pET21b-His6-dNDF and pET21b-His6-hNDF, and the integrity of the sequences was confirmed by DNA sequence analysis. The N-terminally His6-tagged dNDF and hNDF proteins were purified by the following method. Freshly transformed *E. coli* BL21(DE3) was grown in 2 L of LB medium containing 40 µg/mL ampicillin at 37˚C to an A_{600nm} of approximately 0.6, and the synthesis of the recombinant NDF was induced by the addition of IPTG to 0.4 mM final concentration. The culture was then incubated in an 18°C shaking water bath for an additional 16 to 18 h, and the bacteria were collected by centrifugation (Fiberlite F9-4x1000y rotor; 6000 rpm; 10 min; 22˚C). Unless stated otherwise, all subsequent operations were performed at 4°C. The bacterial pellets were resuspended in 30 mL of Lysis Buffer [10 mM Tris-HCl, pH 7.5, 0.2% (v/v) Nonidet P-40, 1 M NaCl, 10% (v/v) glycerol, 15 mM imidazole, 5 mM 2-mercaptoethanol, 0.2 mM PMSF, and 1x Protease Inhibitor Cocktail (Millipore Sigma; P9599)]. The suspension was subjected to sonication on ice (Branson Sonifier 450 with a 0.25 inch microtip; 20% output; 4 x 30 s). Insoluble material was removed by centrifugation (Fiberlite F21S-8x50y rotor; 13,000 rpm; 30 min), and the supernatant was incubated with 3 mL (pre-equilibrated in Lysis Buffer) Ni-NTA Superflow (Qiagen) for 3 h on a nutator. The mixture was transferred into a Bio-Rad Econo-Pac chromatography column, and the flowthrough, which contained ~80% of the His6-NDF proteins, was collected and incubated with 2 mL (pre-equilibrated in Lysis Buffer) Ni-NTA Agarose (Qiagen) for 2 h on a nutator. The mixture was transferred into a Bio-Rad Econo-Pac chromatography column, and the resin was washed with 2 x 10 mL of Buffer A [10 mM Tris-HCl, pH 7.5, 0.2% (v/v) Nonidet P-40, 10% (v/v) glycerol, 15 mM imidazole, and 5 mM 2-mercaptoethanol] containing 1 M NaCl followed by 2 x 10 mL of Buffer A containing 0.2 M NaCl. The His6-NDF was eluted with 4 x 0.5 mL of Elution Buffer [10 mM Tris-HCl, pH 7.5, 0.2% (v/v) Nonidet P-40, 0.2 M NaCl, 10% (v/v) glycerol, 250 mM imidazole, and 5 mM 2-mercaptoethanol]. With His6-dNDF, it was necessary to remove a contaminating nuclease by dialysis into Buffer A containing 0.1 M NaCl followed by chromatography on a 1 mL Source 15Q (GE Life Sciences) anion exchange column with linear gradient elution from 0.1 M to 1.0 M

NaCl. The His6-dNDF elutes at approximately 0.16 M NaCl. The purified NDF proteins were frozen in liquid nitrogen and stored at -80˚C.

Human p300 protein was synthesized in Sf9 cells and purified as described in Kraus and Kadonaga (1998). Recombinant *D. melanogaster* core histones were synthesized in *E. coli* and purified by the method of Khuong et al. (2017). The catalytic domain of *Drosophila* topoisomerase I was synthesized and purified as described by Fyodorov and Kadonaga (2003). The *Saccharomyces cerevisiae* strain that expresses protein A-tagged Rbp3 subunit of RNA polymerase II (Pol II) was the generous gift of Dr. Craig Kaplan (Texas A&M University), and the Pol II was purified as described (Kaplan et al. 2008).

Nucleosome reconstitution

The *Xenopus borealis* 5S rDNA and *Drosophila hsp70* (*87A*) nucleosome positioning sequences were amplified by PCR and purified by 1.5% agarose gel electrophoresis and gel extraction (Qiaquick Gel Extraction Kit; Qiagen). The resulting fragments were then reconstituted into mononucleosomes by salt dialysis (Stein 1989). In the transcription elongation studies, the PCR-amplified DNA was digested with TspRI restriction enzyme (New England Biolabs) for 2 h at 65˚C prior to agarose gel electrophoresis, gel extraction, and mononucleosome reconstitution. Nucleosomes were reconstituted onto CsCl-purified pGIE-0 plasmid DNA (Pazin et al. 1994) by salt dialysis (Stein 1989; Fei et al. 2015) at a core histone to DNA mass ratio of 0.8:1.

H3K56 acetylation assay

The H3K56 acetylation assays were performed with either nucleosomes reconstituted onto plasmid DNA by salt dialysis (1.4 μM concentration of nucleosomes) or free core histones (1.4 μM concentration of core histone octamers) along with 1 μg purified human p300, 30 μM acetyl-CoA, and 1.4 μg bovine serum albumin in Buffer K [25 mM Hepes, K^+, pH 7.6, 100 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, and 10% (v/v) glycerol] in a total volume of 50 μL. The mixture was

incubated at 30˚C for 1 h. The histones were precipitated with 25% (w/v) trichloroacetic acid, subjected to 15% polyacrylamide-SDS gel electrophoresis, and detected by western blot analysis.

Transcription elongation assay

The transcription elongation assay is depicted in Fig. 6A, and is a modified version of the method of Kireeva et al. (2002). In our experiments, the 10 nt 5'-^{32}P-labeled primer RNA (1.6 pmol = 8 µL of a 0.2 µM stock) was annealed to an equimolar amount of template strand (TS) DNA. Purified Pol II (4 pmol) was then added to 1.6 pmol of the RNA-TS complex, and the components were incubated at 30˚C for 10 min. Next, the 5'-biotin-labeled non-template strand (NTS) (1.6 pmol) was added and incubated at 30˚C for 15 min to form the elongation complex (EC), which was then immobilized onto 10 µL of Dynabeads MyOne Streptavidin C1 magnetic beads (Invitrogen). The beads were washed with 2 x 50 µL of EB1000 buffer [20 mM Tris-HCl, pH 7.5, 1 M KCl, 5 mM $MgCl_2$ and 10 mM DTT] followed by 2 x 50 µL of EB40 buffer [20 mM Tris-HCl, pH 7.5, 40 mM KCl, 5 mM $MgCl_2$ and 10 mM DTT], and then resuspended in 16 µL of EB40 buffer. The immobilization efficiency of the EC was estimated with a Geiger counter. Typically, about 60% of the EC is immobilized onto the Dynabeads. Then, either mononucleosomes or the corresponding naked DNA fragments (3 pmol; both with complementary TspRI overhangs) were ligated to the immobilized EC by using 2000 units of T4 DNA ligase (New England Biolabs) in 0.2x (final concentration) T4 DNA Ligase Reaction Buffer (New England Biolabs) in a total volume of 50 µL for 25 min at 22˚C. After ligation, the beads were collected and resuspended in 50 µL EB40 buffer containing 20 µg bovine serum albumin (Thermo Fisher Pierce), 100 units SUPERase•In RNase Inhibitor (Thermo Fisher Invitrogen), and 30 pmol of unlabeled free mononucleosomes on 147 bp DNA. The unlabeled free mononucleosomes were added only to the EC templates that were ligated to a downstream nucleosome, and served to stabilize the template-associated nucleosomes by increasing the overall concentration of nucleosomes in the reaction medium (see, for example, Godde et al.

1995). The beads were then aliquoted into tubes for the individual reactions. Each reaction contained 5 µL of the bead-containing suspension. If the ligation efficiency were 100%, then each tube would contain ~0.08 pmol of EC template. For nucleosome-containing templates, each tube additionally contained 3 pmol of unlabeled free mononucleosomes to stabilize the template-associated nucleosomes. Next, NDF or the corresponding buffer (as a negative control) was added to each tube, EB40 buffer was added to a final volume of 9 µL, and the samples were incubated at 30˚C for 20 min. At this stage, the KCl concentration was 100 mM, unless stated otherwise. Then, transcription elongation was initiated by the addition of 1 µL of 2.5 mM rNTPs, and the reactions were allowed to proceed at 22˚C for the indicated times. The reactions were terminated with 10 µL of Gel Loading Buffer [1 mL of Gel Loading Buffer was prepared by combining 0.9 mL of deionized formamide, 0.1 mL of 0.5 M EDTA, and 2 µL of 4% (w/v) bromphenol blue]. The samples were analyzed by denaturing 8% polyacrylamide-urea gel electrophoresis, and the gel images were collected on a GE Typhoon imager (GE Healthcare).

Analysis of mouse NDF (mNDF) in primary bone marrow-derived macrophages (BMDM)

Bone marrow was isolated from 8- to-10-week-old C57Bl/6 (Harlan Laboratories) female mice. BMDM were generated in DMEM supplemented with M-CSF for 7 days and activated with Kdo2-Lipid A (KLA) for 6 hours as described (Kaikkonen et al. 2013). RNA-seq libraries were prepared from poly(A)-enriched mRNA, as previously described (Kaikkonen et al. 2013). For ChIP-seq, 5 million BMDM were crosslinked in 1% formaldehyde for 10 min. ChIP-seq was performed as described previously (Kaikkonen et al. 2013) with 2 µL of hNDF antisera 7 or 2 µL of H3K36me3 antibodies (Abcam, ab9050). [We also performed ChIP-seq experiments with hNDF antisera 6 and obtained essentially identical results as with hNDF antisera 7 (data not shown).] ChIP-Seq libraries were prepared from ChIP DNA and inputs by blunting, A-tailing, and adapter ligation as previously described (Heinz et al. 2010) by using NextFlex barcode adapters. Libraries were amplified and size selected by gel extraction and subsequently sequenced on the Illumina Hi-Seq 4000 sequencer by using single-end 75 bp reads.

Sequencing data were aligned to the mm10 genome by using Bowtie2, and all data were subsequently analyzed with HOMER and visualized in the UCSC genome browser. Pathway analysis was performed by using Metascape (http://metascape.org).

Additional Materials and Methods are included in the Supplemental Material.

The genome-wide data have been deposited at the Gene Expression Omnibus (GEO; accession number GSE109690), and will be released upon acceptance of the paper for publication.

Acknowledgments

We thank Peter Geiduschek, Grisel Cruz, and Long Vo ngoc for critical reading of the manuscript. We are grateful to Craig Kaplan for the tagged yeast Pol II strain, Gaoyang Liang for CRISPR-Cas9 vectors, and Dong Wang and Jun Xu for their invaluable advice on the elongation assay and Pol II purification. We also thank Yun-Ching Chen for help in the analysis of the RNA-seq data. J.T.K. is the Amylin Chair in the Life Sciences. F. M. was supported by a fellowship from the Deutsche Forschungsgemeinschaft (ME 4713/1-1) and an NIH grant to Karen Oegema (GM074207). This work was supported by funding from the Ludwig Institute for Cancer Research to B.R., the NIH/NIDDK (DK091183) and Foundation Le Ducq to C.K.G, and the NIH/NIGMS (R35 GM118060) to J.T.K.

Author Contributions

This work was carried out in an interactive manner in which all authors made important intellectual contributions to the development of the project. J.F. identified and purified NDF, performed essentially all of the biochemical analyses of the protein, and generated the knockout cell lines. H.I. and B.R. contributed the genome-wide analyses of NDF in human cells. M.A.H. and C.K.G. contributed the analyses of NDF in mouse cells, including the KLA induction experiments. H. I. had a lead role in the analysis of the genome-wide data. F.M. performed the cell cycle studies. G.A.K. contributed to the design and analysis the transcription and footprinting experiments. J.F., G.A.K., and J.T.K. oversaw the overall execution of this work. J.F. and J.T.K. were primarily responsible for writing the manuscript, and received substantial contributions from all other authors in the preparation of figures and the writing and editing of the text.

References

Alhopuro P, Sammalkorpi H, Niittymäki I, Biström M, Raitila A, Saharinen J, Nousiainen K, Lehtonen HJ, Heliövaara E, Puhakka J, et al. 2012. Candidate driver genes in microsatellite-unstable colorectal cancer. *Int J Cancer* **130**: 1558-1566.

Bintu L, Ishibashi T, Dangkulwanich M, Wu YY, Lubkowska L, Kashlev M, Bustamante C. 2012. Nucleosomal elements that control the topography of the barrier to transcription. *Cell* **151**: 738-749.

Fang R, Chen F, Dong Z, Hu D, Barbera AJ, Clark EA, Fang J, Yang Y, Mei P, Rutenberg M et al. 2013. LSD2/KDM1B and its cofactor NPAC/GLYR1 endow a structural and molecular model for regulation of H3K4 demethylation. *Mol Cell* **49**: 558-570.

Fei J, Torigoe SE, Brown CR, Khuong MT, Kassavetis GA, Boeger H, Kadonaga JT. 2015. The prenucleosome, a stable conformational isomer of the nucleosome. *Genes Dev* **29**: 2563-2575.

Flaus A, Luger K, Tan S, Richmond TJ. 1996. Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. *Proc Natl Acad Sci* **93**: 1370-1375.

Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, et al. 2017. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res* **45**: D777-D783.

Formosa T. 2008. FACT and the reorganized nucleosome. *Mol Biosyst.* **4**: 1085-1093.

Fyodorov DV, Kadonaga JT. 2003. Chromatin assembly in vitro with purified recombinant ACF and NAP-1. *Methods Enzymol* **371**: 499-515.

Garcia H, Fleyshman D, Kolesnikova K, Safina A, Commane M, Paszkiewicz G, Omelian A, Morrison C, Gurova K. 2011. Expression of FACT in mammalian tissues suggests its role in maintaining of undifferentiated state of cells. *Oncotarget* **2**: 783-796.

Germond JE, Hirt B, Oudet P, Gross-Bellard M, Chambon P. 1975. Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proc Natl Acad Sci* **72**: 1843-1847.

Godde JS, Wolffe AP. 1995. Disruption of reconstituted nucleosomes. The effect of particle concentration, MgCl2 and KCl concentration, the histone tails, and temperature. *J Biol Chem* **270**: 27399-27402.

Green MH, Buss J, Gariglio P. 1975. Activation of nuclear RNA polymerase by Sarkosyl. *Eur J Biochem* **53**: 217-225.

Hayes JJ, Tullius TD, Wolffe AP. 1990. The structure of DNA in a nucleosome. *Proc Natl Acad Sci* **87**: 7405-7409.

Hodges AJ, Gallegos IJ, Laughery MF, Meas R, Tran L, Wyrick JJ. 2015. Histone sprocket arginine residues are important for gene expression, DNA repair, and cell viability in *Saccharomyces cerevisiae*. *Genetics* **200**: 795-806.

Hsieh FK, Fisher M, Ujvari A, Studitsky VM, Luse DS. 2010. Histone Sin mutations promote nucleosome traversal and histone displacement by RNA polymerase II. *EMBO Rep* **11**: 705-710.

Izban MG, Luse DS. 1991. Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. *Genes Dev* **5**: 683-696.

Kaikkonen MU, Spann NJ, Heinz S, Romanoski CE, Allison KA, Stender JD, Chun HB, Tough DF, Prinjha RK, Benner C, et al. 2013. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol Cell* **51**: 310-325.

Kamakaka RT, Tyree CM, Kadonaga JT. 1991. Accurate and efficient RNA polymerase II transcription with a soluble nuclear fraction derived from *Drosophila* embryos. *Proc Natl Acad Sci* **88**: 1024-1028.

Kaplan CD, Larsson KM, Kornberg RD. 2008, The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Mol Cell* **30**: 547-556.

Kettenberger H, Armache KJ, Cramer P. 2004. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell* **16**: 955-965.

Khuong MT, Fei J, Ishii H, Kadonaga JT. 2015. Prenucleosomes and active chromatin. *Cold Spring Harbor Symp Quant Biol* **80**: 65-72.

Khuong MT, Fei J, Cruz-Becerra G, Kadonaga JT. 2017. A simple and versatile system for the ATP-dependent assembly of chromatin. *J Biol Chem* **292**: 19478-19490.

