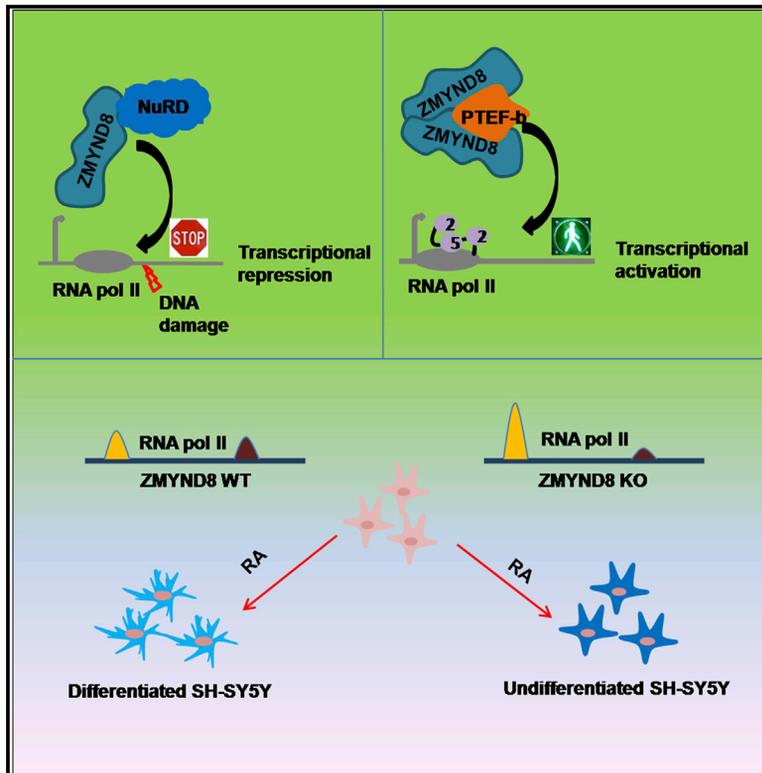


# Cell Reports

## Positive Regulation of Transcription by Human ZMYND8 through Its Association with P-TEFb Complex

### Graphical Abstract



### Authors

Koushik Ghosh, Ming Tang, Nidhi Kumari, ..., Dheerendra Pratap Mall, Kunal Rai, Debabrata Biswas

### Correspondence

dbiswas@iicb.res.in

### In Brief

Ghosh et al. report that coiled-coil domain-dependent dimerization of ZMYND8 positively regulates transcription through P-TEFb complex recruitment, whereas monomeric ZMYND8 associates with NuRD complex for transcriptional repression. ZMYND8-mediated P-TEFb complex recruitment is required for target gene activation during retinoic acid-induced differentiation of SH-SY5Y cells.

### Highlights

- ZMYND8-mediated P-TEFb complex recruitment activates transcription
- Differentiation of SH-SY5Y cells requires transcriptional activation by ZMYND8
- ZMYND8 exists as monomer as well as dimer
- Preferential association of CHD4 with monomer and P-TEFb with dimer form of ZMYND8



# Positive Regulation of Transcription by Human ZMYND8 through Its Association with P-TEFb Complex

Koushik Ghosh,<sup>1</sup> Ming Tang,<sup>2</sup> Nidhi Kumari,<sup>1</sup> Arijit Nandy,<sup>1</sup> Subham Basu,<sup>1</sup> Dheerendra Pratap Mall,<sup>1</sup> Kunal Rai,<sup>2</sup> and Debabrata Biswas<sup>1,3,\*</sup>

<sup>1</sup>Laboratory of Transcription Biology, Molecular Genetics Division, CSIR-Indian Institute of Chemical Biology, 4 Raja S.C. Mullick Road, Kolkata 32, India

<sup>2</sup>Division of Cancer Medicine, Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX 77054, USA

<sup>3</sup>Lead Contact

\*Correspondence: [dbiswas@iicb.res.in](mailto:dbiswas@iicb.res.in)

<https://doi.org/10.1016/j.celrep.2018.07.064>

## SUMMARY

Although human ZMYND8 has been implicated as a transcriptional co-repressor of multiple targets, global association of ZMYND8 with active genes and enhancer regions predicts otherwise. Here, we report an additional function of ZMYND8 in transcriptional activation through its association with the P-TEFb complex. Biochemical reconstitution analyses show that human ZMYND8, through direct association with CyclinT1, forms a minimal ZMYND8-P-TEFb complex. The importance of ZMYND8 in target gene activation, through P-TEFb complex recruitment, is demonstrated on chromosomally integrated reporter gene as well as native target genes *in vivo*. Physiologically, we further show that the ZMYND8-P-TEFb complex-mediated transcriptional activation is required for all-*trans* retinoic acid (ATRA)-mediated differentiation of neuronal precursor cells. Finally, to detail the dual activator and repressor nature, mechanistically we show that, through its putative coiled-coil domain, ZMYND8 forms a homodimer that preferentially associates with the activator P-TEFb complex, whereas the monomer associates with the CHD4 subunit of repressor NuRD complex.

## INTRODUCTION

Human ZMYND8 is a zinc finger, MYND domain-containing protein. Through recognition of H3K4Me1 and H3K14Ac by N-terminal PHD and bromodomains, respectively, ZMYND8 negatively regulates the expression of genes required for cellular metastasis (Li et al., 2016). ZMYND8 has also been shown to be recruited at the DNA damage sites through its DNA binding and H3K14Ac recognition properties (Savitsky et al., 2016). Through its interaction with KDM5C, ZMYND8 also suppresses enhancer

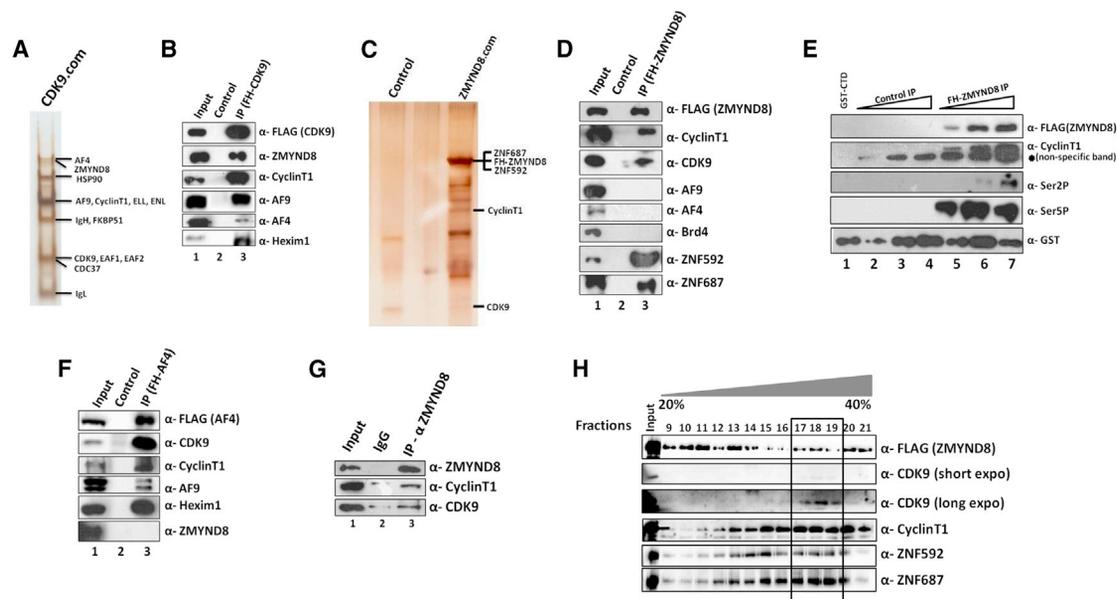
overexpression by maintaining H3K4me1 enhancer signature (Shen et al., 2016). The C-terminal MYND domain of ZMYND8 interacts with the NuRD complex for inhibiting transcription at the DNA repair sites (Gong et al., 2015).

ZMYND8 has also been implicated as a positive regulator for target gene expression by a few studies. Along with an initial report suggesting a positive role of ZMYND8 in regulation of ER $\alpha$ -induced target gene expression (Malovannaya et al., 2011), few recent studies have also indicated a similar role in all-*trans* retinoic acid (ATRA)-induced target gene activation (Adhikary et al., 2016; Basu et al., 2017). Consistent with a positive role, genome-wide association studies showed ZMYND8 occupancy at a majority of the active genes and enhancer regions (Savitsky et al., 2016; Shen et al., 2016). Overall, studies imply that ZMYND8 could regulate target gene expression both positively and negatively. However, the precise role of ZMYND8 in transcriptional activation is completely unknown.

Among all of the positive regulators of transcription, human P-TEFb complex, a heterodimer of CyclinT1/T2 and cyclin-dependent kinase 9 (CDK9) (Peng et al., 1998), plays a major role in activating target gene expression. During transcriptional activation, phosphorylation of serine residues of C-terminal domain (CTD) of RNA polymerase II (hereafter RNA pol II) and DSIF and NELF complexes (Yamaguchi et al., 2013; Zhou et al., 2012) by P-TEFb complex facilitates entry of RNA pol II into productive elongation. The majority of the cellular P-TEFb complex is sequestered in an inactive complex in association with Hexim1/2, LARP7, MePCE2, and 7SK RNA (He et al., 2008; Li et al., 2005; Yang et al., 2001). Although P-TEFb complex has been shown to be recruited by several activator proteins including Brd4, c-Myc, p53, NF- $\kappa$ B, and AF4 (within SEC complex) (Biswas et al., 2011; He et al., 2010; Jang et al., 2005; Lin et al., 2010; Paparidis et al., 2017; Yang et al., 2005), there is intense interest in identifying other factors for further in-depth understanding of gene-specific recruitment of P-TEFb complex for transcriptional activation.

In this study, we report an association and complex formation between ZMYND8 and the P-TEFb complex. We provide evidence of a direct role of ZMYND8-mediated P-TEFb complex recruitment for transcriptional activation of target genes.





**Figure 1. Interaction of ZMYND8 with P-TEFb Complex**

(A) Purification of CDK9-associated protein complex by one-step FLAG(M2 bead)-affinity purification using nuclear extract of a stable cell line that ectopically expresses CDK9 as FLAG-HA tagged. Eluted proteins were run on 4%–12% gradient gel and silver stained for their visualization and were further subjected to mass spectrometry analysis for their identification.

(B) IP and western blotting analysis showing interaction of ectopically expressed CDK9 with the indicated proteins in 293T cells.

(C) Purification of ZMYND8-associated protein complex by one-step FLAG(M2 bead)-affinity purification using nuclear extract of a stable cell line that ectopically expresses ZMYND8 as FLAG-HA tagged. Eluted proteins were run on 4%–12% gradient gel and silver stained for their visualization. Indicated proteins were marked based on their sizes and their presence in the western blot analysis. The parallel control experiment was carried out using nuclear extract from a control 293T cells.

(D) IP and western blotting analysis showing interaction of ectopically expressed ZMYND8 with indicated proteins in 293T cells.

(E) *In vitro* kinase assay for checking phosphorylation activity of purified ZMYND8 complex (Figure 1C) toward GST-CTD (as substrate). After *in vitro* phosphorylation, western blotting was performed using antibodies specific to phospho-Ser2 and -Ser5 for detecting phosphorylated GST-CTD protein.

(F) IP and western blotting analysis showing interaction of ectopically expressed AF4 with the indicated proteins in 293T cells.

(G) IP and western blotting showing interaction of endogenous ZMYND8 with the P-TEFb complex.

(H) Glycerol gradient separation analysis of purified ZMYND8 complex (Figure 1C) showing its co-separation with P-TEFb complex and other associated proteins in a 20%–40% glycerol gradient. Fractions were collected, and individual fractions were tested for the presence of indicated proteins by western blotting.

## RESULTS

### Association between the P-TEFb Complex and ZMYND8

Mass-spectrometric analysis of CDK9-associated protein complex, purified from nuclear extract of a FLAG-HA-CDK9-expressing stable cell line, showed the presence of ZMYND8 along with other known CDK9-interacting proteins (Figures 1A, S1A, and S1B). Further immunoprecipitation and blotting analyses substantiated CDK9 interaction with ZMYND8 (Figure 1B). Similar reciprocal analysis showed the presence of CDK9 in ZMYND8-immunoprecipitated sample along with known ZMYND8 interactors ZNF592 and ZNF687, but not other CDK9 interactors such as AF4, Brd4, and Hsp90 (Figures 1C, 1D, S1C, and S1E). Consistently, the AF4 and Brd4 pulled-down samples also failed to show the presence of ZMYND8 (Figures 1F and S1D). Therefore, we conclude that the ZMYND8 and P-TEFb association is distinct in nature unlike that of other P-TEFb-interacting proteins/complexes that have been reported so far.

Further immunoprecipitation analysis showed that, unlike AF4 (Lu et al., 2014), ZMYND8 does not associate with Hexim1, a component to which P-TEFb interacts and forms an inactive

complex (Figures S1F and S1G). This observation further indicates an association of ZMYND8 with only active form of P-TEFb complex. Consistently, the immunoprecipitated ZMYND8 complex showed strong phosphorylation activity *in vitro* toward purified GST-CTD substrate to generate Ser2 and Ser5 phosphopeptide, whereas control experiment failed to do so (Figure 1E). Although P-TEFb-associated Brd4 protein shows kinase activity (Winter et al., 2017), we failed to observe any kinase activity of our purified recombinant ZMYND8 (Figure S1H), therefore indicating that the phosphorylation activity of ZMYND8-associated protein complex (Figure 1C) reflects activity of associated P-TEFb complex. Since phosphorylation of CDK9 at Thr186 of the T loop region is a hallmark of the active P-TEFb complex (Chen et al., 2004; Li et al., 2005), our observation of association of Thr186-phosphorylated CDK9 (Figure S1I) with ZMYND8 further confirms ZMYND8 association with active form of P-TEFb complex. To rule out artifacts of protein interactions due to overexpression, we immunoprecipitated ZMYND8 using specific antibody and confirmed its interaction with P-TEFb complex in endogenous context in 293T cells (Figure 1G). To substantiate our idea that ZMYND8 and P-TEFb

form a stable complex, we performed complex separation using 20%–40% glycerol gradient. As shown in [Figure 1H](#), although the majority of ZMYND8 is seen in fractions containing CyclinT1, ZNF592, and ZNF687, few fractions also showed co-presence with CDK9. The association of ZMYND8 with CDK9 seems to be very specific since we observed their co-presence only in few fractions (fractions 17–19). Together, we conclude that human ZMYND8 shows a distinct and stable association with active P-TEFb complex.

