4462–4475 *Nucleic Acids Research, 2019, Vol. 47, No. 9* *Published online 13 March 2019*
doi: 10.1093/nar/gkz162

A complex between DYRK1A and DCAF7 phosphorylates the C-terminal domain of RNA polymerase II to promote myogenesis

Dan Yu[1,2], Claudia Cattoglio[1,3], Yuhua Xue[2] and Qiang Zhou[1,*]

[1]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA, [2]State Key Laboratory of Cellular Stress Biology, School of Pharmaceutical Sciences, Xiamen University, Xiamen, China and [3]Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, CA 94720, USA

Received November 20, 2018; Revised February 24, 2019; Editorial Decision February 26, 2019; Accepted March 02, 2019

ABSTRACT

The general transcription factor P-TEFb, a master regulator of RNA polymerase (Pol) II elongation, phosphorylates the C-terminal domain (CTD) of Pol II and negative elongation factors to release Pol II from promoter-proximal pausing. We show here that P-TEFb surprisingly inhibits the myoblast differentiation into myotubes, and that P-TEFb and its two positive complexes are eliminated in this process. In contrast, DYRK1A, another CTD kinase known to control transcription of a subset of genes important for development and tissue homeostasis, is found to activate transcription of key myogenic genes. We show that active DYRK1A exists in a complex with the WD40-repeat protein DCAF7 that stabilizes and tethers DYRK1A to Pol II, so that DYRK1A–DCAF7 can co-migrate with and phosphorylate Pol II along the myogenic gene loci. Thus, DCAF7 modulates the kinase signaling output of DYRK1A on Pol II to stimulate myogenic transcription after active P-TEFb function is shut off.

INTRODUCTION

The phosphorylation of the C-terminal domain (CTD) of the largest subunit of RNA polymerase (Pol) II is essential for the proper control of gene transcription and co-transcriptional pre-mRNA processing in metazoans (1,2). Several kinases are known to phosphorylate the CTD at different stages of the transcription cycle or in a gene-specific manner (3). CDK9, which partners with cyclin T (CycT) to form P-TEFb, is one of the best characterized CTD kinases (4,5). P-TEFb acts at the transcription elongation stage by promoting the transition of Pol II from promoter-proximal pausing to productive elongation to generate long transcripts. Because the inhibition of CDK9 blocks elongation and causes widespread Pol II pausing at numerous gene promoters (6,7), P-TEFb is considered a general transcription factor required for efficient Pol II elongation on the vast majority of genes (4,5,8).

P-TEFb can associate with a number of regulators to form a network of complexes that regulate its biogenesis, activity, gene-target specificity and homeostasis (5,9). For example, the binding of P-TEFb to the bromodomain protein BRD4, a chromatin adapter that recognizes the acetylated histone tails, recruits P-TEFb to primary response genes for transactivation (10–12). Furthermore, the assembly of P-TEFb and another well-known elongation stimulatory factor ELL1/2 into the multi-subunit Super Elongation Complex (SEC) by the scaffolding protein AFF1/4 enables highly efficient elongation of Pol II along the HIV-1 proviral DNA as well as cellular genes involved in leukemogenesis (13,14).

Recently, the Down's syndrome-associated protein kinase DYRK1A has also been reported as a Pol II CTD kinase (15). However, unlike P-TEFb, DYRK1A controls the transcription of only a subset of genes that are key for development and tissue homeostasis (15). Despite this difference, DYRK1A and P-TEFb demonstrate remarkable similarities in two additional aspects besides their hyperphosphorylation of the CTD. First, our earlier work shows that among the >400 human kinases evaluated, DYRK1A and its homolog DYRK1B are the two next best targets of a highly selective CDK9 inhibitor called iCDK9 (6), suggesting that the structure of at least the catalytic site and/or the substrate specificity of DYRK1A/B resemble those of CDK9 to a significant degree. Furthermore, we recently found that both DYRK1A and the CycT1 subunit of P-TEFb contain an evolutionarily conserved histidine-rich region (HRD) that promotes the hyperphosphorylation of Pol II CTD through a phase separation mechanism (16).

*To whom correspondence should be addressed. Qiang Zhou. Email: qzhou@berkeley.edu
Present address: Dan Yu, Beijing Pediatric Research Institute, Beijing Children's hospital, Capital Medical University, National Center for Children's Health, Beijing, China.

Despite these similarities, comparatively little is known up till now about how the DYRK1A transcriptional activity is regulated and how it can recognize Pol II during transcription. Furthermore, it is unclear whether P-TEFb is also important for the genes whose transcription is upregulated by DYRK1A. Here, we show that the WD-repeats family member DCAF7, which can bind to Pol II independently of DYRK1A, forms a stable complex with DYRK1A. Through this interaction, DCAF7 stabilizes and also enables DYRK1A to recognize and hyperphosphorylate the Pol II CTD and activate transcription.

Myogenesis, a multistage process that causes proliferating myoblasts to terminally differentiate into myotubes and then the contractile skeletal muscle fibers (17), is governed by a set of master sequence-specific transcription factors, chromatin modifiers and a core promoter recognition complex (18–21). An irreversible loss of proliferative potential coupled with selective silencing of certain genes and activation of others occur during this transformation.

In C2C12 cells, a widely used mouse myoblast cell line for studying myogenesis (22), we found that the DYRK1A–DCAF7 complex, but not DYRK1A alone, is able to promote the differentiation into myotubes by co-occupying with and phosphorylating Pol II at several key myogenic gene loci to stimulate transcription. Surprisingly, unlike DYRK1A, CDK9 is found to be inhibitory to myoblast differentiation and myogenic gene expression. In fact, contrary to the conventional view of P-TEFb as an essential general transcription factor, the CDK9-CycT1 heterodimer as well as two positive P-TEFb complexes, the BRD4-P-TEFb and the SEC, were eliminated in differentiating cells to likely enable robust myogenic transcriptional activation by DYRK1A–DCAF7.

MATERIALS AND METHODS

Plasmids and antibodies

Plasmids for expressing Flag-tagged F-DYRK1A (754 amino acids form) and F-DCAF7 were constructed by ligating the corresponding cDNAs into the pcDNA3 vector (Thermo Fisher Scientific). Mutations or deletions in DYRK1A and DCAF7 were generated by using the KAPA HiFi PCR kit (KR0368, Roche). The 6 × Gal4 UAS-HIV-1 LTR-luciferase reporter construct was constructed previously (16). All Gal4 fusion proteins containing the DNA-binding domain (DBD) of Gal4 (aa1–147) attached to the N-termini of the various DYRK1A and DYRK1B proteins were expressed from the pcDNA3 expression vector. PCR primers used in the plasmid constructions are listed in Supplemental Table S1. The MyoD construct and MYOG promoter-luciferase reporter construct were a gift from Drs Tapscott and Nissen (23,24).

Antibodies used in this study are listed as follows. CyclinT1 (sc-10750, Santa Cruz Biotechnology); Anti-RNA polymerase II subunit B1 (phospho CTD Ser-5) clone 3E8 (04-1572-I, Millipore); Anti-RNA polymerase II subunit B1 (phospho CTD Ser-2) clone 3E10 (04-1571-I, Millipore); Anti-RNA polymerase II subunit B1 (sc-56767, Santa Cruz Biotechnology); Anti-DYRK1A (A303-802A, Bethyl Laboratories); Anti-DCAF7 (ab138490, Abcam); Anti-MYH2 (Santa Cruz Biotechnology); normal rabbit IgG (sc-2027, Santa Cruz Biotechnology); Anti-BRD4, HEXIM1, LARP7 and CDK9 were made and stored in our own lab as previously described (25).

C2C12 cell culture and differentiation

The C2C12 myoblast cell line was obtained from Robert Tjian's laboratory at the University of California, Berkeley and cultured at low density in DMEM + 10% fetal bovine serum as the growth medium (GM). For induction of differentiation, the cells were first grown to ∼90% confluency and then shifted to the differentiation medium (DM), which is DMEM plus 2% horse serum for the indicated number of days.

Affinity-purification and mass spectrometry analysis of F-DYRK1A and associated proteins

Eight 15 cm dishes of HeLa cells were transfected with either an empty vector or the F-DYRK1A-expressing plasmid. Nuclei were collected and extracted with 0.35 M NaCl for 2 h at 4°C, which was followed by centrifugation three times at 14 000 rpm for 10 min each. The nuclear extracts (NE; 1000 μl) containing 0.2% NP-40 were added to 45 μl anti-Flag mAb beads (Sigma) and incubated overnight with constant rotation. The immunoprecipitates were washed with buffer D [20 mM HEPES (pH 7.9), 15% Glycerol, 0.2 mM EDTA, 0.4% NP-40] plus 0.5 M NaCl for three times and eluted with Flag peptide (60 μl; 0.5 mg/ml) at room temperature for 0.5 h. Then, the eluted products were subjected to Western-blotting analysis to confirm the expression of F-DYRK1A and SDS-PAGE/silver staining to examine the purity. The samples were digested with trypsin and submitted to the Berkeley QB3 Mass Spectrometry Facility for further analysis.

In vitro kinase assay

F-DYRK1A and its associated DCAF7 were immunoprecipitated with the anti-Flag beads under high salt plus detergent (0.6 M KCl + 0.3% NP-40) conditions as described above. The purified proteins from 4 × 15-cm plates cells still attached to the beads were incubated for 30 min with 100 ng recombinant GST-CTD containing all 52 repeats of the Pol II CTD at 30°C in a 25 μl reaction that also contained 50 mM HEPES (pH 7.4), 50 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$, 10 mM $MnCl_2$ and 0.1 mM ATP. The reaction products were analyzed by Western-blotting as indicated.

For reactions containing GST-DYRK1A and F-DCAF7, the recombinant GST-DYRK1A was purchased from Life technologies (PV3785) and F-DCAF7 was purified from the DYRK1A KO 293T cells. GST-DYRK1A (40 ng) was first incubated with 40 ng F-DCAF7 at room temperature for 10 min to form the complex. The kinase buffer containing 50 mM HEPES (pH 7.4), 50 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$, 10 mM $MnCl_2$ with or without 100 ng GST-CTD was then added to start the kinase reactions. The total volume of the reaction was 30 μl. The reactions were stopped at different time points by adding 10 μl of the SDS-PAGE sample-loading buffer. After heating at 95°C for 10 min, the samples were analyzed by SDS-PAGE and Western-blotting with the indicated antibodies.