Kireeva ML, Walter W, Tchernajenko V, Bondarenko V, Kashlev M, Studitsky VM. 2002. Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. *Mol Cell* **9**: 541-552.

Kraus WL, Kadonaga JT. 1998. p300 and estrogen receptor cooperatively activate transcription via differential enhancement of initiation and reinitiation. *Genes Dev* **12**: 331-342.

Kruger W, Peterson CL, Sil A, Coburn C, Arents G, Moudrianakis EN, Herskowitz I. 1995. Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. *Genes Dev* **9**: 2770-2779.

Muthurajan UM, Bao Y, Forsberg LJ, Edayathumangalam RS, Dyer PN, White CL, Luger K. 2004. Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions. *EMBO J* **23**: 260-271.

Mylonas C, Tessarz P. 2018. Transcriptional repression by FACT is linked to regulation of chromatin accessibility at the promoter of ES cells. bioRxiv doi: 10.1101/251611.

Noll M. 1974. Internal structure of the chromatin subunit. *Nucleic Acids Res* **1**: 1573-1578.

Oishi Y, Spann NJ, Link VM, Muse ED, Strid T, Edillor C, Kolar MJ, Matsuzaka T, Hayakawa S, Tao J, et al. 2017. SREBP1 contributes to resolution of pro-inflammatory TLR4 signaling by reprogramming fatty acid metabolism. *Cell Metab* **25**: 412-427.

Pazin MJ, Kamakaka RT, Kadonaga JT. 1994. ATP-dependent nucleosome reconfiguration and transcriptional activation from preassembled chromatin templates. *Science* **266**: 2007-2011.

Qin S, Min J. 2014. Structure and function of the nucleosome-binding PWWP domain. *Trends Biochem Sci* **39**: 536-547.

Raetz CR, Garrett TA, Reynolds CM, Shaw WA, Moore JD, Smith DC, Jr., Ribeiro AA, Murphy RC, Ulevitch RJ, Fearns C, et al. 2006. Kdo2-Lipid A of *Escherichia coli*, a defined endotoxin that activates macrophages via TLR-4. *J Lipid Res* **47**: 1097-1111.

Reinberg D, Sims RJ, 3rd. 2006. deFACTo nucleosome dynamics. *J Biol Chem* **281**: 23297-23301.

Rhodes D. 1985. Structural analysis of a triple complex between the histone octamer, a *Xenopus* gene for 5S RNA and transcription factor IIIA. *EMBO J* **4**: 3473-3482.

Samkurashvili I, Luse DS. 1996. Translocation and transcriptional arrest during transcript elongation by RNA polymerase II. *J Biol Chem* **271**: 23495-23505.

Scheer U. 1978. Changes of nucleosome frequency in nucleolar and non-nucleolar chromatin as a function of transcription: an electron microscopic study. *Cell* **13**: 535-549.

Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, et al. 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**: 84-87.

Stein A.1989. Reconstitution of chromatin from purified components. *Methods Enzymol* **170**: 585-603.

Sternberg PW, Stern MJ, Clark I, Herskowitz I. 1987. Activation of the yeast HO gene by release from multiple negative controls. *Cell* **48**: 567-577.

Torigoe SE, Patel A, Khuong MT, Bowman GD, Kadonaga JT. 2013. ATP-dependent chromatin assembly is functionally distinct from chromatin remodeling. *eLife* **2**: e00863.

Torigoe SE, Urwin DL, Ishii H, Smith DE, Kadonaga JT. 2011. Identification of a rapidly formed non-nucleosomal histone-DNA intermediate that is converted into chromatin by ACF. *Mol Cell* **43**: 638-648.

True JD, Muldoon JJ, Carver MN, Poorey K, Shetty SJ, Bekiranov S, Auble DT. 2016. The modifier of transcription 1 (Mot1) ATPase and Spt16 histone chaperone co-regulate

transcription through preinitiation complex assembly and nucleosome organization. *J Biol Chem* **291**: 15307-15319.

Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. 2015. Tissue-based map of the human proteome. *Science* **347**: 1260419.

Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, Benfeitas R, Arif M, Liu Z, Edfors F, et al. 2017. A pathology atlas of the human cancer transcriptome. *Science* **357**: eaan2507.

Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, et al. 2010. Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* **142**: 967-980.

Wang CI, Alekseyenko AA, LeRoy G, Elia AE, Gorchakov AA, Britton LM, Elledge SJ, Kharchenko PV, Garcia BA, Kuroda MI. 2013. Chromatin proteins captured by ChIP-mass spectrometry are linked to dosage compensation in *Drosophila*. *Nat Struct Mol Biol* **20**: 202-209.

Weber CM, Ramachandran S, Henikoff S. 2014. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell* **53**: 819-830.

Wechser MA, Kladde MP, Alfieri JA, Peterson CL. 1997. Effects of Sin– versions of histone H4 on yeast chromatin structure and function. *EMBO J* **16**: 2086-2095.

Figure Legends

Figure 1. Purification of *Drosophila* NDF (dNDF). (**A**) Identification of NDF as an activity in a *Drosophila* nuclear extract that facilitates nucleosomal H3K56 acetylation by purified p300. The absence (–) or presence of lower (+) or higher (++) concentrations of the indicated components is shown. Chromatin was reconstituted with plasmid DNA and purified core histones by the salt dialysis method. H3K56 acetylation was monitored by western blot analysis. Both bands in the doublets are derived from histone H3. The heat-treated extract was subjected to incubation at 75˚C for 5 min. (**B**) Scheme for the purification of dNDF. (**C**) Purification of dNDF by Source 15Q chromatography. The Source 15Q column fractions were analyzed by 10% polyacrylamide-SDS gel electrophoresis and staining by Coomassie Brilliant Blue R-250 (upper panel), the H3K56 acetylation activity assay, as in A (middle panel), and western blot with antibodies that were raised against recombinant dNDF.

Figure 2. Destabilization of nucleosomes by purified recombinant dNDF and human NDF (hNDF). (**A**) Schematic diagram of hNDF and dNDF. The drawing is roughly to scale. Specific features are described in the text. (**B**) Purification of bacterially synthesized, N-terminally His6-tagged dNDF and hNDF. (**C**) Recombinant dNDF and hNDF facilitate the acetylation of nucleosomal H3K56 by p300. This assay was performed as in Fig. 1A. (**D**) Loss of DNA supercoiling in chromatin reveals destabilization of nucleosomes by hNDF. Chromatin was reconstituted with plasmid DNA and purified core histones by the salt dialysis method and then incubated (at 0.3 μM nucleosome concentration) with purified recombinant hNDF and topoisomerase I (Topo I), as indicated. Samples were deproteinized and subjected to 0.8% agarose gel electrophoresis in the absence or presence of 10 ng/mL chloroquine. DNA was visualized by staining with ethidium bromide. As a control, reactions were carried out in parallel with plasmid DNA.

Figure 3. NDF is enriched in transcribed regions of protein-coding genes with a preference for longer genes relative to shorter genes. (**A**) Generation of human cell lines lacking NDF. By using CRISPR-Cas9 methodology, four knockout cell lines were generated in HeLa cells by disruption of either exon 1 (KO-1.1 and KO-1.2) or exon 2 (KO-2.1 and KO-2.2) of the gene encoding NDF. The loss of NDF was shown by western blot analysis. The largest subunit of RNA polymerase II (Pol II) was used as a loading control. (**B**) The loss of NDF reduces the growth rate of cells. On Day 0, 1 x 10^4 cells were seeded in triplicate in 6-cm culture dishes, and attached cells were counted on the indicated days. (**C**) The loss of NDF increases the fraction of cells in G1, as seen by FACS analysis. (**D**) NDF ChIP-seq data at the *PCBP2* gene. (**E**) NDF co-localizes with H3K36me3 over transcribed regions. Metagene analysis of NDF and H3K36me3 occupancy at 20,232 protein-coding genes. (**F**) NDF preferentially localizes to longer protein-coding genes relative to shorter protein-coding genes. By ChIP-seq analysis, approximately 4500 protein-coding genes were found to be enriched in NDF in HeLa cells. This figure shows overlaid bar graphs of the size distributions of all protein-coding genes (blue) versus NDF-enriched protein-coding genes (pink). The purple color indicates regions of overlap of the two bar graphs.

Figure 4. NDF contributes to the expression of many genes. Four clusters of genes were generated by k-means clustering with three variables: $\log_{10}$[RNA-seq signal of genes in RPKM], $\log_{10}$[(NDF ChIP-seq signal over gene body)/(Input signal over gene body)], $\log_{10}$[(H3K36me3 ChIP-seq signal over gene body)/(Input signal over gene body)]. (**A**) NDF is enriched in active gene cluster 1, but not in active gene cluster 2. RNA-seq signal (RPKM) versus NDF ChIP/Input signal for gene clusters 1 to 4. (**B**) Heat maps of NDF and H3K36me3 occupancies and gene expression in wild-type HeLa cells. Genes within each cluster were ranked by the (NDF ChIP-seq signal)/(Input signal) ratio. (**C**) RNA-seq analysis of wild-type versus hNDF KO-1.2 cells. The plot shows $\log_{10}$[(RPKM KO)/(RPKM WT)] versus $\log_{10}$[RPKM WT] for genes in clusters 1 to 4. Genes with RPKM = 0, such as those in cluster 4, were not included in this analysis. The

upper and lower blue lines correspond to (RPKM KO)/(RPKM WT) ratios of 2 and 0.5, respectively.

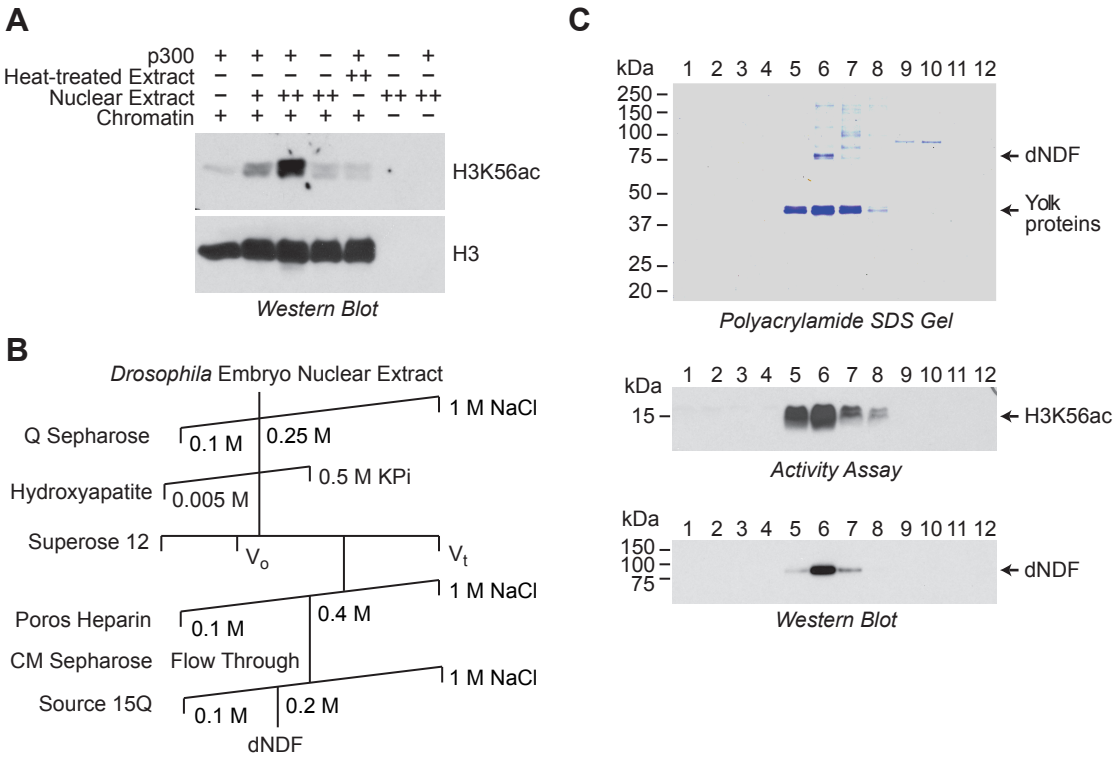
Figure 5. NDF is recruited to the transcribed regions of thousands of genes in mouse primary bone marrow-derived macrophages (BMDM) upon transcriptional induction by KLA. (**A**) Genome-browser view of H3K36me3 ChIP-seq, mouse NDF (mNDF) ChIP-seq, and RNA-seq with or without induction by KLA for 6 h. The region in the vicinity of the *Tmem67* gene is indicated by the box. (**B**) Recruitment of mNDF to gene bodies upon activation of BMDM by KLA. Metagene analysis of NDF and H3K36me3 ChIP-seq signals at 2249 genes in which NDF occupancy increases by at least two-fold upon 6 h induction with KLA. (**C**) The KLA-induced increase in NDF occupancy correlates with an increase in H3K36me3. The scatter plot (35,933 total genes) shows 2249 genes in which NDF occupancy increases (up arrow) by at least two-fold upon KLA induction for 6 h and 4097 genes in which NDF occupancy decreases (down arrow) by at least two-fold upon KLA induction for 6 h. (**D**) The KLA-induced increase in NDF occupancy generally correlates with an increase in transcript levels. The genes at which NDF levels change upon KLA induction are as in C.

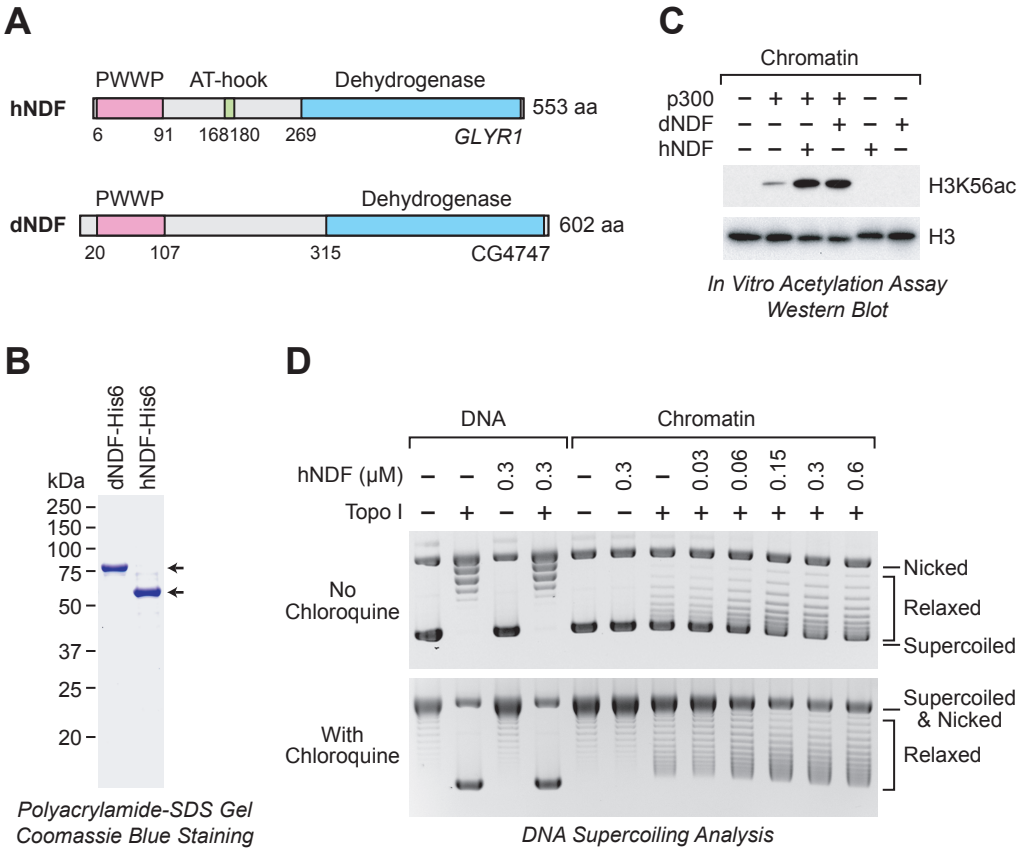
Figure 6. NDF facilitates transcription elongation through a nucleosome. (**A**) Schematic diagram of Pol II transcription through a positioned nucleosome. A transcriptionally-engaged elongation complex (Pol II EC) was assembled with purified yeast RNA polymerase II (Pol II) and a 5'-end-labeled 10 nt RNA primer. This complex was attached to streptavidin beads, and then ligated to a downstream mononucleosome positioned on the 5S rDNA sequence from *Xenopus borealis*. Transcription elongation was initiated by the addition of ribonucleoside 5'-triphosphates (rNTPs). Where indicated, NDF was added after the ligation step but before the addition of the rNTPs. In parallel experiments, naked DNA (same 5S rDNA sequence) was ligated downstream of the elongation complex instead of a mononucleosome. The distance (~15 nt) from the leading edge of Pol II to the 3' end of the transcript is also shown. (**B**) NDF reduces the inhibition of Pol II