### ZMYND8 Directly Associates with P-TEFb and Forms a Functionally Active Minimal ZMYND8-P-TEFb Complex

Immunoprecipitation analysis of ZMYND8 when ectopically expressed with either CyclinT1 or CDK9 in heterologous Sf9 expression system showed ZMYND8 association with CyclinT1 but not CDK9 ([Figures S2A](#) and [S2B](#)). This observation indicated CyclinT1-mediated minimal ZMYND8-P-TEFb complex formation. As shown in [Figure 2A](#), through co-expression of all of the factors in Sf9 cells, a minimal ZMYND8-P-TEFb complex can be reconstituted, as evidenced by Coomassie staining (left panel) and immunoblotting (right panel). Further *in vitro* kinase assay using GST-CTD substrate confirmed reconstitution of functionally active ZMYND8-P-TEFb complex ([Figure 2B](#), lanes 5–7), whereas a parallel control experiment showed no activity ([Figure 2B](#), lanes 2–4). Therefore, we conclude that ZMYND8 directly associates with P-TEFb to form a functionally active minimal ZMYND8-P-TEFb complex.

### CyclinT1 Mediates Formation of ZMYND8-P-TEFb Complex that Is Distinct from Other P-TEFb-Containing Complexes

To define subunit interactions among the reconstituted ZMYND8-P-TEFb complex, we co-infected Sf9 cells with various combinations of baculoviruses as indicated in [Figure 2C](#). Immunoprecipitation and subsequent blotting analyses showed that (1) ZMYND8 associates with the CyclinT1 and CDK9 subunits of the P-TEFb complex ([Figure 2C](#), lane 4), (2) ZMYND8 fails to show an interaction with CDK9 alone ([Figure 2C](#), lane 2), and (3) only in the presence of CyclinT1, ZMYND8 interacts with CDK9. These experiments clearly demonstrate that CyclinT1 acts as bridge between ZMYND8 and CDK9 for their interaction. Consistently, an *in vitro* interaction analysis using purified recombinant ZMYND8 and CyclinT1 further confirmed their direct interactions ([Figures S2C](#) and [S2D](#)). Notably, although the AF4 protein forms a minimal AF4-P-TEFb complex ([Biswas et al., 2011](#)), its absence in the purified recombinant minimal ZMYND8-P-TEFb complex ([Figure S2E](#)) further confirmed the distinct nature of the ZMYND8-P-TEFb complex and explained our inability to observe AF4 in our ZMYND8 immunoprecipitated samples ([Figure 1D](#)). Taken together, along with our observation from [Figure 1](#), we conclude that the ZMYND8-P-TEFb complex is distinct in nature and resides in a mutually exclusive manner in association with other P-TEFb interacting proteins such as AF4 and Brd4.

### Domain Analysis of ZMYND8 and CyclinT1 for Their Association and Higher-Order Complex Formation

Further interaction analyses through expression of various ZMYND8 domains and CyclinT1 in Sf9 cells showed that a deriv-

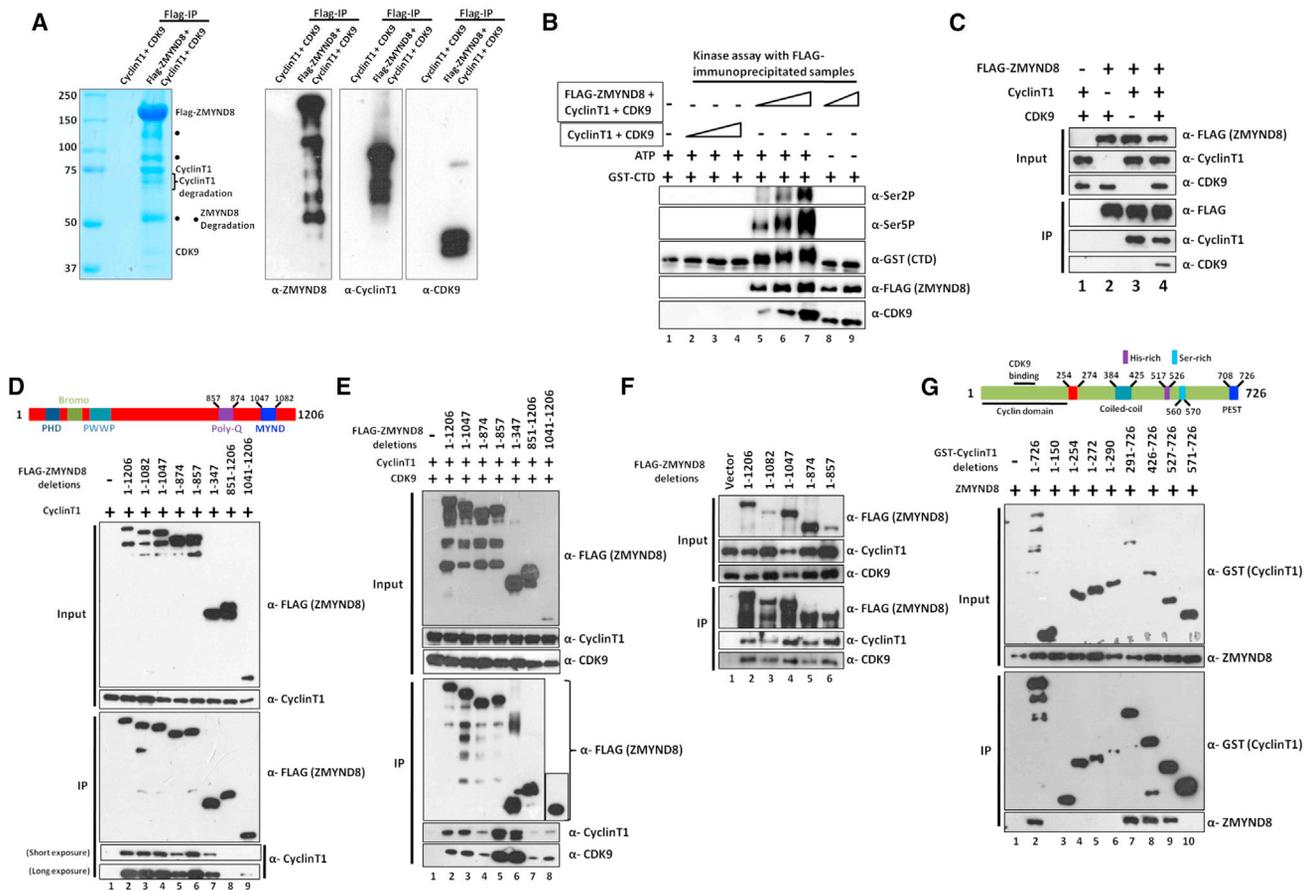
ative harboring deletion of 18 aa (857–874 aa, glutamine-rich, hereafter referred as poly-Q region) drastically enhanced ZMYND8 interaction with CyclinT1 ([Figure 2D](#), compare lane 5 with lane 6). This observation strongly indicates a negative role of the poly-Q region in regulating ZMYND8 interaction with CyclinT1. The C-terminal 851–1,206 aa failed to show an interaction, whereas further deletion of 190 aa weakly restored C-terminal interaction with ZMYND8, indicating a potential negative role of the deleted 190 aa (potentially through the poly-Q region) for its interaction with CyclinT1.

Subsequent complex formation analyses by expressing various ZMYND8 domains along with CDK9 and CyclinT1 in Sf9 cells further confirmed a negative role of the poly-Q region in higher-order ZMYND8-P-TEFb complex assembly ([Figures 2E](#) and [S2F](#), compare lane 4 versus lane 5). Deletion of N-terminal 850 aa greatly impaired ZMYND8-P-TEFb complex formation ([Figures 2E](#) and [S2F](#), lane 7) and thus emphasized a role for this domain in overall assembly. Taken together, we conclude that, whereas the poly-Q region of ZMYND8 inhibits, the N-terminal 850 aa of ZMYND8 is critical for ZMYND8-P-TEFb complex formation. Consistent with our *in vitro* associations, transfection of 293T cells with ZMYND8 deletion constructs and subsequent immunoprecipitation analyses reproducibly pulled down higher amount of the P-TEFb complex when the poly-Q region was absent in the ZMYND8 construct ([Figures 2F](#) and [S2G](#), compare lane 5 with lane 6).

A reciprocal analyses through expression of various CyclinT1 domains ([Young et al., 2003](#)) and ZMYND8 in Sf9 cells further showed that CyclinT1 region spanning 527–570 aa is important for its interaction with ZMYND8 ([Figure 2G](#)). Since CDK9 interacts with further N-terminal region of CyclinT1, we conclude that two non-overlapping domains of CyclinT1 are used for ZMYND8-P-TEFb complex assembly. Interestingly, ZMYND8-interacting region of CyclinT1 is in close proximity with Brd4-interacting region (426–516 aa) ([Jang et al., 2005](#)), which could potentially explain the mutually exclusive existence of ZMYND8 with Brd4.

### ZMYND8 Positively Regulates Transcription from a Reporter Construct

Toward understanding potential positive role of ZMYND8 in transcriptional regulation—presumably through ZMYND8-mediated P-TEFb complex recruitment—we initially tested the effect of ZMYND8 on expression of a chromosomally integrated luciferase reporter construct that is expressed downstream of five tandem Gal4-binding sites ([Figure S3A](#)). We reasoned that, since ZMYND8 contains multiple chromatin reader domains at the N terminus, a chromosomally integrated reporter gene construct—as opposed to plasmid DNA construct—would reflect ZMYND8 function close to actual native target *in vivo*. In our analysis, we did not observe any effect of ZMYND8 on the basal level expression of reporter gene ([Figure S3B](#)). However, in the presence of Gal-VP16 activator, ZMYND8 strongly activated transcription from this reporter gene construct ([Figure 3A](#)). Subsequent RNA analysis ([Figure 3B](#)) further confirmed a positive role of ZMYND8 in transcriptional regulation of chromosomally integrated reporter gene. Failure to observe an effect on activator GAL-VP16 expression ([Figure S3C](#)) ruled out indirect



**Figure 2. ZMYND8 Directly Interacts with P-TEFb Complex to Form Minimal ZMYND8-P-TEFb Complex**

(A) SDS-PAGE Coomassie staining of reconstituted ZMYND8-P-TEFb complex and immunoblot analyses for confirming presence of the indicated proteins (right panel). Sf9 cells were co-infected with baculoviruses expressing indicated proteins, and protein complexes were purified using FLAG(M2)-affinity purification method. Eluted proteins were loaded onto a 4%–12% gradient gel and were Coomassie stained for their visualization (left panel). Presence of ZMYND8 and P-TEFb complex components were further confirmed by western blotting (right panel).

(B) *In vitro* CTD kinase assay showing phosphorylation activity of the purified reconstituted ZMYND8-P-TEFb complex (Figure 2A) using GST-CTD as substrate (lanes 5–7). Parallel control experiment was carried out using control purified proteins from Sf9 cells that were co-infected with non-tag CyclinT1 and CDK9 proteins (lanes 2–4).

(C) Immunoblot analyses showing subunit interactions among reconstituted ZMYND8-P-TEFb complex through their co-expression in Sf9 cells. After IP (using M2 beads), ZMYND8 and other target proteins were identified by western blot analysis using indicated antibodies.

(D) Immunoblot analyses showing domains of ZMYND8 that are important for its interaction with CyclinT1 through their co-expression in Sf9 cells. After IP (using M2 beads), ZMYND8 deletion fragments and CyclinT1 were identified by western blot analysis using indicated antibodies. Top panel shows cartoon diagram of important domains of ZMYND8.