Lentiviral vector production and infection

The procedures for shRNA-mediated gene knockdown have been described previously (25) with some modifications. The shRNA sequences are listed in Supplemental Table S1. To produce lentiviral particles expressing the specific shRNA in a well of a six-well dish, the lentiviral vector pLKO.1 (Life Technologies) with the inserted shRNA-expressing cassette (0.5 μg), together with the pCMV-VSV-G envelope (0.5 μg; Addgene #8454) and the psPAX2 packaging (0.5 μg; Addgene #12260) plasmids, were co-transfected into 293T cells. Sixteen hrs later, cells were changed into fresh medium (DMEM + 20% FBS) and cultured for additional 24 h. The virus-containing medium was then collected, filtered through a 0.45 micron filter and added to the cells to be infected. Spin-infection was performed by incubation of the virus-containing medium with the cells at 25°C for 1.5 h at 1200 rpm. The infected cell pool was obtained after the selection with puromycin (2 μg/ml) for 5 days for C2C12 cells and 3 days for Hela and 293T cells.

DYRK1A and DCAF7 knockout by CRISPR–Cas9

The 293T-based DYRK1A or DCAF7 KO cell line was generated by using CRISPR–Cas9 with a single guide RNA (sgRNA) 5′- GCCAAACATAAGTGACCAAC-3′ that targets exon 1 of DYRK1A and 5′-GCCAAGCGAAAGCGCTTATC-3′ that targets exon 1 of DCAF7. The plasmid vector pSpCas9(BB)-2A-Puro (PX459), which expresses Cas9 and sgRNA, was from Addgene (Plasmid #48139). Two days after transfection, cells were challenged with 2 μg/ml puromycin. The drug-resistant cells were diluted and allowed to grow into single colonies, which were subsequently examined for the loss of DYRK1A or DCAF7 expression by Western-blotting. The positive KO clones were verified by Sanger sequencing of the genomic amplicons obtained with the TA cloning kit (Life Technologies).

Quantitative RT-PCR analyses

qRT-PCR reactions were performed as described previously (25). All reactions were run in triplicates. The primers used for the analyses are listed in Supplemental Table S1.

Immunofluorescence staining

The procedure was performed as described previously (26) with some modifications. Briefly, C2C12 cells grown on the surface of a cover glass in a six-well plate were fixed in 4% paraformaldehyde (PFA)/PBS for 10 min at room temperature and permeabilized in 0.5% Triton X-100/PBS solution for 5 min at room temperature. The cells were blocked in 5% BSA/PBS blocking solution, after which primary antibodies (0.5 μg/ml) prepared in blocking solution were added and incubated at 4°C overnight. This was followed by three washes in PBS and incubation with Alexa fluor-conjugated antibodies (0.5 μg/ml; Life Technologies) in blocking solution for 1 h at room temperature. After another three washes in PBS, the slides were mounted with a coverslip and VECTASHIELD mounting medium for fluorescence with DAPI (VECTOR Laboratories, H-1200). Immunofluorescent signals from stained cells were captured by using a Zeiss AxioImager M1 fluorescence microscope equipped with a high-resolution digital CCD camera (Hamamatsu Orca 03) and analyzed with the imaging software iVision.

Co-immunoprecipitation (Co-IP)

The co-IP assay was performed as described previously (27) with minor changes. Briefly, for co-IP of endogenous proteins, 4 μg of a specific antibody were added to 700 μl nuclear extracts (NE) from 293T cells and incubated overnight. Then, 20 μl protein A agarose beads (15918-014, Sigma-Aldrich) were added to the mixture for another 7 h to collect the immunoprecipitated complex. For anti-Flag IP to precipitate F-DYRK1A or F-DCAF7, 20 μg expression plasmids were transfected into HeLa or 293T cells. 20 μl Flag beads (Sigma) were used to collect the IP products. After washing and elution, the denatured protein lysates were analyzed by SDS-PAGE and Western-blotting with the indicated antibodies.

ChIP-qPCR assay

The assay conditions were described previously (11) with some modifications. Briefly, one 15 cm dish of C2C12 cells grown in GM at 90% confluency or DM at 100% confluency was harvested and subjected to the ChIP procedure. Cells were lysed and sonicated by using a Covaris-S2 sonicator (Covaris, Inc., Woburn, MA, USA) for a total processing time of 30 min (30 s on and 30 s off). For each IP, 5 μg anti-DYRK1A antibody, 5 μg anti-DCAF7 antibody, 5 μg anti-Pol II antibody, 5 μg anti-BRD4 antibody or 5 μg total rabbit IgG were incubated with the diluted sheared chromatin DNA. 20 μl of protein G Dynabeads (10003D, Life Technologies) or protein A Dynabeads (10002D, Life Technologies) were added to each tube and incubated at 4°C for 3 h. After washing and elution, DNA was purified by using the PCR purification kit (Qiagen) and analyzed by quantitative PCR.

RESULTS

DYRK1A and DCAF7 form a nuclear complex that promotes stability of both proteins

To investigate the factor(s) and mechanism that regulate the transcriptional activity of DYRK1A, we performed anti-Flag immunoprecipitation (IP) followed by mass spectrometry to identify proteins that are specifically associated with the Flag-tagged DYRK1A (F-DYRK1A) in HeLa nuclear extract (NE). Among the few proteins that were present in the anti-Flag IP but not the control sample derived from non-F-DYRK1A-expressing cells, we discovered DCAF7, two Hsp70 isoforms (HSPA1A and HSPA8) and two hnRNP proteins (HNRNPH1 and HNRNPF) as potential DYRK1A-binding partners (Figure 1A). While the function and significance of the bindings of DYRK1A to the latter two groups of proteins are still under investigation and may be a subject of a future report, we decided to focus on DCAF7 in the current study because this protein has already been reported as a partner of DYRK1A but its role

A

Mass spectrometry analysis of DYRK1A-binding partners

Protein name	Amino acids	Molecular weight	Peptides
DYRK1A	763	85.5	22
DCAF7	342	38.9	5
HSPA1A	641	70	5
HSPA8	646	70.9	4
HNRNPH1	449	49.2	3
HNRNPF	415	45.6	3

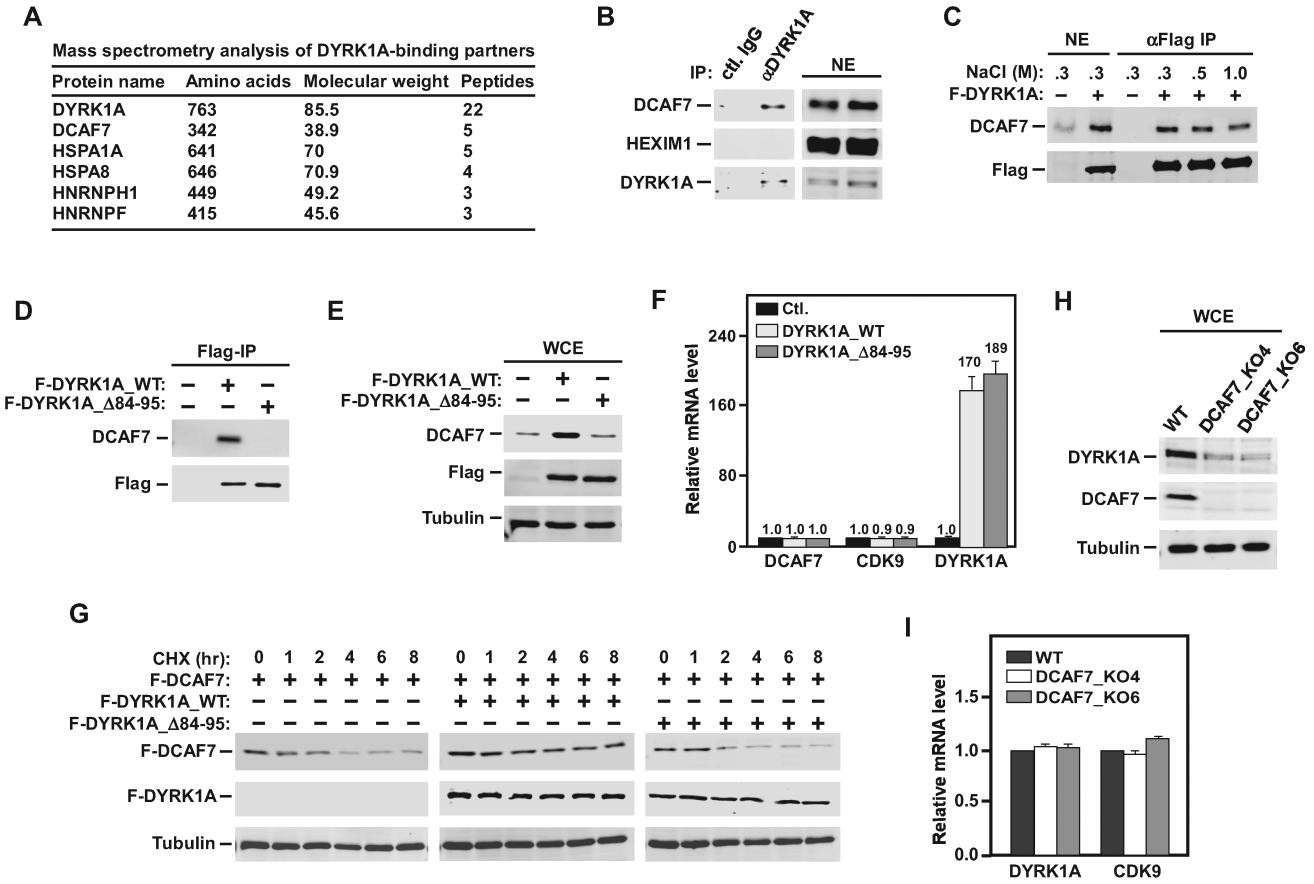

Figure 1. DYRK1A and DCAF7 form a nuclear complex that promotes stability of both proteins. (**A**) Analysis by mass spectrometry of proteins co-immunoprecipitated with F-DYRK1A from HeLa nucleus extracts (NE) with anti-Flag mAb beads. Only the proteins specifically isolated from the F-DYRK1A-expressing cells but not the cells containing an empty vector were shown. (**B**) NE of HeLa cells were subjected to immunoprecipitation (IP) with either the anti-DYRK1A antibody or non-specific rabbit total IgG as a negative control. The precipitates and the NE were examined by Western-blotting (WB) for the indicated proteins. (**C**) NE from Hela cells expressing F-DYRK1A were adjusted to the indicated NaCl concentrations and subjected to anti-Flag IP followed by WB. (**D**) NE from Hela cells expressing WT or Δ84–95 F-DYRK1A were analyzed by IP/WB as in B. All IP experiments were repeated at least three times and representatives are shown. (**E**, **F**) HeLa cells expressing WT or Δ84–95 F-DYRK1A were analyzed for the presence of the indicated proteins in whole cell extracts (WCE) by WB (E) or the indicated mRNAs by qRT-PCR (F), with the signals from the empty vector (Ctl.) cells set to 1. (**G**) HeLa cells co-transfected with F-DCAF7 plus WT or Δ84–95 F-DYRK1A were cultured for 40 hours and then treated with cycloheximide (CHX; 50 μg/ml) for the indicated hours. WCE were analyzed by WB for the indicated proteins. (**H**, **I**) WT 293T or the 293T-based DCAF7 KO clones were analyzed by WB (H) and qRT-PCR (I) as in E and F. Error bars in F and I represent mean ± standard deviations (SD).

in DYRK1A's transcriptional regulation remains unknown (28,29).