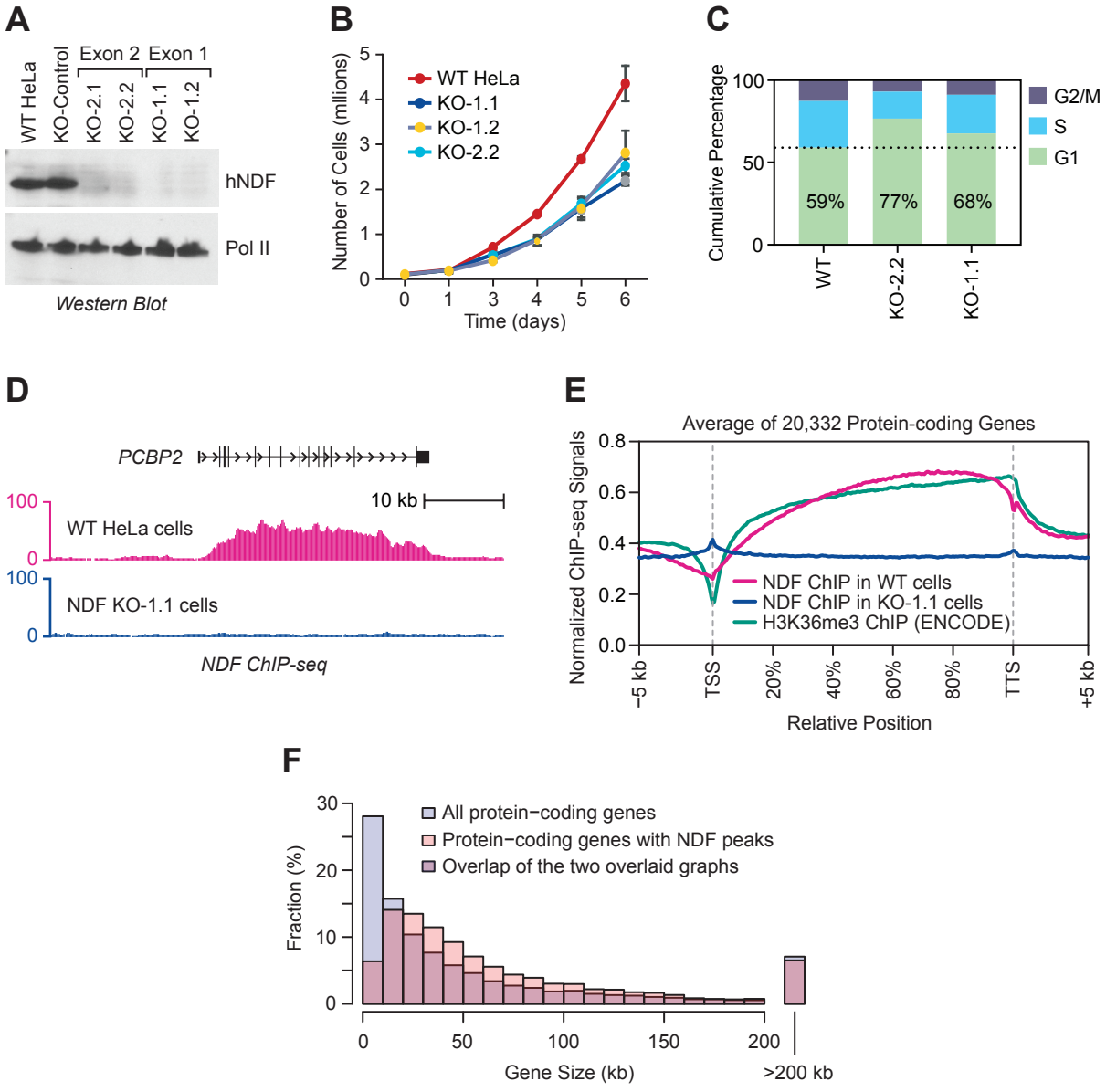
elongation by a nucleosome. Transcription elongation reactions were performed as described in A in the presence or absence of purified hNDF for the indicated times. Experiments with nucleosomal templates included a 75-fold molar excess (0.3 μM) of free unligated mononucleosomes, which stabilize the low concentration of immobilized nucleosomes. hNDF was included, where indicated, at a concentration of 1.5 μM. The reaction products were resolved by 8% polyacrylamide-urea gel electrophoresis. The diagram shows the locations of the positioned nucleosome, the sites of Pol II pausing, the run-off product, the ligation junction, and the 10 nt primer RNA. The sizes of the RNA species were estimated by comparison to a 25 nt radiolabeled DNA ladder.

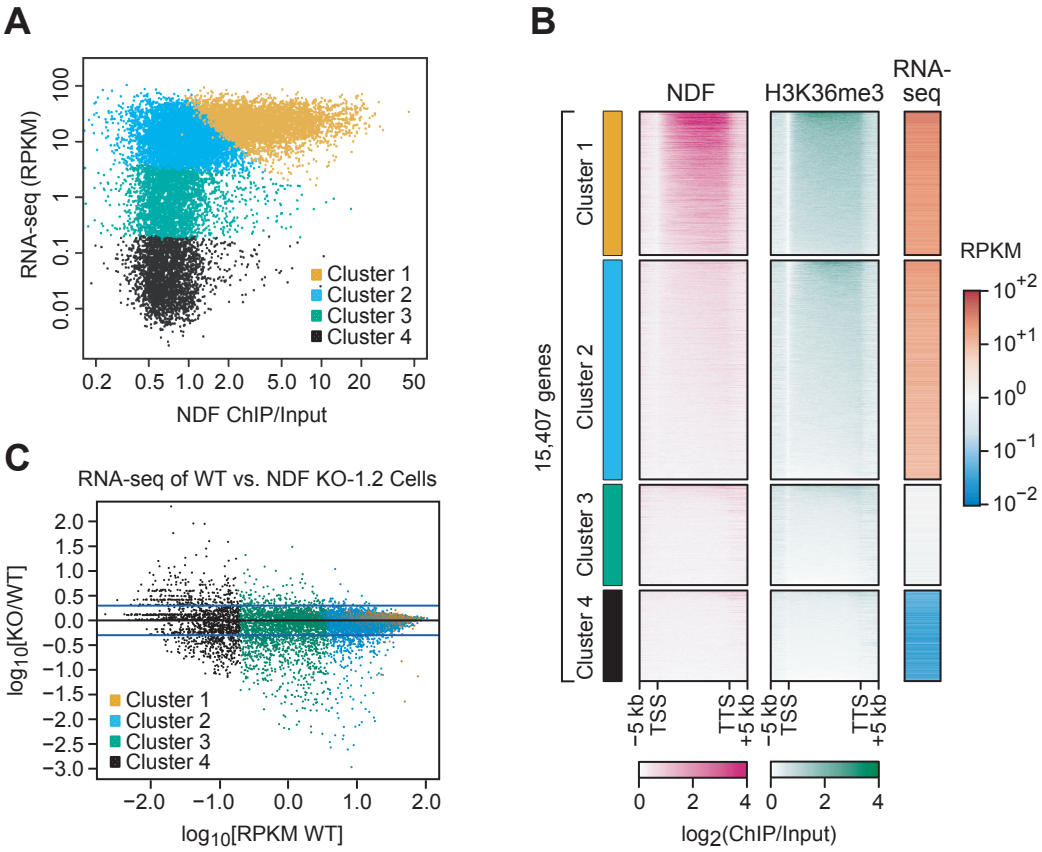
Figure 7. NDF interacts with nucleosomes near the dyad. (**A**) NDF binds to mononucleosomes as well as to free DNA. Gel mobility shift assays were performed with the indicated concentrations of dNDF and 50 nM of 5S rDNA (147 bp) as either naked DNA or mononucleosomes. The resulting species were separated by 4% nondenaturing polyacrylamide gel electrophoresis and visualized by staining with ethidium bromide. The positions of the shifted complexes are indicated. (**B**) DNase I footprinting analysis reveals an interaction of NDF near the nucleosome dyad. DNase I footprinting reactions were carried out with the forward strand of the 5S rDNA sequence (183 bp) as either naked DNA or mononucleosomes. Purified hNDF (73 nM or 150 nM concentration) was included where indicated. The Maxam-Gilbert A sequencing ladder was used to map the locations of the DNase I digestion products. The pink dot shows an NDF-protected nucleotide at +2 relative to the nucleosomal dyad at position 0. The blue squares designate bands that are enhanced in the presence of hNDF. The forward strand has the same sense as the template strand of the 5S rRNA gene. (**C**) Comparison of the effects of NDF and the H4R45A Sin mutation on Pol II elongation through a nucleosome. Transcription elongation experiments were performed as in Fig. 6 with either wild-type histone H4 or mutant histone H4R45A. The red arrow denotes the −20 Pol II pause site that is lost upon mutation of H4R45. The black dot shows the location of a pause site downstream of the −20

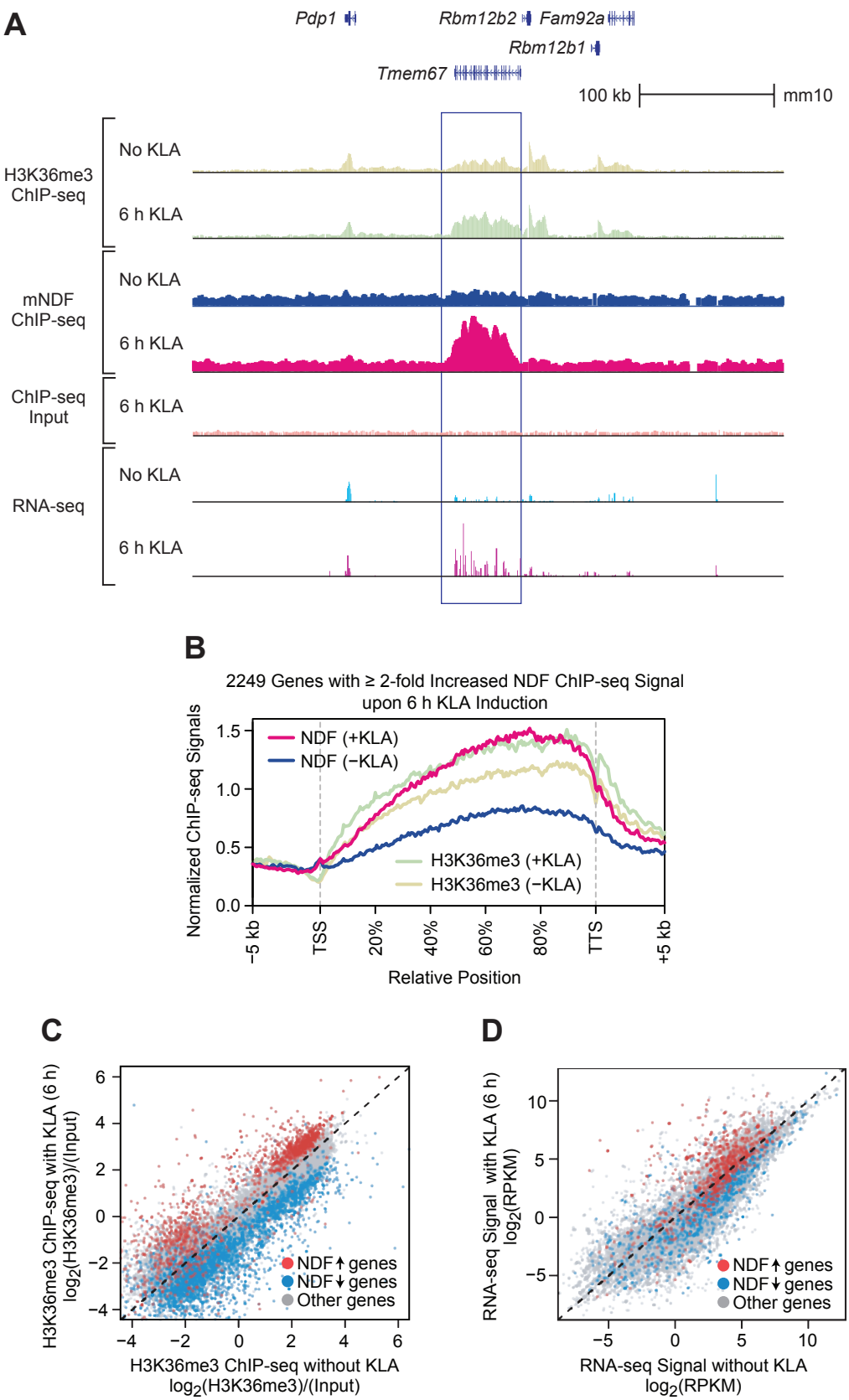
position that accumulates with H4R45A nucleosomes. The asterisk corresponds to a pause site that likely derives from the presence of some downstream naked DNA, as in the left panel of Fig. 6B. (**D**) Model for the function of NDF.

