MOLECULAR BIOLOGY

LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells

Gary LeRoy[1,2]*, Ozgur Oksuz[1,2]*, Nicolas Descostes[1,2,3]*, Yuki Aoi[4], Rais A. Ganai[1,2], Havva Ortabozkoyun Kara[1,2], Jia-Ray Yu[1,2], Chul-Hwan Lee[1,2], James Stafford[1,2], Ali Shilatifard[4], Danny Reinberg[1,2†]

FACT (facilitates chromatin transcription) is a protein complex that allows RNA polymerase II (RNAPII) to overcome the nucleosome-induced barrier to transcription. While abundant in undifferentiated cells and many cancers, FACT is not abundant or is absent in most tissues. Therefore, we screened for additional proteins that might replace FACT upon differentiation. We identified two proteins, lens epithelium-derived growth factor (LEDGF) and hepatoma-derived growth factor 2 (HDGF2), each containing two high mobility group A (HMGA)–like AT-hooks and a methyl-lysine reading Pro-Trp-Trp-Pro (PWWP) domain that binds to H3K36me2 and H3K36me3.LEDGF and HDGF2 colocalize with H3K36me2/3 at genomic regions containing active genes. In myoblasts, LEDGF and HDGF2 are enriched on most active genes. Upon differentiation to myotubes, LEDGF levels decrease, while HDGF2 levels are maintained. Moreover, HDGF2 is required for their proper expression. HDGF2 knockout myoblasts exhibit an accumulation of paused RNAPII within the transcribed region of many HDGF2 target genes, indicating a defect in early elongation.

INTRODUCTION

RNA polymerase II (RNAPII) transcription is regulated at the level of initiation, pause release, promoter escape, +1 nucleosome release, and elongation (*1*–*3*). After promoter escape, RNAPII must overcome a nucleosome-induced barrier to transcription (*4*, *5*). Two decades ago, we identified FACT (facilitates chromatin transcription) as a protein complex composed of suppressor of Ty elements 16 (SPT16), and structure specific recognition protein 1 (SSRP1) that alleviates this nucleosome-induced barrier to transcription (*6*, *7*). More recently, we mapped the genomic binding of FACT in stem cells using chromatin immunoprecipitation sequencing (ChIP-seq) with an SPT16 antibody and found that FACT occupancy varies, being associated at high levels with only a subset of active genes (Fig. 1A). In addition, we and others have found that FACT is only expressed at high levels in some progenitor and transformed cells (fig. S1A) (*8*, *9*). These observations imply the existence of alternative FACT-like chaperones in differentiated cells. We first hypothesized that the BET family proteins (Brd2, Brd3, and Brd4), which also have FACT-like activity in vitro, might replace FACT at these genes (*10*, *11*). However, bromodomain and extra-terminal domain (BET) proteins predominately localize with acetylated nucleosomes proximal to the transcription start site (TSS) of genes, while a chaperone capable of fulfilling the function of FACT should localize in gene bodies (fig. S1B) (*11*, *12*).

RESULTS

Exploiting a similar biochemical strategy that first led us to identify FACT, we fractionated HeLa cell nuclear extract and identified a fraction that was depleted of FACT, BET proteins, and nucleolin (RNAPI-specific FACT-like chaperone), yet able to support transcription through nucleosomes (fig. S1, C and D) (*6*, *10*, *11*, *13*). Through further chromatographic fractionation and mass spectrometry (MS), we identified LEDGF long isoform (p75) (also known as PSIP1) (Fig. 1B, fig. S1E, and table S1) (*14*, *15*). Semiquantitative proteomics and RNA sequencing (RNA-seq) data suggest that this protein and its family member, HDGF2, are relatively abundant and expressed in most tissues, unlike the restricted expression of FACT. LEDGF and HDGF2 each contain a methyl-lysine reading Pro-Trp-Trp-Pro (PWWP) domain that has been shown to recognize H3K36me2/3 (histone H3 lysine 36 di- and trimethylation), two HMGA-like AT-hooks (similar to nucleolin) and an integrase binding domain (IBD) (Fig. 1C) (*16*–*20*). The short isoform of LEDGF (p52) and another PWWP-domain containing protein, HDGF, lack the IBD, and the latter contains a high mobility group box domain instead of high mobility group A (HMGA)–like AT-hooks (*21*). LEDGF and HDGF2 have generated considerable interest given their requirement for lentiviral integration, which favors integration into the body of transcribed genes by concurrently binding directly to lentiviral integrase and H3K36me2/3 in the host cell chromatin (*17*, *22*). Retroviruses, which favor integration near the TSS of transcribed genes, use BET proteins in an analogous manner (*23*, *24*). LEDGF and HDGF2 share another similarity with BET proteins in that they remain bound to mitotic chromosomes, suggesting that they might contribute to transcriptional memory (*25*–*27*).

To validate the FACT-like activity of LEDGF and its related proteins, we generated and tested purified recombinant versions of these proteins in a fully in vitro reconstituted chromatin transcription assay (Fig. 1D). As depicted, both isoforms of LEDGF (p75 and p52) and HDGF2 allow RNAPII to transcribe through nucleosomes in a manner similar to FACT. HDGF, which lacks the HMGA-like AT-hooks, did not substitute for FACT, suggesting that the FACT-like activity of these proteins probably lies within the region that contains the AT-hooks (*21*). Increasing amounts of FACT, LEDGF, or HDGF2 led to a linear increase in transcription activity such that at least two or more molecules of LEDGF and HDGF2 per nucleosome appeared to be required for efficient transcription in the assay (Fig. 1, E and F).

[1]Howard Hughes Medical Institute, New York University School of Medicine, New York, NY 10016, USA. [2]Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY 10016, USA. [3]EMBL Rome, Adriano Buzzati-Traverso Campus, Via Ramarini 32, 00015 Monterotondo (RM), Italy. [4]Department of Biochemistry and Molecular Genetics, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: danny.reinberg@nyumc.org