The specific interaction between endogenous DYRK1A and DCAF7 in HeLa nuclei was confirmed by performing anti-DYRK1A IP followed by Western-blotting (Figure 1B). The interaction was found to be very stable and no obvious decrease was detected even when 1.0 M NaCl was present in NE during the IP (Figure 1C). Consistent with the previous demonstrations that the DYRK1A N-terminal region between amino acids 84 and 95 is required for the DYRK1A–DCAF7 binding (29), the F-DYRK1A_Δ84–95 mutant lacking this region failed to bind DCAF7 (Figure 1D).

Interestingly, the overexpression of WT but not the Δ84–95 form of DYRK1A significantly elevated the protein but not mRNA level of endogenous DCAF7 (Figure 1E and F). Furthermore, pre-treating cells with cycloheximide (CHX) for various durations to inhibit new protein synthesis allowed WT but not Δ84–95 DYRK1A to prevent the degradation of existing DCAF7 (Figure 1G). On the other hand, when the DCAF7 gene in HEK293T cells was knocked out by CRISPR–Cas9 (Supplemental Figure S1), the protein (Figure 1H) but not mRNA level of DYRK1A (Figure 1I) significantly decreased. Together, these results reveal a stable interaction between DYRK1A and DCAF7, which is required to maintain stability of both binding partners.

DCAF7 is required for optimal kinase and transcriptional activities of DYRK1A

To determine whether the association with DCAF7 could affect DYRK1A's transcriptional activity, we adopted the Gal4-tethering luciferase system used by Di Vona *et al.* (15) to investigate the abilities of Gal4-fused WT and Δ84–95 DYRK1A to activate transcription of an HIV-1 LTR-luciferase reporter gene containing upstream Gal4-binding sites. While WT DYRK1A robustly activated luciferase expression, Δ84–95 was largely inactive in this assay (Fig-

ure 2A). Correlating with this difference, affinity-purified WT but not Δ84–95 F-DYRK1A efficiently phosphorylated GST-CTD containing all 52 repeats of the Pol II CTD on both the Ser2 and Ser5 positions as revealed by Western-blotting with the epitope-specific antibodies (Figure 2B).

To further investigate the role of DCAF7 in promoting DYRK1A's activity, we created two HeLa-based doxycycline (Dox)-inducible DCAF7 knockdown (KD) clones expressing short hairpin (sh)RNAs (shDCAF7-1 and shDCAF7-2) that target two separate regions of the DCAF7 mRNA. Compared to the Dox(-) conditions, Gal4-DYRK1A displayed decreased transcriptional activity once the KD of DCAF7 was induced in both clones (Figure 2C). Correlating with this decrease, F-DYRK1A immunoprecipitated from shDCAF7-2 cells showed significantly lower CTD kinase activity compared to the precipitate derived from normal, non-KD cells (Figure 2D; 50% more DYRK1A-expressing plasmid was transfected into the KD cells to make up for the reduced stability of DYRK1A in the absence of DCAF7). Similarly, recombinant GST-DYRK1A displayed relatively low CTD kinase activity unless F-DCAF7 affinity-purified from the DYRK1A KO cells (Supplemental Figure S2) was also present in the reaction, which resulted in a marked enhancement of CTD phosphorylation (Figure 2E). Together, these results confirmed from multiple angles that DCAF7 is required for the efficient CTD-kinase and transcriptional activities of DYRK1A.

DCAF7 binds to RNA Pol II independently of DYRK1A and promotes DYRK1A-Pol II interaction

To test the hypothesis that DCAF7 promotes the DYRK1A activities through facilitating the interaction of DYRK1A with Pol II, we first performed a co-IP followed by Western analysis and found that in addition to DYRK1A, transfected F-DCAF7 also pulled down Pol II but not the P-TEFb subunits CDK9 and CycT1 or the P-TEFb regulator HEXIM1 (Figure 3A). The interaction between endogenous DCAF7 and Pol II was also confirmed by performing anti-DCAF7 or anti-Pol II co-IP/Western analyses (Figure 3B and C). Importantly, the DCAF7-Pol II interaction was independent of DYRK1A as the CRISPR–Cas9-mediated DYRK1A KO did not affect the interaction (Figure 3D).

Another piece of information that could be deduced from the above data is that the DCAF7-DYRK1A complex is physically separate from any CDK9-containing complexes. This notion was further supported by the anti-F-DYRK1A co-IP/Western analysis, which shows that F-DYRK1A pulled down DCAF7 and Pol II but not CDK9 and CycT1 (Supplemental Figure S3A).

Given the salt-resistant binding between DCAF7 and DYRK1A (Figure 1C) and the above observations that DCAF7 could also interact with Pol II independently of DYRK1A, we next asked whether DCAF7 promotes the interaction of DYRK1A with Pol II. Indeed, compared to WT DYRK1A, DYRK1A_Δ84–95, which failed to bind to DCAF7, displayed a markedly decreased interaction with Pol II (Figure 3E). Furthermore, upon the Dox-induced KD of DCAF7, less Pol II was found to co-precipitate with F-DYRK1A (Figure 3F). *In vitro*, Flag-DYRK1A affinity-purified from DCAF7 KO cells pulled down only 37% of GST-CTD compared to 100% bound to Flag-DYRK1A isolated from WT cells (Supplemental Figure S3B). Together, these results indicate that DCAF7, a WD-repeats-containing protein, promoted DYRK1A's kinase and transcriptional activities likely through serving as a scaffold to tether DYRK1A to its substrate Pol II.

Notably, the above data reveal that the absence of DCAF7 decreased but not completely eliminated the DYRK1A-Pol II binding. Given our recent finding that the DYRK1A histidine-rich region (HRD) also contributes to the binding (16), we created a double deletion mutant, F-DYRK1A_Δ84–95&Δ566-621, which removes both the DCAF7-binding and the HRD regions, and investigated its Pol II-binding ability by co-IP/Western. The result shows that the combined deletions indeed further disrupted binding of DYRK1A to Pol II (Figure 3G). Notably, the double deletion mutant likely still retained folding in the kinase domain as indicated by its nearly WT ability to autophosphorylate to produce an ATP-dependent upshift (Figure 3H). Together, these results indicate that the DYRK1A-Pol II binding through the HRD alone is weak and that DCAF7 greatly enhances this binding.

DYRK1A and DCAF7 are required for optimal myoblast differentiation and expression of key myogenic genes

With the above demonstrations that DCAF7 promoted DYRK1A's activities in vitro and in HeLa and 293T cells, we would like to extend the analysis to a different cell system that can display induced and well-coordinated physiological changes. Previously, DCAF7 and DYRK1B, a DYRK1A homologue, have been separately reported as required for the expression of key genes essential for the differentiation of C2C12 myoblast cells into myotubes. In light of these findings, we set out to determine whether DYRK1A exerts a similar control in this process and whether DCAF7 also upregulates DYRK1A's activities as it does in HeLa and 293T cells.

We first examined by Western-blotting and qRT-PCR the expression of DYRK1A, DCAF7 and MYH2, a key myoblast differentiation marker gene that encodes the motor protein myosin heavy chain 2, in C2C12 cells before and after the induction of differentiation (Materials and Methods). Comparing to undifferentiated cells kept in the growth medium (GM), both the protein and mRNA levels of DYRK1A and MYH2 gradually increased in cells cultured in the differentiation medium (DM) for various days (Figure 4A–C). In contrast, only the protein but not mRNA level of DCAF7 showed a significant increase in this process (Figure 4A and B), while the expressions of two control proteins CDK9 and α-Tubulin remained unchanged. It was shown earlier that the interaction with DYRK1A markedly stabilized DCAF7 (Figure 1). Thus, the elevated DYRK1A protein level in differentiated C2C12 cells likely allowed more DCAF7 to become protected once the two formed a stable complex (confirmed by anti-DCAF7 IP followed by Western blotting in Figure 4D), which can explain the increased protein but not mRNA level of DCAF7.