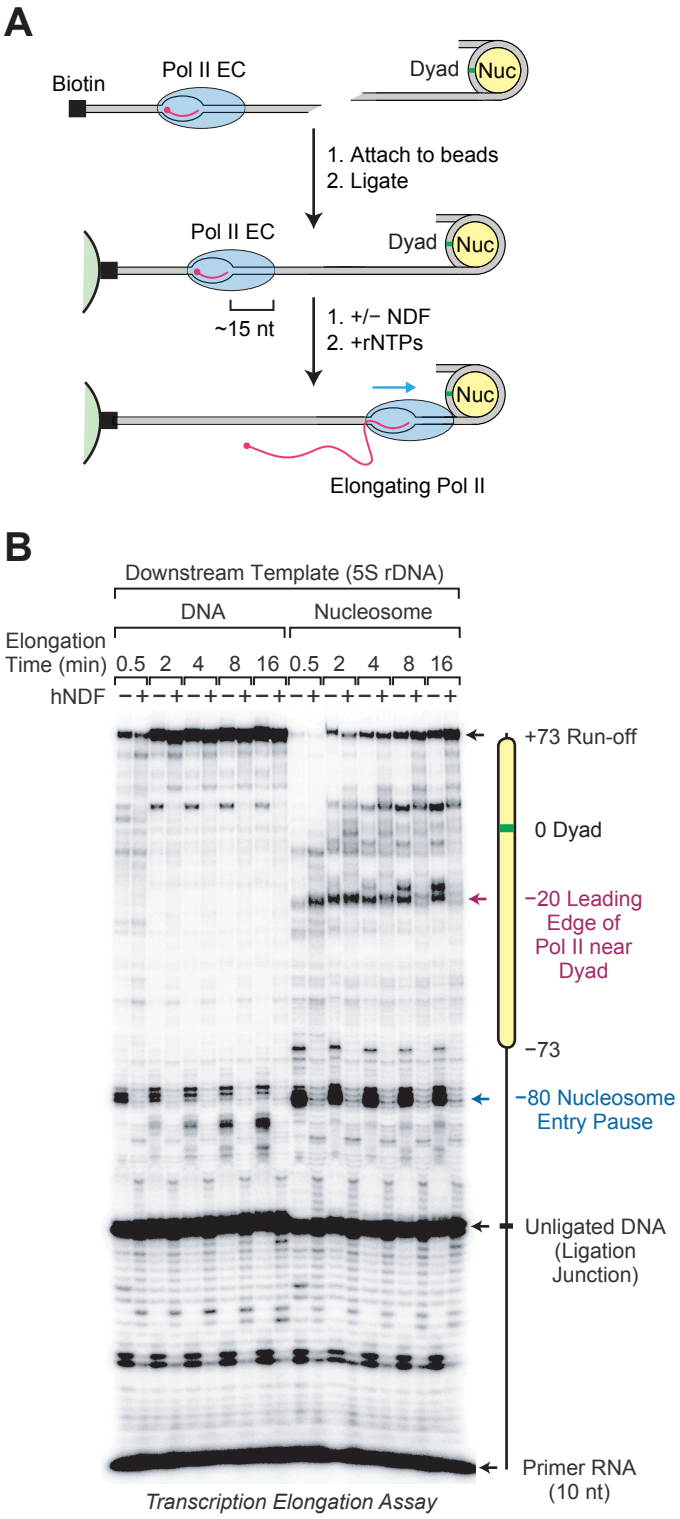

A
p300 + + + − + − +
Heat-treated Extract − − − − ++ − −
Nuclear Extract − + ++ ++ − ++ ++
Chromatin + + + + + − −
H3K56ac
H3
Western Blot
B
Drosophila Embryo Nuclear Extract
Q Sepharose
0.1 M
0.25 M
1 M NaCl
Hydroxyapatite
0.005 M
0.5 M KPi
Superose 12
Vo
Vt
Poros Heparin
0.1 M
0.4 M
1 M NaCl
CM Sepharose
Flow Through
Source 15Q
0.1 M
0.2 M
1 M NaCl
dNDF
C
kDa 1 2 3 4 5 6 7 8 9 10 11 12
250
150
100
75
50
37
25
20
dNDF
Yolk proteins
Polyacrylamide SDS Gel
kDa 1 2 3 4 5 6 7 8 9 10 11 12
15
H3K56ac
Activity Assay
kDa 1 2 3 4 5 6 7 8 9 10 11 12
150
100
75
dNDF
Western Blot

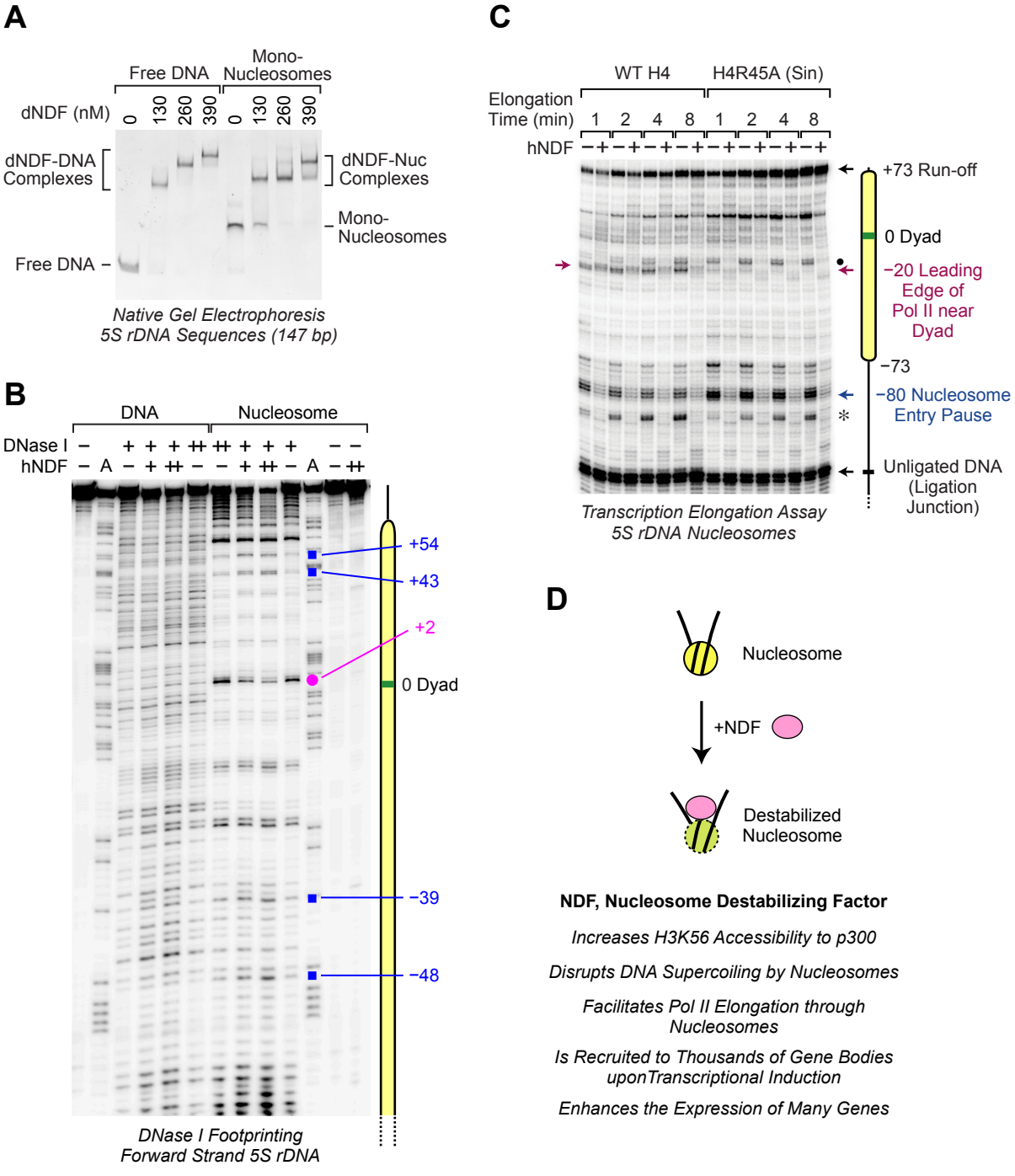

A
PWWP
AT-hook
Dehydrogenase
hNDF
553 aa
6
91
168
180
269
GLYR1
PWWP
Dehydrogenase
dNDF
602 aa
20
107
315
CG4747
B
dNDF-His6
hNDF-His6
kDa
250
150
100
75
50
37
25
20
Polyacrylamide-SDS Gel
Coomassie Blue Staining
C
Chromatin
p300
dNDF
hNDF
H3K56ac
H3
In Vitro Acetylation Assay
Western Blot
D
DNA
Chromatin
hNDF (μM)
Topo I
No Chloroquine
With Chloroquine
Nicked
Relaxed
Supercoiled
Supercoiled & Nicked
Relaxed
DNA Supercoiling Analysis


A
WT HeLa
KO-Control
Exon 2
Exon 1
KO-2.1
KO-2.2
KO-1.1
KO-1.2
hNDF
Pol II
Western Blot
B
Number of Cells (millions)
WT HeLa
KO-1.1
KO-1.2
KO-2.2
Time (days)
C
Cumulative Percentage
G2/M
S
G1
59%
77%
68%
WT
KO-2.2
KO-1.1
D
PCBP2
10 kb
WT HeLa cells
NDF KO-1.1 cells
NDF ChIP-seq
E
Average of 20,332 Protein-coding Genes
Normalized ChIP-seq Signals
NDF ChIP in WT cells
NDF ChIP in KO-1.1 cells
H3K36me3 ChIP (ENCODE)
-5 kb
TSS
20%
40%
60%
80%
TTS
+5 kb
Relative Position
F
All protein−coding genes
Protein−coding genes with NDF peaks
Overlap of the two overlaid graphs
Fraction (%)
Gene Size (kb)
>200 kb


A
RNA-seq (RPKM)
100
10
1
0.1
0.01
0.2 0.5 1.0 2.0 5.0 10 20 50
NDF ChIP/Input
Cluster 1
Cluster 2
Cluster 3
Cluster 4
B
NDF
H3K36me3
RNA-seq
15,407 genes
Cluster 1
Cluster 2
Cluster 3
Cluster 4
RPKM
10^{+2}
10^{+1}
10^{0}
10^{-1}
10^{-2}
−5 kb TSS
TTS +5 kb
−5 kb TSS
TTS +5 kb
0 2 4 0 2 4
log2(ChIP/Input)
C
RNA-seq of WT vs. NDF KO-1.2 Cells
log10[KO/WT]
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
Cluster 1
Cluster 2
Cluster 3
Cluster 4
−2.0 −1.0 0.0 1.0 2.0
log10[RPKM WT]


A
Pdp1
Rbm12b2
Fam92a
Rbm12b1
Tmem67
100 kb
mm10
H3K36me3 ChIP-seq
No KLA
6 h KLA
mNDF ChIP-seq
No KLA
6 h KLA
ChIP-seq Input
6 h KLA
RNA-seq
No KLA
6 h KLA
B
2249 Genes with ≥ 2-fold Increased NDF ChIP-seq Signal upon 6 h KLA Induction
Normalized ChIP-seq Signals
NDF (+KLA)
NDF (−KLA)
H3K36me3 (+KLA)
H3K36me3 (−KLA)
−5 kb
TSS
20%
40%
60%
80%
TTS
+5 kb
Relative Position
C
H3K36me3 ChIP-seq with KLA (6 h) log2(H3K36me3)/(Input)
H3K36me3 ChIP-seq without KLA log2(H3K36me3)/(Input)
NDF↑ genes
NDF↓ genes
Other genes
D
RNA-seq Signal with KLA (6 h) log2(RPKM)
RNA-seq Signal without KLA log2(RPKM)
NDF↑ genes
NDF↓ genes
Other genes

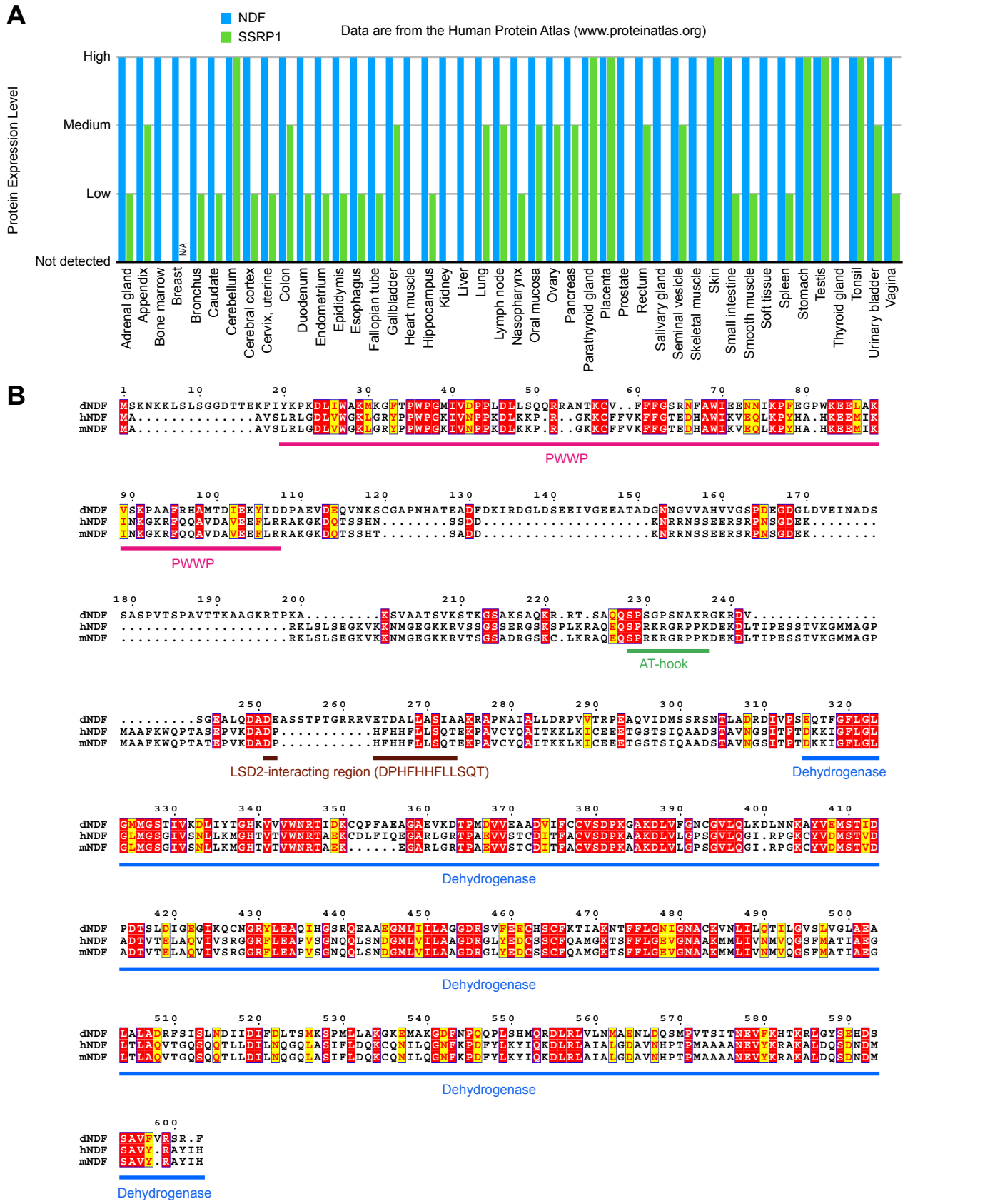

A
Biotin
Pol II EC
Dyad
Nuc
1. Attach to beads
2. Ligate
Pol II EC
Dyad
Nuc
~15 nt
1. +/− NDF
2. +rNTPs
Nuc
Elongating Pol II
B
Downstream Template (5S rDNA)
DNA
Nucleosome
Elongation
Time (min) 0.5 2 4 8 16 0.5 2 4 8 16
hNDF − + − + − + − + − + − + − + − +
+73 Run-off
0 Dyad
−20 Leading Edge of Pol II near Dyad
−73
−80 Nucleosome Entry Pause
Unligated DNA (Ligation Junction)
Primer RNA (10 nt)
Transcription Elongation Assay


A
Free DNA
Mono-Nucleosomes
dNDF (nM)
0 130 260 390 0 130 260 390
dNDF-DNA Complexes
dNDF-Nuc Complexes
Mono-Nucleosomes
Free DNA
Native Gel Electrophoresis
5S rDNA Sequences (147 bp)
B
DNA
Nucleosome
DNase I
hNDF
+54
+43
+2
0 Dyad
−39
−48
DNase I Footprinting
Forward Strand 5S rDNA
C
WT H4
H4R45A (Sin)
Elongation Time (min)
1 2 4 8 1 2 4 8
hNDF
+73 Run-off
0 Dyad
−20 Leading Edge of Pol II near Dyad
−73
−80 Nucleosome Entry Pause
Unligated DNA (Ligation Junction)
Transcription Elongation Assay
5S rDNA Nucleosomes
D
Nucleosome
+NDF
Destabilized Nucleosome
NDF, Nucleosome Destabilizing Factor
Increases H3K56 Accessibility to p300
Disrupts DNA Supercoiling by Nucleosomes
Facilitates Pol II Elongation through Nucleosomes
Is Recruited to Thousands of Gene Bodies uponTranscriptional Induction
Enhances the Expression of Many Genes