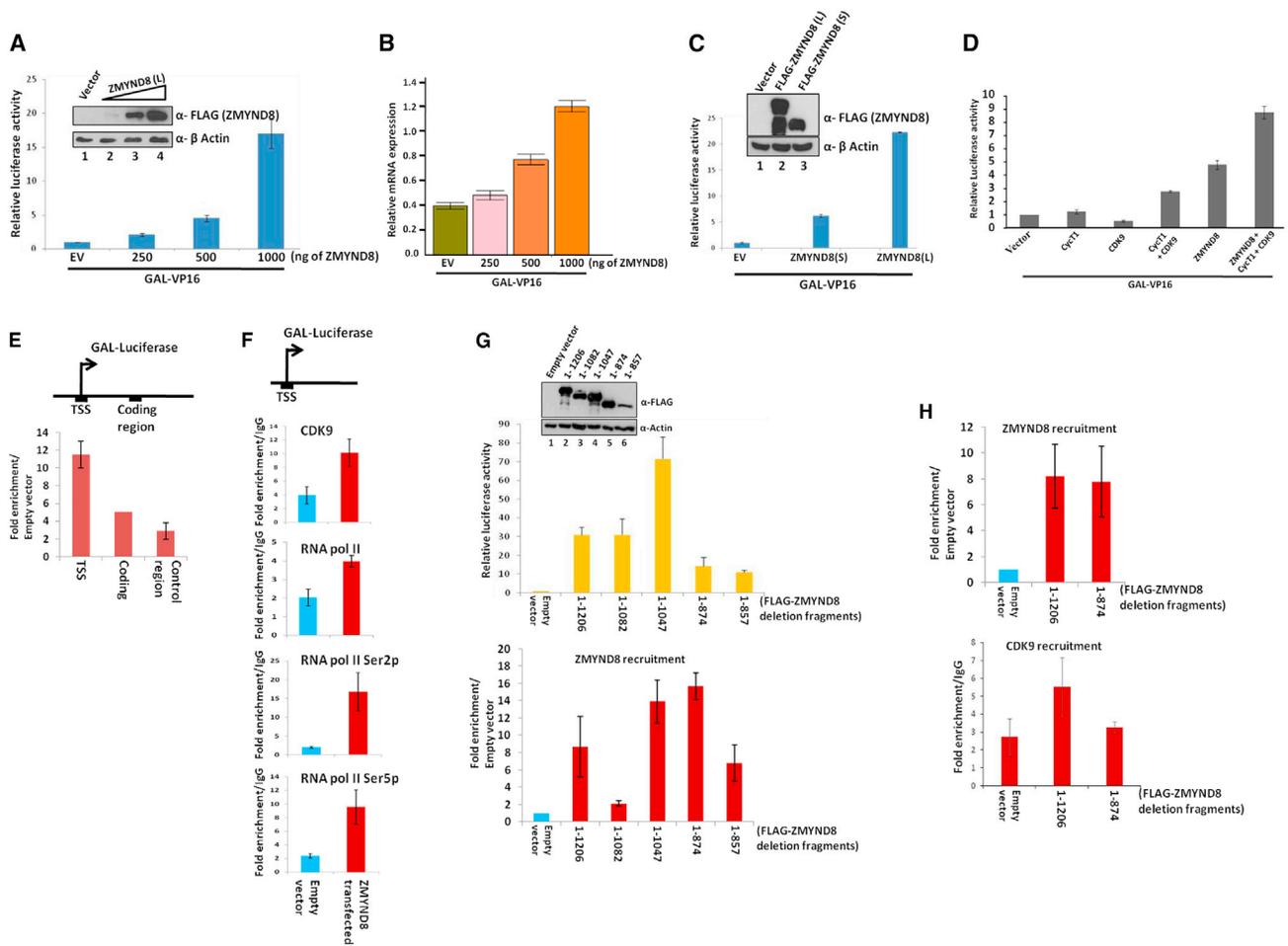
(E) Immunoblot analysis showing domains of ZMYND8 that are important for the minimal ZMYND8-P-TEFb complex formation through their co-expression in Sf9 cells. After IP (using M2 beads), ZMYND8 deletion fragments, CyclinT1, and CDK9 were identified by western blot analysis using indicated antibodies. The topmost band in each lane represents the actual ZMYND8 fragment in this assay. The other bands represent either ZMYND8 degradation or non-specific proteins in the reactions. Because of its small size (1,041–1,206), for better resolution of this ZMYND8 fragment in our assay, we have run the IP sample in a higher-percentage gel and it is represented as a separate band in the ZMYND8 IP panel.

(F) Immunoblot analysis showing ability of ZMYND8 domains for their interaction with P-TEFb complex in mammalian cells. 293T cells were transfected with plasmids expressing indicated ZMYND8 domains, and ZMYND8-interacting proteins were purified by IP using FLAG-agarose beads. Interacting proteins were identified by immunoblot analyses using specific antibodies.

(G) Immunoblot analysis showing domains of CyclinT1 that are important for its interaction with ZMYND8 through their co-expression in Sf9 cells. CyclinT1-interacting proteins were purified using glutathione-agarose beads. Further western blotting was performed for identifying the interacting proteins. Top panel shows a cartoon diagram of important domains of CyclinT1.

effect of ZMYND8 overexpression on activator expression that would thus affect transcription of the reporter gene. Our further analysis showed that, despite a similar level of expression, the shorter isoform of ZMYND8 (1–1,054, isoform 17, Uniprot designation; Figure S3D) showed reduced

transcriptional activation of the reporter gene (Figure 3C) than the longer isoform (1–1,206, isoform 10, Uniprot designation; Figure S3D). Consistently, immunoprecipitation analysis showed reduced



**Figure 3. Positive Regulation of Transcription of Reporter Gene by ZMYND8 through Recruitment of P-TEFb Complex**

(A) Luciferase assay showing positive role of ZMYND8 in regulating luciferase reporter gene expression. Chromosomally integrated Gal-luciferase cells were transfected with increasing amount of ZMYND8 along with activator GAL-VP16, and the expression of luciferase reporter gene was assayed. The inset panel shows the expression of transfected ZMYND8.

(B) RNA analysis by qRT-PCR showing positive role of ZMYND8 in transcription of reporter gene. mRNAs isolated from samples of Figure 3A were subjected to qRT-PCR analysis.

(C) Luciferase assay showing increased potency of ZMYND8 longer isoform (L) (1–1,206 aa, isoform 10, based on Uniprot designation) in transcriptional activation than the shorter one (S) (1–1,054 aa, isoform 17 based on Uniprot designation) in the presence of activator GAL-VP16. The inset panel shows expressions of ZMYND8 isoforms in this study. The top bands in each lane represent expression of ZMYND8 isoforms in this assay. The other bands represent the degradation product since ZMYND8 showed its propensity to degrade in our assay system.

(D) Luciferase reporter assay showing additive role of ZMYND8 and the P-TEFb complex in regulating reporter gene expression in the presence of activator GAL-VP16.

(E) ChIP analysis showing recruitment of ZMYND8 at different regions of target luciferase gene. Gal-Luciferase cells were transfected with empty vector (EV) and FLAG-ZMYND8-expressing plasmid along with activator GAL-VP16. FLAG-ZMYND8 was immunoprecipitated using anti-FLAG M2 magnetic beads. Primers for different regions of GAL-luciferase construct were used for addressing the presence of FLAG-ZMYND8 at the target locus. PCR analysis using primers for control (non-target) region shows specificity of ZMYND8 recruitment in our assay.

(F) ChIP analysis showing recruitment of CDK9, total RNA pol II, and phosphorylated Ser2 and Ser5 forms of RNA pol II at the TSS region of target luciferase gene.

(G) Luciferase assay showing the effect of different ZMYND8 domains on activating target luciferase gene expression. Gal-luciferase cells were transfected with empty vector (EV) and other ZMYND8 domains expressing plasmids along with activator GAL-VP16. One part of the transfected cells was used for luciferase gene expression, and the other part was used for ChIP analysis for identifying recruitment of ZMYND8 domains at the TSS. The inset immunoblot shows the level of expression of each ZMYND8 fragment that has been used in this assay (upper inset panel). ChIP analysis shows recruitment of different ZMYND8 domains at the TSS region of target luciferase gene (lower panel).

(H) ChIP analysis showing recruitment of full-length FLAG-ZMYND8 (1–1,206), FLAG-ZMYND8 domain (1–874) (upper panel) and CDK9 (lower panel) at the TSS region of target luciferase gene.

Data represent mean  $\pm$  SD, a minimum of two biological replicates.

interaction of shorter isoform with P-TEFb complex than the longer one in our heterologous Sf9 expression-based interaction assays (Figure S3E).

We reasoned that, if the interaction between ZMYND8 and P-TEFb complex has a functional consequence, they may show additive effect for transcriptional activation. Indeed, ZMYND8 and P-TEFb activated transcription in an additive manner when co-transfected together in reporter assay (Figure 3D).

### ZMYND8 Regulates Transcriptional Activation via P-TEFb Complex

For deeper understanding of the role for ZMYND8 in transcriptional activation of reporter gene through P-TEFb complex recruitment, our initial chromatin immunoprecipitation (ChIP) analysis also showed the presence of ZMYND8 both at the transcription start site (TSS) and the gene body of our reporter construct (Figure 3E), without showing significant enrichment at a control intergenic region (Figure 3E). Based on the observation of genome-wide association studies, we further focused on ZMYND8-mediated factor recruitment at the TSS regions for transcriptional regulation. Interestingly and consistent with ZMYND8 interaction with P-TEFb complex, we observed a significant increase in CDK9 recruitment with concomitant increase in RNA pol II as well as Ser2 and Ser5 phosphorylation of the CTD of RNA pol II at the TSS region upon overexpression of ZMYND8 (Figure 3F). Notably, normalization of the amount of Ser2P and Ser5P form of RNA pol II over relative amount of RNA pol II present at the target locus clearly showed a positive role of ZMYND8 in recruiting P-TEFb complex, resulting in increased Ser2P and Ser5P phosphorylation at the TSS region (Figure S3F). A greater effect of ZMYND8 on increasing Ser2P than Ser5P further indicates involvement of P-TEFb complex in this phosphorylation event.

Our subsequent analysis using different ZMYND8 domains showed that, although for the majority of ZMYND8 fragments reporter gene expression correlated well with its recruitment, the ZMYND8 fragment (1–874) that showed reduced P-TEFb interaction (Figures 2E and 2F) also failed to activate reporter gene transcription (Figure 3G, upper panel) despite being recruited at the same level to that of full-length as measured by our ChIP analysis (Figure 3G, lower panel). Further ChIP analysis confirmed the inability of the ZMYND8 fragment (1–874) to recruit CDK9/P-TEFb complex at the target TSS locus (Figure 3H). Together, these experiments demonstrate an additional role of ZMYND8 in positive regulation of transcription through recruitment of the P-TEFb complex.

### Global Co-occupancy of ZMYND8 and P-TEFb Complex

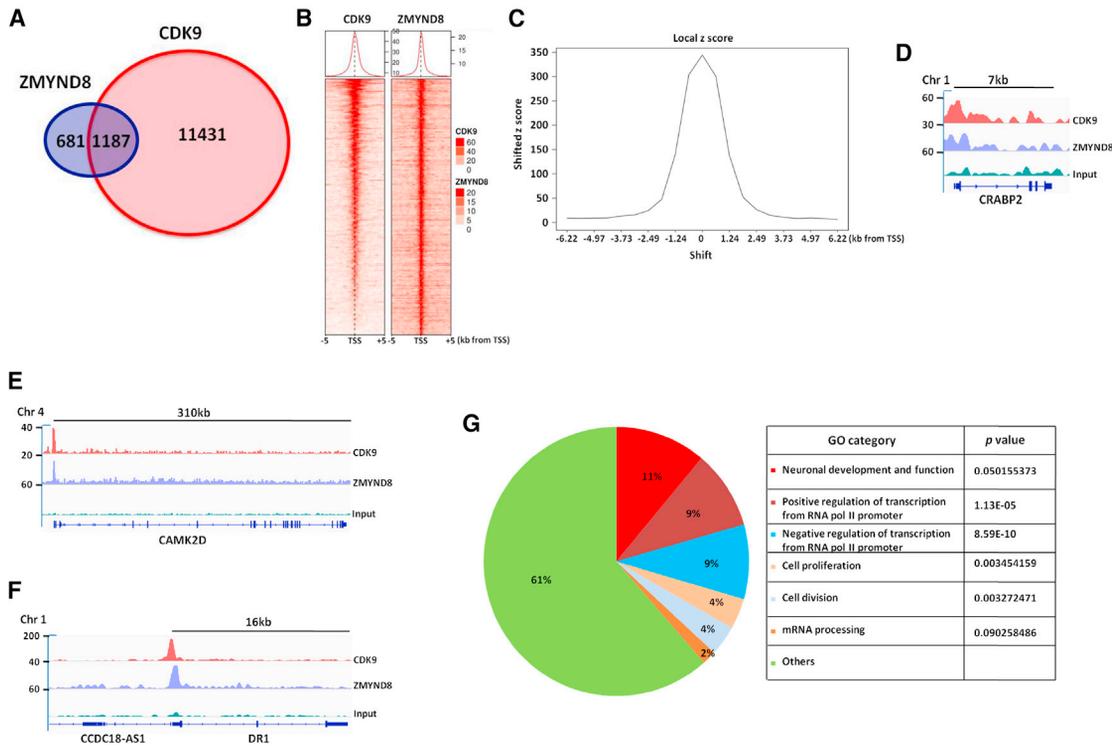
For identifying the direct targets for ZMYND8-P-TEFb complex in cellular context, we sought to identify genes that are being co-occupied by both ZMYND8 and P-TEFb complex. To this end, we exploited already published ChIP-sequencing (ChIP-seq) studies for ZMYND8 and CDK9 in 293T cells (Liu et al., 2013; Savitsky et al., 2016) (GEO: GSE81696 and GSE51633). Further analysis showed that ZMYND8 occupied 1,868 loci (at  $p < 10^{-6}$ ), whereas CDK9 bound to 12,618 loci (at  $p < 10^{-6}$ ) (Figures 4A and 4B). Comparative analysis of these two datasets showed that 1,187 sites were occupied by both ZMYND8 and

CDK9 ( $p = 0.002$ ), which was >60% of total ZMYND8 occupied sites (Figures 4A and S4A). The observed high degree of co-occupancy strongly implies a role of ZMYND8 in P-TEFb-mediated transcriptional regulation. Importantly, the majority of ZMYND8 and CDK9 co-localized sites (1,023/1,187, 86%) were in the vicinity of ( $\pm 5$  kb) of TSS, suggesting preferential localization of this complex at promoters (Figures 4B, 4C, and S4A). Representative examples of their co-occupancy are illustrated in few selected genes as shown in Figures 4D–4F and S4B–S4H. Specific presence of both ZMYND8 and P-TEFb complex at the TSS region compared to the gene bodies, irrespective of the gene length, further signifies a specific functional role of these factors in regulating transcription around the TSS region. Gene ontology (GO) analysis of these co-occupied target genes showed their preferential enrichment at the genes regulating multiple cellular events including cell proliferation and cell division (Figure 4G). Interestingly, a substantial number of these genes are also involved in both positive and negative regulation of transcription by RNA pol II. Finally, we also note that a large set of genes (~11%) that are involved in regulating several aspects of neuronal development and functions. Therefore, we conclude that both ZMYND8 and P-TEFb complex are co-occupied at the TSS region of selected target genes in a genome-wide manner for transcriptional regulation. Importance of this ZMYND8-P-TEFb complex-mediated transcriptional regulation may have an implication in neuronal development and functions.

### ZMYND8 and P-TEFb Complex Positively Regulate Expression of Diverse Target Genes in Mammalian Cells

Next, to validate our high-throughput data, we generated stable ZMYND8 knockout and CDK9 knockdown cells (Figures 5A and 5B) and subjected them to RNA expression analysis by qRT-PCR. Consistent with positive regulation of ZMYND8 in transcriptional activation through P-TEFb complex recruitment, knockdown or knockout of either CDK9 or ZMYND8 affected expression of several target genes without affecting non-target genes as measured by qRT-PCR analyses (Figures 5C and S5A–S5D).

