

UNDERSTANDING THE FUNCTIONAL REGULATION OF THE RNA POLYMERASE II ELONGATION FACTOR, ELL



**A THESIS SUBMITTED FOR THE FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY (SCIENCE)
OF
JADAVPUR UNIVERSITY
2022**

BY

SUBHAM BASU

**MOLECULAR GENETICS DIVISION
CSIR- INDIAN INSTITUTE OF CHEMICAL BIOLOGY
KOLKATA-700032, INDIA**



सी.एस.आई.आर-भारतीय रासायनिक जीवविज्ञान संस्थान
वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद की एक इकाई
विज्ञान एवं प्रौद्योगिकी मंत्रालय के अधीन, एक स्वायत्त निकाय, भारत सरकार
4, राजा एस. सी. मल्लिक रोड, यादवपुर, कोलकाता - 700 032



CSIR - INDIAN INSTITUTE OF CHEMICAL BIOLOGY

A Unit of Council of Scientific & Industrial Research
An Autonomous Body, under Ministry of Science & Technology, Government of India
4, Raja S. C. Mullick Road, Jadavpur, Kolkata-700 032

CERTIFICATE FROM THE SUPERVISOR

This is to certify that the thesis entitled “ **Understanding the Functional regulation of the RNA Polymerase II Elongation Factor, ELL**” submitted by **Sri Subham Basu** who got his name registered on **11/05/2017 (11th May, 2017)** for the award of Ph. D. (Science) degree of Jadavpur University, is absolutely based upon his own work under the supervision of **Dr. Debabrata Biswas, Senior Principal Scientist, Molecular Genetics Division, CSIR-Indian Institute of Chemical Biology** and that neither this thesis nor any part of it has been submitted for either any degree/diploma or any other academic award anywhere before.

Dr. Biswas 19/04/22

Dr. Debabrata Biswas
Senior Principal Scientist,
CSIR-IICB, Kolkata-700032



Debabrata Biswas, Ph.D.
Senior Principal Scientist,
CSIR - Indian Institute of Chemical Biology
Professor, Academy of Scientific and
Innovative Research (AcSIR)
4, Raja S. C. Mullick Road, Kolkata-700032, India

Preface

The studies conducted for obtaining the findings presented in this thesis titled “**Understanding the Functional regulation of the RNA Polymerase II Elongation Factor, ELL**” were initiated in December, 2014. The results provide a deep mechanistic insight into the functional regulation of a factor essential for RNA Polymerase II-mediated transcriptional elongation in mammalian cells.

The thesis has been divided into seven chapters. The first chapter, **Chapter 1**, deals with some basic concepts of eukaryotic gene regulation which forms the basis of the work carried out in the lab. **Chapter 2** comprises the rationale behind the study and involves details on why understanding the topics covered in this work are necessary to be examined to further our knowledge of transcriptional regulation and understand the cause behind the development of diseases. The detailed methodology of all the experiments performed is described in **Chapter 3**. The results are divided into two parts: Chapters 4 and 5. **Chapter 4** delves into the mechanistic regulation of ELL stability and functions by the action of multiple proteins like DBC1, p300, HDAC3, and Siah1, while **Chapter 5** elucidates the interplay between ELL, DBC1, EAF1, and EAF2, which are part of a regulatory network for affecting each other's levels in a context-dependent manner. The results obtained from this vast study are finally summarized in **Chapter 6**. A brief discussion has been included as **Chapter 7** for the future prospects of the studies as well as the implications the results obtained may indicate for the overall regulation of transcription and the development of diseases.

Due acknowledgments have been made for reporting the observations and findings made by other scientists. The thesis represents original research work that has been conducted. I take the responsibility for any kind of unintentional error in this thesis.

Subham Basu

Molecular Genetics Division

CSIR- Indian Institute of Chemical Biology

Kolkata

Dedication

For My Family

They made me who I am today

Acknowledgments

As I write my Ph.D. thesis in 2022, I look back to 10 years ago, when I first set foot to work in a research lab, for the purpose of my Master's dissertation. The last decade has been a roller coaster ride, with highs and lows, much like a sine curve!! And yet, as I stand at this juncture of my life, I am humbled and ever so grateful for this journey, and even more towards the people who were there by my side and helped me traverse this arduous path. I am feeling so overwhelmed that I have written this paragraph itself about four times, trying to put my thoughts into words in the best possible way. However, I have numerous people to thank for their support, guidance, advice and love, and so, without dilly-dallying, let me get right to it.

At the start, I wish to express my heartfelt thanks to my supervisor, **Dr. Debabrata Biswas**. Of all the people I have interacted with in relation to science, I am yet to come across a more enthusiastic person. His passion and love for the subject know no bounds, and it is for this reason and this reason only that I applied to him, hoping that he will give me an opportunity to work under his guidance. Over the last 7+ years, I have gone to him with doubts on topics from all diversities. I have approached him with doubts regarding things as simple as a vector-insert ligation, or as complex as setting up previously unexplored assays and experiments. His zeal and enthusiasm for exploring the unknown is infectious, and I was particularly struck by how involved he has always been in our work, remaining updated about every single step. I think, on some occasions, he has himself been more excited to know the results of an experiment than many of us. The best part I admire about him is that he is always available for a discussion, no matter what time of the day it is. A very important thing I have learnt from him is that "Do your work, and don't bother about what the results will be. If you do good work, your efforts will surely be recognized and appreciated." So, I thank you, Sir, for believing in me to undertake such a massive project, and teaching me all I now know which certainly has helped me in my evolution as a researcher.