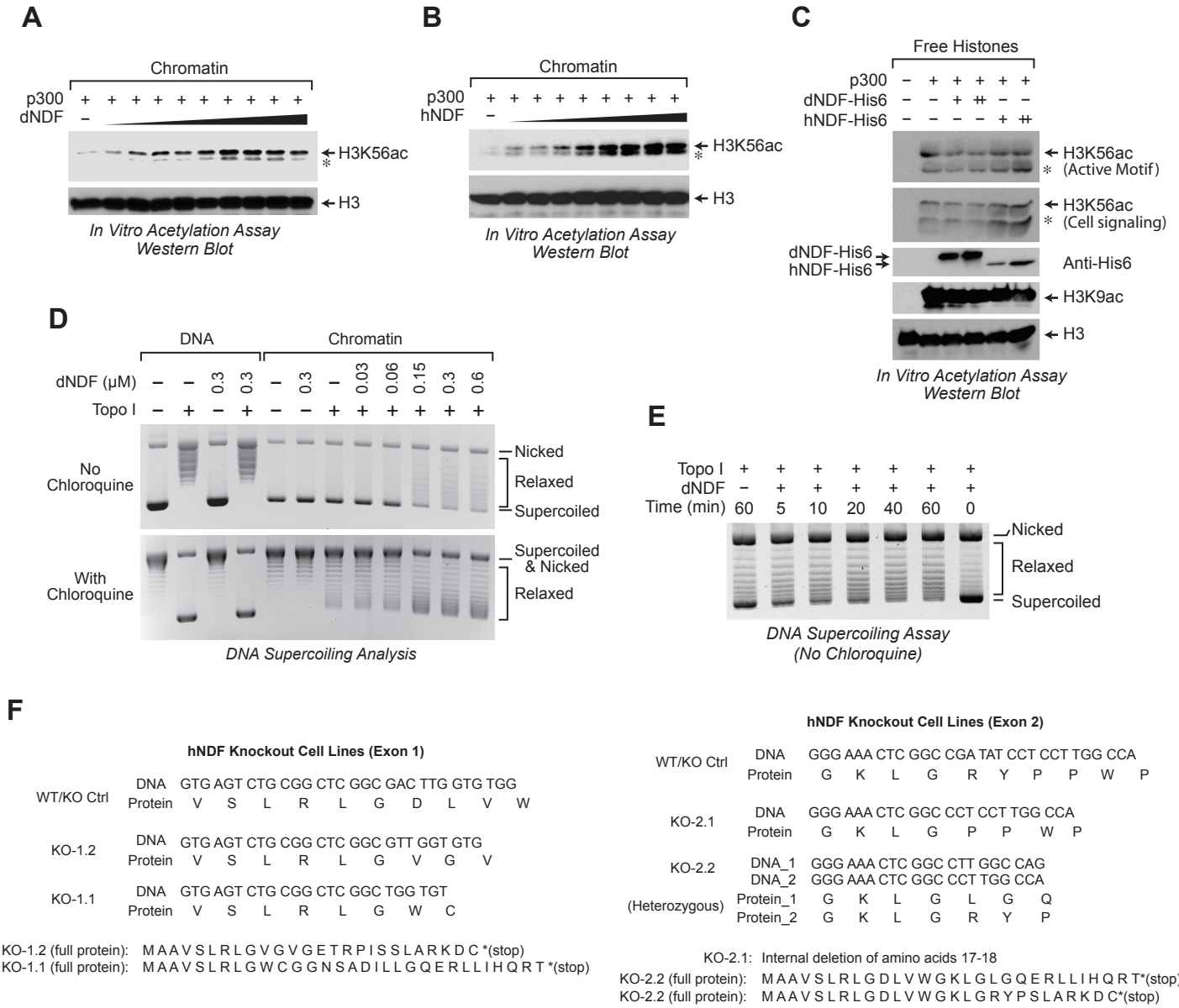
SUPPLEMENTAL MATERIAL

NDF, a Nucleosome Destabilizing Factor that Facilitates Transcription through Nucleosomes

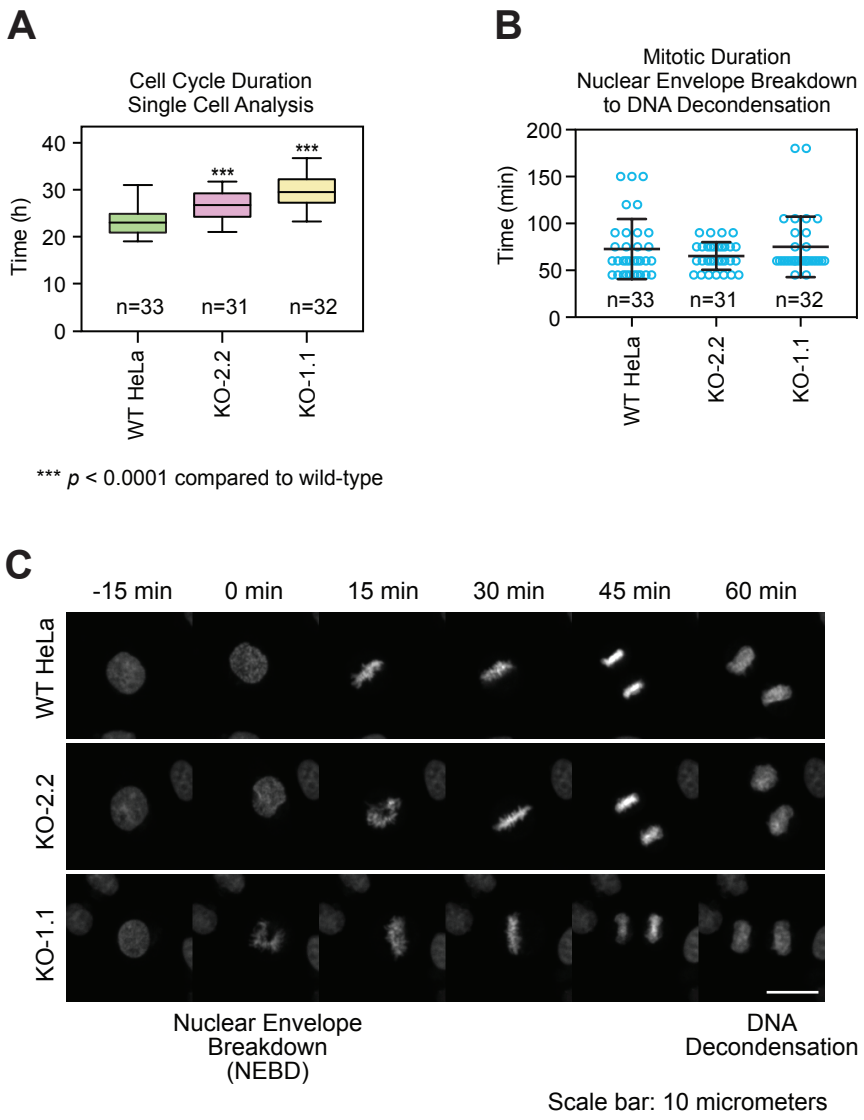
Jia Fei, Haruhiko Ishii, Marten A. Hoeksema, Franz Meitinger, George A. Kassavetis, Christopher K. Glass, Bing Ren, and James T. Kadonaga



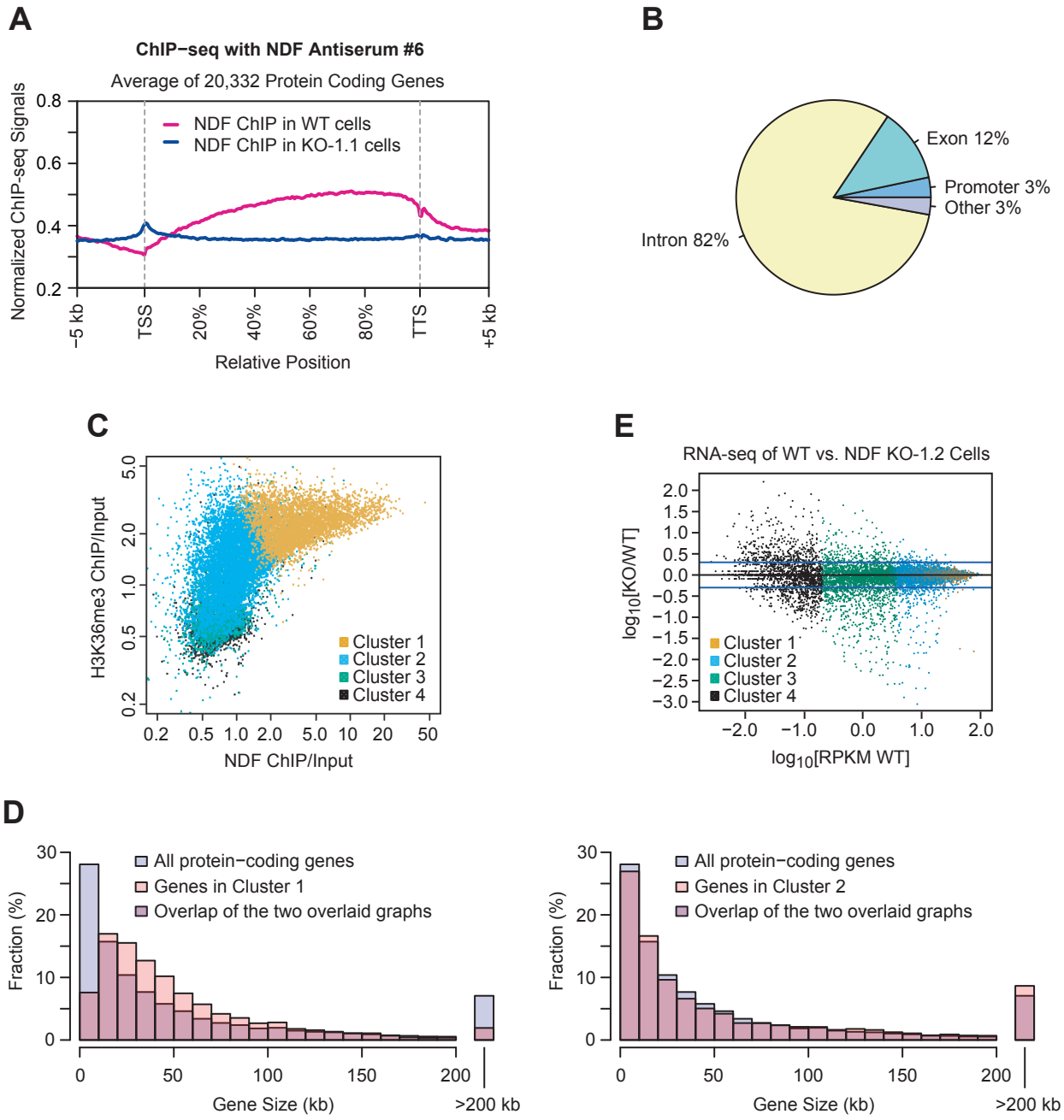
Supplemental Figure S1. Sequence and protein expression of NDF. (**A**) Protein expression levels of hNDF (GLYR1) and SSRP1 in 44 normal human adult tissues. DAB (3,3'-diaminobenzidine)-labeled antibodies were used for the immunohistochemistry experiments. The data were from www.proteinatlas.org. (**B**) Sequence alignment of *Drosophila*, human, and mouse NDF. These are the three versions of NDF that were examined in this study. The full-length sequences of dNDF (CG4747), hNDF, and mNDF (Glyr1) were aligned with T-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:regular). The resulting ClustalW file was converted into a PostScript file with ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/). Amino acid residues that are identical in the three proteins are highlighted in red, and amino acid residues with similar properties in all three proteins are highlighted in yellow. The locations of the PWWP domain, AT-hook motif, LSD2-interacting region, and dehydrogenase domain are shown. All three proteins contain the PWWP and dehydrogenase domains. The AT-hook and LSD2-interaction region are present in hNDF and mNDF, but not in dNDF. mNDF differs from hNDF by 14 amino acid residues.



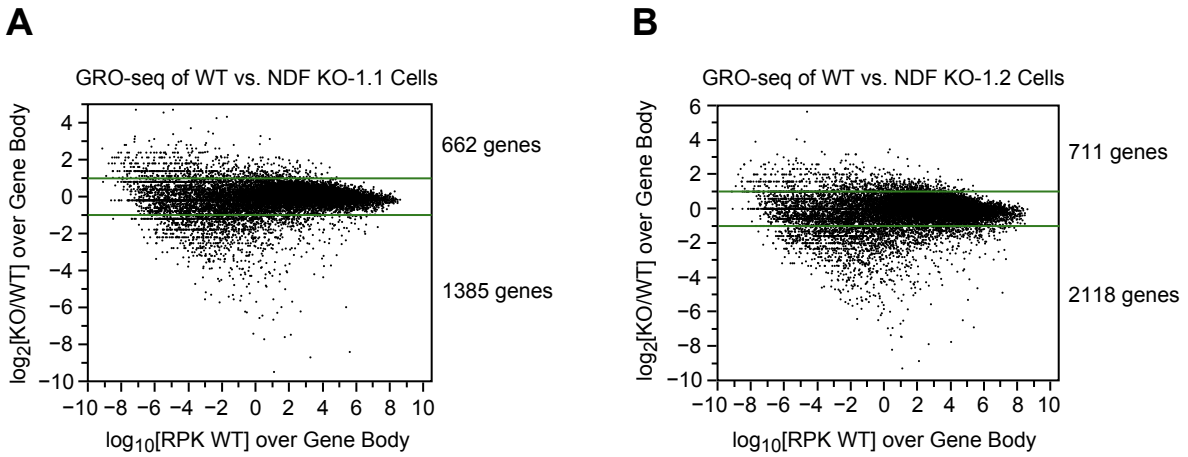
Supplemental Figure S2. Characterization of the activities of purified recombinant human and *Drosophila* NDF and generation of NDF knockout cell lines. (**A**) Analysis of the effects of different concentrations of recombinant dNDF in the acetylation of nucleosomal histone H3K56 by p300. Reactions were performed as in Fig. 2B of the main text in the presence of 0, 0.3, 0.6, 0.8, 1.1, 1.3, 1.4, 1.5, 1.7 or 1.8 μM dNDF with a nucleosome concentration of 1.4 μM. The asterisk denotes a nonspecific band. (**B**) Analysis of the effects of different concentrations of recombinant hNDF in the acetylation of nucleosomal histone H3K56 by p300. Reactions were performed as in Fig. 2B of the main text in the presence of 0, 0.3, 0.6, 0.8, 1.1, 1.3, 1.4, 1.5 or 1.7 μM hNDF. The asterisk denotes a nonspecific band. (**C**) Recombinant NDF does not stimulate the acetylation of H3K56 with free histones. Reactions were performed as in Fig. 2B of the main text, except that free histones were used instead of chromatin, with 1.4 μM core histone octamers, 0.7 μM (+) and 1.4 μM (++) dNDF, and 0.35 μM (+) and 0.7 μM (++) hNDF. (**D**) Destabilization of nucleosomes by dNDF. DNA supercoiling analysis of partial nucleosome disassembly by dNDF. Reactions with dNDF were carried out as with hNDF in Fig. 2D of the main text. (**E**) Time course of dNDF-mediated destabilization of chromatin. Reactions were performed as in **d** for the indicated times. (**F**) Genomic sequences of the human NDF (*GLYR1*) gene in wild-type and knockout (KO) cell lines. By using CRISPR-Cas9 methodology, four knockout cell lines were generated in HeLa cells by disruption of either exon 1 (KO-1.1 and KO-1.2) or exon 2 (KO-2.1 and KO-2.2) of the gene encoding NDF (*GLYR1*). For each cell line, the genomic DNA was purified, amplified by PCR, and sequenced. The proteins encoded by each of the cell lines are also indicated.