Our subsequent ChIP analyses were directed toward addressing the effect of ZMYND8 knockout on P-TEFb complex recruitment at TSS region of *CRABP2*, *DR1* or *NC2*, *TK2*, *CEBPG*, and *CAMK2D* target genes (Figure 5C). As shown in Figures 5D–5H, the absence of ZMYND8 significantly reduced CDK9 recruitment at the TSS region of all of these target genes. Furthermore, ZMYND8 knockout did not show any effect on CDK9 expression (Figure 5I) and thus ruled out an indirect effect of ZMYND8 knockout on reduced CDK9 recruitment through its reduced expression. Owing to the reduced CDK9 recruitment, further ChIP analysis showed significant reduction of phosphorylated Ser2 and Ser5 form of RNA pol II. Interestingly, our observation of increased total RNA pol II at the TSS region strongly indicated increased pausing upon ZMYND8 knockout. For further addressing of the pausing event, we quantitated total RNA pol II present at the indicated coding regions of these genes (Figure S5E, upper panel) and calculated pausing index (ratio of RNA pol II present at TSS/coding region). As shown in the lower panel of Figure S5E, ZMYND8 knockout strongly increased pausing index of RNA pol II at all of the genes that we have



**Figure 4. Genome-wide Co-occupancy of ZMYND8 and P-TEFb Complex on Target Genes**

(A) Publicly available ChIP-seq studies of ZMYND8 (GEO: GSE81696) and CDK9 occupancy (GEO: GSE51633) were further analyzed (see STAR Methods for details) to show genome-wide co-occupancy of ZMYND8 and CDK9 on target genes. Approximately 64% of genes that are occupied by ZMYND8 also showed concurrent occupancy with CDK9. (B) Heatmap showing genome-wide co-occupancy of ZMYND8 and CDK9 at the TSS region of target genes. (C) Local z-score analysis showing abundance of co-presence of ZMYND8 and CDK9 at different locations from the TSS region of target genes. (D–F) ChIP-seq tracks showing co-occupancy of ZMYND8 and CDK9 at the target *CRABP2* (D), *CAMK2D* (E), and *DR1* or *NC2* (F) genes. (G) GO analysis of ZMYND8-CDK9 co-occupied genes showing pathways that are potentially being regulated by the ZMYND8-P-TEFb complex.

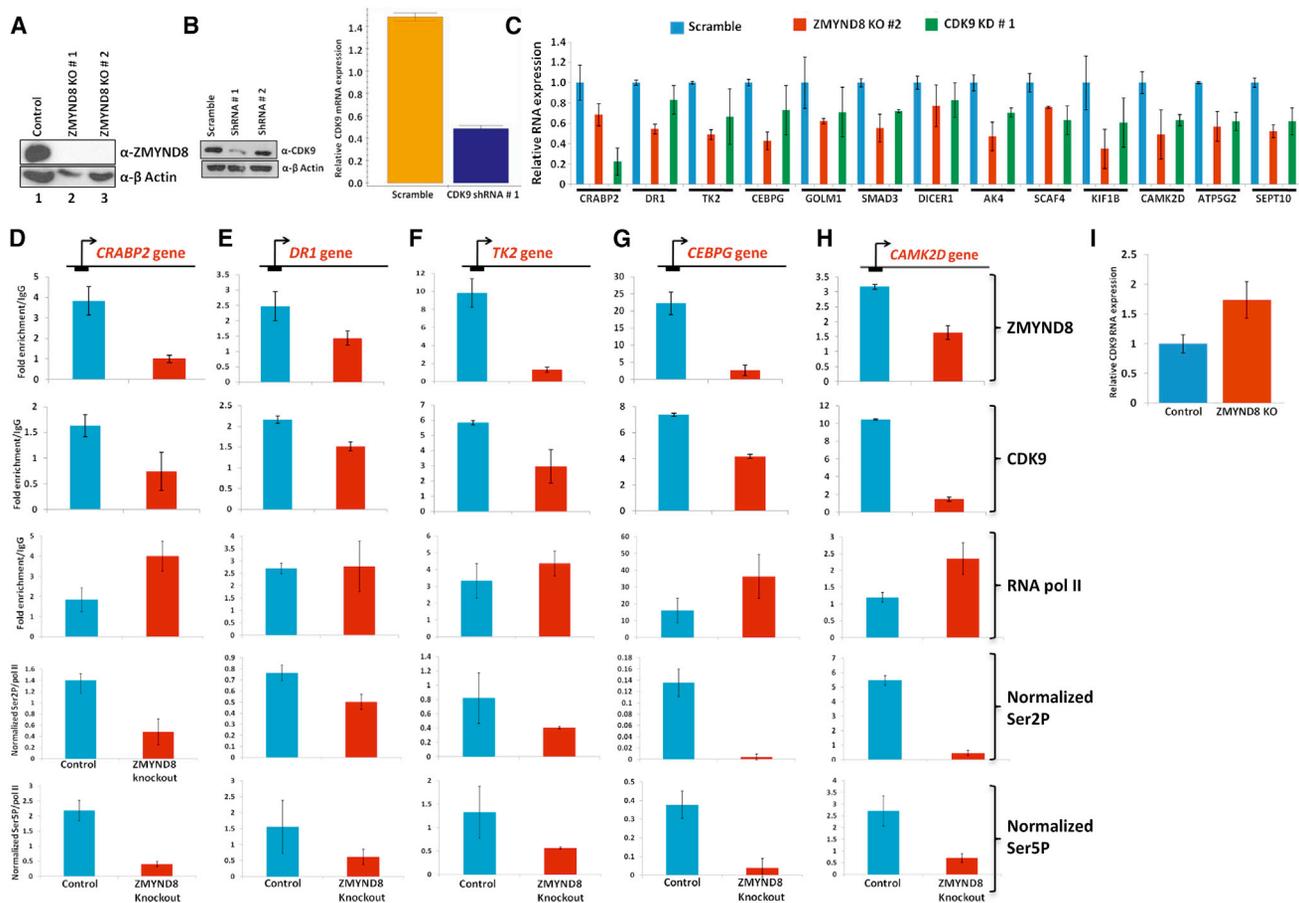
tested. This observation is very much consistent with the role of P-TEFb complex in overcoming the pausing event for RNA pol II and its entry into productive elongation. Therefore, based on our collective biochemical as well as cell-based assays, we conclude that human ZMYND8 is involved in activation of target genes through recruitment of P-TEFb complex at the promoter proximal region. This P-TEFb complex recruitment further helps in the release of paused RNA pol II into productive elongation step.

### ZMYND8-Mediated Target Gene Expression Regulates Retinoic Acid-Induced Neuronal Differentiation of SH-SY5Y Cells

To explore the role of ZMYND8-P-TEFb complex-mediated transcriptional activation in regulation of biological processes, we focused our studies on CRABP2-mediated differentiation of neuronal precursor cells (SH-SY5Y) upon retinoic acid (RA) induction. Human SH-SY5Y is a bone marrow-derived neuroblastoma cell line that, upon RA treatment, differentiates into neuronal progeny (Kovalevich and Langford, 2013; Xicoy et al., 2017). We knocked out ZMYND8 in SH-SY5Y cell line using the CRISPR-CAS9 method (Figure 6A) and subjected the knocked-out cells to RA-induced differentiation. As shown in

Figure 6B, compared to control (DMSO), upon RA treatment, the SH-SY5Y cells differentiate into neuronal progeny with branched neurite growth (Figure 6B, top right panel). Interestingly, ZMYND8 knockout SH-SY5Y cell line showed significantly reduced differentiation capabilities and increased proliferation compared to the control (Figures 6B, bottom panel, and 6C) as well as impaired expression of neuronal differentiation marker genes such as *DCX* and *MAP2* (Figure 6D). Therefore, we conclude that ZMYND8 regulates RA-induced differentiation of SH-SY5Y cells to neuronal progeny.

Consistent with the defect in RA-induced differentiation processes, subsequent RNA analysis further showed impaired expression of key target genes such as *FGS2*, *CRABP2*, and *NAV2* upon ZMYND8 knockout (Figure 6E) in SH-SY5Y cells after RA treatment. Consistent with our observation in 293T cells, ZMYND8 knockout also showed significant reduction in recruitment of CDK9 as well as reduced Ser2 and Ser5 form of RNA pol II at the TSS region of *CRABP2* gene (Figure 6F). Importantly, and consistent with our earlier observation in 293T cells, we also observed dramatic increase in total RNA pol II at the TSS region of the *CRABP2* gene. We conclude that ZMYND8 knockout causes a defect in P-TEFb recruitment and CTD phosphorylation of RNA pol II at the promoter proximal region of *CRABP2* gene



**Figure 5. Positive Regulation of Expression of Diverse Set of Genes *In Vivo* by ZMYND8 through P-TEFb Complex Recruitment**

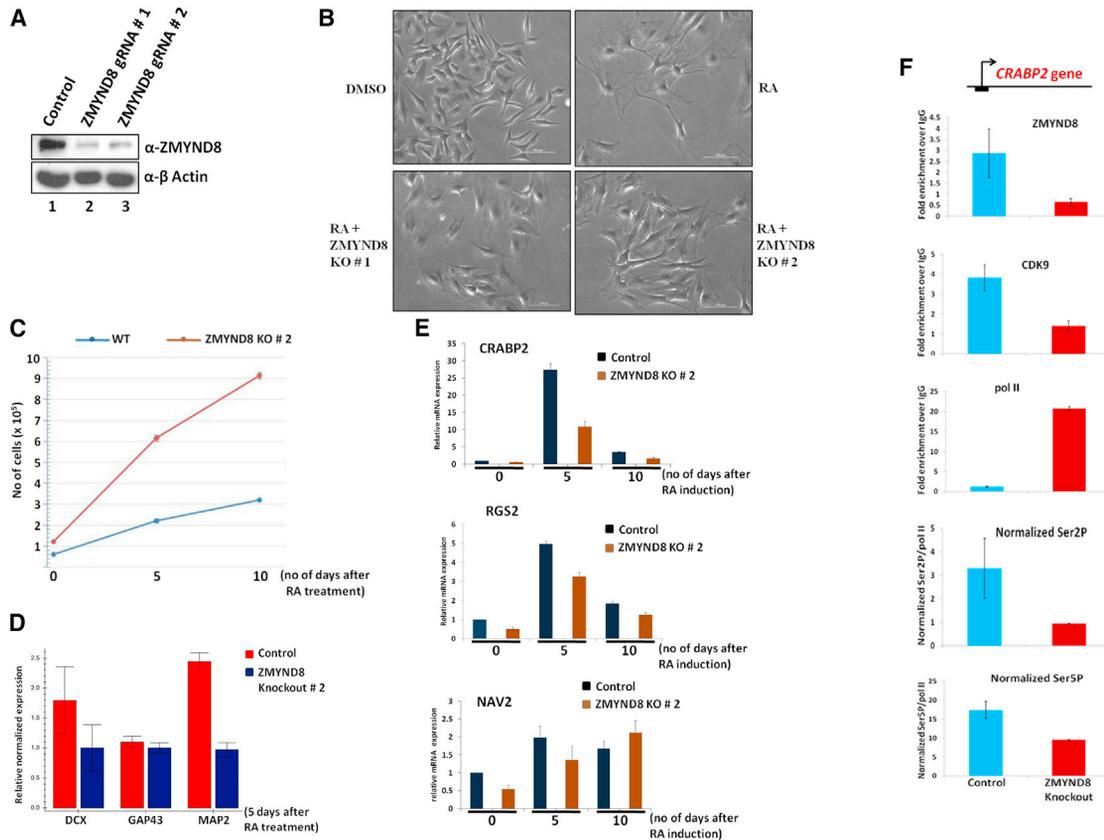
(A) Generation of CRISPR-CAS9-mediated knockout of ZMYND8 in 293T cells. Western blot analysis was performed for confirming ZMYND8 knockout. (B) Lentiviral short hairpin RNA (shRNA)-mediated stable knockdown of CDK9 in 293T cells. Western blot (left panel) and RNA analysis (right panel) were performed for confirming CDK9 knockdown. (C) qRT-PCR analysis showing the effect of ZMYND8 knockout and CDK9 knockdown on target gene expression. RNA isolated from ZMYND8 knockout and CDK9 knockdown cells (Figures 5A and 5B) was subjected to qRT-PCR analysis to check the expression of target genes. (D–H) ChIP analyses showing the effect of ZMYND8 knockout on recruitment of CDK9 and RNA pol II as well as the presence of phosphorylated Ser2 and Ser5 forms of RNA pol II at the TSS regions of target *CRABP2* (D), *DR1* (E), *TK2* (F), *CEBPG* (G), and *CAMK2D* (H) genes. (I) qRT-PCR analysis showing the effect of ZMYND8 knockout on expression of endogenous CDK9 RNA. Data represent mean  $\pm$  SD, a minimum of two biological replicates.

during RA-induced neuronal differentiation of SH-SY5Y cells. A defect in these key events results in increased paused RNA pol II and ultimately affects expression of key target genes and differentiation processes.