It is said, in life, often things fall into place unexpectedly. Such has been my journey here at IICB. I got my first experience of life in research thanks to **Dr. Asish K. Banerjee**, Senior Principal Scientist and Project Director (Retd), NIPER- Kolkata, who was adamant I apply to **Dr. Susanta Roychoudhury**, Chief Scientist, IICB, for pursuing my Master's dissertation under his guidance. I express my sincerest gratitude to Dr. Roychoudhury for accepting me to be a part of his lab temporarily, and also for putting me under the able guidance of **Dr. Arindam Datta** (Arindam da to me). I could not have asked for a better guide during my Master's dissertation. Most of the techniques I know today, I have learnt from

him. Planning new experiments, working long nights, and experiencing the joy upon finally seeing the successful results of our efforts really motivated me to join this career path. I would also like to thank other members of SRC Sir's lab, like **Dr. Kayum Alam** (Kayum Da), **Dr. Sangeeta Ghuwalewala** (Sangeeta Di), who have helped and guided me on innumerable occasions on multiple things. Even today, when I need some advice, I approach them and they are always ready to help. For this, I am ever so grateful to them. I also wish to thank the other members of the SRC Lab who have made my stay there memorable. I wish to thank **Dr. Somsubhra Nath** (Somsubhra Da), **Dr. Sumana Bhattacharjya** (Sumana Di), **Dr. Tania Das** (Tania Di), **Debrup Sen** (Debrup Da), and **Dr. Kumar Singha Roy** (Kumar Da). The collegial atmosphere in the lab, with people always ready to help out each other, and our regular lab outings are memories that I will always cherish to my last day. I thank you all for making my stay in your lab a very pleasant one. I am also very much thankful to SRC Sir for two other things: Advising me to read "Molecular Cloning" by Maniatis, which enabled me to build a strong conceptual base in Molecular Biology, and for suggesting me to apply to Dr. D. Biswas for pursuing my Ph.D. under his guidance. Both of these suggestions have proven to be a big boon for me. So, I thank you, Sir!

CSIR-Indian Institute of Chemical Biology has become like my home. Over the last seven and a half years, I have spent about twelve hours daily (almost) in this institute. Hence, my sincerest thanks to this home away from home for shaping who I am today. I would also like to thank all the Directors whose tenure I have worked under. **Dr. Siddhartha Roy**, **Dr. Chitra Mandal**, **Dr. Samit Chattopadhyay**, and currently **Dr. Arun Bandhopadhyay**. Thank you all for allowing me to be a part of this esteemed institution. I would also like to acknowledge the big help I have received from IICB in terms of instrumentation and infrastructural facilities, without which I would not have been able to carry out so much of my work. In this regard, I especially wish to thank **Mr. Tapas Chowdhury** (Sr. Technician II) and **Dr. Ardhendu K. Mandal** (Principal Technical Officer) for use of the Azure and iBright Chemi Documentation Systems which forms the backbone of getting the results of most of our experiments. My acknowledgment to my institution would not be complete if I did not thank **UGC**, who provided me with a Research Fellowship for five years.

The scientific community is a small, well-knit entity, with its members always ready to help each other out. I am no exception to this rule. I express my sincerest thanks to **Dr. Suvendranath Bhattacharya**, Head, Molecular Genetics Division, CSIR-IICB, for his kind co-operation during my entire tenure. Whether it has been allowing us the use of his cell culture facility, or the nano-drop spectrophotometer, or chemicals and reagents from his lab, he has always been ready to help. I especially acknowledge his patience whilst signing numerous documents on multiple occasions so that I could receive my fellowship. I am grateful to **Dr. Partha Chakrabarti**, Principal Scientist, CSIR-IICB for

lending us reagents, and allowing the use of the Nanodrop Spectrophotometer at odd hours on multiple occasions, each time with a smile on his face. I believe the entire student fraternity would also be thankful to him for allowing us access to his lab's chemi-documentation system so that our work would not stop when the machine at the Central Instrumentation Facility was out of order. My sincerest thanks also go to him for collaborating with us on our project and helping out in the collection of samples for our physiological experiments, a part without which our paper would not be complete. In this context, I also wish to thank **Dr. Satinath Mukhopadhyay**, Endocrinologist and Professor, Department of Endocrinology and Metabolism, Post Graduate Institute of Medical Education and Research, Kolkata, for agreeing to be our primary clinical collaborator and ensuring we get the best and most appropriate samples for our *in vivo* experiments. His able guidance was monumental in the success of our experiments. It will be unfair if I do not mention **Dr. Umesh P. Singh**, Principal Scientist, CSIR-IICB, for his guidance at a critical juncture for some *in vitro* experiments. Our discussions on topics ranging from science to philosophy and everything in between (sometimes for hours, while standing in the middle of the corridor itself!!) have been highly motivational and a welcome change for freshening the mind. Here, I should also mention **Dr. N. Rajendran**, **Dr. Kavita Thirumurugan** and **Dr. V. Pragasam** who motivated me during my Master's course after seeing in me the potential to pursue a Ph.D. Your efforts have led to my present position today.