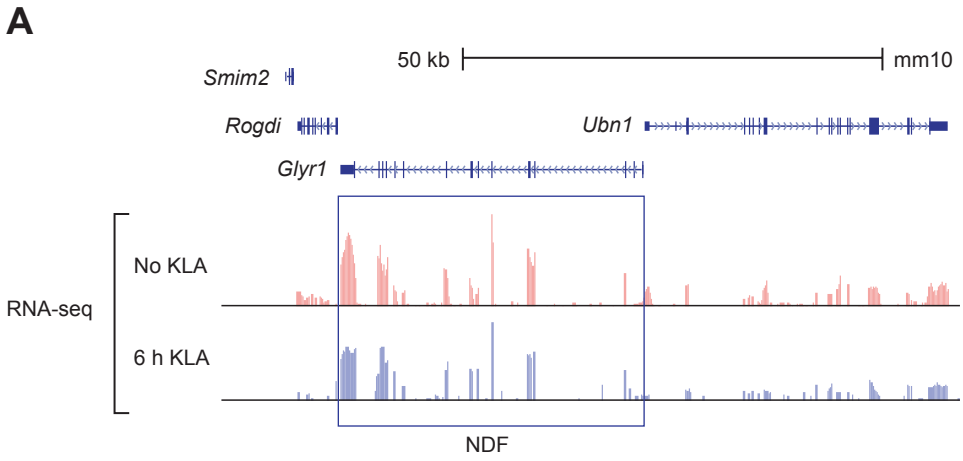
Supplemental Figure S3. Live-cell imaging analysis of wild-type and NDF knockout (KO) HeLa cells. (**A**) Single-cell cell cycle duration analysis. Both knockout cells exhibit an increase in the duration of the cell cycle relative to that in wild-type cells. (**B**) The duration of mitosis in NDF knockout cells is similar to that seen in wild-type cells. (**C**) Cell cycle imaging analysis reveals no apparent mitotic defects in NDF knockout cells.

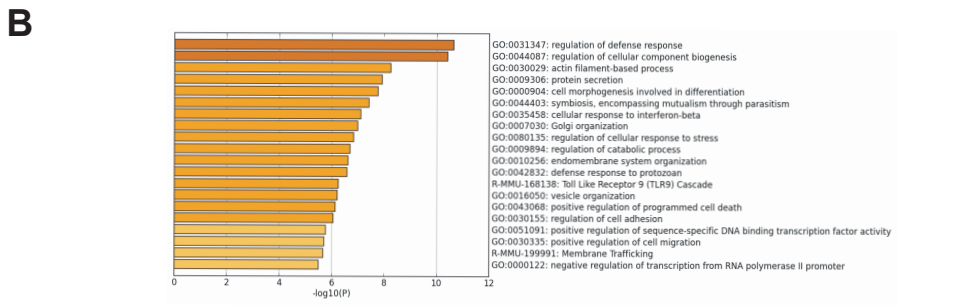


Supplemental Figure S4. Human NDF (hNDF) is enriched in transcribed regions of protein-coding genes with a preference for longer genes relative to shorter genes. (**A**) NDF is enriched in transcribed regions of protein-coding genes. Metagene analysis of NDF occupancy at 20,232 protein-coding genes with wild-type or NDF knockout (KO-1.1) HeLa cells. This experiment is similar to that shown in Fig. 3E of the main text, except that the data in this figure were generated with hNDF antiserum #6, whereas the data in main text Fig. 3E were generated with hNDF antiserum #7. (**B**) Genomic distribution of NDF ChIP-seq peaks in HeLa cells. Promoters were defined as regions that are +/– 200 bp from the transcription start site (TSS). (**C**) H3K36me3 is enriched in both active gene clusters 1 and 2. H3K36me3 ChIP/Input versus NDF ChIP/Input for gene clusters 1 to 4. The gene clusters were generated as described in Fig. 4 of the main text. (**D**) Active genes containing NDF (Cluster 1; left panel) are generally longer than active genes lacking NDF (Cluster 2; right panel). The gene clusters were generated as described in Fig. 4 of the main text. (**E**) RNA-seq analysis of wild-type versus hNDF KO-1.1 cells. The plot shows $\log_{10}$[(RPKM KO)/(RPKM WT)] versus $\log_{10}$[RPKM WT] for genes in clusters 1 to 4. Genes with RPKM = 0, such as those in cluster 4, were not incuded in this analysis. The upper and lower blue lines correspond to (RPKM KO)/(RPKM WT) ratios of 2 and 0.5, respectively.



Supplemental Figure S5. GRO-seq analyses of wild-type and hNDF knockout (KO) cells reveal changes in Pol II dynamics upon loss of NDF. (**A**) Scatter plot of $\log_2$[(RPK KO)/(RPK WT)] versus $\log_{10}$[RPK WT] for genes in wild-type versus KO-1.1 cells. RPK is defined to be the reads per kilobase of gene body length (from the TSS to the TTS) per 10 million reads. The upper and lower green lines correspond to (KO/WT) ratios of 2 and 0.5, respectively. Out of 21,510 genes analyzed, there are 662 genes with at least a two-fold increase in the GRO-seq signal over the gene body and 1385 genes with at least a two-fold decrease in the GRO-seq signal over the gene body. (**B**) Scatter plot of $\log_2$[(RPK KO)/(RPK WT)] versus $\log_{10}$[RPK WT] for genes in wild-type versus KO-1.2 cells. Out of 21,287 genes analyzed, there are 711 genes with at least a two-fold increase in the GRO-seq signal over the gene body and 2118 genes with at least a two-fold decrease in the GRO-seq signal over the gene body.

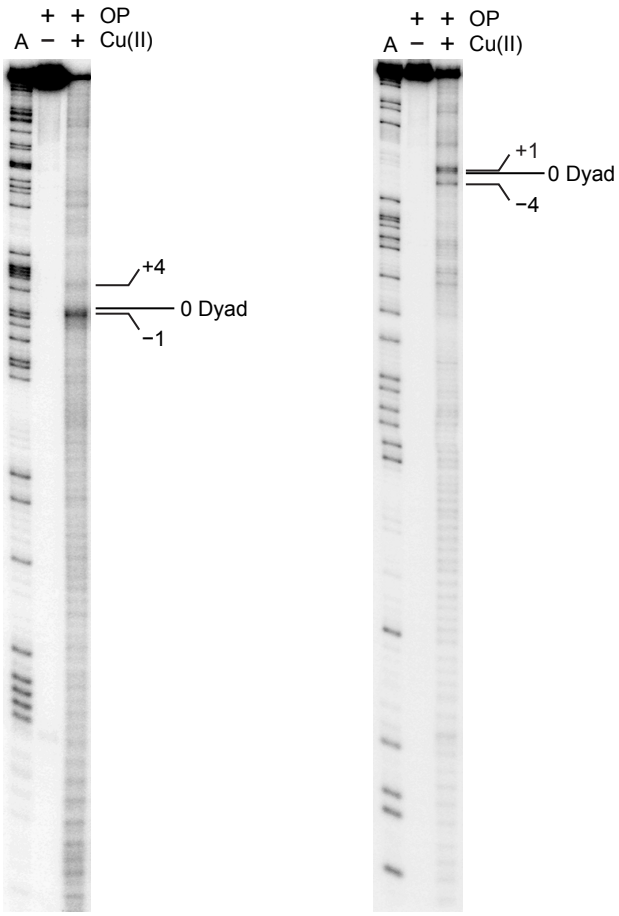




GO analysis of KLA-induced NDF peaks (6 h)

Supplemental Figure S6. Induction of mouse primary bone marrow derived macrophages (BMDM) by KLA. (**A**) Induction BMDM by KLA does not affect the expression of the *Glyr1* gene, which encodes mNDF. Genome-browser view of RNA-seq signal at the *Glyr1* gene with or without induction by KLA for 6 h. Transcripts in the region of the *Glyr1* gene are indicated by the box. (**B**) GO analysis of KLA-induced NDF peaks.

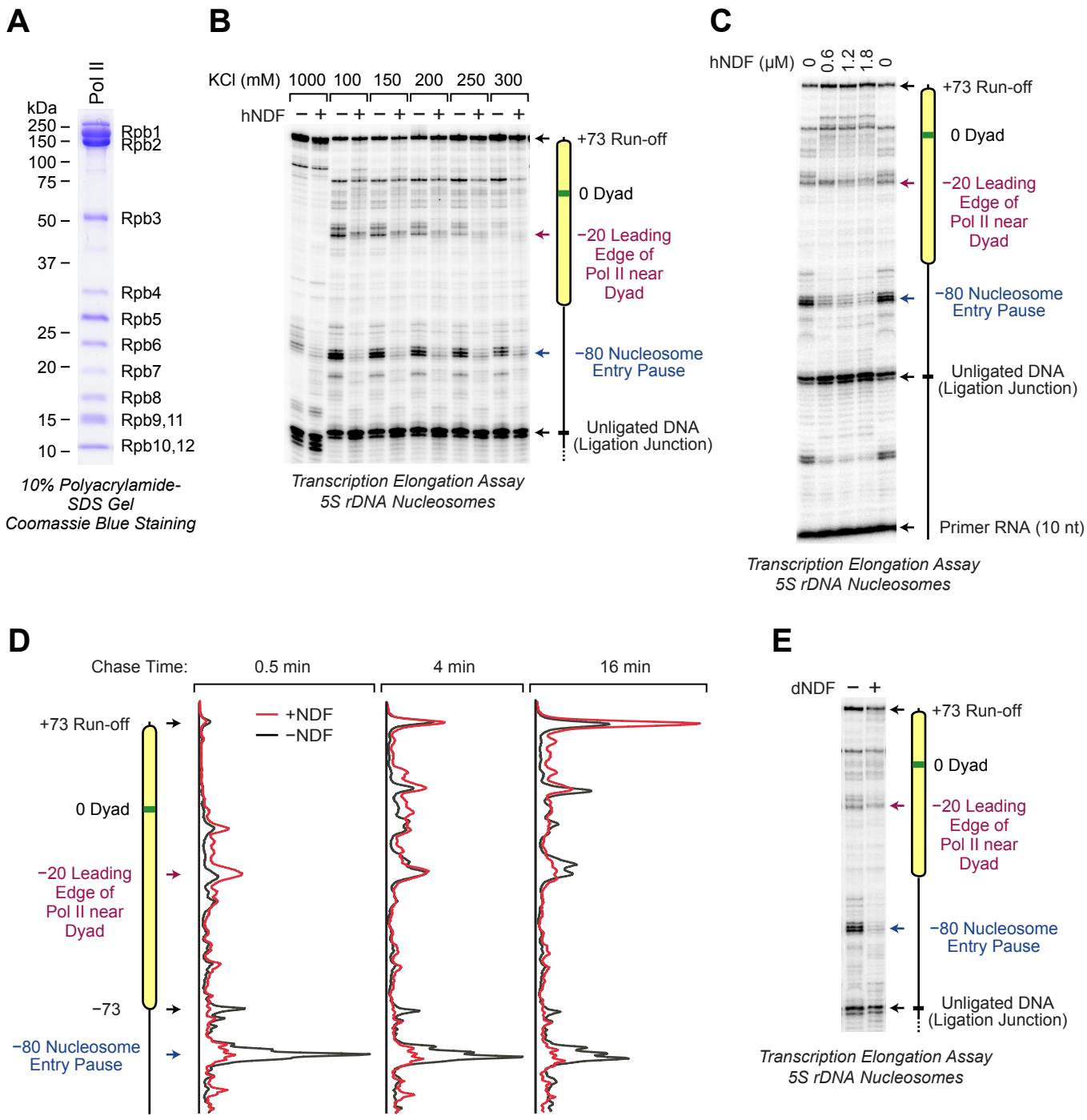
DNA Cleavage Catalyzed by Cu(II) Bound to an o-Phenanthroline Derivative (OP) Linked to Histone H4 S47C at the Nucleosome Dyad



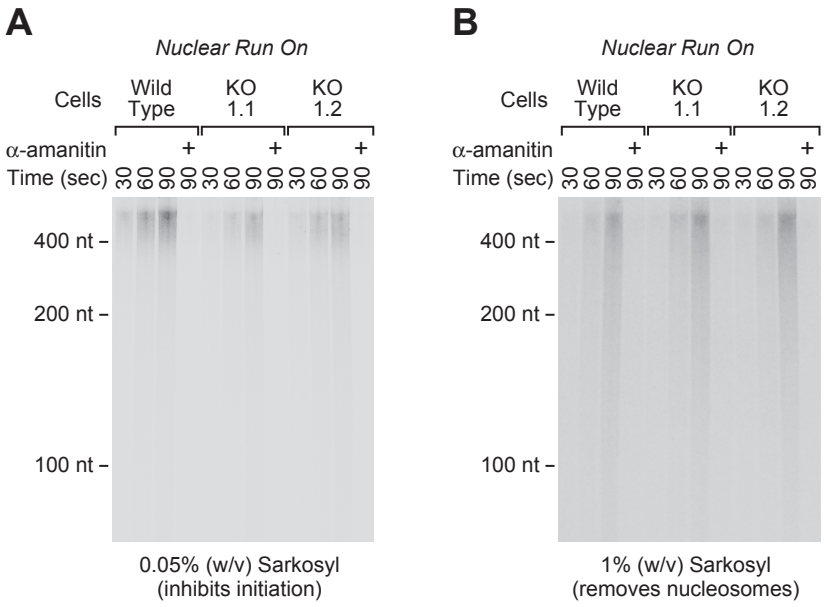
8% Polyacrylamide-Urea Gel
Forward Strand 5S rDNA

8% Polyacrylamide-Urea Gel
Reverse Strand 5S rDNA

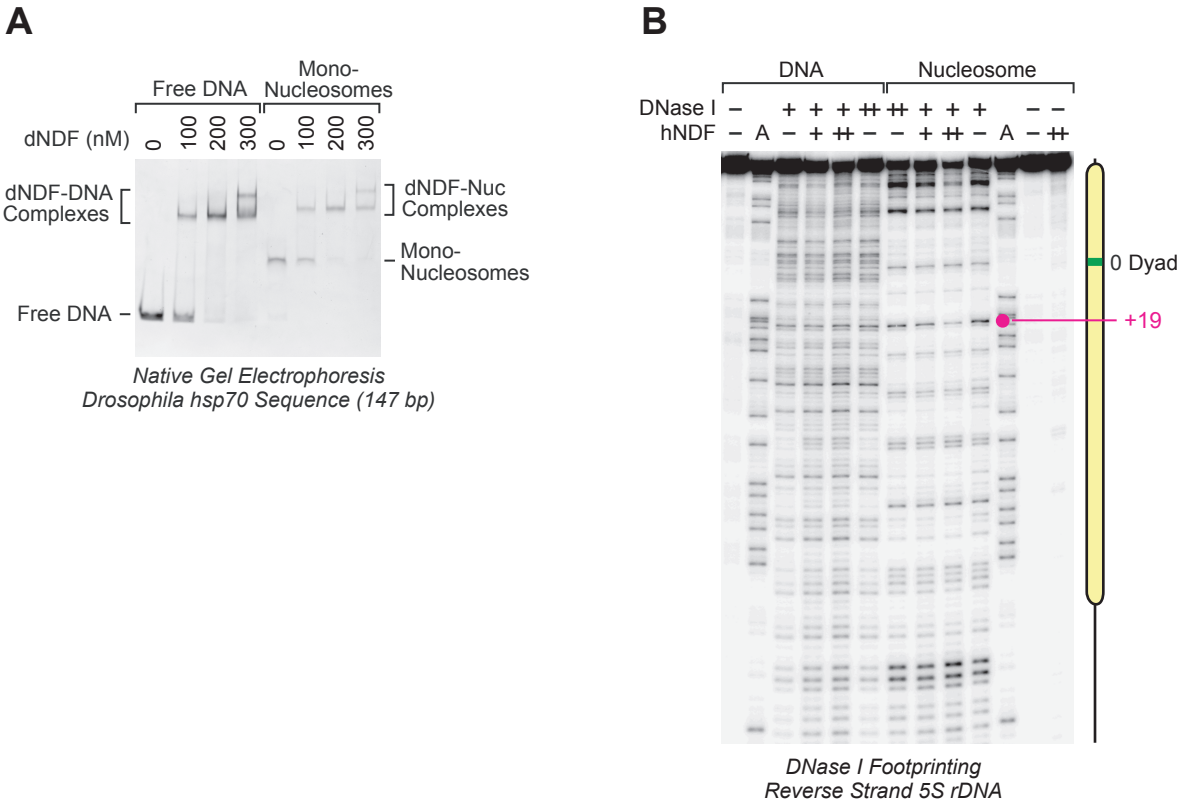
Supplemental Figure S7. Mapping the location of the dyad in nucleosomes positioned with the *Xenopus borealis* 5S rDNA sequence. Core histones containing H4S47C were modified with an o-phenanthroline derivative (OP), which links an o-phenanthroline moiety onto H4S47C via alkylation of the thiol group. The resulting derivatized histones were reconstituted into mononucleosomes by salt dialysis onto the forward or reverse strands of the *Xenopus borealis* 5S rDNA nucleosome positioning sequence. Hydroxyl radical cleavage reactions were initiated by the addition of Cu(II), hydrogen peroxide, and mercaptopropionic acid. The cleavage products were purified and analyzed by electrophoresis on an 8% polyacrylamide-urea gel. The Maxam-Gilbert A ladder was used to identify the OP cleavage products. The nucleosome dyad is located between the major and minor cleavage sites.