### ZMYND8 Interacts Directly with CHD4 Subunit of NuRD Complex

To understand the dual positive and negative roles of ZMYND8 in transcriptional regulation, we initially addressed, based on earlier report (Gong et al., 2015), whether CHD4 subunit of the NuRD complex would directly interact with ZMYND8. To this end, Sf9 expression-based interaction analysis indeed showed CHD4 and ZMYND8 interaction in this heterologous system (Figure 7A) and therefore indicates the existence of direct interactions between these proteins. Furthermore, *in vitro* interaction analysis using purified ZMYND8 and GST-CHD4, and GST

alone, indeed showed direct interactions between these proteins (Figure S6A). MYND domain of ZMYND8 has been implicated for its role in NuRD complex interaction. However, our deletion analysis showed that a ZMYND8 fragment without MYND domain (1–874 aa) also retained CHD4 interaction, albeit weakly, when compared to MYND domain-containing fragments (Figure 7B, compare lane 2 versus lanes 4 and 5). Notably, as shown in Figure 7B, MYND domain-containing C-terminal fragment (851–1,206 aa) failed to show an interaction (lane 6), whereas further deletion of 190 aa (1,041–1,206) fully restored ZMYND8 interaction with CHD4 (compare lane 6 versus lane 7). Therefore, we conclude that direct CHD4 and ZMYND8 interaction involves multiple regions and is not exclusively MYND domain dependent as reported earlier (Gong et al., 2015). Since deletion of C-terminal 851–1,040 region of ZMYND8 regains its interaction with CHD4, we conclude



**Figure 6. ZMYND8 and P-TEFb Complex-Mediated Gene Activation Is Required for Retinoic Acid-Induced Neuronal Differentiation of SH-SY5Y Cells**

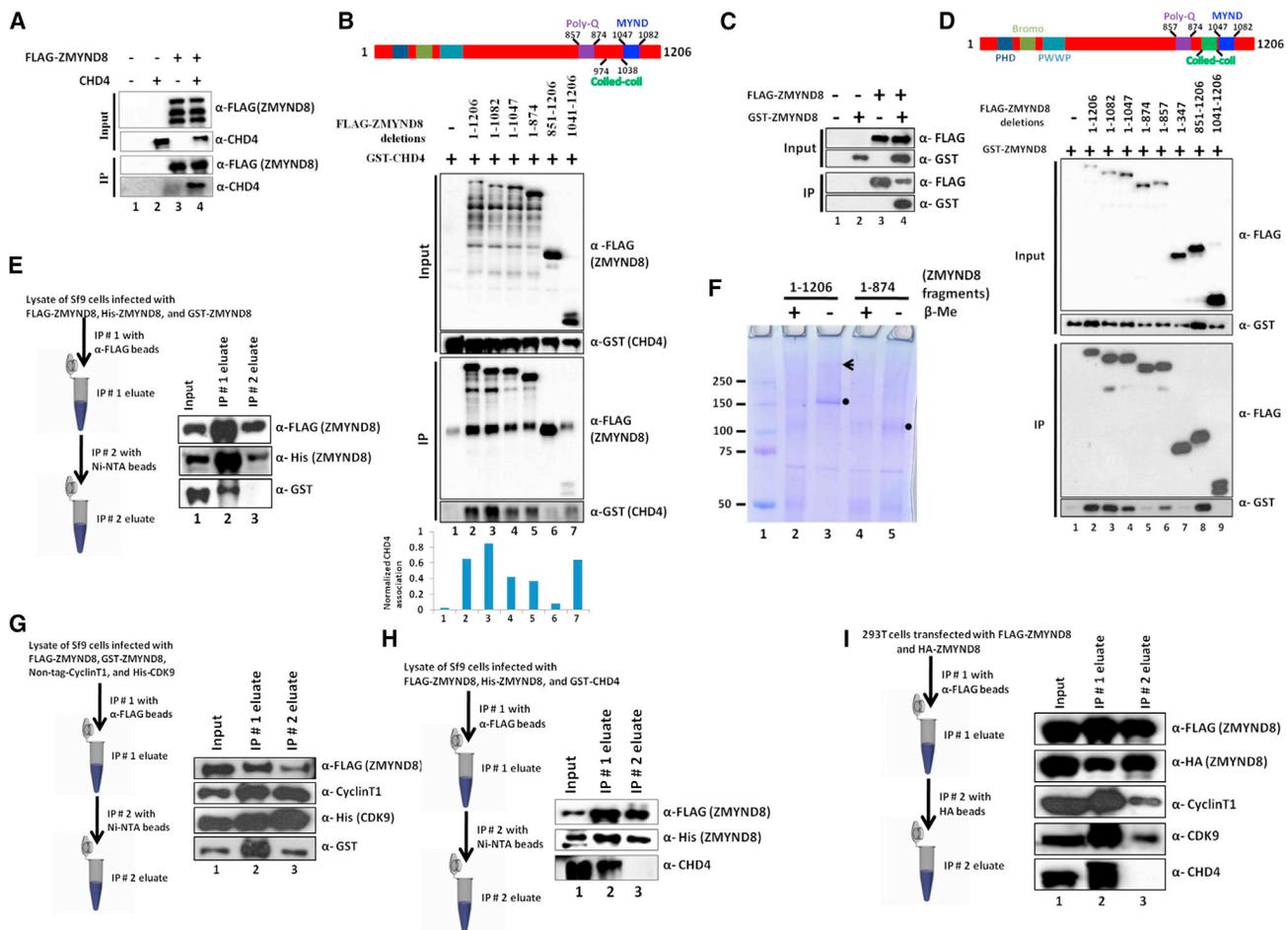
(A) CRISPR/CAS9-mediated ZMYND8 knockout in SH-SY5Y cells as confirmed by western blotting.  
 (B) Effect of ZMYND8 knockout on retinoic acid (RA)-induced differentiation of SH-SY5Y cells. Control DMSO addition did not cause any differentiation (upper left panel), whereas RA addition caused differentiation of the SH-SY5Y cells (upper right panel). Knockout of ZMYND8 severely impaired RA-induced differentiation of this SH-SY5Y cells as evidenced in the bottom left and right panels.  
 (C) Proliferation analysis of control and ZMYND8 knockout SH-SY5Y cells during RA-induced differentiation. The control and ZMYND8-knocked-out SH-SY5Y cells were treated with RA for inducing their differentiation. Cells were harvested after the indicated time periods, and their numbers were counted.  
 (D) qRT-PCR analysis showing the effect of ZMYND8 knockout on the expression of neuronal differentiation-specific marker genes in SH-SY5Y cells during RA-induced differentiation process.  
 (E) qRT-PCR analysis showing the effect of ZMYND8 knockout on the kinetics of expression of differentiation-specific target genes in SH-SY5Y cells after RA treatment. Control and ZMYND8-knocked-out SH-SY5Y cells were treated with RA for the indicated time periods, and cells were harvested to check the expression of target genes as measured by RNA analysis through qRT-PCR.  
 (F) ChIP analysis showing recruitment of ZMYND8, CDK9, and RNA pol II, and abundance of Ser2 and Ser5 phosphorylated forms of RNA pol II on *CRABP2* gene upon ZMYND8 knockout in SH-SY5Y during RA-induced differentiation.  
 Data represent mean  $\pm$  SD, a minimum of two biological replicates.

that the C-terminal 851–1,040 region inhibits MYND domain-dependent CHD4 association.

### ZMYND8 Remains Both as Monomer and Dimer Proteins

Since coiled-coil domain-mediated homodimerization plays an important role in determining positive and negative roles of several transcription factors (Amoutzias et al., 2008), we initially addressed the presence of putative coiled-coil domain structure within ZMYND8. Using publicly available software (McDonnell et al., 2006), we indeed observed the presence of amino acid sequences within the ZMYND8 protein that could potentially form coiled-coil structure (Figure S6B). For gaining further insights into this regulation, our initial Sf9

expression-based interaction analysis showed homooligomerization of ZMYND8 (Figure 7C). Further analyses showed requirement of 875–1,047 aa for ZMYND8 homooligomerization (Figure 7D). Like its interaction with P-TEFb complex (Figure 2F), further deletion of poly-Q region restored ZMYND8 oligomerization slightly (Figure 7D, compare lane 6 versus lane 5). Interestingly, whereas ZMYND8 C-terminal fragment (851–1,206) fully retained its ability to oligomerize, further deletion of 190 aa (encompassing the coiled-coil region) completely abolished ZMYND8 oligomerization (Figure 7D, compare lane 8 versus lane 9). Similar requirement of coiled-coil domain-dependent ZMYND8 oligomerization is also observed in 293T cells (Figure S6C, compare lanes 5 and 6



**Figure 7. ZMYND8 Is Also a Dimerized Protein, and Dimerized ZMYND8 Preferentially Associates with the P-TEFb Complex**

(A) Immunoblot analysis showing the interaction between ZMYND8 and CHD4 subunit of NuRD complex in heterologous Sf9 expression system through their co-expression in Sf9 cells. After IP (using M2 beads), ZMYND8 and CHD4 proteins were identified by western blot analysis using indicated antibodies.

(B) Immunoblotting analysis showing domains of ZMYND8 that interact with CHD4 through their co-expression in Sf9 cells. After IP (using M2 beads), ZMYND8 domains and CHD4 proteins were identified by western blot analysis using the indicated antibodies. The topmost band in each lane represents the actual ZMYND8 fragment in this assay. The other bands represent either ZMYND8 degradation or non-specific proteins in the reactions.

(C) Immunoblot analysis showing oligomerization of ZMYND8 protein through co-expression of FLAG-ZMYND8 and GST-ZMYND8 in Sf9 cells. After IP (using M2 beads), specific ZMYND8 proteins were identified by western blot analysis using indicated antibodies.

(D) Immunoblot analysis showing domains of ZMYND8 that are important for ZMYND8 oligomerization through their co-expression in Sf9 cells. After IP (using M2 beads), FLAG-ZMYND8 domains and GST-ZMYND8 were identified by western blot analysis using indicated antibodies.

(E) Immunoblot analysis showing dimerization of ZMYND8 proteins in heterologous Sf9 expression system. The left panel shows the experimental strategy that was used for this assay, and the right panel shows immunoblot analysis for identifying interacting proteins in each step.

(F) SDS-PAGE Coomassie staining of purified ZMYND8 proteins (1–1,206 and 1–874) showing ZMYND8 monomer and dimer populations in the presence or absence of reducing reagent (β-mercaptoethanol). ZMYND8 protein was expressed and purified from Sf9 cells. As shown in lane 3, full-length ZMYND8 (1–1,206) showed both monomer and dimer populations (lane 3) in absence of reducing reagent, whereas the coiled-coil deletion mutant (1–874) failed to show any dimerized population (lanes 4–5).

(G) Immunoblot analysis showing the interaction of P-TEFb complex with dimerized ZMYND8 in heterologous Sf9 expression system. The left panel shows the experimental strategy that was used for this assay, and the right panel shows immunoblot analysis for identifying interacting proteins in each step.

(H) Immunoblot analysis showing association of CHD4 subunit with monomeric ZMYND8. The left panel shows the experimental strategy that was used for this assay, and the right panel shows the immunoblot analysis for identifying interacting proteins in each step.

(I) Immunoblot analysis showing association of P-TEFb and NuRD complexes with different forms of ZMYND8 in 293T cells. The left panel shows the experimental strategy that was used for this assay, and the right panel shows the immunoblot analysis for identifying interacting proteins in each step.

with lane 4). Therefore, we conclude that through the putative coiled-coil domain, ZMYND8 protein is oligomerized both in heterologous Sf9 expression system as well as in mammalian 293T cells.

Next, to understand the nature of oligomerization, we co-infected Sf9 cells with FLAG-ZMYND8, GST-ZMYND8, and His-ZMYND8 (Figure 7E, left panel for experimental strategy). First round of immunoprecipitation using FLAG-ZMYND8 showed

the presence of both GST-ZMYND8 and His-ZMYND8 (Figure 7E, right panel, lane 2). A subsequent second round of immunoprecipitation using His-ZMYND8, present in first-round eluate, showed the presence of only FLAG-ZMYND8 (Figure 7E, right panel, lane 3). Absence of GST-ZMYND8 in the second-round eluate clearly indicates that ZMYND8 forms homodimer and not any other form of oligomer. Our further analysis with purified ZMYND8 showed the presence of both monomer as well as dimer (as evidenced by their size) in absence of reducing reagent in our assay (Figure 7F, lane 2 versus lane 3) in a coiled-coil domain-dependent manner (Figure 7F, lane 3 versus lane 5). Consistently, cell extract from 293T cells also showed the presence of monomer and dimer population of ZMYND8 in absence of reducing reagent (Figure S6D). Further direct evidence came from atomic force microscopy (AFM) image analysis using purified protein clearly showing the presence of both monomer and dimer populations of ZMYND8 (Figure S6E). Therefore, based on all of this collective evidence, we conclude that human ZMYND8 remains as a population of both monomers and dimers in mammalian as well as in heterologous Sf9 cells (when ectopically expressed).

#### **Dimeric ZMYND8 Preferentially Associates with P-TEFb Complex, Whereas the Monomer Associates with CHD4 Subunit of NuRD Complex**

For further addressing differential ability of monomer and dimer forms of ZMYND8 to interact CHD4 and P-TEFb complex, we ectopically expressed FLAG and GST-tagged ZMYND8, CyclinT1, and His-CDK9 (Figure 7G, left panel for experimental strategy). First round of immunoprecipitation using FLAG-ZMYND8 showed the presence of all of the components, i.e., GST-ZMYND8, His-CDK9, and CyclinT1 (Figure 7G, right panel, lane 2). Interestingly, subsequent second round of immunoprecipitation using His-CDK9, present in the first-round eluate, also showed the presence of both the FLAG-ZMYND8 and GST-ZMYND8 in the immunoprecipitation (IP) samples (Figure 7G, right panel, lane 3). This observation clearly indicates that the P-TEFb complex preferentially associates with dimerized form of ZMYND8.

Subsequent similar analysis with CHD4 subunit of NuRD complex (Figure 7H, left panel for experimental strategy) showed the presence of His-ZMYND8 and GST-CHD4 in the first round of FLAG-ZMYND8 immunoprecipitated sample (Figure 7H, right panel, lane 2). Interestingly, the second round of IP using His-ZMYND8, present in the eluate of the first round, showed the presence of only FLAG-ZMYND8 (Figure 7H) and thus confirmed CHD4 association with monomeric ZMYND8 only. Consistent with this hypothesis, our *in vitro* binding assay with purified proteins showed preferential CHD4 association with ZMYND8 that contains both monomer and dimer than dimer population alone (Figure S7A).

Similar analysis in 293T cells by co-transfection of FLAG and HA-ZMYND8 showed the presence of both the P-TEFb complex and CHD4 in the first-round eluate (Figure 7I, right panel, lane 2). However, a second round of IP using HA-ZMYND8, present in the first eluate, showed presence of only P-TEFb complex and not CHD4. This observation thus indicates an ability of dimerized ZMYND8 to interact with P-TEFb complex and not the CHD4 in

mammalian cells. Overall, we conclude that through its coiled-coil domain-mediated dimerization, ZMYND8 preferentially associates with the P-TEFb complex for activation, whereas the monomer associates with CHD4 subunit of the NuRD complex for repression of transcription.