The lab is a home away from home. And once you spend more than twelve hours a day at a place for years on end, the people in that place become like your family. Such has been the case for me as well. I joined this lab in November, 2014, and from the first day itself, never once have I felt that I am in some foreign place. As I sit here today to express my gratitude to all these people, flashback upon flashback of so many incidents are just passing by my eyes. From being mistaken as a salesperson on my first day, losing a bet and treating the entire lab to chips and snacks, our impromptu plan to go watch an IPL match (which incidentally got rained out), birthday parties where we bombarded each other with cake, lab outings to some restaurant or movies, the many pranks we played on each other, celebrating festivals, visiting the nearby tea shop or our canteen for academic and non-academic discussions, and of course working together on so many projects, talking about our results or different hypotheses, getting feedback, I am really going to miss all of these things. Coming to the members of the lab, whom I have perhaps driven crazy while they have been, my first thanks goes to **Dipika (di)**, who made me feel welcome and comfortable in the lab, and helped me to find my niche in this ecosystem. I especially recall her advice to me on multiple occasions, when I was really confused about which road to take. **Mahesh (da)** was the senior most member when I joined. With a strict and rigid appearance, he seemed a hard taskmaster. Ironically, we got the opportunity to work together, and I would perhaps not be exaggerating if I said that he was very much the reason I was involved in the biggest project of the lab, which gave me my first first-author paper, and another one after that. **Nidhi**

(di) had always been the silent one in the lab, but always efficient in terms of planning and working. I learned a lot from her, especially how to carry out *in vitro* experiments, which helped me a lot. **Koushik** and **Dheerendra** joined just before me, just a few months prior, which meant that often we were doing similar experiments and facing the same challenges. It was always fun to sit and talk to them and try to hash out a solution, either in the lab or in the canteen, or outside at the tea shop. I admired Koushik for his ability to think and plan experiments, and I can never forget that his was the first paper from our lab, which inspired the rest of us, especially me, to believe that we could also publish in standard reputable journals. Dheerendra has always been the carefree fun-loving guy, very much into sports, and several extra-curricular activities. He taught me how important it is to have a work-life balance, something I have tried to imbibe in myself. Incidentally, both of us are submitting our theses together at the same university!! **Sunny** and I did not interact much, but he was always ready to help when any of us needed assistance with “heavy lifting” or issues pertaining to instruments. I connected with **Arijit** on many levels when he joined. Our common interests include music, movies, and sitcom series. We have worked together on all projects, and worked well!! We’ve had our rough patches, but successfully left them behind as we work on yet another project, which hopefully gets published soon. **Sujay** joined about a year after I did, and initially appeared to be very reticent. However, with time, we got talking and I realized he has a very different and fresh perspective towards life. On multiple occasions, he has treated the entire lab to small snacks and tidbits, which have definitely brightened up the day. I will surely miss his quirkiness. We call **Prathama** the “grandmother” of the lab, as she behaves just like one. Taking an active part in arranging lab celebrations and extra-curricular events, she behaves as if we all are her grandchildren and admonishes us if she feels its necessary!! **Sushrita** had a brief stay in the lab, but I must thank her for lending me plasmid constructs on multiple instances. She also has suggested to all of us movies and series, which come as a welcome respite. I got my first junior who would work alongside me in an extension of a project of mine in the form of **Avik**. His enthusiasm to know the unknown, and knowledge of a wide variety of things under the sun really impressed me. I should also mention he is an avid bibliophile like me, and we have often struck up conversations on topics ranging from Egypt, to stamps, psychology, books, and whatnot, during lunchtime. **Poushali** is just a year old in the lab now, but she has taken up a lot of the administrative burden, and her enthusiasm coupled with her ever readiness to help others is worth mentioning. My sincerest thanks also go to the short-term members of our lab, our dissertation trainees: **Mehuli**, **Nivedita**, **Anchal**, and **Rupsha**. In one way or another, I have learnt from you all, and so a big thanks from my end.

The contributions of the student fraternity at IICB has also been instrumental in me completing my Ph.D.. My first order of thanks goes to Jit from Dr. Partha Chakraborti’s lab and Bidisha from Dr. Satinath Mukhopadhyay’s lab for their help in our collaborative project. I also wish to thank the other members