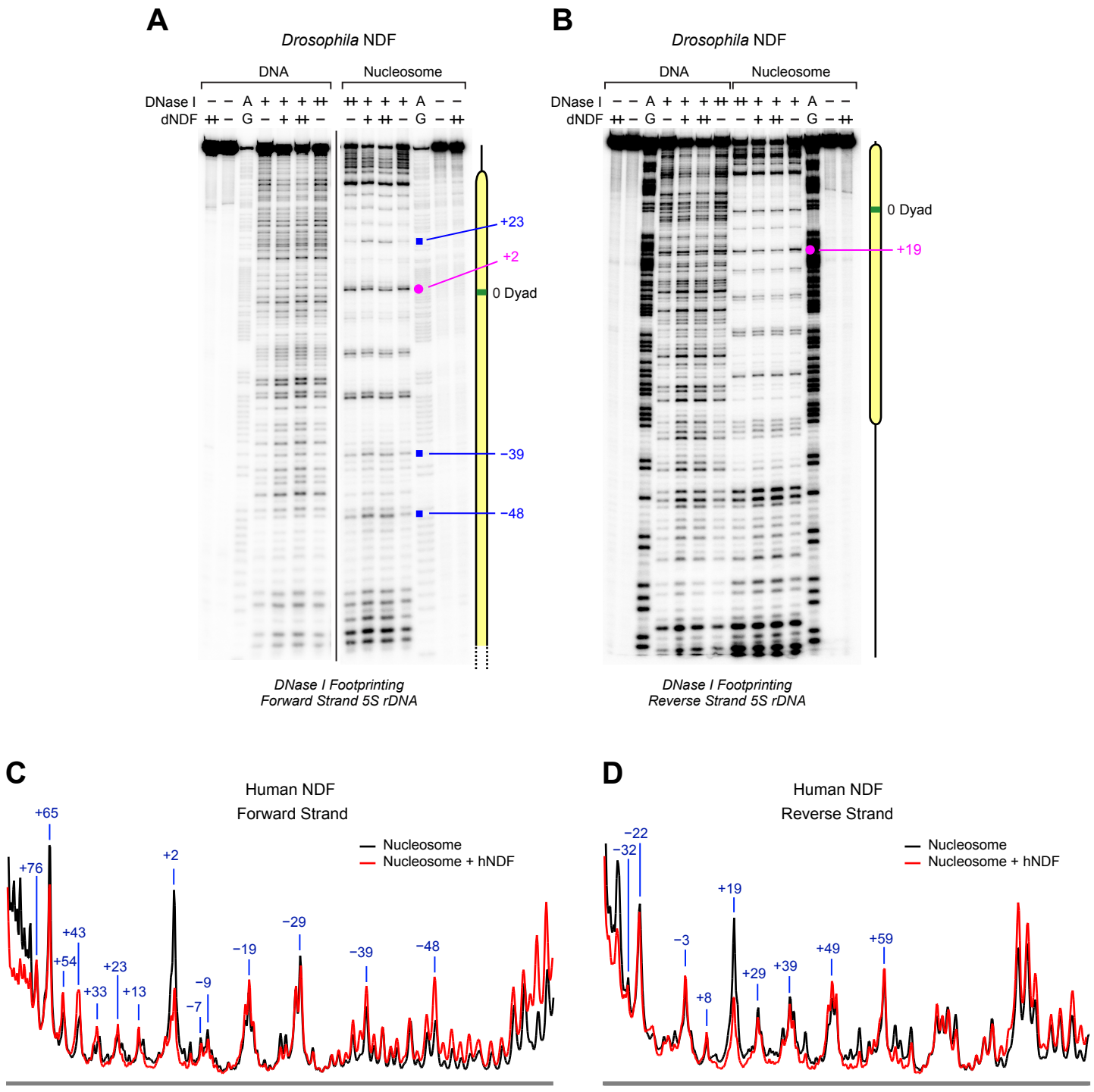
Supplemental Figure S8. Analysis of NDF in transcriptional elongation of Pol II through a nucleosome. (**A**) Purification of yeast RNA polymerase II (Pol II). The purified protein was analyzed by 10% polyacrylamide-SDS gel electrophoresis and staining with Coomassie Brilliant Blue R-250. (**B**) NDF stimulation of transcriptional elongation at different KCl concentrations. Reactions were performed with nucleosomal templates as in Fig. 6 of the main text (in which the KCl concentration is 100 mM) at the indicated concentrations of KCl for 6 min. hNDF was used at a concentration of 1.5 μM. As a reference/control, there is minimal inhibition of Pol II elongation upon dissociation of nucleosomes at 1000 mM KCl. (**C**) Effect of different concentrations of hNDF upon Pol II elongation through a nucleosome. Reactions were performed with nucleosomal templates as in Fig. 5 of the main text (in which the hNDF concentration is 1.5 μM) with the indicated concentrations of hNDF for 6 min. (**D**) Gel intensity scanning profile of Fig. 6B of the main text. Specifically, the 0.5 min, 4 min, and 16 min time points of Pol II elongation through nucleosomal templates are shown. (**E**) *Drosophila* NDF (dNDF) facilitates transcription elongation through nucleosomes. Reactions were performed with nucleosomal templates as in Fig. 5 of the main text for 4 min with dNDF (5 μM), where indicated.



Supplemental Figure S9. NDF facilitates Pol II transcription elongation in chromatin in cell nuclei. (**A**) The loss of NDF results in a decrease in Pol II elongation in HeLa cell nuclei. Pol II run-on analysis was performed with nuclei from wild-type and NDF knockout cells. Nuclei were prepared from the indicated cells and incubated with α-^{32}P-labeled UTP and unlabeled ATP, CTP, and GTP in the presence of 0.05% (w/v) Sarkosyl to inhibit transcription initiation. The same number of nuclei were analyzed in each lane. Where indicated, α-amanitin was included at 4 μg/mL concentration to inhibit Pol II transcription. (**B**) The loss of NDF does not affect Pol II elongation in nuclei after removal of the histones from the DNA with 1% (w/v) Sarkosyl. Nuclear run-on experiments were performed as in A with 1% (w/v) Sarkosyl. Where indicated, α-amanitin was included at 4 μg/mL concentration to inhibit Pol II transcription.

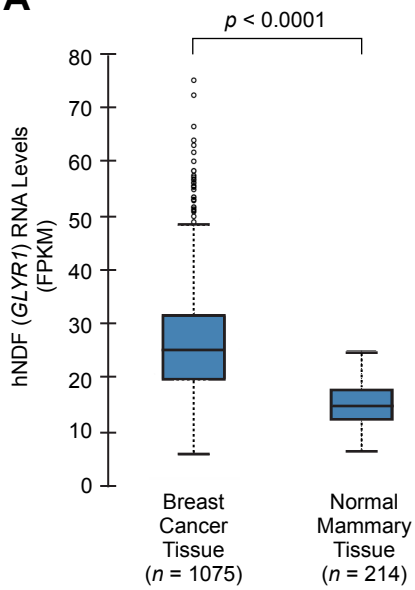


Supplemental Figures S10. Analysis of the binding of NDF to naked DNA and mononucleosomes. (**A**) *Drosophila* NDF (dNDF) binds to a *Drosophila hsp70* promoter DNA fragment as either mononucleosomes or free DNA. Gel mobility shift assays were performed with the indicated concentrations of dNDF and 50 nM of *Drosophila hsp70* promoter DNA (147 bp) as either naked DNA or mononucleosomes. The samples were incubated at 30˚C for 30 min and then subjected to electrophoresis on a 4% nondenaturing polyacrylamide gel. The DNA was visualized by staining with ethidium bromide. The positions of the shifted complexes are indicated. (**B**) DNase I footprinting analysis of hNDF with the reverse strand of *Xenopus borealis* 5S rDNA as either free DNA or mononucleosomes. DNase I footprinting reactions were performed as in Fig. 7B of the main text with the reverse strand of the 5S rDNA sequence (183 bp) as either naked DNA or mononucleosomes. Purified hNDF (73 nM or 150 nM concentration) was included where indicated. The Maxam-Gilbert A sequencing ladder was used to map the locations of the DNase I digestion products. The pink dot shows an NDF-protected nucleotide at +19 relative to the nucleosomal dyad at position 0.



Supplemental Figure S11. DNase I footprinting analysis of the binding of NDF to naked DNA and mononucleosomes. (**A**) DNase I footprinting analysis reveals an interaction of dNDF near the nucleosome dyad. DNase I footprinting reactions were carried out with the forward strand of the 5S rDNA sequence (183 bp) as either naked DNA or mononucleosomes. Purified dNDF (0.35 μM or 0.7 μM concentration) was included where indicated. The Maxam-Gilbert A+G sequencing ladder was used to map the locations of the DNase I digestion products. The pink dot shows an NDF-protected nucleotide at +2 relative to the nucleosomal dyad at position 0. The blue squares designate bands that are enhanced in the presence of hNDF. (**B**) DNase I footprinting reactions, as in A, with the reverse strand of the 5S rDNA sequence. The pink dot shows an NDF-protected nucleotide at +19 relative to the nucleosomal dyad at position 0. (**C**) Gel intensity scanning profiles of the DNase I footprints of hNDF on mononucleosomes reconstituted with 5S rDNA. DNase I footprints with the forward strand, as shown in Fig. 7B of the main text. (**D**) Gel intensity scanning profiles of the DNase I footprints of hNDF with the reverse strand of 5S rDNA, as shown in Supplemental Fig. S10B.

A

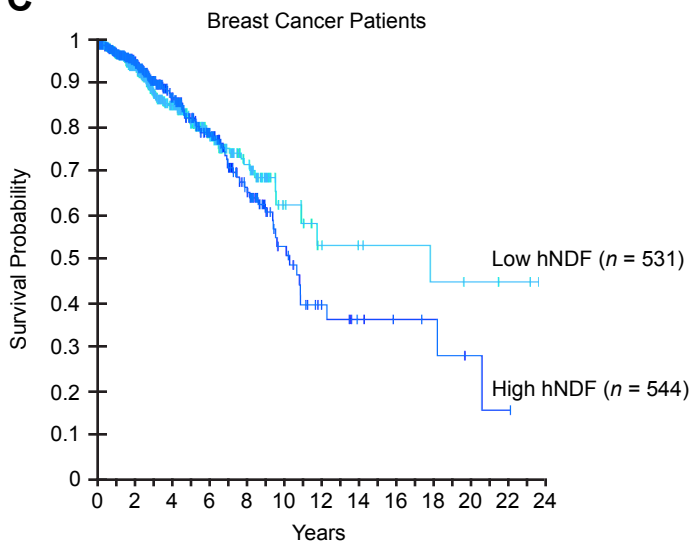


B

Expression of hNDF (*GLYR1*)

Tissue	% Upregulated	% Downregulated	Number of patients tested
Adrenal gland	15.19	N/A	79
Breast	21.19	3.62	1104
Central nervous system	3.87	4.88	697
Cervix	7.17	3.91	307
Endometrium	3.32	3.65	602
Haematopoietic & lymphoid	2.71	1.81	221
Kidney	5.67	1.83	600
Large intestine	6.56	2.62	610
Liver	2.41	9.65	373
Lung	4.61	6.38	1019
Oesophagus	4	N/A	125
Ovary	4.89	7.14	266
Pancreas	6.15	5.59	179
Prostate	1	5.82	498
Skin	5.29	4.65	473
Soft tissue	4.94	0.38	263
Stomach	3.16	4.21	285
Thyroid	2.92	3.9	513
Upper aerodigestive tract	4.02	6.13	522
Urinary tract	2.45	8.09	408

C



Supplemental Figures S12. hNDF (*GLYR1*) overexpression is frequently observed in breast cancer patients and appears to be an unfavorable prognostic marker. (**A**) hNDF expression levels (mRNA level, FPKM) in mammary tissues. The data are from The Human Protein Atlas (www.proteinatlas.org). (**B**) Altered expression of hNDF in different types of cancer. The data are curated by the Catalogue of Somatic Mutations in Cancer (COSMIC), Sanger Institute (cancer.sanger.ac.uk/cosmic). (**C**) Kaplan-Meier plot of breast cancer patients with high versus low levels of hNDF. The data are from The Human Protein Atlas (www.proteinatlas.org).

Supplemental Materials and Methods

To ensure the reproducibility of the results, each experimental condition was performed independently at least two times.

Nucleic acid sequences

The sequences of the nucleic acids used in this study are as follows.

CRISPR-Cas9 targeting sequences for hNDF (GLYR1)

Exon 1: GAGTCTGCGGCTCGGCGACT

Exon 2: AACTCGGCCGATATCCTCCT

Scramble/Control (from origene.com) : GCACTACCAGAGCTAACTCA

Transcription elongation experiments

Non-template strand (NTS) for transcription elongation experiments:

5’ Biotin-TGTCATCCCTTATGTGATGGACCCTATACGAATTCAGATATCGAGAGGACAGAAG GCGAAGAGCCAACCCAAGAGACACCGGCACTGGG-3’ [The 9 nt overhang for ligation of the TspRI-digested downstream sequence is highlighted in blue.]

Template strand (TS) for transcription elongation experiments:

5’-GGTGTCTCTTGGGTTGGCTCTTCGCCTTCTGTCCTCTCGATATCTGAATTCGTATAGGG TCCATCACATAAGGGATG-3’

RNA primer for transcription elongation experiments: 5’-AUCGAGAGGA-3’

[The Pol II elongation complex is assembled as follows. The 5' ^{32}P-labeled RNA primer is hybridized to the template strand (TS), purified Pol II is added, and then 5'-biotinylated non-template strand (NTS) is added to give the Pol II elongation complex, which is then immobilized

onto streptavidin magnetic beads.]

Downstream sequence with the 5S rDNA nucleosome positioning sequence that is digested with TspRI and then ligated to the Pol II elongation complex:

CACATTAGGGCACTGGGTGGCACCGGCAAGGTCGCTGGGCCCGACCCTGCTTGGCTTCC
GAGATCAGACGATATCGGGCACTTTCAGGGTGGTATGGCCGTAGGCGAGCACAAGGCTGA
CTTTTCCTCCCCTTGTGCTGCCTTCTGGGGGGGGCCCAGCCGGATCCCCGGGCGAGCTC
GAATT

[The strand that is shown has the same sense as the non-template strand in the Pol II transcription experiments. This sequence is amplified by PCR, digested with TspRI, gel purified, reconstituted into a mononucleosome, and then ligated to the Pol II elongation complex. The underlined sequence is the central 146 bp of the 5S rDNA positioning sequence. The blue sequence is the TspRI restriction site.]