## **DISCUSSION**

The major highlight of our study is identifying an additional (co) activator function of ZMYND8 through regulation of P-TEFb complex recruitment at the target genes. Beyond its simple interaction, we show minimal ZMYND8-P-TEFb complex formation through direct interaction with CyclinT1. Chromosomally integrated reporter gene expression analysis further shows requirement of domain-dependent P-TEFb complex recruitment by ZMYND8 dictates transcriptional activation. Strong correlation of genome-wide co-occupancy of ZMYND8 and CDK9 at the TSS further substantiates our hypothesis. This is further supported by the evidence that ZMYND8 knockout inhibits target gene expression *in vivo* through impaired P-TEFb complex recruitment at the TSS. Physiologically, a role for ZMYND8-P-TEFb complex-mediated transcriptional activation has been shown to be important for RA-induced differentiation of neuroblastoma cells. Finally, our analyses also showed preferential association of P-TEFb complex and CHD4 subunit of NuRD complex with dimerized and monomeric ZMYND8, respectively, and thus provided mechanistic insight into the dual positive and negative regulation of transcription by ZMYND8. A proposed model for this mechanism of action is illustrated in Figure S7C.

#### **Additional Recruiter of P-TEFb Complex at the Promoter Proximal Region for Transcriptional Activation**

Beyond already known factors, our study has identified ZMYND8 as another recruiter of P-TEFb complex at the TSS region of target genes. Our evidence that ZMYND8-P-TEFb complex does not show any association with known P-TEFb interactors, such as AF4 and Brd4, by multiple analyses (Figures 1D, 1F, S1D, and S2E) corroborates the idea of presence of a distinct ZMYND8-P-TEFb complex for transcriptional activation. Although they are distinct in nature, questions remain as to whether a specific set of genes would be more attracted toward ZMYND8-P-TEFb complex for their transcriptional activation than the other P-TEFb-containing complexes? It may also be possible that during high transcriptional activity, a given gene may require multiple ways of recruiting P-TEFb complex for optimal transcriptional output. Consistent with this hypothesis, a recent study has shown the requirement of multiple ways of P-TEFb complex recruitment for optimal transcriptional activation of target genes (Lu et al., 2016). Therefore, further studies are required for understanding the complex nature of transcriptional regulation by several P-TEFb-containing complexes.

#### **Chromatin Signatures and Association of ZMYND-P-TEFb Complex at Promoters, Coding Regions, and Enhancers**

Global ChIP-seq analyses showed a strong presence of ZMYND8 at the TSS region of the transcriptionally active genes (Li et al., 2016; Savitsky et al., 2016; Shen et al., 2016). These

associations are also correlated well with transcriptionally active histone marks such as histone H3K14Ac and H4K16Ac marks. However, mechanistic understanding of the importance of these associations in target gene activation is still elusive. Our study provides an explanation of these observations wherein ZMYND8 activates transcription through recruitment of P-TEFb complex at the promoter proximal region.

Based on the reported association of ZMYND8 and CDK9 with active enhancer and super-enhancer region by several independent studies (Lovén et al., 2013; Malovannaya et al., 2011; Shen et al., 2016; Whyte et al., 2013), it is tempting to speculate that beyond the proposed role of ZMYND8 in downregulating enhancer activation, it could also play an important role in enhancer activation through P-TEFb complex recruitment in a ZMYND8-KDM5C-independent manner. Like ZMYND8, another bromodomain-containing protein, Brd4, has also been proposed to recruit P-TEFb complex for enhancer and super-enhancer activation (Lovén et al., 2013; Whyte et al., 2013). To address whether a combined action of both the Brd4 and ZMYND8 would regulate P-TEFb recruitment for optimal enhancer and super-enhancer activation would require further studies.

### The Dual Roles of ZMYND8 as an Activator and Repressor of Target Gene Expression

Earlier studies have shown that quite a few other transcription factors have both positive and negative roles in regulating transcription. For example, ETO protein is homooligomerized to recruit both the repressor Sin3 and activator AETFC complex (Sun et al., 2013; Zhang et al., 2001). Furthermore, heterodimerization between Myc and Max as well as Mad and Max also determines their association with the cognate activators and repressor proteins/complexes (Grandori et al., 2000; Lüscher, 2001). Coiled-coil domain-mediated dimerization plays an important role in determining activator or repressor role of the majority of these proteins.

Along the similar lines of regulation, we have also observed homodimerization of ZMYND8 dictates preferential association with activator P-TEFb complex, whereas the monomer form associates with CHD4 subunit of the NuRD complex (Figure 7). ZMYND8 is found in multiple isoforms. Interestingly, multiple sequence alignment shows absolute presence coiled-coil region among all the ZMYND8 isoforms excepting isoform 22 (Figure S7B). This observation clearly emphasizes an importance of this domain in critical functional regulation. Considering the high number of isoforms present within mammalian cells (a total of 23, based on Uniprot annotation), ZMYND8 may create multiple different surfaces through homodimerization for binding of factors that are yet to be explored.

How do the monomer and dimer forms of ZMYND8 differentially interact with CHD4 and P-TEFb complex? From our multiple analyses, it is clear that human ZMYND8 can remain both as monomer and dimer forms to interact with specific complexes under normal cellular growth. We propose that depending on the external stimuli, the coiled-coil domain-mediated dimerization would induce a conformational change leading to the adjacent MYND domain inaccessible for CHD4 interaction. Although CHD4 interacts with the other regions of ZMYND8, this interaction is weaker compared to the MYND domain-dependent inter-

action (Figure 7B, lanes 4 and 5). Since P-TEFb interaction with the ZMYND8 protein is primarily at the N terminus of ZMYND8 (1–857, Figures 2E and 2F), the coiled-coil domain-mediated dimerization shifts the cellular pool of ZMYND8 to be more accessible to the P-TEFb complex binding than the NuRD complex for transcriptional activation. We further propose that this regulation is dynamic in nature and would depend on the environmental stimuli. Further studies are required for detailed mechanistic understanding of the stimuli-driven conformational change of ZMYND8 for its preferential association with different complexes for regulating several biological processes.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Cell Culture and Transfection Experiments
  - Construction of Plasmids Used in This Study
  - Generation of FLAG-HA-ZMYND8 and FLAG-HA-CDK9 Stable Cell Lines
  - Mass Spectrometry Analysis of CDK9-Associated Proteins
  - Generation of Knockdown and Knockout Cells
  - Nuclear Extract Preparation
  - Immunoprecipitation Analysis
  - Purification of CDK9 and ZMYND8 Protein Complexes from Nuclear Extract
  - Kinase Assays for P-TEFb Activity
  - Glycerol Gradient Analysis
  - Baculovirus Expression-Based Reconstitution Analyses
  - Baculovirus Expression-Based Interaction Analyses
  - Luciferase Reporter Assays
  - qRT-PCR Analysis for RNA Expression
  - ChIP Analyses
  - ChIP-Seq Analyses
  - GO Analysis
  - Differentiation of SH-SY5Y Cells upon RA Treatment
  - Cell Proliferation Assay
  - Atomic Force Microscopy (AFM) Image Capturing
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

### SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and three tables and can be found with this article online at <https://doi.org/10.1016/j.celrep.2018.07.064>.

### ACKNOWLEDGMENTS

This work was supported in part by Ramalingaswami Re-entry Fellowship and Wellcome Trust DBT Intermediate Fellowship (IA/14/1/501287) awarded to D.B. K.G. is a recipient of CSIR-Senior Research Fellowship, and N.K. and S.B. are the recipients of the UGC Senior Research Fellowship. D.P.M. is a

recipient of a DBT Fellowship. We also would like to thank Dipika Yadav and Mahesh Barad for their contribution in generating a few plasmid constructs used in this study. Furthermore, we would also like to thank Dr. Smarajit Polley from Bose Institute for his technical help. The contribution of CSIR-IICB AFM facility towards generating the AFM images for ZMYND8 is also acknowledged. We also acknowledge initial funding from CSIR-IICB during the early stage of this work.

## AUTHOR CONTRIBUTIONS

K.G. performed the majority of the experiments in consultation with D.B. M.T. performed all of the analysis pertaining to ChIP-seq data in consultation with K.R. and D.B. N.K., S.B., A.N., and D.P.M. contributed toward generating important constructs and baculoviruses that were used in this study. K.G., K.R., and D.B. wrote the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 23, 2017

Revised: July 2, 2018

Accepted: July 18, 2018

Published: August 21, 2018

## REFERENCES

- Adhikary, S., Sanyal, S., Basu, M., Sengupta, I., Sen, S., Srivastava, D.K., Roy, S., and Das, C. (2016). Selective recognition of H3.1K36 dimethylation/H4K16 acetylation facilitates the regulation of all-trans-retinoic acid (ATRA)-responsive genes by putative chromatin reader ZMYND8. *J. Biol. Chem.* *291*, 2664–2681.
- Amoutzias, G.D., Robertson, D.L., Van de Peer, Y., and Oliver, S.G. (2008). Choose your partners: dimerization in eukaryotic transcription factors. *Trends Biochem. Sci.* *33*, 220–229.
- Basu, M., Sengupta, I., Khan, M.W., Srivastava, D.K., Chakrabarti, P., Roy, S., and Das, C. (2017). Dual histone reader ZMYND8 inhibits cancer cell invasion by positively regulating epithelial genes. *Biochem. J.* *474*, 1919–1934.
- Biswas, D., Milne, T.A., Basur, V., Kim, J., Elenitoba-Johnson, K.S., Allis, C.D., and Roeder, R.G. (2011). Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc. Natl. Acad. Sci. USA* *108*, 15751–15756.
- Chen, R., Yang, Z., and Zhou, Q. (2004). Phosphorylated positive transcription elongation factor b (P-TEFb) is tagged for inhibition through association with 7SK snRNA. *J. Biol. Chem.* *279*, 4153–4160.
- Gel, B., Díez-Villanueva, A., Serra, E., Buschbeck, M., Peinado, M.A., and Malinverni, R. (2016). regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics* *32*, 289–291.
- Gong, F., Chiu, L.Y., Cox, B., Aymard, F., Clouaire, T., Leung, J.W., Cammarata, M., Perez, M., Agarwal, P., Brodbelt, J.S., et al. (2015). Screen identifies bromodomain protein ZMYND8 in chromatin recognition of transcription-associated DNA damage that promotes homologous recombination. *Genes Dev.* *29*, 197–211.
- Grandori, C., Cowley, S.M., James, L.P., and Eisenman, R.N. (2000). The Myc/Mad network and the transcriptional control of cell behavior. *Annu. Rev. Cell Dev. Biol.* *16*, 653–699.
- He, N., Jahchan, N.S., Hong, E., Li, Q., Bayfield, M.A., Maraja, 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.
- 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.
- Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* *4*, 44–57.
- 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.
- Jiang, H., Lu, X., Shimada, M., Dou, Y., Tang, Z., and Roeder, R.G. (2013). Regulation of transcription by the MLL2 complex and MLL complex-associated AKAP95. *Nat. Struct. Mol. Biol.* *20*, 1156–1163.
- Kovalevich, J., and Langford, D. (2013). Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. *Methods Mol. Biol.* *1078*, 9–21.
- Laubert, S.M., Nakayama, T., Wu, X., Ferris, A.L., Tang, Z., Hughes, S.H., and Roeder, R.G. (2013). H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* *152*, 1021–1036.
- Li, Q., Price, J.P., Byers, S.A., Cheng, D., Peng, J., and Price, D.H. (2005). Analysis of the large inactive P-TEFb complex indicates that it contains one 7SK molecule, a dimer of HEXIM1 or HEXIM2, and two P-TEFb molecules containing Cdk9 phosphorylated at threonine 186. *J. Biol. Chem.* *280*, 28819–28826.
- Li, N., Li, Y., Lv, J., Zheng, X., Wen, H., Shen, H., Zhu, G., Chen, T.Y., Dhar, S.S., Kan, P.Y., et al. (2016). ZMYND8 reads the dual histone mark H3K4me1-H3K14ac to antagonize the expression of metastasis-linked genes. *Mol. Cell* *63*, 470–484.
- 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.
- Liu, W., Ma, Q., Wong, K., Li, W., Ohgi, K., Zhang, J., Aggarwal, A., and Rosefield, M.G. (2013). Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell* *155*, 1581–1595.
- Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* *153*, 320–334.
- Lu, H., Li, Z., Xue, Y., Schulze-Gahmen, U., Johnson, J.R., Krogan, N.J., Alber, T., and Zhou, Q. (2014). AFF1 is a ubiquitous P-TEFb partner to enable Tat extraction of P-TEFb from 7SK snRNP and formation of SECs for HIV transactivation. *Proc. Natl. Acad. Sci. USA* *111*, E15–E24.
- Lu, X., Zhu, X., Li, Y., Liu, M., Yu, B., Wang, Y., Rao, M., Yang, H., Zhou, K., Wang, Y., et al. (2016). Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. *Nucleic Acids Res.* *44*, 6853–6867.
- Lüscher, B. (2001). Function and regulation of the transcription factors of the Myc/Max/Mad network. *Gene* *277*, 1–14.
- Malovannaya, A., Lanz, R.B., Jung, S.Y., Bulynko, Y., Le, N.T., Chan, D.W., Ding, C., Shi, Y., Yucer, N., Krenciute, G., et al. (2011). Analysis of the human endogenous coregulator complexome. *Cell* *145*, 787–799.
- McDonnell, A.V., Jiang, T., Keating, A.E., and Berger, B. (2006). Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* *22*, 356–358.
- Papadidis, N.F., Durvale, M.C., and Canduri, F. (2017). The emerging picture of CDK9/P-TEFb: more than 20 years of advances since PITALRE. *Mol. Biosyst.* *13*, 246–276.
- Peng, J., Zhu, Y., Milton, J.T., and Price, D.H. (1998). Identification of multiple cyclin subunits of human P-TEFb. *Genes Dev.* *12*, 755–762.
- Savitsky, P., Krojer, T., Fujisawa, T., Lambert, J.P., Picaud, S., Wang, C.Y., Shanle, E.K., Krajewski, K., Friedrichsen, H., Kanapin, A., et al. (2016). Multivalent histone and DNA engagement by a PHD/BRD/PWWP triple reader cassette recruits ZMYND8 to K14ac-rich chromatin. *Cell Rep.* *17*, 2724–2737.
- Shen, H., Xu, W., Guo, R., Rong, B., Gu, L., Wang, Z., He, C., Zheng, L., Hu, X., Hu, Z., et al. (2016). Suppression of enhancer overactivation by a RACK7-histone demethylase complex. *Cell* *165*, 331–342.