of Partha Sir's lab, the seniors: **Dipsikha Di, Titli Di, Mainak da** and **Moumita Di** for their advice and help on multiple occasions. **Debojyoti** also took out time to plan animal experiments. Though these experiments never saw the light of day, I appreciate his time nonetheless. Thanks are also in order to the other members of the lab for helping with the use of the chemi-doc system and nano-drop spectrophotometer. Our next-door neighbor, Dr. Suvendranath Bhattacharya's lab has also helped us on many many instances. I thank **Mainak Da** for inspiring me, and the others, particularly **Saikat, Susanta, Sourav**, and **Shreya** for their help in real-time PCR experiments when our machine was out of order for a long period of time. Not to mention the multiple times I have needed to use the centrifuge in their lab. I also appreciate **Pampa Di** from Dr. S.C. Biswas' lab for fluorescent imaging in the initial stages of my tenure. As our lab is highly dependent on Western blotting, and there have been many days when more than twenty SDS-PAGE gels have been cast and run, we have often needed to borrow casting and running apparatus from other labs. In this regard, **Tanaya, Ayesha, Sudeshna, Kamran** and **Sonali** have been a boon. I can confidently say that on many days, I have been able to run gels because of their help, particularly in the last year. I thank **Rajeev** from DR. Saumen Dutta's lab for providing the *E. coli* XL1-Blue® competent cells which were absolutely essential for performing our site-directed mutagenesis experiments. It would not have been possible for doing the Luciferase experiments had it not been for **Sayanika Di** and **Jeet**. I also wish to thank Jeet for advising me on which laptop to buy, as his advice resulted in me purchasing the laptop on which I am typing now. Had it not been for **Himadri**, I would not have gotten the information for my synopsis presentation, something I will be forever grateful for. **Vivek** has helped out on several occasions by lending us plastic wares for lab use when our stock had run out. My sincerest thanks also go out to **Ashmita** and **Shivangi**, two of the most genuine people I have met in this institute. We have had tête-à-têtes on so many topics, I have simply lost count! I will reminisce and think back to all our crazy conversations. I cannot also forget though, our stage performance on the Foundation Day of CSIR-IICB, when **Sourav Da, Apabrita Da, Sayantan Da, Avijit, Sambit** and myself set on "The Journey" with our Eastern Classical fusion music performance, which is appreciated to this day. That day was a highlight of my stay at this institute and I will cherish it always!

I also wish to express my gratitude to the Security Officer, **Mr. Sabyasachi Karmakar**, and the security staff at CSIR-IICB. They have been highly cooperative when their help has been necessary. There have been instances when we have needed to access the central instrumentation facility, or required coordination for the delivery of some items, and they have been always ready to oblige our requests. I also wish to thank our canteen staff for keeping us well fed during our stay here.

Our work would not have been possible had it not been for the timely supply of reagents and consumables from the suppliers. In this regard, I really wish to express my gratitude towards all the

suppliers, **Patel da**, **Biswajit Mishraji**, **Proloy da**, **Abhishek Da**, and **Amarnath da**. In this section, I will take the opportunity to separately thank **Soumen Patra**, for being so cooperative while working with me on the delivery of plastic wares, being prompt and courteous, and delivering items on multiple instances at a notice of a few hours.

I could not possibly miss out on thanking a very important person in my life, **Sreeparna**. You have simply been instrumental in what I have been able to accomplish in the most critical years of my Ph.D. tenure. You kept me grounded, gave me inspiration and motivation, and made me believe in myself at a stage when most get easily frustrated and look for ways out. I consider it a blessing to have you in my life, and though I may not always be able to define in words how appreciative I am of you, I know you can read my mind. Aunty and Uncle, both of you have been my well-wishers for a very long time now. I thank you both for your advice, support, and constant motivation to be the best I can be.

My relatives, namely my **Mashi**, **sister Roma**, **brother Rupam**, **Pishi**, **Mama** and **Maami** have looked at me, blessed me and prayed for me all their lives. It is due to their good wishes and blessings that I have been able to come so far. Though my thanking them here is a minimal gesture from my end, I know they desire nothing but happiness and success for me. I appreciate you all!!

It would probably be the biggest lie if I said I could have managed this long and arduous journey without the support of my family members. My **grandmother** battling cancer was the major inspiration for me to join the field of biological research. I am sure she is looking down at me and keeps on showering me with love and blessing from wherever she is now. **Mom** and **Dad**, though it appears that I tend to dismiss your words, to tell you the truth I could not move an inch forward had you both not been in my life. I may not express this in as many words, but day and night I thank my lucky stars that I have the best parents one could hope for. I know I do not need to write here how much I respect and love you both, but all I can say is, if I can become even 1% of what you both are, I would consider myself complete. Thank you so so so so much.

Thank you all!!

Subham

Abbreviations

DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
TFIIA	Transcription factor II A
TFIIB	Transcription factor II B
TFIIC	Transcription factor II C
TFIID	Transcription factor II D
TFIIE	Transcription factor II E
TFIIF	Transcription factor II F
TFIIH	Transcription factor II H
TBP	TATA-binding protein
UBF	Upstream Binding Factor
SL1	Selectivity factor 1
CPSF	Cleavage and Polyadenylation Specificity Factor
CSTF	Cleavage Stimulation Factor
TSS	Transcription Start Site
Inr	Initiator
DPE	Downstream Promoter Element
3C	Chromosome Conformation Capture
4C	Circular chromosome conformation capture
5C	Chromosome Conformation Capture Carbon Copy
PIC	Pre-Initiation Complex
ChIP-Seq	Chromatin Immunoprecipitation Sequencing
EMSA	Electrophoretic Mobility Shift Assay
UV	Ultraviolet
HDAC	Histone Deacetylase
SWI/SNF	SWItch/Sucrose Non-Fermentable
INO80/SWR1	INOsitol requiring/ Sick With Rat8 ts
ISWI	Imitation switch
CHD	Chromodomain Helicase DNA-binding
NURD	Nucleosome Remodeling Deacetylase
TRIM28	Tripartite Motif Containing 28
P-TEFb	Positive Transcription Elongation Factor-b