Next, the requirement of DYRK1A and DCAF7 for optimal myogenesis was investigated first by immunofluo-

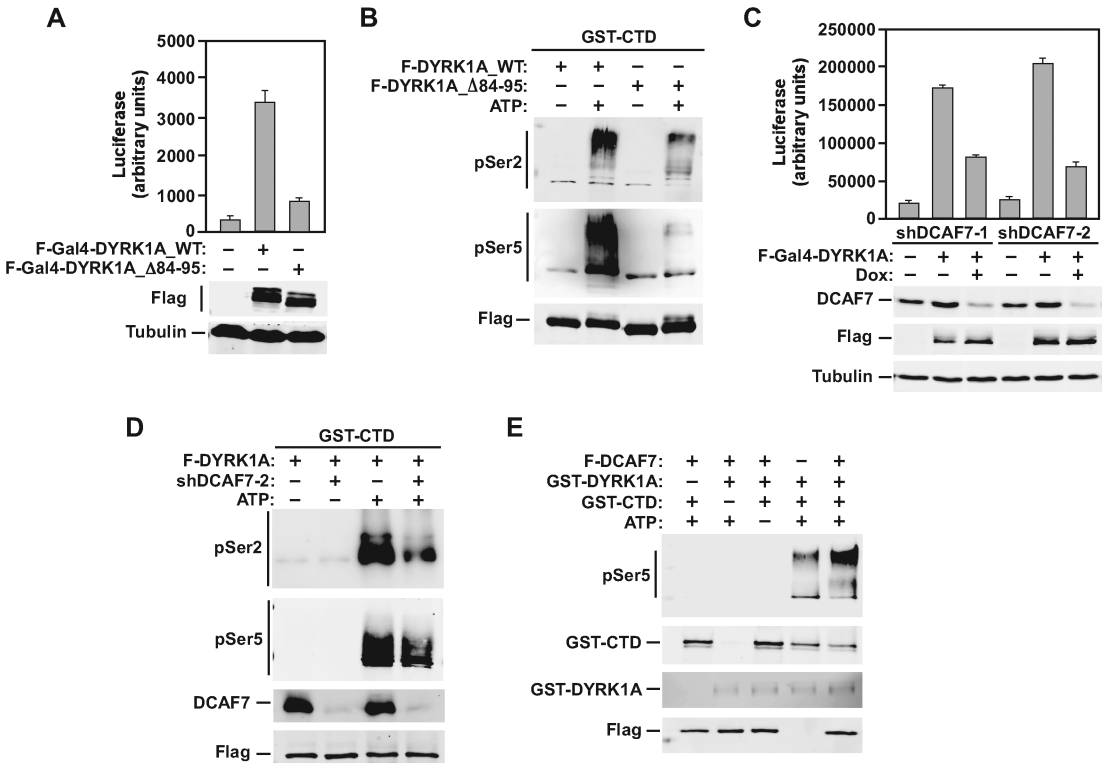

Figure 2. DCAF7 is required for optimal kinase and transcriptional activities of DYRK1A. (**A**) HeLa cells were co-transfected with the HIV-1 LTR-luciferase reporter construct containing six copies of the Gal4 upstream activation sequence (UAS) and the plasmid expressing the indicated Gal4-DYRK1A fusion. Luciferase activities in cell extracts were measured. Error bars represent mean ± SD from three independent experiments. (**B**) Flag-tagged WT and Δ84–95 DYRK1A were affinity-purified from NE of transfected 293T cells and analyzed in kinase reactions containing GST-CTD as the substrate. The reaction products were examined by WB for the Ser2- and Ser5-phosphorylated CTD. (**C**) Two independent and doxycycline (Dox)-inducible DCAF7 KD clones expressing the indicated shRNAs were co-transfected with the F-Gal4-DYRK1A-expressing plasmid (50% more was used for the Dox-plus conditions) and the 6× Gal4 UAS-HIV-1 LTR-luciferase reporter construct. Luciferase activities were measured and analyzed as in A. (**D**) WT F-DYRK1A affinity-purified from either the parental HeLa cells or the HeLa-based shDCAF7-2 KD cells were tested in kinase reactions as in B. (**E**) Recombinant GST-DYRK1A (Life Technologies) was incubated with or without F-DCAF7, which was affinity-purified from the 293T-based DYRK1A KO cells, in kinase reactions containing GST-CTD as the substrate. The reaction products were analyzed as in B. All in vitro kinase assays were repeated at least three times and representatives are shown.

rescence staining of MYH2 in engineered C2C12 cells, in which the expression of DYRK1A or DCAF7 was silenced by two different shRNAs that target separate regions of the relevant genes. Compared to the control cells expressing a scrambled shRNA, silencing either DYRK1A or DCAF7 with the specific shRNAs largely prevented the cells from forming fused and elongated myotubes containing multiple nuclei (Figure 4E and F) and at the same time decreased the number of MYH2(+) cells in the population (Figure 4G and H). Furthermore, the Western-blotting and qRT-PCR analyses indicate that the KD markedly decreased the MYH2 expression at both the protein (Figure 4I and J) and mRNA levels (Supplemental Figure S4).

In addition to MYH2, the differentiation-induced mRNA production from two other myogenic genes MYOG (Myogenin), which encodes a muscle-specific transcription factor involved in the coordination of myogenesis (30), and CAV3 (Caveolin 3), which encodes the principal structural protein component of caveolae membrane domains in skeletal muscle cells (31), also depended on DYRK1A and DCAF7 (Supplemental Figure S4). Finally, by measuring the luciferase gene expression driven by the MYOG promoter and activated by the master transcriptional regulator of skeletal myogenesis MYOD (32), DYRK1A and DCAF7 were also found to be essential for the MYOD-dependent MYOG promoter activity in DM (Figure 4K and L). Collectively, these data indicate that both DYRK1A and DCAF7 are required for proper myoblast differentiation and expression of key myogenic genes.

DYRK1A–DCAF7 interaction is required for efficient expression of key myogenic genes

To determine whether the specific interaction between DYRK1A and DCAF7 is required for the efficient expression of the three myogenic genes analyzed above, we tested the abilities of ectopically expressed WT and DYRK1A_Δ84–95, which failed to bind to DCAF7, to rescue the expression of these genes in the DYRK1A KD cells. When expressed to a similar level (Figure 5A), only WT but not Δ84–95 DYRK1A was able to enhance the MYH2, CAV3 and MYOG mRNA production in DM to levels that were significantly above those in control KD cells containing an empty vector (Figure 5B–D). Furthermore, the simple overexpression of DYRK1A_Δ84–95 in normal C2C12

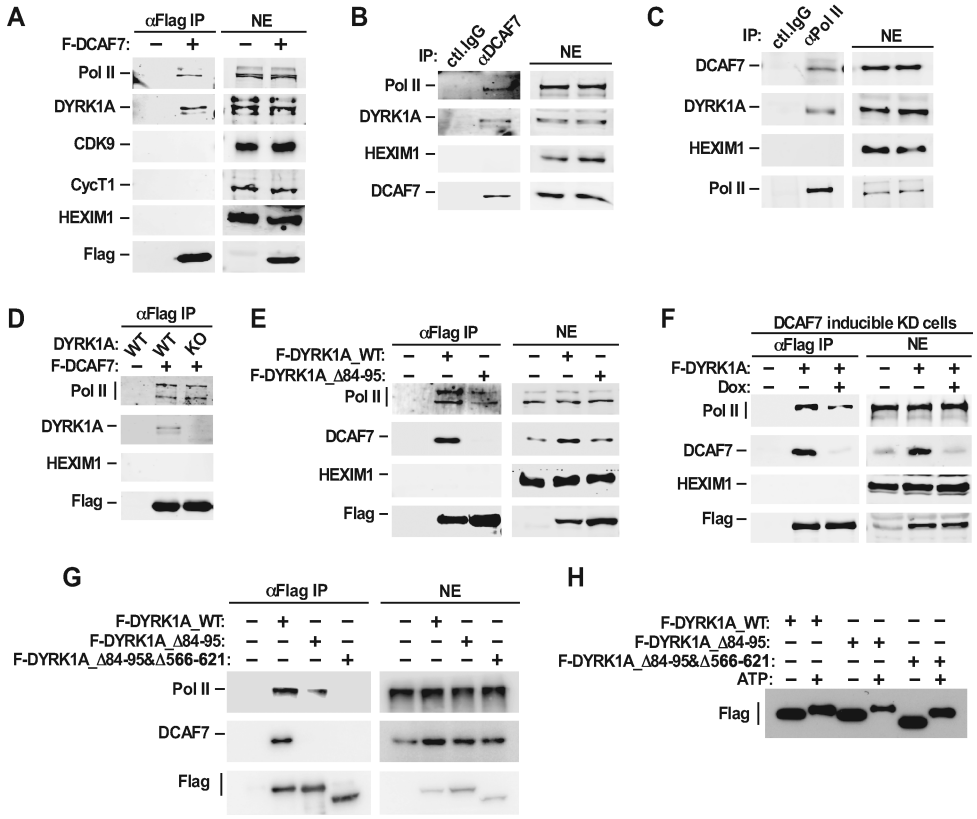

Figure 3. DCAF7 interacts with RNA Pol II independently of DYRK1A and is required for the DYRK1A-Pol II interaction. (**A**–**C**) NE and anti-Flag (A), anti-DCAF7 (B) or anti-Pol II (C) immunoprecipitates (IP) derived from NE of normal 293T cells (B, C) or 293T cells transfected with F-DCAF7 (A) were analyzed by WB to detect the indicated proteins. (**D**) WT 293T or the 293T-based DYRK1A KO cells were transfected with the F-DCAF7-expressing plasmid. Anti-Flag IP from NE of these cells were analyzed by WB as in A. (**E**) NE and the anti-Flag IP from 293T cells transfected with the WT or Δ84–95 F-DYRK1A-expressing plasmid were analyzed by WB as in A. (**F**) The 293T-based inducible DCAF7 KD cells expressing F-DYRK1A were treated with DMSO or Dox to induce the KD. NE and anti-Flag IP derived from NE were examined by WB as in A. (**G**) NE and anti-Flag IP from 293T cells transfected with WT or Δ84–95 F-DYRK1A, or Δ84–95&Δ566-621 F-DYRK1A-expressing plasmid were analyzed by WB. All IP experiments were repeated at least three times. (**H**) The anti-Flag IP examined in G were incubated with or without ATP in kinase reactions and analyzed by SDS-PAGE followed by WB.

cells to a level about 7-fold higher than that of endogenous DYRK1A (Figure 5E) also produced a dominant-negative effect on the expression of the three myogenic genes (Figure 5F–H). Together, these results indicate that optimal myoblast differentiation and expression of key myogenic genes not only require DYRK1A and DCAF7 but also their stable interaction.

DYRK1A–DCAF7 complex exerts direct transcriptional control of key myogenic genes through binding and phosphorylating Pol II at the gene loci

Having established that the DYRK1A–DCAF7 complex acts as a Pol II CTD kinase and promotes mRNA production from key myogenic genes, we hypothesized that this complex is directly involved in the transcriptional control of these genes. Indeed, in cells kept in DM, our ChIP-qPCR analysis detected robust DCAF7 (from 38.1- to 61-fold enrichment) and DYRK1A (from 5.3- to 7.3-fold enrichment) signals at the promoter regions of MYH2, CAV3 and MYOG that showed a significant enrichment over the signals obtained with a control, non-specific IgG (Figure 6A and B). In contrast, <2.5-fold enrichment was observed for the control genes SOX2 and IgH under the same conditions. We also noticed that in undifferentiated C2C12 cells in GM, even though the expression of the three myogenic genes was very low compared to that in DM (Supplemental Figure S4 and Figure 4I–J), substantial levels of DCAF7 (from 21.8- to 28-fold enrichment), and to a lesser degree, DYRK1A (1.7- to 3.0-fold enrichment) were already present at the gene promoters (Figure 6A & B), suggesting that their mere presence was insufficient to induce robust transcription in GM.