- Sun, X.J., Wang, Z., Wang, L., Jiang, Y., Kost, N., Soong, T.D., Chen, W.Y., Tang, Z., Nakadai, T., Elemento, O., et al. (2013). A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. *Nature* *500*, 93–97.
- Tang, M. (2017). pyflow-ChIPseq: a snakemake based ChIP-seq pipeline. *Zenodo*. <https://doi.org/10.5281/zenodo.819971>.
- Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* *153*, 307–319.
- Winter, G.E., Mayer, A., Buckley, D.L., Erb, M.A., Roderick, J.E., Vittori, S., Reyes, J.M., di Iulio, J., Souza, A., Ott, C.J., et al. (2017). BET bromodomain proteins function as master transcription elongation factors independent of CDK9 recruitment. *Mol. Cell* *67*, 5–18.e19.
- Xicoy, H., Wieringa, B., and Martens, G.J. (2017). The SH-SY5Y cell line in Parkinson's disease research: a systematic review. *Mol. Neurodegener.* *12*, 10.
- Yamaguchi, Y., Shibata, H., and Handa, H. (2013). Transcription elongation factors DSIF and NELF: promoter-proximal pausing and beyond. *Biochim. Biophys. Acta* *1829*, 98–104.
- Yang, Z., Zhu, Q., Luo, K., and Zhou, Q. (2001). The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* *414*, 317–322.
- 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.
- Young, T.M., Wang, Q., Pe'ery, T., and Mathews, M.B. (2003). The human I-mfa domain-containing protein, HIC, interacts with cyclin T1 and modulates P-TEFb-dependent transcription. *Mol. Cell. Biol.* *23*, 6373–6384.
- Zhang, J., Hug, B.A., Huang, E.Y., Chen, C.W., Gelmetti, V., Maccarana, M., Minucci, S., Pelicci, P.G., and Lazar, M.A. (2001). Oligomerization of ETO is obligatory for corepressor interaction. *Mol. Cell. Biol.* *21*, 156–163.
- Zhou, Q., Li, T., and Price, D.H. (2012). RNA polymerase II elongation control. *Annu. Rev. Biochem.* *81*, 119–143.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
ZMYND8	Bethyl Lab	Cat# A302-090A
CyclinT1	Santa Cruz Biotechnology	Cat# sc-10750
CDK9	Santa Cruz Biotechnology and Bethyl Lab	Cat# sc-8338 and Cat# A303-492A
FLAG epitope	Sigma	Cat# F7425
GST	Santa Cruz Biotechnology	Cat# sc-53909
His epitope	Santa Cruz Biotechnology	Cat# sc-8036
AF4	Abcam	Cat# ab31812
AF9	Bethyl Lab	Cat# A300-596
ZNF592	Bethyl Lab	Cat# A301-531
ZNF687	Bethyl Lab	Cat# A303-279
Rpb1 (4H8 clone)	Cell Signaling Technology	Cat# 2629
Phospho Rpb1 CTD (Ser2)	Cell Signaling Technology	Cat# 13499
Phospho Rpb1 CTD (Ser5)	Cell Signaling Technology	Cat# 13523
Beta Actin	Santa Cruz Biotechnology	Cat# sc47778
Hexim1	Cell Signaling Technology	Cat# 9064
CHD4	Cell Signaling Technology	Cat# 11912
HRP conjugated secondary antibody	Bio-Rad	Cat# 1706515
<b>Bacterial and Virus Strains</b>		
DH5 $\alpha$ <i>E.Coli</i>	This lab	N/A
BI21 <i>E.Coli</i>	This lab	N/A
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
dNTP	BioBharatiLifeScience	Cat#BB-C0020B
ATP	BioBharatiLifeScience	Cat#BB-C0030
Phusion high fidelity DNA polymerase	NEB	Cat# M0530S
T4DNA ligase	NEB	Cat# M0202S
3X Flag peptide	Sigma	F4799-25MG
L-Glutathione reduced	Sigma	G425-25G
Retinoic acid	Sigma	R2625-50MG
Verso c-DNA synthesis kit	Life Technologies	Cat# AB1453A
iTaq Universal SYBR Green Supermix	Bio-Rad	Cat# 172-5201AP
TRIzol	Life Technologies	Cat# 15596026
Penicillin and Streptomycin	Life Technologies	Cat# 1510163
Gentamycin	Gold Biotechnology	Cat# G-400-100
Purified recombinant GST-CTD	This study	N/A
Purified recombinant P-TEFb complex	This study	N/A
Purified recombinant ZMYND8 protein	This study	N/A
<b>Critical Commercial Assays</b>		
Dual-glow luciferase assay	Promega	Cat# E2920
<b>Deposited Data</b>		
Unprocessed raw western data	This study (Mendeley data)	DOI: <a href="https://doi.org/10.17632/gx863ffxc.1#file-027fd3f7-139f-4940-8412-6c01f2f28159">https://doi.org/10.17632/gx863ffxc.1#file-027fd3f7-139f-4940-8412-6c01f2f28159</a>
Genes that show ZMYND8 and CDK9 co-occupancy in 293T cells	This study (Mendeley data)	DOI: <a href="https://doi.org/10.17632/hvgpbgpjfw.1#file-e6cb491d-40d8-4d94-b247-96daf4d673a">https://doi.org/10.17632/hvgpbgpjfw.1#file-e6cb491d-40d8-4d94-b247-96daf4d673a</a>

(Continued on next page)

<b>Continued</b>		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Cell Lines		
Flp-In <sup>TM</sup> -293	Invitrogen	Cat# R75007
Flp-In <sup>TM</sup> T-REx <sup>TM</sup> 293	Invitrogen	Cat# R78007
Luciferase reporter cell line	<a href="#">Jiang et al., 2013</a>	N/A
Human SH-SY5Y neuroblastoma cell line	ATCC	CRL-2266
Sf9 insect cell line	Invitrogen	B-82501
HEK293T	ATCC	CRL-3216
Oligonucleotides		
gRNA # 1 (5'-GTGATGTGTCCTGCGGCGAG-3') for knocking out ZMYND8 cloned in LentiCRISPRv2	<a href="#">Shen et al., 2016</a>	N/A
gRNA # 2 (5'-GACTTAGCGTGATAAACCCG-3') for knocking out ZMYND8 cloned in LentiCRISPRv2	<a href="#">Shen et al., 2016</a>	N/A
CDK9 KD #1 Upper oligo (CCGGCCCGCTGCAAGG GTAGTATATACTCGAGTATATACTACCCTTG CAG CGTTTTTG)	This paper	N/A
CDK9 KD #1 lower oligo (AATTCAAAAACCGCTG CAAGGGTAGTATATACTCGAGTATATACTACCC TTGCAGCGG) for CDK9 knockdown cloned into pLKO.1-TRC	This paper	N/A
CDK9 KD #2 Upper oligo (CCGGAGGGACATGAA GGCTGCTAATCTCGAGATTAGCAGCCTTCATG TCCTTTTTTG)	This paper	N/A
CDK9 KD #2 lower oligo(AATTCAAAAAAGGGACA TGAAGGCTGCTAATCTCGAGATTAGCAGCCTTC ATGTCCCT) for CDK9 knockdown cloned into pLKO.1-TRC	This Paper	N/A
Primers for RNA analyses	This paper	<a href="#">Table S2</a>
Primers for ChIP analyses	This paper	<a href="#">Table S3</a>
Recombinant DNA		
Plasmid constructs used in this study	This paper	<a href="#">Table S1</a>
LentiCRISPRv2	Addgene	Plasmid # 52961
pLKO.1-TRC	Addgene	Plasmid #10878
ZMYND8 cDNA pGEMT vector	Sino-Biological Inc.	Cat # HG16278-G
Software and Algorithms		
Paircoil2	McDonnell., 2006	N/A
Excel	Excel software	N/A
IGV	Broad Institute	N/A
pyflow-ChIPseq	<a href="#">Tang, 2017</a>	N/A
DAVID	<a href="#">Huang et al., 2009</a>	N/A
Pico view version 1.10.1	Agilent Technologies	N/A
Other		
Protein A agarose	Thermo fisher Scientific	Cat# 20333
Sure beads <sup>TM</sup> Protein G magnetic beads	Bio-Rad	Cat# 161-4013
Anti-flag M2 agarose beads	Sigma	Cat# 8823-5ML
Anti-flag M2 magnetic beads	Sigma	Cat# F2426
GST agarose beads	GOLDBIOM	Cat#G-250-10
Ni-NTA agarose beads	QIAGEN	Cat# 30210
ZMYND8 ChIP-seq data	<a href="#">Savitsky et al., 2016</a>	GEO: GSE81696
CDK9 ChIP-seq data	<a href="#">Liu et al., 2013</a>	GEO: GSE51633

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PCR purification kit	QIAGEN	Cat# 28104
Gel extraction kit	QIAGEN	Cat# 28706
Plasmid miniprep kit	QIAGEN	Cat# 27104

**CONTACT FOR REAGENT AND RESOURCE SHARING**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Debra Biswas ([dbiswas@iicb.res.in](mailto:dbiswas@iicb.res.in)).

**EXPERIMENTAL MODEL AND SUBJECT DETAILS**

For all of our experiments, we have used either HEK293T or SHSY-5Y or Sf9 insect cell culture. HEK293T and SHSY-5Y cells were maintained in DMEM media supplemented with high glucose and 10% FBS and 1X penicillin-streptomycin solution. Sf9 cells were cultured in Grace's insect media supplemented with 10% FBS and gentamicin at 26°C.

**METHOD DETAILS****Cell Culture and Transfection Experiments**

Unless otherwise mentioned, all mammalian cell lines used in this study, were grown in DMEM media with high glucose (Invitrogen inc.) supplemented with 10% FBS, penicillin and streptomycin. Sf9 insect cells were grown in Grace's insect media supplemented with 10% FBS and gentamicin (7ug/ml). Transfections of mammalian cells were performed using Fugene Transfection Reagent following manufacturer's protocol. For protein-protein interaction analyses, cells were harvested 48 hr after transfection for further immunoprecipitation analyses.

**Construction of Plasmids Used in This Study**

The cDNA for ZMYND8 long (1-1206 amino acid, isoform 10 as described in Uniprot) was purchased from SinoBiological Inc. ZMYND8 cDNA was sequenced in its entirety before being used for further cloning purposes. Full-length ZMYND8 was cloned into pcDNA5-FRT-TO vector. All the deletion constructs of ZMYND8 were made in 1x epitope tag-containing pcDNA5-FRT-TO as well as pFASTBac vectors using appropriate restriction enzymes. Details of the restriction enzymes used for cloning is available upon request from the corresponding author of this study. Plasmid constructs cloned in pcDNA5-FRT-TO vector were used for expression of target factors in mammalian cells whereas, the constructs cloned in pFASTBac vector were used for expression of target factors in Sf9-cell-based baculovirus expression system.

**Generation of FLAG-HA-ZMYND8 and FLAG-HA-CDK9 Stable Cell Lines**

FLAG-HA-ZMYND8 and FLAG-HA-CDK9 stable cells were made using Flp-In methods in 293T cells as per manufacturer's protocol. After transfection of the Flp-In 293 cells with the appropriate pCND5-FRT-TO vector containing the ZMYND8 and CDK9 ORFs, cells were subjected to hygromycin selection (200ug/ml) for several weeks. Individual colonies obtained by this method was further amplified and screened for expression of target proteins. Harvested cells were lysed in RIPA buffer containing 20mMTris-HCl (pH 7.5), 150mM NaCl, 1mM EDTA 1% NP-40, and 1% SDS. Ectopic expression of target proteins was checked by western blotting using epitope tag-specific antibodies.

**Mass Spectrometry Analysis of CDK9-Associated Proteins**

Mass Spectrometry analysis of CDK9-associated proteins were essentially carried out using protocol as mentioned earlier ([Biswas et. al, 2011](#)). Purified CDK9 protein complexes were separated on a 4%–12% SDS-PAGE and proteins visualized using Coomassie stain. Visible bands were excised and subsequently destained with 20% methanol for 6 hr. After necessary steps of reduction and alkylation of the cysteines, proteins present in the gel pieces were digested overnight with sequencing grade, modified trypsin. Resulting peptides were resolved on a nanocapillary reverse phase column using a 1% acetic acid/acetonitrile gradient at 400nl/min and separated proteins were directly injected in to a linear ion-trap mass spectrometer (LTQ XL, ThermoFisher). MS/MS spectra of the most intense ion were collected and proteins were identified by searching the data against Human IPI database (v 3.41) using X! Tandem/Trans-Proteomic Pipeline (TPP) software suite.

### Generation of Knockdown and Knockout Cells

For stable knockdown of CDK9, target shRNAs were cloned into lentiviral pLKO.1 puro vector. Lentiviral particles were generated through co-transfection of 250ng target shRNAs, 500ng of psPAX2 (packaging plasmid), and 500ng of pMD2.G plasmid (envelope plasmid) in  $3 \times 10^5$  cells in a well of 6 well plate. 72 hr after transfection, virus particles were collected and stored in  $-80^{\circ}\text{C}$  for further use. 100  $\mu\text{l}$  of virus particles along with 8  $\mu\text{g}/\text{mL}$  of polybrene was used for transduction of  $3 \times 10^5$  cells in each well of 6 well plate. 24 hr after transduction, cells were subjected to puromycin selection (3 $\mu\text{g}/\text{ml}$ ). Puromycin positive cells were further checked for knocking down of target factors through western blot analysis using target factor specific antibodies and by RNA analysis using qRT-PCR.

For generating stable ZMYND8 knockout cells, protocols were followed as described earlier by Shen et al., (Shen et al., 2016). Target gRNAs were cloned into LentiCRISPRv2 vector. Lentiviral particles were generated using the same protocol as described above. 48hrs after transduction, transduced cells were plated in low dilution in a 100mm dish in puromycin-containing media. Cells were allowed to grow till individual visible colonies come up. Several colonies were screened for checking knockout of ZMYND8 through western blot analysis using ZMYND8 antibody.