CDK9	Cyclin Dependent Kinase 9
CTD	C-terminal domain
ELL	Eleven Nineteen Lysine Rich Leukemia
ENL	Eleven Nineteen Leukemia
AFF1	AF4/FMR2 Family Member 1
AF9	ALL1-fused gene from chromosome 9
MLL	Mixed Lineage Leukemia
SEC	Super Elongation Complex
LEC	Little Elongation Complex
EAF1	ELL associated factor 1
EAF2	ELL associated factor 2
DBC1	Deleted in Breast Cancer 1
SIRT1	Sirtuin-1
LZ	Leucine Zipper
HEK-293T	Human Embryonic Kidney-293T
DMEM	Dulbecco's Modified Eagle Media
FBS	Fetal Bovine Serum
ORF	Open Reading Frame
GST	Glutathione S-transferase
IPTG	Isopropyl β -D-1-thiogalactopyranoside
X-Gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
EDTA	Ethylene Diaminetetracetic Acid
PMSF	Phenylmethylsulfonyl fluoride
NP40	Nonidet P-40
PICs	Protease Inhibitor Cocktails
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis
CHX	Cycloheximide
RPM	revolutions per minute
DAVID	Database for Annotation, Visualization and Integrated Discovery
PBMC	Peripheral blood mononuclear cells
PANTHER	Protein ANALysisTHrough Evolutionary Relationships

Abstract

Abstract

Understanding the functional regulation of the RNA Polymerase II elongation factor, ELL

Over the last decade, studies on eukaryotic transcription have focused on the role of the Super Elongation Complex (SEC) in transcription control. The SEC is important as the expression of several of its members are found to be misregulated in terminal diseases such as mixed lineage leukemia (MLL). Hence the detailed study of SEC components becomes essential and critical. Of all the SEC components identified and described, ELL1 (also known as ELL) is the only factor capable of stimulating RNA Pol II-mediated transcription elongation *in vitro*. However, the mechanisms of regulation of ELL are completely unknown. In our studies, we report a novel interaction between ELL and the protein DBC1, which acts to regulate ELL stability and its functions within mammalian cells. Mechanistically, we show that ELL is acetylated by the lysine acetyltransferase, p300, which provides it stability and makes it functionally active. Upon deacetylation, mediated by HDAC3, ELL is targeted for polyubiquitination by the E3 ubiquitin ligase Siah1 and undergoes proteasomal degradation. DBC1 competes with HDAC3 for binding to ELL by virtue of possessing the ability to interact with the same region of ELL, thereby preventing ELL deacetylation and providing it stability. A similar mode of ELL regulation is seen in the case of EAF1/EAF2 as well, wherein, in the presence of EAF1/2, HDAC3-ELL interaction is displaced. The ELL-DBC1 axis serves to regulate the expression of SEC-affected genes, influencing genes involved in metabolism, cell cycle regulation, and immunity. Our studies show that dysregulation of these genes can lead to the development of diseases like cancer or type 2 diabetes, a phenomenon confirmed by studies in PBMCs from diabetic patients. In contrast, ELL-EAF1 functions are more prevalent in the genotoxic stress response. Our assays show that two independent factors, DBC1 and EAF1 work in tandem to maintain the protein levels of ELL for cellular response to stimuli in a context-dependent manner.

Contents

Chapter 1: Review of Literature

	<u>Page</u>
1.1	The importance of gene regulation
1.2	Transcription
1.3	Factors controlling Transcription
1.3.1	Cis-regulatory elements
1.3.2	Trans-regulatory elements
1.4	Steps in Eukaryotic Transcription
1.4.1	Initiation
1.4.2	Elongation
1.4.3	Termination
1.5	The Super Elongation Complex in Transcription Regulation
1.6	ELL (Eleven Nineteen Lysine Rich Leukemia)
1.7	DBC1 (Deleted in Breast Cancer 1)

Chapter 2: Rationale of the study	32
-----------------------------------	----

Chapter 3: Materials and Methods

	<u>Page</u>
3.1 Experimental models	36
3.2 Cell Culture and Transfection	36
3.3 Generation of plasmid constructs for expression in mammalian, insect and bacterial expression systems	37
3.4 Generation of plasmids by site-directed mutagenesis	40
3.5 Bacmid and baculovirus preparation	41
3.6 Generation of Stable Overexpressing Cell Lines	41
3.7 Generation of Stable Knockdown Cell Lines	42
3.8 Mass Spectrometry Analysis of DBC1 associated proteins	43
3.9 Nuclear Extract Preparation from Plated Cells	44
3.10 Immunoprecipitation Analyses	44
3.11 Luciferase based Reporter Gene assay	45
3.12 Cycloheximide Chase Assay	46
3.13 Purification of recombinant proteins	47
3.14 Western Blotting	48
3.15 <i>In vitro</i> interaction Assays	50
3.16 <i>In vitro</i> Acetylation Assay	51
3.17 <i>In vitro</i> Deacetylation Assay	51
3.18 RNA Extraction and qRT-PCR	52
3.19 Chromatin Immunoprecipitation Assay	55
3.20 Cell Proliferation Assay	57
3.21 Clonogenic Assay	57