Consistent with our earlier observations that DCAF7 was required for DYRK1A to interact with Pol II and stimulate transcription, a significant reduction in the DYRK1A ChIP signals was detected at the three myogenic gene promoters after DCAF7 expression was knocked down by the specific shRNA (Figure 6C). Furthermore, the Ser5-phosphorylated Pol II (pSer5) also displayed a marked decrease at the promoters in the DCAF7 KD cells (Figure 6D), correlating with the reduced myogenic gene transcription under these conditions (Supplemental Figure S4E–H). Finally, by conducting ChIP-qPCR analysis with multiple pairs of primers across the MYH2 gene locus (Figure 6E), we noticed that DYRK1A, DCAF7, as well as the target

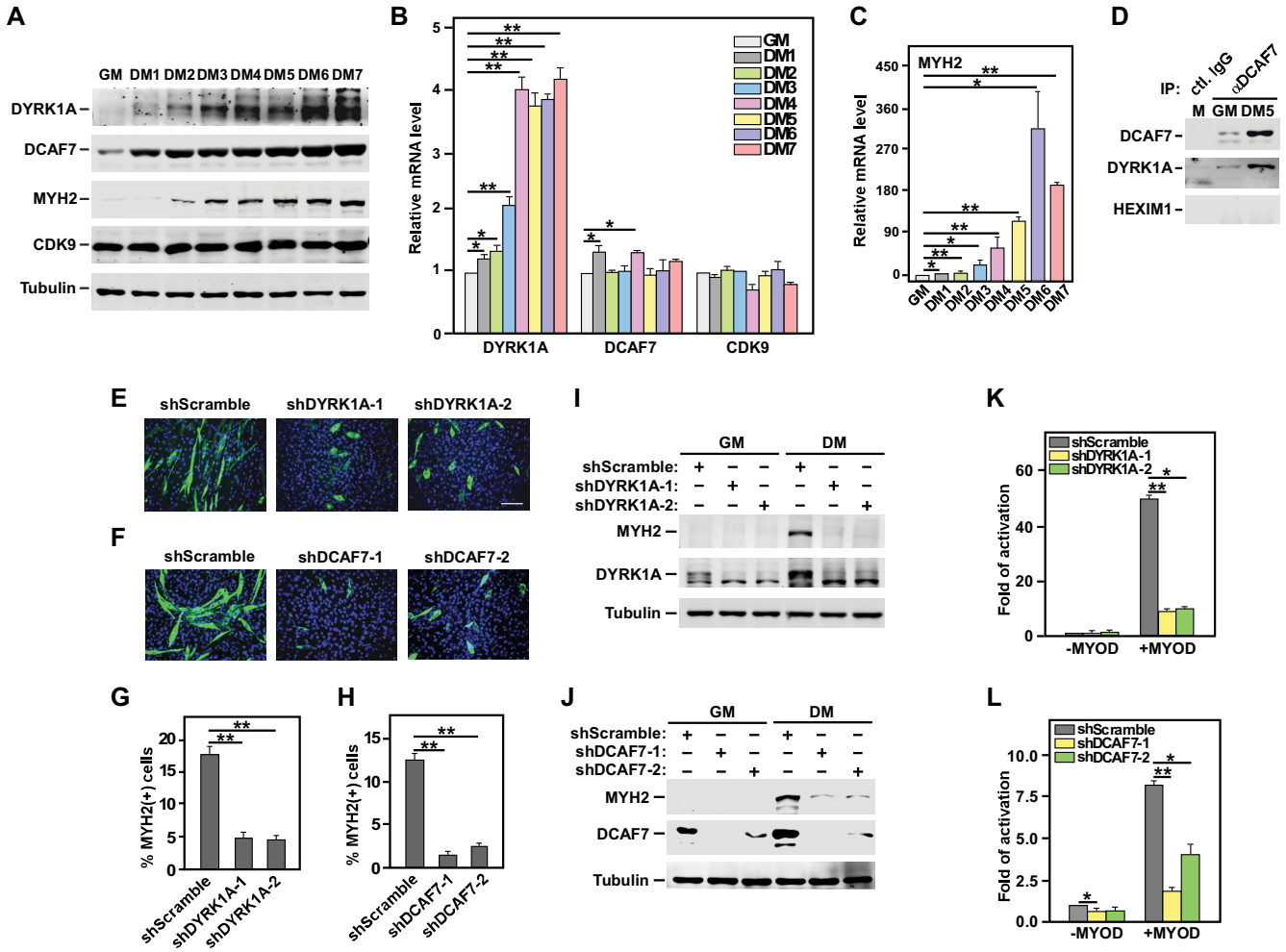

Figure 4. DYRK1A and DCAF7 are required for optimal myoblast differentiation and expression of key myogenic genes. (**A**) The levels of the indicated proteins in 1×10^5 C2C12 cells grown in the growth medium (GM) or differentiation medium for 1-7 days (DM1-7) were examined by WB. (**B**, **C**) The mRNA levels of the indicated genes in C2C12 cells cultured in GM or DM1-7 were measured by qRT-PCR and shown. (**D**) Extracts of C2C12 cells grown in GM, DM5, or containing an equal mixture of the two populations (M) were subjected to IP with the anti-DCAF7 Ab or a non-specific control IgG. The precipitates were analyzed by WB for the indicated proteins. (**E**, **F**) C2C12 cells stably expressing the indicated shRNAs were cultured in GM for 5 days and subjected to immunofluorescence staining with the mouse anti-MYH2 mAb plus goat anti-mouse AF680 (green). Nuclei are counterstained with DAPI (blue). Scale bars represent 100 μm. (**G**, **H**) The percentages of MYH(+) C2C12 cells expressing the indicated shRNAs were counted and shown. (**I**, **J**). The expression of the proteins marked on the left in cells expressing the indicated shRNAs and grown in GM or DM for 5 days was detected by WB. (**K**, **L**) The MYOG promoter-driven luciferase reporter construct was co-transfected with the indicated shRNA- and/or the MYOD-expressing plasmids. After 48 hr of transfection, luciferase activities were measured in the cell extracts. All error bars represent mean ± SD from at least three independent experiments/measurements. The asterisks indicate different levels of statistical significance as calculated by two-tailed Student's *t*-test. $***P < 0.001$; $**P < 0.01$; $*P < 0.05$.

of the DYRK1A–DCAF7 complex, RNA Pol II, displayed nearly identical binding profiles (Figure 6F–H). In comparison, the unrelated BRD4 protein showed a very different distribution (Figure 6I). Taken together, these results implicate a direct control of the myogenic gene transcription by the DYRK1A–DCAF7 complex through binding and phosphorylating Pol II at these gene loci.

CDK9 is inhibitory to myoblast differentiation and expression of key myogenic genes

Given that DYRK1A and CDK9 are both CTD kinases that display striking similarities as revealed in our recent studies (6,16), we next asked whether CDK9, a well-known master regulator of eukaryotic transcriptional elongation (5), could play a role similar to that of DYRK1A during myoblast differentiation. We first determined whether the shRNA-mediated KD of CDK9 would affect the ability of C2C12 to differentiate into myotubes and express key myogenic genes. Surprisingly, the KD by using either of the two shRNAs that target different regions of CDK9 (KD efficiency analyzed at the protein and mRNA level in Figure 7B and Supplemental Figure S5, respectively) produced a stimulatory effect to promote differentiation and production of the MYH2 protein (Figure 7A and B) and mRNA (Figure 7D) in DM. Furthermore, the KD also increased the MYOG promoter-driven luciferase expression especially under the MYOD(+) conditions (Figure 7C). Finally, the KD was found to enhance the mRNA produc-