DNase I footprinting and gel mobility shift experiments

Xenopus borealis 5S rDNA sequence:

CATCCAAGTACTAACCAGGCCCGACCCTGCTTGGCTTCCGAGATCAGACGATATCGGGCA
CTTTCAGGGTGGTATGGCCGTAGGCGAGCACAAGGCTGACTTTTCCTCCCCTTGTGCTGC
CTTCTGGGGGGGGCCCAGCTCCTCCCCATGCCAGGGTCTTTTCCCCCAGGCAGGAAAACA
AGC

[This is the 183 bp sequence that was used in the DNase I footprinting experiments. The underlined sequence is the central 147 bp that was used in the gel mobility shift experiments.]

Drosophila hsp70 (87A) sequence:

AAATTTCTCTGGCCGTTATTCTCTATTCGTTTTGTGACTCTCCCTCTTTGTACTATTGCTCTC
TCACTCTGTCACACAGTAAACGGCGCACTGTTCTCGTTGCTTCGAGAGAGCGCGCCTCGA
ATGTTCGCGAAAAGAGCGCCGGAG

[This is the 147 bp sequence used in gel mobility shift experiments.]

DNA supercoiling assay

DNA supercoiling analyses were performed with either nucleosomes that were reconstituted onto plasmid DNA by salt dialysis (0.3 μM concentration of nucleosomes) or naked DNA as a reference along with the indicated concentrations of NDF, topoisomerase I (0.1 Units, as defined in ref. 42), and bovine serum albumin (1.4 μg) in Buffer S [25 mM Hepes, K^+, pH 7.6, 100 mM KCl, 0.2 mM $MgCl_2$, 0.1 mM EDTA, and 10% (v/v) glycerol] in a total volume of 50 μL. The mixture was incubated at 30˚C for 1 h. The reaction products were deproteinized, subjected to 0.8% agarose gel electrophoresis in the absence or presence of 10 ng/mL chloroquine, and stained with ethidium bromide.

CRISPR-Cas9 knockout experiments

Guide RNA sequences that target hNDF (*GLYR1*) exons 1 and 2 were designed by using crispr.mit.edu. The corresponding DNA sequences were subcloned into PX458 vector (gift from Gaoyang Liang, Salk Institute). The plasmids were transfected into HeLa cells by using Lipofectamine 2000 (Invitrogen) according the recommendations of the manufacturer. GFP-positive cells were collected and distributed into single cell clones in 96-well plates with a BD FACSAria II Sorter at the UCSD Stem Cell Core Facility. Individual clones were characterized by western blot analysis with anti-hNDF antibodies as well as by DNA sequencing (Ran et al. 2013).

Cell cycle analysis

For cell cycle analysis by flow cytometry, the cells were grown in 10 cm dishes. After dissociation with trypsin, 1 to 2 million cells were washed with PBS and resuspended in 300 μL of cold (4˚C) PBS. Cells were fixed by the dropwise addition of 0.8 mL of cold (4˚C) ethanol followed by incubation at 4˚C for a minimum of 24 h. After fixation, the cells were pelleted and

resuspended in 1 mL of PBS containing 50 µg/mL of propidium iodide and 10 µg/mL of RNase A. The cells were incubated for 30 min at 22˚C and then subjected to flow cytometry analysis. For each cell line, more than 50,000 cells were analyzed. For live cell imaging, 10,000 cells/well were plated in 96-well polystyrene plates 24h prior to imaging. The cells were stained with SiR-DNA dye (Spirochrome; used as recommended by the manufacturer) and imaged on the CQ1 spinning disk confocal system (Yokogawa Electric Corporation) with a 40x (0.95 NA) U-PlanApo objective and 2560x2160 pixel sCMOS camera (Andor) at 37°C and 5% CO_2. Image acquisition and data analysis were performed with CellVoyager software and ImageJ, respectively. 25 fields/well were imaged. 5 x 2 µm z-sections in far-red (25% power, 100 ms, binning 2) were acquired in each field at 15-minute intervals for 30 h.

Nuclear run-on experiments

Nuclei were isolated from wild-type and NDF knockout HeLa cells as follows. Except for the 1 x PBS, all buffers were at 4˚C. Cells in 15 cm dishes were washed three times with 1 x PBS and incubated with 10 mL of Swelling Buffer [10 mM Tris-HCl, pH 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$, and 2 Units/mL Superase-in) on ice for 5 min. Cells were collected and resuspended in 0.5 mL of Swelling Buffer containing 10% (v/v) glycerol and 4 Units/mL Superase-In. One volume of IGEPAL Swelling Buffer [Swelling Buffer containing 10% (v/v) glycerol and 1% (v/v) IGEPAL CA-630 (Sigma)] was added, and the mixture was incubated on ice for 5 min. Cells were then washed once with 10 mL Lysis Buffer [Swelling Buffer with 0.5% (v/v) IGEPAL CA-630, 10% (v/v) glycerol, and 2 Units/mL Superase-In] and once with 1 mL Freezing Buffer [50 mM Tris-HCl, pH 7.5, 5 mM $MgCl_2$, 0.1 mM EDTA, 40% (v/v) glycerol, and 2 Units/mL Superase-In). The nuclei were resuspended at a concentration of 5 x 10^7 nuclei/mL in Freezing Buffer.

The run-on experiments, which detect the elongation of transcriptionally-engaged Pol II in nuclei, were performed as follows. In each reaction, 10 µL of 3x Run-on Buffer [16.5 mM Tris-HCl, pH 7.5, 496 mM KCl, 8.25 mM $MgCl_2$, 1.5 mM DTT, 0.15% (w/v) Sarkosyl, 0.6 Units/µL Superase-In, 375 µM ATP, 375 µM GTP, 1.2 µM CTP and 27 nM α-^{32}P-labeled UTP] was

added to 20 μL (1 x 10^6) nuclei. The sample was gently mixed and incubated at room temperature for the indicated length of time (30, 60, or 90 s). The 0.05% (w/v) Sarkosyl was included to inhibit the initiation of Pol II transcription (Hawley and Roeder 1987).

To test Pol II elongation activity in the absence of nucleosomes, run-on reactions were performed in the presence of 1% (w/v) Sarkosyl, which removes the histones from the chromatin (Green et al. 1975; Scheer 1978). In these experiments, 3x Run-on Buffer containing 3% (w/v) Sarkosyl was used.

To test whether transcription was carried out by Pol II, α-amanitin was added, where indicated, to the nuclei to a final concentration of 4 ng/μL before the addition of the nucleotides. The run-on reactions were terminated by the addition of 150 μL Trizol LS (Invitrogen). The resulting radiolabeled transcripts were purified by chloroform extraction, precipitated with isopropanol, and analyzed by denaturing 6% polyacrylamide-urea gel electrophoresis. The gel images were collected on a GE Typhoon imager (GE Healthcare).

DNase I footprinting

The DNase I footprinting probes were generated by PCR of a 183 bp DNA fragment containing the 5S rDNA nucleosome positioning sequence with a 5'-^{32}P-labeled primer and an unlabeled opposite strand primer. The resulting single end-labeled DNA fragments were purified on a nondenaturing 6% polyacrylamide gel, and then reconstituted into nucleosomes by salt dialysis (Stein 1989; Fei et al. 2015). The efficiency of nucleosome reconstitution was estimated by nondenaturing 4% polyacrylamide gel electrophoresis, and samples with ≥95% nucleosome reconstitution were used for footprinting experiments. In a typical DNase I footprinting experiment, either mononucleosomes or corresponding naked DNA fragments (final concentration in 20 μL binding reaction is 14 nM) were incubated with the indicated concentrations of NDF in Buffer F [10 mM Hepes, Na^+, pH 7.8, 0.2 mM EDTA, 5% (v/v) glycerol, 100 mM NaCl, 5 mM $MgCl_2$ and 1 mM DTT] in a total volume of 20 μL at 30°C for 30 min. DNase I (Roche) was prepared freshly in DNase Dilution Buffer [40 mM Tris-HCl, pH 8.0, 5 mM

$MgCl_2$, 5 mM $CaCl_2$, and 1 mM DTT], and 2 µL was added to the reaction. The samples were quickly mixed and then incubated at 30˚C for 30 s. A final concentration of 1.0 or 1.5 mU/µL DNase I was used for the digestion of naked DNA, whereas a final concentration of 10 or 15 mU/µL was used for the digestion of chromatin. Reactions were terminated by the addition of 158 µL of Stop Buffer [10 mM Tris-HCl, pH 7.8, 3 mM EDTA and 0.2% (w/v) SDS]. The reaction products were extracted with phenol-chloroform, precipitated with ethanol, resuspended in formamide loading buffer, and subjected to electrophoresis on a denaturing 8% polyacrylamide-urea gel. The gel images were collected on a GE Typhoon imager (GE Healthcare).

Mapping of the nucleosome dyad by H4S47C-directed cleavage

To analyze the positioning of the nucleosomes, we used a modified version of the histone-directed DNA cleavage method of Flaus et al. (1996) with N-(1,10-phenanthrolin-5-yl)iodoacetamide (OP) as the histone H4S47C-alkylating reagent (Brogaard et al. 2012; Henikoff et al. 2014; Fei et al. 2015). We reconstituted nucleosomes with histone H4S47C, which is located near the nucleosome dyad, and the H4S47C was alkylated with OP, which resulted in a covalent linkage between the H4S47C thiol and an o-phenanthroline moiety. Then, upon addition of Cu(II) and hydrogen peroxide, the o-phenanthroline group mediated the generation of hydroxyl radicals that cleaved the nucleosomal DNA in the vicinity of the dyad. Specfiic reaction conditions are described in Fei et al. (2015).

ChIP-seq analysis with HeLa cells

For human NDF ChIP-seq, approximately 15 to 16 million cells (wild-type HeLa S3 cells or the indicated hNDF knockout cells) were collected and crosslinked with 1% formaldehyde (Sigma). The crosslinked cells were resuspended in 5 mL of Cell Lysis Buffer [5 mM Hepes, K^+, pH 8.0, 85 mM KCl, 0.5% (v/v) Nonidet P-40, and cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche)]. The cells were incubated on ice for 5 min, and then pelleted by centrifugation at 500 x *g* at 4˚C for 5 min. The supernatant was removed, and the nuclear pellet was resuspended in

Nuclear Lysis Buffer [50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.2% (w/v) SDS, and cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche)] and then adjusted to a final volume of 1.2 mL. The resuspended sample was divided into four aliquots of ~300 µL in polystyrene tubes (Evergreen Scientific, 214-3710-010). Next, the samples were sonicated with a Qsonica Q800R sonicator with 70% amplitude and cycles of 30s on and 30s off for a total on time of 2 h. The samples were spun at 15,000 rpm (21130 x *g*) at 4˚C for 10 min, and the supernatant was collected. The samples were mixed with 1 x volume of Dilution Buffer [16.7 mM Tris-HCl, pH 8.0, 167 mM NaCl, 1.2 mM EDTA, 0.01% (w/v) SDS, 1.1% (v/v) Triton X-100, and cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche)] and 0.5 x volumes 50% glycerol and then frozen in liquid nitrogen.

For immunoprecipitation, 300 µL of the sonicated sample and 200 µL of Dilution Buffer containing 0.25% (w/v) sodium deoxycholate were added to 20 µL of Dynabeads Protein A that were pre-bound with 1 µL of NDF antiserum. [Note that 1 µL of hNDF antiserum was optimal for hNDF antiserum #7. For hNDF antiserum #6, 2 µL was the optimal volume.] After incubation at 4˚C for 2 h on a nutating mixer, the tubes were placed on a magnetic stand, and the supernatant was removed. The beads were washed with 3 x 500 µL of Wash Buffer I [20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM EDTA, 0.1% (w/v) SDS, 1% (v/v) Triton X-100], 3 x 500 µL of Wash Buffer II [10 mM Tris-HCl, pH 7.5, 250 mM LiCl, 1 mM EDTA, 0.7% (w/v) sodium deoxycholate, 1% (v/v) Nonidet P-40], and 1 x 500 µL of TET Buffer [10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.1% (v/v) Tween-20] by resuspending the beads in the buffer and removing the supernatant on a magnetic stand. Next, the beads were resuspended in 100 µL of TET Buffer and transferred to new tubes. The supernatant was removed again, and the beads were resuspended in 50 µL of TE [10 mM Tris-HCl, pH 8, 1 mM EDTA] containing 1% (w/v) SDS for elution. The tubes were incubated at room temperature for 20 min and were placed on a magnetic stand to collect the eluate. Elution with 50 µL of TE containing 1% (w/v) SDS was repeated one more time. The eluates from both rounds were combined, and 6.4 µL of 5 M NaCl was added to adjust the final NaCl concentration to ~0.3 M. The samples were incubated at

65°C to reverse the crosslinks. After digestion with RNase A and proteinase K, the resulting DNA was isolated with MinElute columns (Qiagen). ChIP-seq libraries were prepared by end-repair, A-tailing, and ligation of Illumina TruSeq adaptors, which was then followed by PCR amplification and size selection with AMPure beads (Beckman Coulter). The libraries were sequenced with an Illumina Hi-Seq 4000 sequencer by using single-end 50 bp reads. The sequencing data were aligned to the hg19 reference genome by using Bowtie, and PCR duplicates were removed with the Picard tool (http://broadinstitute.github.io/picard/). Gencode Release 19 was used as gene annotation for the analysis.

RNA-seq analysis with HeLa cells

Total RNA was purified by using the PureLink RNA Mini RNA Kit (Thermo Fisher Invitrogen). The integrity of the RNA was measured by using the RNA 6000 Nano Kit with the Agilent Bioanalyzer 2100 system (Agilent Technologies). RNA-seq libraries were prepared with the TruSeq Stranded mRNA Library Prep Kit (Illumina). The libraries were sequenced with the Illumina Hi-Seq 4000 sequencer by using single-end 50 bp reads. The sequencing data were aligned to the hg19 reference genome by using the STAR RNA-seq aligner (Dobin et al. 2013), and PCR duplicates were removed with the Picard tool (http://broadinstitute.github.io/picard/). The reads that were assigned to genes were analyzed by using featureCounts (Liao et al. 2014) and gene annotation according to Gencode Release 19. The comparison of the wild-type cells versus the KO-1.1 and KO-1.2 cells was carried out three independent times.

Global run-on sequencing (GRO-seq)

GRO-seq was performed as described by Core et al. (2008). The data were mapped to the hg19 reference genome by using Bowtie2 (Langmead and Salzberg 2012). The resulting bam files were by analyzed by using HOMER (Heinz et al. 2010).

Supplemental References

Brogaard K, Xi L, Wang JP, Widom J. 2012. A map of nucleosome positions in yeast at base-pair resolution. *Nature* **486**: 496-501.

Core LJ, Waterfall JJ, Lis JT. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**: 1845-1848.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.

Fei J, Torigoe SE, Brown CR, Khuong MT, Kassavetis GA, Boeger H, Kadonaga JT. 2015. The prenucleosome, a stable conformational isomer of the nucleosome. *Genes Dev* **29**: 2563-2575.

Flaus A, Luger K, Tan S, Richmond TJ. 1996. Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. *Proc Natl Acad Sci* **93**: 1370-1375.

Green MH, Buss J, Gariglio P. 1975. Activation of nuclear RNA polymerase by Sarkosyl. *Eur J Biochem* **53**: 217-225.

Hawley D K, Roeder RG. 1987. Functional steps in transcription initiation and reinitiation from the major late promoter in a HeLa nuclear extract. *J Biol Chem* **262**: 3452-3461.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Lasko P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**: 576-589.

Henikoff S, Ramachandran S, Krassovsky K, Bryson TD, Codomo CA, Brogaard K, Widom J, Wang JP, Henikoff JG. 2014. The budding yeast Centromere DNA Element II wraps a stable Cse4 hemisome in either orientation in vivo. *eLife* **3**: e01861.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357-359.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923-930.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc* **8**: 2281-2308.

Scheer U. 1978. Changes of nucleosome frequency in nucleolar and non-nucleolar chromatin as a function of transcription: an electron microscopic study. *Cell* **13**: 535-549.

Stein A.1989. Reconstitution of chromatin from purified components. *Methods Enzymol* **170**: 585-603.