3.22	Analysis of High Throughput NGS Data and Gene Ontology Analyses	58
3.23	Peripheral Blood Mononuclear Cell Isolation	58

Chapter 4: Interplay between DBC1, HDAC3, p300 and Siah1 for maintenance of ELL stability and functions

		<u>Page</u>
4.1	Identification of DBC1 as a novel interactor of ELL	61
4.2	DBC1 associates with ELL as well as other human SEC components	62
4.3	DBC1 and ELL interaction occurs through specific structural domains	64
4.4	ELL and DBC1 have an additive effect on RNA Pol II-mediated transcription	66
4.5	DBC1 stabilizes ELL protein expression within mammalian cells	69
4.6	ELL interacts with HDAC3 within mammalian cells	70
4.7	ELL is degraded specifically by HDAC3 within cells	72
4.8	DBC1 acts oppositely to HDAC3 for stabilization of ELL	76
4.9	HDAC3 and DBC1 compete with each other for binding to ELL	77
4.10	The lysine acetyltransferase p300 acetylates ELL	79
4.11	ELL is stabilized upon p300 mediated acetylation	81

4.12	HDAC3 can act as the deacetylase for ELL	85
4.13	DBC1 and p300 act in opposition to HDAC3 for maintenance of ELL acetylation levels	86
4.14	ELL is subjected to proteasome mediated degradation	88
4.15	DBC1 and p300 decrease ELL polyubiquitination by increasing its acetylation	90
4.16	Two lysine residues at the N-terminus of ELL, K5 and K29, are targets for ubiquitination	93
4.17	The lysines, K5 and K29, are also the predominant targets of p300 mediated ELL acetylation	96
4.18	Siah1 acts as the E3 ligase for ELL	98
4.19	DBC1 and ELL co-regulate diverse sets of genes	100
4.20	Depletion of DBC1 and ELL results in the inefficient formation of the transcription complex	102
4.21	The ELL-DBC1 axis regulates key genes involved in glucose metabolism	106
4.22	Dysregulation of the ELL-DBC1 axis is a potential cause of Type 2 Diabetes	108

Chapter 5: DBC1 and EAF1/EAF2 maintain ELL stability and functions via a negative feedback loop mechanism

	<u>Page</u>
5.1 The presence of EAF1 and EAF2 increases the level of ELL protein	113
5.2 The stabilizing effect of EAF1 is limited only to ELL	114
5.3 EAF1 and EAF2 protect ELL from ubiquitin mediated degradation	116
5.4 HDAC3 mediated ELL degradation is prevented by EAF1 and EAF2	118
5.5 EAF1 and EAF2 bind to similar regions on ELL protein	120
5.6 ELL-HDAC3 interaction is reduced in the presence of EAF1 and EAF2	123
5.7 An interplay exists between DBC1 and EAF1 for regulating each other's levels	125
5.8 EAF1 is polyubiquitylated in the presence of DBC1	127
5.9 TRIM28 interacts with and ubiquitinates EAF1	129
5.10 DBC1 forms a bridge between TRIM28 and EAF1	131
5.11 DBC1 and EAF1 function to maintain ELL protein levels within cells	133
5.12 DBC1 and EAF1 modulate ELL levels and function in a context-dependent manner	136
 Chapter 6: Summary	 142
 Chapter 7: Discussion	 146
 References	 150

Index of Figures and Tables

Review of Literature

	<u>Page</u>
Figure 1.1 The central dogma of molecular biology	3
Figure 1.2 The structure of a simple eukaryotic promoter	5
Figure 1.3 Schematic of a transcriptional regulatory module found in metazoans	8
Figure 1.4 An overview of the initial steps of transcription in eukaryotes	17
Figure 1.5 Structure and functions of the ELL protein	25
Figure 1.6 DBC1 domains and its molecular architecture	29

Materials and Methods

	<u>Page</u>
Table 1	List of plasmid constructs
	37
Table 2	Oligonucleotides used for generating shRNA constructs
	43
Table 3	Antibodies used for Immunoblots and Immunoprecipitation Assays
	49
Table 4	Primers used for qRT-PCR analyses for mRNA expression of target genes
	53
Table 5	Primers used for ChIP-qRT-PCR
	56

Results

Interplay between DBC1, HDAC3, p300, and Siah1 for the maintenance of ELL stability and functions

	<u>Page</u>
Figure 4.1 DBC1 is a novel ELL interacting protein	62
Figure 4.2 DBC1 forms stable associations with several SEC components	64
Figure 4.3 DBC1 and ELL interact with each other through specific structural domains	66
Figure 4.4 DBC1 and ELL together can stimulate RNA Pol II mediated transcription	68
Figure 4.5 DBC1 stabilizes ELL protein within mammalian cells	70
Figure 4.6 A novel interaction is seen between ELL and HDAC3 within mammalian cells	72
Figure 4.7 HDAC3 causes degradation of ELL within mammalian cells	75
Figure 4.8 DBC1 protects ELL from being degraded by HDAC3	77
Figure 4.9 DBC1 and ELL competitively bind to the same region of ELL	79
Figure 4.10 p300 acetylates ELL	81