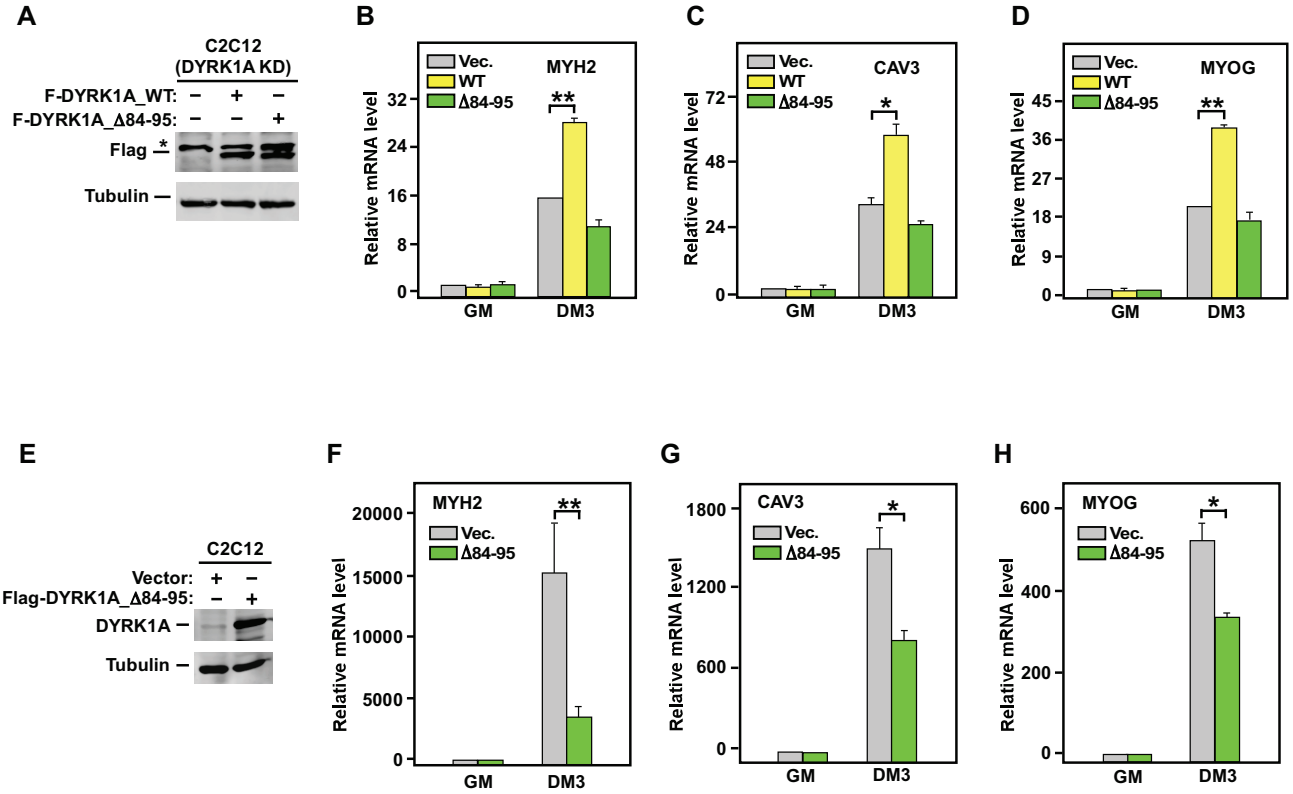

Figure 5. DYRK1A–DCAF7 interaction is essential for efficient expression of key myogenic genes. (**A**) The DYRK1A KD C2C12 cells were transfected with an empty vector (-) or the vector expressing WT or Δ84–95 F-DYRK1A and then analyzed by WB for the indicated proteins. A non-specific band is denoted with an asterisk. (**B**–**D**) The cells in A were cultured in GM or DM for 3 days and the mRNAs produced from the indicated genes were detected by qRT-PCR and shown. (**E**) WB was used to assess the levels of endogenous DYRK1A in normal C2C12 cells and the overexpressed Δ84–95 F-DYRK1A in transfected cells. (**F**–**H**) The cells in E were grown in GM and DM for 3 days and the mRNAs produced from the indicated genes were detected by qRT-PCR and shown. All error bars represent mean +/− SD from three independent measurements. The asterisks indicate different levels of statistical significance as calculated by two-tailed Student's t-test. $***P < 0.001$; $**P < 0.01$; $*P < 0.05$.

tion from two additional myogenic genes CAV3 and MYOG (Figure 7E and F).

Loss of P-TEFb and two positive P-TEFb complexes during myoblast differentiation causes myogenic genes to rely on DYRK1A–DCAF7 for transcription

If CDK9 inhibits the myogenic gene expression, how does myoblast differentiation proceed in the presence of the interfering P-TEFb? To answer this question, we first performed an IP/Western analysis to examine the abilities of CDK9 to pair with its cyclin partner CycT1 to form P-TEFb as well as to bind to several well-known P-TEFb regulators before and after C2C12 cells were induced to differentiate. Unexpectedly, as the cells spent progressively longer time in DM, there was a gradual but significant decrease in the amount of CycT1 bound to CDK9, although the total CycT1 level in the whole cell extracts (WCE) did not show much change during this process (Figure 7G). Consistent with the loss of the CDK9-CycT1 heterodimer in the differentiated cells, CDK9 immunoprecipitated from WCE displayed a markedly decreased kinase activity toward GST-CTD (Figure 7H).

In addition to the loss of P-TEFb, the level of BRD4, a bromodomain protein that binds and recruits P-TEFb to chromatin templates through recognizing the acetylated histone tails (10,11), also gradually decreased in WCE after cells were shifted into DM, leading to a lower level of the BRD4-P-TEFb complex in differentiated cells (Figure 7G). Furthermore, AFF4, the scaffolding subunit that assembles P-TEFb and other proteins into the multisubunit Super Elongation Complex (SEC) (33), also showed a substantially decreased binding to P-TEFb upon differentiation (Figure 7G). In contrast to CycT1, BRD4 and AFF4, the associations with CDK9 by HEXIM1 and LARP7, two components of the 7SK snRNP that sequesters the cellular excess P-TEFb into an inactive form (5), were little affected by the differentiation process (Figure 7G). Together, these data indicate that during myoblast differentiation, there was a substantial reduction in the cellular levels of P-TEFb and the two positive P-TEFb-containing complexes, the BRD4-P-TEFb and the SEC.

Consistent with the above observations, robust CDK9 signals were detected by ChIP-qPCR at the MKI67 gene locus, which is strongly associated with cell proliferation, only when C2C12 cells were cultured in GM but not DM (Figure 7I). On the other hand, even though the three myogenic genes MYH2, CAV3 and MYOG were highly expressed under the DM conditions, only low levels of CDK9 were detected at these gene loci irrespective of the differentiation

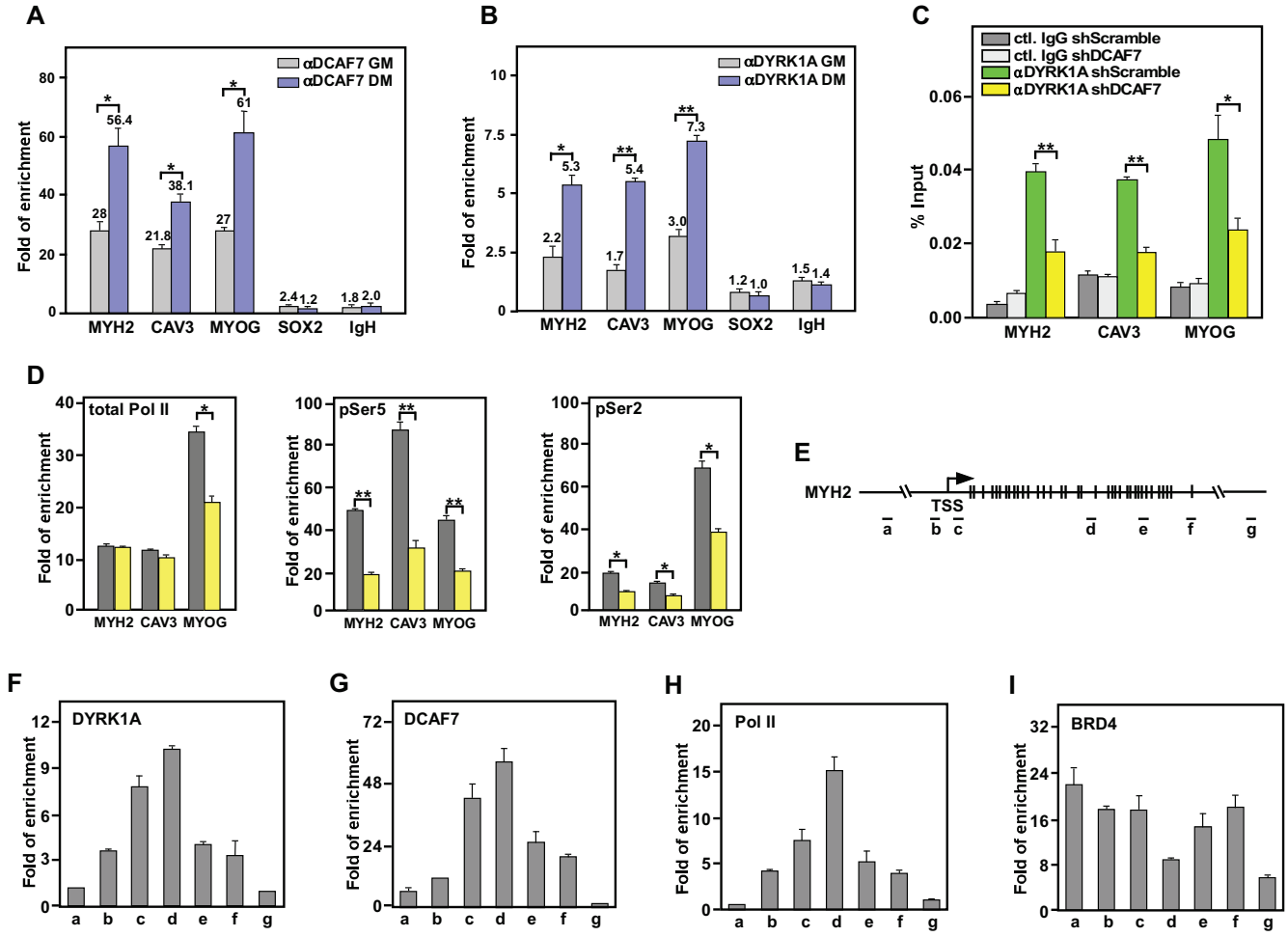

Figure 6. DYRK1A–DCAF7 complex exerts direct transcriptional control of key myogenic genes through binding and phosphorylating Pol II at the gene loci. (**A**, **B**) ChIP-qPCR analyses were conducted to assess the occupancy of DCAF7 (A) and DYRK1A (B) at the indicated gene promoters in C2C12 cells cultured in GM or DM for 3 days. The ChIP-qPCR signals were normalized to the input DNA, divided by the values obtained with the rabbit total IgG in control reactions and shown. (**C**, **D**) C2C12 cells stably expressing the indicated shRNAs were grown in DM for 3 days and then analyzed by ChIP-qPCR to detect the occupancy of DYRK1A (C), total Pol II (D), and the Ser5- or Ser2-phosphorylated Pol II (pSer5 or pSer2; D) at the indicated gene loci. Rabbit total IgG was used as a negative control. (**E**) A diagram showing the positions of the primer sets used in ChIP-qPCR reactions in F-I to map the bindings across the MYH2 gene locus. TSS: transcription start site. The vertical lines/blocks represent exons. (**F**–**I**) C2C12 cells in DM for 3 days were subjected to ChIP-qPCR analysis using the primer sets in E to detect the bindings of DYRK1A (F), DCAF7 (G), total Pol II (H) and BRD4 (I) to the various positions (a to g) across the MYH2 gene locus. The ChIP-qPCR signals were normalized to the input DNA and divided by those obtained with the rabbit total IgG in control reactions. All error bars represent mean ± SD from three independent reactions. The asterisks indicate different levels of statistical significance as calculated by two-tailed Student's t-test. $***P < 0.001$; $**P < 0.01$; $*P < 0.05$.

status of the cells. In addition, for a gene that is not normally expressed in a non-immune cell type such as C2C12, low levels of CDK9 were detected at the IgH locus both before and after the induction of myoblast differentiation (Figure 7I).

Lending further support to the notion that P-TEFb is more important in undifferentiated, highly proliferative C2C12 cells, whereas DYRK1A–DCAF7 is critical after the induction of differentiation, the shRNA-mediated KD of CDK9 but not DYRK1A significantly inhibited the growth of C2C12 cells in GM as indicated by the decreased doubling time of the CDK9 KD but not DYRK1A KD cells (Figure 7J and K). Taken together, these results suggest that CDK9 inhibits myoblast differentiation, which must be suppressed through the disruption of P-TEFb and positive P-TEFb complexes in differentiated cells. This in turn likely allows the DYRK1A–DCAF7 complex to take over and activate myogenic gene transcription through hyperphosphorylating the Pol II CTD.