### Nuclear Extract Preparation

For nuclear extract preparation, harvested cells were spun down and packed cell volume (PCV) was measured. Cells were further resuspended in 2X PCV of NE1 buffer containing 10mM Tris-Cl (pH 7.3), 1.5mM  $\text{MgCl}_2$ , 10mM NaCl, and 0.7  $\mu\text{l}/\text{ml}$   $\beta$ -mercaptoethanol and incubated for 15min on ice. The swollen cells were subsequently passaged through a 23 gauge needle for 8 times. The resultant cell lysate was spun down at 6000 RPM for 5min at  $4^{\circ}\text{C}$  to obtain the nuclear pellet. Nuclear pellet was subsequently resuspended by vortexing in 0.5X volume of NPV using pre-chilled NE2 buffer containing 20mM Tris-Cl (pH7.3), 1.5mM  $\text{MgCl}_2$ , 20mM NaCl, 0.2mM EDTA, 25% glycerol + 0.7  $\mu\text{l}/\text{ml}$   $\beta$ -mercaptoethanol, and protease inhibitor cocktail from Roche. Subsequently, equal amount of NE3 (20mM Tris-Cl (pH7.3), 1.5mM  $\text{MgCl}_2$ , 1.2M NaCl, 0.2mM EDTA, 25% glycerol + 0.7  $\mu\text{l}/\text{ml}$   $\beta$ -mercaptoethanol, and protease inhibitor cocktail) was added and incubated on ice for 45min. Samples were vortexed after every 3min for efficient mixing of cells and nuclear extract buffer. Finally, nuclear extract was separated by spinning down the sample at 18000 RPM at  $4^{\circ}\text{C}$  for 20min. The resulting supernatant containing the nuclear extract was collected and either used directly for experimental analysis or stored in  $-80^{\circ}\text{C}$  for future use.

### Immunoprecipitation Analysis

To test the endogenous interaction between ZMYND8 and P-TEFb, nuclear extract (approximately 500 $\mu\text{g}$ ) prepared from the 293T cells was pre-cleared using protein-A agarose beads for 2-3 hr at  $4^{\circ}\text{C}$ . The pre-cleared extract was subsequently used for immunoprecipitation with anti- ZMYND8 antibody for overnight. Antibody-bound proteins were pulled-down using protein-G magnetic beads. Beads were washed rigorously with the same buffer that was used for immunoprecipitation. Bound proteins were eluted by boiling in 1X SDS loading dye for 3 min at  $95^{\circ}\text{C}$ . The sample was analyzed by western blotting using the indicated antibodies.

For all other epitope-tagged mammalian proteins, after pre-clearing, target proteins were pulled-down using commercially available tag-specific antibody-coated magnetic/agarose beads. After rigorous washing, bound proteins were eluted by boiling in 1X SDS loading dye for 3 min at  $95^{\circ}\text{C}$ . The sample was analyzed by western blotting using the indicated antibodies.

### Purification of CDK9 and ZMYND8 Protein Complexes from Nuclear Extract

Nuclear extract isolated from large volume of cells, stably expressing target proteins, were initially precleared using protein-A agarose beads for 2hrs at  $4^{\circ}\text{C}$ . The precleared nuclear extract was incubated with FLAG antibody-coated (M2) magnetic beads for overnight. Magnetic bead bound protein complexes were isolated by spinning down the extract at 4000 RPM for 5min at  $4^{\circ}\text{C}$ . Subsequently, beads were rigorously washed using washing buffer (20mM Tris-Cl (pH7.3), 1.5mM  $\text{MgCl}_2$ , 500M KCl, 0.2mM EDTA, 25% glycerol, 0.7 $\mu\text{l}/\text{ml}$   $\beta$ -mercaptoethanol, and protease inhibitor cocktail). Bound proteins were eluted by incubating with 3X FLAG peptide solution (250ng/ $\mu\text{l}$ ) for 30min at  $4^{\circ}\text{C}$ . Eluted proteins were separated on a 4%–12% gel and silver stained for their visualization.

### Kinase Assays for P-TEFb Activity

Kinase assays were performed using ZMYND8-P-TEFb complexes that were purified from either mammalian or Sf9 cells. For the purpose of kinase assays, protein complexes were eluted using 3X FLAG peptide (250ng/ $\mu\text{l}$ ). Bacterially expressed and purified recombinant GST-CTD was used in the kinase assay as substrate. Reactions were carried out in 1X kinase assay buffer (50 mM Tris-cl pH8.0, 2mM  $\text{MgCl}_2$ , 500  $\mu\text{M}$  ATP) at  $30^{\circ}\text{C}$  for 2 hr. Phosphorylated CTD was analyzed by western blotting using indicated Ser2 and Ser5 phospho-specific antibodies as mentioned.

### Glycerol Gradient Analysis

20%–40% glycerol gradient was prepared manually and was left overnight at  $4^{\circ}\text{C}$  for uniform gradient formation. Subsequently, protein complexes purified from nuclear extract of FLAG-HA-ZMYND8 stable cell line was loaded on the top of the gradient. Protein complexes were fractionated by centrifuging the sample at 40000 RPM for 12 hr at  $4^{\circ}\text{C}$ . Multiple fractions were collected carefully from the top of the gradient and target proteins were identified in each fractions through western blot using factor specific antibodies.

### Baculovirus Expression-Based Reconstitution Analyses

For reconstitution analyses of protein complexes, Sf9 cells were grown around 90% confluency and were subsequently co-infected with baculoviruses expressing indicated proteins. 48 hours after infection, cells were collected and lysate was prepared by resuspending the cells in BC300 buffer (20mM Tris-Cl, 300mM KCl, 2mM EDT, 20% Glycerol, 0.7ul/ml  $\beta$ -mercaptoethanol, 0.1% NP40 and freshly added protease inhibitor cocktail) and rotating for 2hrs at 4°C. The Crude lysate was subsequently spun down at 15000 rpm for 30 min. Proteins were pulled-down using epitope-tag specific agarose beads (for FLAG tag, M2 agarose beads from Sigma). Bound proteins were competitively eluted using 3X FLAG peptide at a concentration of 250ng/ $\mu$ l. The entire purification protocol was carried out at either at 4°C or on ice.

### Baculovirus Expression-Based Interaction Analyses

Baculovirus expression-based interaction analyses were performed essentially following the same protocol as mentioned earlier (Biswas et al., 2011). In brief, Sf9 cells were co-infected with the combination of viruses as indicated. 48 hours after infection, crude lysate was prepared in BC300 buffer using the protocol mentioned above. Cleared lysate obtained after centrifugation was subsequently used for pulling down the target proteins using epitope-tag specific antibody-coated magnetic or agarose beads. After extensive washes, interacting proteins were identified by western blot analysis using factors specific antibodies as indicated.

### Luciferase Reporter Assays

For reporter assays, stably integrated luciferase reporter cell line was used (Jiang et al., 2013). In this cell line, luciferase reporter gene is expressed downstream of five tandem Gal4-binding sites. Addition of activator, Gal-VP16, strongly stimulates luciferase gene expression. For the reporter assay,  $3 \times 10^5$  cells were seeded in a 6 well plate and transfected with indicated constructs along with synthetic activator Gal-VP16 (10ng). 48 hr after transfection, cells were harvested and reporter activity was checked using dual-glow luciferase assay kit from Promega. Expression of each target construct was analyzed by western blot with indicated antibodies.

### qRT-PCR Analysis for RNA Expression

Total RNA was extracted using TRIzol Reagent (Invitrogen Inc.) following manufacturer's protocol. 1  $\mu$ g of the total RNA was used for cDNA synthesis using verso cDNA synthesis kit from Thermo Scientific following the manufacturer's protocol. The synthesized cDNA was diluted 10X before using for qRT-PCR analysis. qRT-PCR analyses were performed using iTaq Universal SYBR Green Supermix (from BioRad Laboratories) and primers specific for the target genes. Signal from GAPDH was used for normalization of target gene expression.

### ChIP Analyses

ChIP assays were performed essentially following the same protocol as mentioned earlier (Laubert et al., 2013). Target cells for ChIP analyses were cross-linked using 1% formaldehyde (Sigma) for 10 min at room temperature and was subsequently quenched by addition of 125mM glycine. Harvested cells were washed by cold PBS for 3X. Subsequently cells were spun down at 2500 RPM for 10 min. For the purpose of enriching the ChIP signal, nuclear fractions were prepared using ice cold lysis buffer containing 0.5% NP40, 1% Triton X-100, 300mM NaCl, 20mM Tris (pH 7.5), 2mM EDTA and freshly added protease inhibitor cocktail. Nuclear pellet obtained by this method was resuspended in shearing buffer containing 1% SDS, 50mM Tris (pH 8.1), 10mM EDTA, and freshly added protease inhibitor cocktail. The resuspended samples were subsequently sonicated using Bioruptor™ UCD-200 (Diagenode) sonicator for 20 min (30 s on and 30 s off pulse). Before setting up immunoprecipitation, the sonicated samples were centrifuged and the lysates were pre-cleared using 20  $\mu$ L protein-A agarose beads. The cleared lysate was further diluted 10X in dilution buffer containing 0.01% SDS, 1.1% Triton X-100, 1.1mM EDTA, 20mM Tris-Cl (pH 8.0) and 167mM NaCl. Immunoprecipitations were carried out for overnight at 4°C using 2  $\mu$ g of indicated target antibodies for approximately 65  $\mu$ g of sonicated DNA. The antibody-bound factors were immunoprecipitated using 25  $\mu$ L of protein G magnetic beads and rotating for 1 hr at 4°C. Precipitated Protein-G complex was rigorously washed with low salt buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-Cl (pH 8.0), 150mM NaCl, and freshly added Protease inhibitor cocktail); high salt buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-Cl (pH 8.0), 500mM NaCl, and freshly added Protease inhibitor cocktail) and lithium chloride buffer (0.5M LiCl, 1% NP40, 1% deoxycholate, 20 mM Tris-Cl (pH8.0) and 1mM EDTA). The immunoprecipitated DNA was eluted using elution buffer (1% SDS, 0.1M NaHCO<sub>3</sub>). Pulled-down DNA was purified using QIAGEN PCR purification kit. The eluted DNA were either used as such or diluted for the subsequent qRT-PCR analyses (BIORAD CFX96™ Real-Time-System) for quantifying the amount of DNA. IgG and empty vector corresponding to the target antibodies were used as control as mentioned in figure.

### ChIP-Seq Analyses

Publicly available ChIP-seq data for ZMYND8 and CDK9 in HEK293 cells were downloaded from GEO datasets (GSE81696 and GSE51633) and processed by pyflow-ChIPseq, a snakemake-based ChIP-seq pipeline. Overall, raw reads were mapped by bowtie1, duplicated reads were removed, only uniquely mapped reads were retained. RPKM normalized bigwigs were generated by deepTools. Peaks were called using macs1.4 with a *p*-value of 1e-8. Heatmap was generated by R package Enriched Heatmap. Total 1878 peak regions were used in the heatmap. CDK9 and ZMYND8 overlapping peaks were determined by bedtools with a minimal

1 base overlap. The overlapping peaks were then annotated by Homer with the nearest genes. Track for different sets of gene was taken using Integrative genome viewer (IGV) from Broad institute.

### GO Analysis

Genes harboring a strong co-occupancy of ZMYND8 and CDK9 were identified through ChIP-seq analysis. Gene Ontology (GO) analysis of this gene set was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang et al., 2009) tool. ENSEMBL terms were converted to ENTREZ IDs and a functional classification of the genes was carried out on the basis of the biological process they were associated with. All GO terms selected for representation had at least four associated genes with a  $p$ -value of  $< 0.05$ .

### Differentiation of SH-SY5Y Cells upon RA Treatment

Human SH-SY5Y cells were grown in DMEM media containing 10% FBS and 1% penicillin and streptomycin. Differentiation of SH-SY5Y was performed using 10  $\mu$ M All Trans Retinoic Acid (ATRA) dissolved in DMSO (Sigma) for 7-10 days with changing retinoic acid containing media on each alternative day. Microscopic picture of differentiated SHSY-5Y cells were taken using LICA DMI 4000b microscope.

### Cell Proliferation Assay

$6 \times 10^4$  SH-SY5Y cells were seeded in 6 well plate. Cells were treated with 10  $\mu$ M ATRA-containing media for differentiation. Media was changed with fresh ATRA-containing media on each alternative day. Cell number was counted on 5 days and 10 days after ATRA addition by using hemocytometer.

### Atomic Force Microscopy (AFM) Image Capturing

For the purpose of AFM experiment, recombinant ZMYND8 protein was purified from Sf9 cells. 150nM of purified recombinant protein was air-dried on freshly cleaved MICA for 30sec. The mica was washed with autoclaved MilliQ-filtered water. After air drying for 3 min, sample was imaged at room temperature with Pico plus 5500 ILM AFM (Agilent Technologies USA). The images were analyzed with the help of Pico view version 1.10.1 software (Agilent Technologies).

## QUANTIFICATION AND STATISTICAL ANALYSIS

All of our experiments dealing with statistical analysis represent a minimum of two biological replicates. Other pertaining information related to statistical analysis of each figure is mentioned in the legend section of relevant figures. The Z-score and the permutation test to test the significance of overlapping between CDK9 and ZMYND8 was carried out by regioneR bioconductor package (Gel et al., 2016).  $p$ -values in our ChIP-Seq and GO analyses were calculated with embedded software in macs1.4 and DAVID programs.

## DATA AND SOFTWARE AVAILABILITY

Data showing the genes that showed strong overlap of ZMYND8 and CDK9 occupancy in our ChIP-Seq analyses has been uploaded in Mendeley database and is available using following doi: <https://doi.org/10.17632/hvgpbjgpfjw.1#file-e6cb491d-40d8-4d94-b247-96daf4d673a>

For raw images of western blot analyses have been uploaded in the Mendeley database and is available using following doi: <https://doi.org/10.17632/gx863ffxc.1#file-027fd3f7-139f-4940-8412-6c01f2f28159>