Figure 4.11	ELL is stabilized by p300 within cells	84
Figure 4.12	HDAC3 acts as the deacetylating enzyme for ELL	86
Figure 4.13	DBC1 and p300 prevent HDAC3 action and maintain acetylated ELL within cells	88
Figure 4.14	ELL is degraded by the proteasome within mammalian cells	90
Figure 4.15	DBC1 and p300 increase ELL acetylation and decrease its ubiquitylation within cells	93
Figure 4.16	ELL is ubiquitylated at two lysine residues at its extreme N terminus	95
Figure 4.17	ELL K5 and K29 are the predominant targets for p300 mediated acetylation	97
Figure 4.18	ELL is degraded by the action of the E3 ubiquitin ligase, Siah1	99
Figure 4.19	DBC1 and ELL regulate the expression of a wide variety of target genes within cells	102
Figure 4.20	DBC1 and ELL regulate the formation of a functional transcription complex	105
Figure 4.21	DBC1 and ELL regulate the expression of key genes involved in glucose metabolism	107
Figure 4.22	Dysregulation in the DBC1-ELL axis potentially causes Type 2 Diabetes	109
Figure 4.23	Model of functional regulation of ELL by the concerted action of multiple proteins	110

DBC1 and EAF1/EAF2 maintain ELL stability and functions via a negative feedback loop mechanism

	<u>Page</u>
Figure 5.1	Overexpression of EAF1 and EAF2 increases ELL levels within cells
	114
Figure 5.2	EAF1 stabilizes only ELL and CDK9 in the SEC
	115
Figure 5.3	ELL is protected from ubiquitin mediated degradation by EAF1 and EAF2
	118
Figure 5.4	EAF1 and EAF2 prevent HDAC3 mediated ELL degradation
	120
Figure 5.5	Interaction of EAF1, EAF2 and HDAC3 occur predominantly at the N-terminus of ELL
	122
Figure 5.6	A competition exists between EAF1/EAF2 and HDAC3 for binding to the N-terminus of ELL and maintaining its cellular pool
	125
Figure 5.7	DBC1 and EAF1 antagonize each other within the cell
	127
Figure 5.8	EAF1 undergoes ubiquitin mediated degradation in presence of DBC1
	129
Figure 5.9	TRIM28 can interact with EAF1 and ubiquitinate it
	131
Figure 5.10	DBC1 is essential for TRIM28 mediated EAF1 ubiquitylation
	133
Figure 5.11	DBC1 and EAF1 work in tandem to maintain ELL stability and thus its function within mammalian cells
	136

Figure 5.12	ELL protein levels and its cellular functions are regulated by DBC1 and ELL in a context dependent manner	139
Figure 5.13	Model depicting the regulation of ELL by the action of EAF1/EAF2 and DBC1 within cells for efficient transcription of target genes	140