DISCUSSION

The interaction between DYRK1A and DCAF7 was first reported in 2011, and its then proposed function was to facilitate the nuclear entry of DCAF7 by the associated DYRK1A (28). A subsequent study in 2016 showed that DCAF7 can also serve to recruit DYRK1A and another protein kinase HIPK2 to their substrate the adenovirus E1A protein (29). In the current study, we provide additional evidence in support of DCAF7's substrate recruitment function by demonstrating that this protein is also required for DYRK1A to target RNA Pol II so that

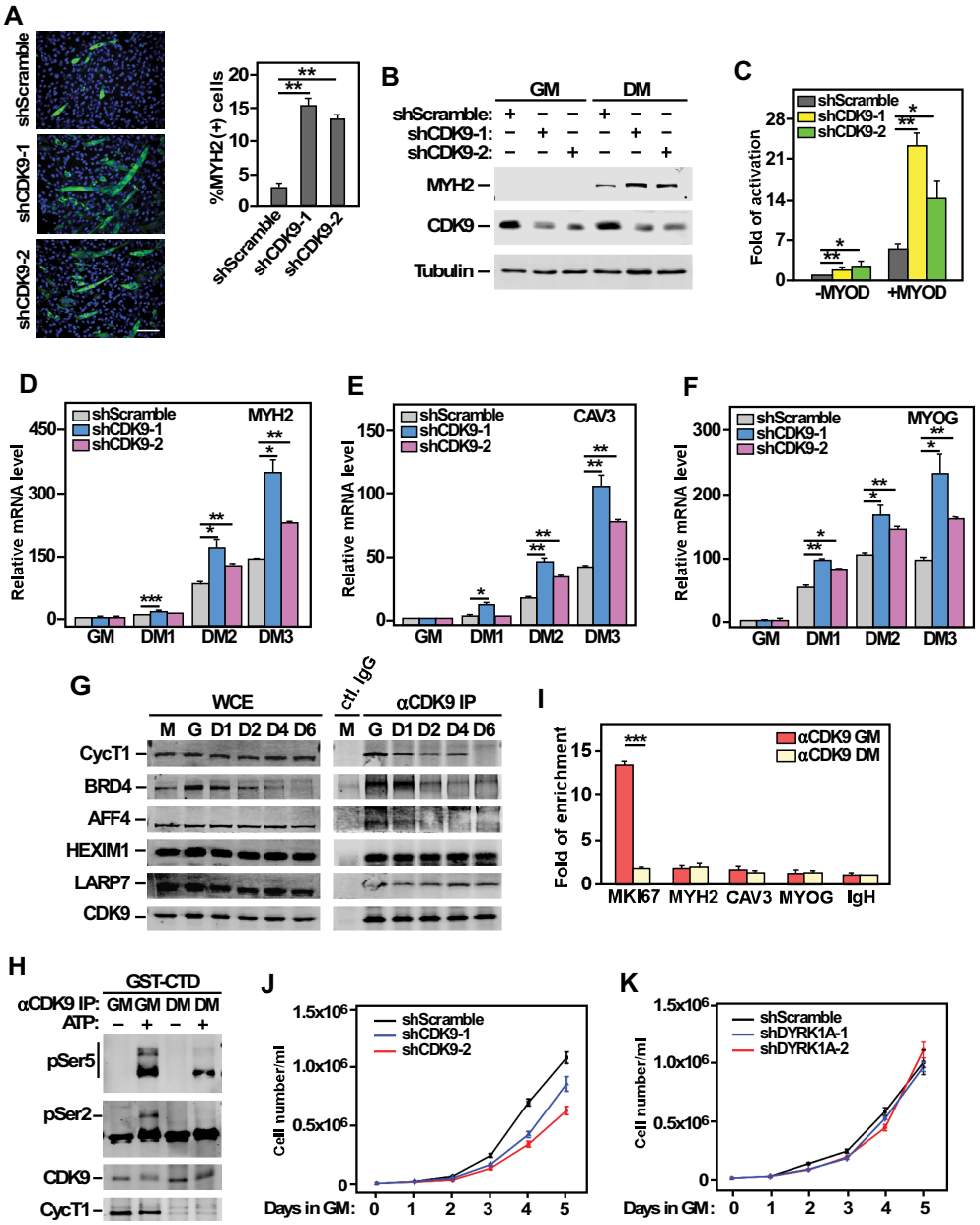

Figure 7. CDK9 inhibits expression of key myogenic genes and is inactivated during myoblast differentiation to likely allow DYRK1A–DCAF7 to activate transcription of these genes. (**A**) Left: C2C12 cells stably expressing the indicated shRNAs were cultured in DM for 4 days and subjected to anti-MYH2 immunofluorescence staining as in Figure 4E. Right: The percentages of MYH2(+) cells detected by immunofluorescence were counted and shown. (**B**) Analysis by WB of the indicated proteins in cells expressing the various shRNAs and cultured in GM or DM for 4 days. (**C**) The MYOG promoter-driven luciferase reporter construct was co-transfected with the indicated shRNA- and/or MYOD-expressing plasmid into C2C12 cells. Luciferase activities in cell extracts were measured and shown. (**D**–**F**) qRT-PCR analyses were conducted to assess the expression of MYH2 (D), CAV3 (E) and MYOG (F) in C2C12 cells expressing the indicated shRNAs and cultured in GM or DM for the various days. (**G**) WCE of C2C12 cells cultured in GM (labeled as G) or DM (labeled as D) for various days were subjected to IP with either the anti-CDK9 antibody or a non-specific control IgG. The IP products and WCE were examined by WB. M represents an equal mixture of the GM and DM4 WCE. (**H**) Anti-CDK9 immunoprecipitates from cells grown in GM or DM for 6 days were incubated with GST-CTD in kinase reactions. The reaction mixtures were examined by WB. (**I**) ChIP-qPCR analysis was conducted to assess the occupancy of CDK9 at the indicated gene promoters in cells cultured in GM or DM for 3 days. The signals were processed as in Figure 6A. (**J**, **K**) C2C12 cells expressing the various shRNAs were grown in GM for the indicated number of days and the cell concentrations were determined and shown. All error bars represent mean ± SD from three independent measurements. The asterisks indicate different levels of statistical significance as calculated by two-tailed Student's *t*-test. ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$.

DYRK1A can hyperphosphorylate the CTD and stimulate transcription of key myogenic genes. In addition, we have identified another function of the DYRK1A–DCAF7 interaction in promoting the stability of the two binding partners.

DCAF7 is a WD40-repeat protein and members of this family often use the WD domain to fold into a β-propeller structure that serves as a molecular scaffold to assemble various components into multi-subunit complexes for diverse regulations (34). The demonstrated substrate recruitment function of DCAF7 that enables DYRK1A to target Pol II and E1A, as well as the previously revealed role of DCAF7 in tethering several kinases together (35), are in agreement with the predicted function of a WD40-repeat protein.

It is worth noting that DCAF7 has previously been speculated to play a role in transcription, although the precise manner by which it contributes to this process was unknown (24). The present discovery that DCAF7 is essential for DYRK1A to bind and phosphorylate Pol II reveals the mechanism used by this protein to regulate transcription and also explains the purpose of DCAF7's nuclear entry that is mediated by DYRK1A (28).

The genome-wide analysis conducted by Di Vona *et al.* (15) in T98G and HeLa cells reveal that the genes targeted by DYRK1A usually contain a promoter-proximal TCTCGCGAGA motif required for transcriptional activation. In the present study, we have presented multiple lines of evidence indicating that the three myogenic genes, MYH2, CAV3 and MYOG, are likely directly controlled at the transcription level by DYRK1A–DCAF7 in differentiating C2C12 cells. However, examination of the promoter-proximal sequences and upstream regions up to 3000-bp away from the transcription start site (TSS) indicates that the three genes lack the reported consensus motif, suggesting that they are likely targeted by DYRK1A–DCAF7 through a different mechanism or a yet-to-be-determined myogenic specific factor.

Indeed, the lack of a distinctive and prominent DYRK1A enrichment peak at the MYH2 promoter (Figure 6F-I), the co-localization of DYRK1A–DCAF7 with Pol II across the MYH2 gene locus (Figure 6F-I), and the dependence on DCAF7 for DYRK1A to occupy the three myogenic gene loci (Figure 6C) agree with the model that the association of DYRK1A with the three genes is mediated by the binding of DCAF7 to the elongating Pol II. The idea that DYRK1A may use different gene-targeting modes to regulate transcription in different cell types and/or in response to distinct signals is further supported by a very recent study showing that DYRK1A interacts with histone acetyl transferases P300 and CBP to localize to gene enhancers and influence transcription in T98G cells (36).

The transcriptional control of myogenesis has been studied extensively with an initial main focus on the central roles of MYOD and other members of the MYOD family (MYF5, MRF4 and MYOG) that serve as master regulators of this process (19,20,32). Subsequent studies have revealed a far more complicated regulatory network and mechanism that involve additional key transcription factors, chromatin-modifying activities, microRNAs and non-coding RNAs acting upstream and/or in conjunction with the MYOD family of proteins to further control the myogenic transcriptional program (18). Although the present data have implicated a critical role of the DYRK1A–DCAF7 complex in the transcriptional control of myogenesis, the mere association of this complex with a myogenic gene locus is not enough to turn on transcription of the gene (see Figure 6A and B). It is likely that the complex occupies a fairly downstream position to interpret and integrate the signals sent out by the upstream network of regulators so that the signals can be properly executed to result in robust Pol II phosphorylation and transcriptional activation. Future studies are necessary to reveal the functional relationship and interactions between DYRK1A–DCAF7 and the rest of the myogenic transcription network to better understand how the signals can be accurately and efficiently transduced during myogenesis.

The general transcription factor P-TEFb is considered a master regulator of Pol II elongation. However, our data show that it plays an inhibitory role during myoblast differentiation, and that likely due to this inhibition, the level of the CDK9-CycT1 heterodimer as well as the two positive P-TEFb complexes, the Brd4-P-TEFb and the SEC, were found to be reduced in differentiated cells to enable strong myogenic transcriptional activation by DYRK1A–DCAF7. Our findings are consistent with a recent report showing that CDK9 is involved in epigenetic silencing of tumor suppressor genes and that CDK9 inhibition restores expression of these genes and promotes differentiation of cancer cells (37), suggesting that CDK9 likely plays an inhibitory role during cell differentiation.

Notably, an earlier report that relied almost exclusively on transient overexpression and *in vitro* experiments has come to the opposite conclusion by showing that the CDK9–CycT2a heterodimer binds to and phosphorylates MyoD to promote myogenesis (38). Contrary to these findings, we failed to detect a stable interaction between CDK9 and MyoD in vivo, and moreover, we found that the endogenous T2a levels were undetectable by standard Western-blotting both before and after the differentiation (data not shown).

Although much is yet to be learnt about how CycT1 is downregulated during C2C12 differentiation and whether CDK9 may gain a new partner in this process, our results are consistent with a model that a switch from a general CTD kinase functioning in proliferative myoblast to a myogenesis-specific CTD kinase working in terminally differentiated myotubes likely occurs during myoblast differentiation. This is reminiscent of the switch from the canonical core promoter recognition complexes and chromatin remodelers to specialized transcription machineries during several differentiation processes (PMID: 23326641, 24082143, 15774719, 17920018). Switching of core transcriptional machinery is thus probably used more frequently by organisms as a simple yet effective mechanism to control cell type- and developmental stage-specific gene expression than previously thought.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Rongdiao Liu for technical assistance with an auto-phosphorylation experiment, Dr Susana de la Luna and Dr Chiara Di Vona of the Center for Genomic Regulation in Spain for valuable advice on conducting the DYRK1A *in vitro* kinase assay, Dr Stephen J. Tapscott of Fred Hutchinson Cancer Research Center and Dr Robert M. Nissen of California State University for providing the expression or reporter plasmids.

Authors Contributions: D.Y. performed most of the experiments and analyzed the data. C.C. performed part of the ChIP experiments and provided technical support with the ChIP-qPCR and lentiviral infection experiments. Y.X. helped with the muscle cell differentiation experiments. D.Y. and Q.Z. conceived and designed experiments and wrote the manuscript. Q.Z. provided support and supervised the project.

FUNDING

National Institutes of Health [R01AI041757 to Q.Z.]; Tang Distinguished Scholarship (to D.Y.); National Key R&D Program of China [2018YFA0107303] and National Natural Science Foundation of China [81672955 to Y.X.].

Conflict of interest statement. None declared.

REFERENCES

1. Hsin,J.P. and Manley,J.L. (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.*, **26**, 2119–2137.
2. Phatnani,H.P. and Greenleaf,A.L. (2006) Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.*, **20**, 2922–2936.
3. Corden,J.L. (2013) RNA polymerase II C-terminal domain: tethering transcription to transcript and template. *Chem. Rev.*, **113**, 8423–8455.
4. Price,D.H. (2000) P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol. Cell. Biol.*, **20**, 2629–2634.
5. Zhou,Q., Li,T. and Price,D.H. (2012) RNA polymerase II elongation control. *Annu. Rev. Biochem.*, **81**, 119–143.
6. Lu,H., Xue,Y., Yu,G.K., Arias,C., Lin,J., Fong,S., Faure,M., Weisburd,B., Ji,X., Mercier,A. *et al.* (2015) Compensatory induction of MYC expression by sustained CDK9 inhibition via a BRD4-dependent mechanism. *eLife*, **4**, e06535.
7. Chao,S.H. and Price,D.H. (2001) Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. *J. Biol. Chem.*, **276**, 31793–31799.
8. Peterlin,B.M. and Price,D.H. (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell*, **23**, 297–305.
9. Bartholomeeusen,K., Xiang,Y., Fujinaga,K. and Peterlin,B.M. (2012) Bromodomain and extra-terminal (BET) bromodomain inhibition activate transcription via transient release of positive transcription elongation factor b (P-TEFb) from 7SK small nuclear ribonucleoprotein. *J. Biol. Chem.*, **287**, 36609–36616.
10. Jang,M.K., Mochizuki,K., Zhou,M., Jeong,H.S., Brady,J.N. and Ozato,K. (2005) The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol. Cell*, **19**, 523–534.
11. Yang,Z., Yik,J.H., Chen,R., He,N., Jang,M.K., Ozato,K. and Zhou,Q. (2005) Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol. Cell*, **19**, 535–545.
12. Hargreaves,D.C., Horng,T. and Medzhitov,R. (2009) Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell*, **138**, 129–145.
13. He,N., Liu,M., Hsu,J., Xue,Y., Chou,S., Burlingame,A., Krogan,N.J., Alber,T. and Zhou,Q. (2010) HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol. Cell*, **38**, 428–438.
14. Lin,C., Smith,E.R., Takahashi,H., Lai,K.C., Martin-Brown,S., Florens,L., Washburn,M.P., Conaway,J.W., Conaway,R.C. and Shilatifard,A. (2010) AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell*, **37**, 429–437.
15. Di Vona,C., Bezdan,D., Islam,A.B., Salichs,E., Lopez-Bigas,N., Ossowski,S. and de la Luna,S. (2015) Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. *Mol. Cell*, **57**, 506–520.
16. Lu,H., Yu,D., Hansen,A.S., Ganguly,S., Liu,R., Heckert,A., Darzacq,X. and Zhou,Q. (2018) Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature*, **558**, 318–323.
17. Pownall,M.E., Gustafsson,M.K. and Emerson,C.P. Jr (2002) Myogenic regulatory factors and the specification of muscle progenitors in vertebrate embryos. *Annu. Rev. Cell Dev. Biol.*, **18**, 747–783.
18. Palacios,D. and Puri,P.L. (2006) The epigenetic network regulating muscle development and regeneration. *J. Cell. Physiol.*, **207**, 1–11.
19. Sartorelli,V. and Caretti,G. (2005) Mechanisms underlying the transcriptional regulation of skeletal myogenesis. *Curr. Opin. Genet. Dev.*, **15**, 528–535.
20. Berkes,C.A. and Tapscott,S.J. (2005) MyoD and the transcriptional control of myogenesis. *Semin. Cell Dev. Biol.*, **16**, 585–595.
21. Deato,M.D. and Tjian,R. (2007) Switching of the core transcription machinery during myogenesis. *Genes Dev.*, **21**, 2137–2149.
22. Blau,H.M., Chiu,C.P. and Webster,C. (1983) Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell*, **32**, 1171–1180.
23. Berkes,C.A., Bergstrom,D.A., Penn,B.H., Seaver,K.J., Knoepfler,P.S. and Tapscott,S.J. (2004) Pbx marks genes for activation by MyoD indicating a role for a homeodomain protein in establishing myogenic potential. *Mol. Cell*, **14**, 465–477.
24. Wang,B., Doan,D., Roman Petersen,Y., Alvarado,E., Alvarado,G., Bhandari,A., Mohanty,A., Mohanty,S. and Nissen,R.M. (2013) Wdr68 requires nuclear access for craniofacial development. *PLoS One*, **8**, e54363.
25. He,N., Jahchan,N.S., Hong,E., Li,Q., Bayfield,M.A., Maraia,R.J., Luo,K. and Zhou,Q. (2008) A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol. Cell*, **29**, 588–599.
26. Hayashi,S., Manabe,I., Suzuki,Y., Relaix,F. and Oishi,Y. (2016) Klf5 regulates muscle differentiation by directly targeting muscle-specific genes in cooperation with MyoD in mice. *eLife*, **5**, e17462.
27. He,N., Chan,C.K., Sobhian,B., Chou,S., Xue,Y., Liu,M., Alber,T., Benkirane,M. and Zhou,Q. (2011) Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, E636–E645.
28. Miyata,Y. and Nishida,E. (2011) DYRK1A binds to an evolutionarily conserved WD40-repeat protein WDR68 and induces its nuclear translocation. *Biochim. Biophys. Acta*, **1813**, 1728–1739.
29. Glenewinkel,F., Cohen,M.J., King,C.R., Kaspar,S., Bamberg-Lemper,S., Mymryk,J.S. and Becker,W. (2016) The adaptor protein DCAF7 mediates the interaction of the adenovirus E1A oncoprotein with the protein kinases DYRK1A and HIPK2. *Sci. Rep.*, **6**, 28241.
30. Moncaut,N., Rigby,P.W. and Carvajal,J.J. (2013) Dial M(RF) for myogenesis. *FEBS J.*, **280**, 3980–3990.
31. Yi,J.S., Park,J.S., Ham,Y.M., Nguyen,N., Lee,N.R., Hong,J., Kim,B.W., Lee,H., Lee,C.S., Jeong,B.C. *et al.* (2013) MG53-induced IRS-1 ubiquitination negatively regulates skeletal myogenesis and insulin signalling. *Nat. Commun.*, **4**, 2354.
32. Weintraub,H. (1993) The MyoD family and myogenesis: redundancy, networks, and thresholds. *Cell*, **75**, 1241–1244.
33. Chou,S., Upton,H., Bao,K., Schulze-Gahmen,U., Samelson,A.J., He,N., Nowak,A., Lu,H., Krogan,N.J., Zhou,Q. *et al.* (2013) HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E123–E131.
34. Stirnimann,C.U., Petsalaki,E., Russell,R.B. and Muller,C.W. (2010) WD40 proteins propel cellular networks. *Trends Biochem. Sci.*, **35**, 565–574.

35. Ritterhoff,S., Farah,C.M., Grabitzki,J., Lochnit,G., Skurat,A.V. and Schmitz,M.L. (2010) The WD40-repeat protein Han11 functions as a scaffold protein to control HIPK2 and MEKK1 kinase functions. *EMBO J.*, **29**, 3750–3761.
36. Li,S., Xu,C., Fu,Y., Lei,P.J., Yao,Y., Yang,W., Zhang,Y., Washburn,M.P., Florens,L., Jaiswal,M. *et al.* (2018) DYRK1A interacts with histone acetyl transferase p300 and CBP and localizes to enhancers. *Nucleic Acids Res.*, **46**, 11202–11213.
37. Zhang,H., Pandey,S., Travers,M., Sun,H., Morton,G., Madzo,J., Chung,W., Khowsathit,J., Perez-Leal,O., Barrero,C.A. *et al.* (2018) Targeting CDK9 reactivates epigenetically silenced genes in cancer. *Cell*, **175**, 1244–1258.
38. Simone,C., Stiegler,P., Bagella,L., Pucci,B., Bellan,C., De Falco,G., De Luca,A., Guanti,G., Puri,P.L. and Giordano,A. (2002) Activation of MyoD-dependent transcription by cdk9/cyclin T2. *Oncogene*, **21**, 4137–4148.