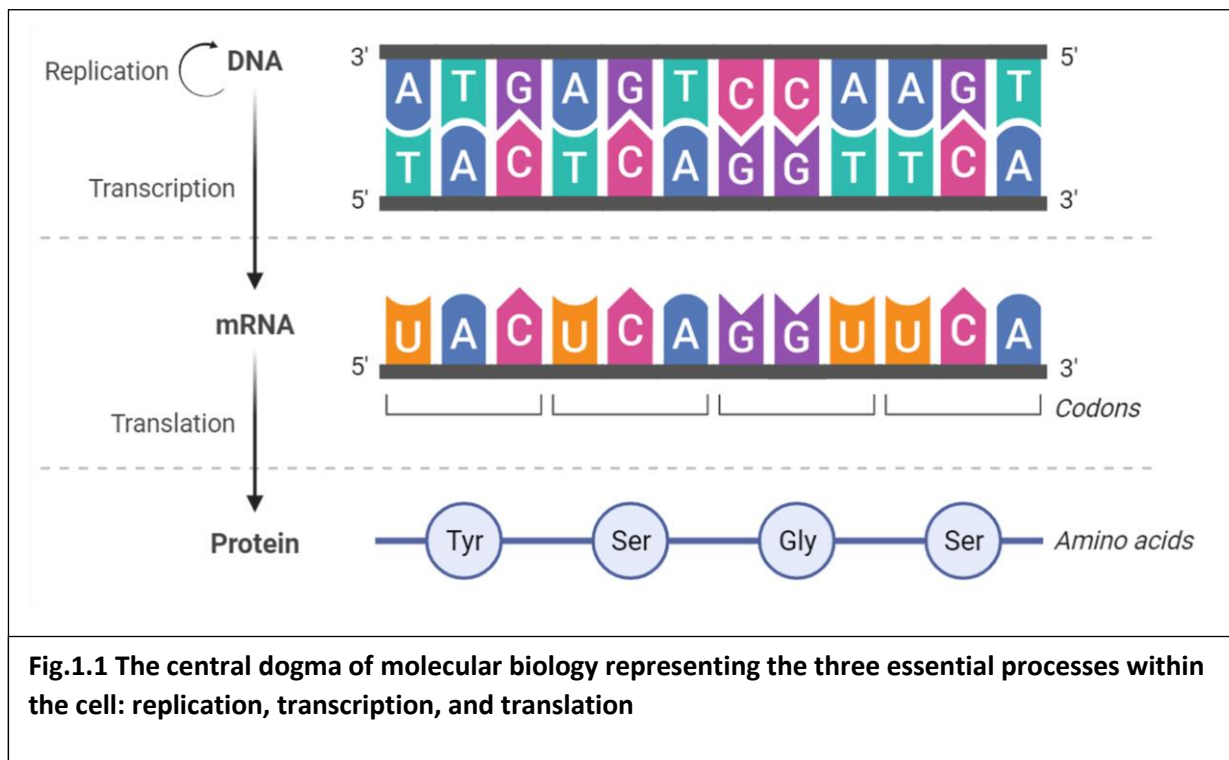
Chapter 1

Review of Literature

1. Review of Literature

1.1 The importance of gene regulation

The most comprehensive definition of a gene is that it is a functional unit of heredity, carrying within it specific information for the synthesis of RNA or proteins, and located on a specific position on a chromosome [1]. Genes consist of stretches of DNA, comprised of four nitrogenous bases, termed nucleotides, arranged in a specific sequence. The mechanism by which the genetic information stored in the DNA is converted into usable products within the cell is termed as the central dogma of molecular biology [2]. The central dogma states that the three essential processes which affect all cellular functions are: replication, transcription, and translation (Figure 1.1). DNA replication involves the generation of an exact copy of the genome, which occurs during the cell cycle and is mediated by enzymes termed as DNA polymerases [3]. During the course of development, and furthermore, for maintenance of cellular processes, specific genes need to be “turned on” and some genes are required to be kept “off”. This process, by which certain specific stretches of DNA i.e. genes, are used as a template for the synthesis of RNA, is termed as transcription and is carried out by the action of RNA Polymerases [4]. The final step, resulting in protein synthesis, is called translation and involves the use of the newly transcribed mRNA as the template, occurring within the cytoplasm carried out by ribosomes [5, 6]. Of these three processes mentioned, perhaps the most stringently controlled process is transcription. The reason for this can be best illustrated with a simple observation: the microscopically small worm, *Caenorhabditis elegans* has about 80% genes similar as compared to the most complex organism on earth, *Homo sapiens* [7]. However, looking at the functional processes in both these organisms, it is fascinating to note that the output from these similar numbers of genes is so drastically different. This observation is one of many that ascertains the urgent need to deeply understand transcription and its regulation. Researchers, over the last few decades, have devoted significant time and efforts to delve into the complex, intricate, and tightly regulated world of transcriptional regulation, both in eukaryotes as well as prokaryotes. Looking at it from a broader perspective, the overall transcription machinery is similar in all organisms. However, the more complex the organizational structure, more are the numbers of regulatory factors involved in these processes. Understanding the intricacies of this process would thus enable a better understanding of development, as well as provide extensive details about the implications of misregulation, and their contribution to the development of various disease conditions [8].



1.2 Transcription

The development of an organism and determining how it responds to a variety of cellular and extracellular stimuli is dependent on the expression of its genes. All genes in a cell are not expressed continuously, but in a context-dependent manner, the expression of certain genes are increased whilst others are hindered. Several factors play important yet very distinct roles in determining which genes are to be turned “on” and which stay “off”. One of the major processes by which this is carried out is regulation at the level of transcription [9]. The textbook definition of transcription states that it is a process by which the genetic information in a DNA sequence is copied onto a strand of RNA, called messenger RNA (mRNA), and used as a template for the synthesis of functional proteins. Essential for transcription are enzymes known as RNA Polymerases, as well as accessory proteins termed simply as transcription factors. In prokaryotes, the genetic information is tightly packed into an irregularly shaped protein-DNA complex, termed as the nucleoid [10]. Only one RNA polymerase is present for the transcription of all genes, enabling faster transcription. The most extensively studied prokaryotic transcription model is in *E. coli*, where RNA polymerase, comprising of five subunits including two copies of α , β , β' , and ω associate with the σ factor, and initiates transcription from specific sites [11]. Two consensus sequences, TATAAT and TTGACA (located 10 and 35 bases upstream respectively from the transcription start site) play key roles in the unwinding of DNA and binding of RNA Polymerase σ factor, leading to transcription initiation (reviewed in [12]). mRNA is synthesized at a rate of about 40 nucleotides per second. Transcription termination occurs by two methods, one dependent on the Rho

protein, and the other independent of it, resulting in the formation of a hairpin structure of RNA, leading to the stalling of RNA Polymerase (reviewed in [13]). In eukaryotes, due to their more complex organization, the transcription process is more elaborate. Three different RNA Polymerases are present, each having different functions[14]. RNA Polymerase I, which is responsible for transcribing ribosomal RNA, binds about 150 bases upstream of the transcription start site, recruited by UBF (upstream binding factor) and SL1 (selectivity factor 1). SL1 contains TATA-binding protein (TBP), which is essential for TATA box recognition for RNA Pol II-mediated transcription of certain genes [15]. Genes for smaller RNAs, tRNAs, and 5s rRNA are transcribed by RNA Polymerase III. For these genes, the binding of TFIIIA occurs first, followed by TFIIIC, TFIIIB, and the polymerase. The recruitment of the TFIIIC and the polymerase is carried out by TFIIIC. TFIIIB, in this case, contains the TBP[16]. In the context of gene regulation in eukaryotes though, the majority of studies have focussed on the transcription of mRNA from protein coding genes. This process, mediated by RNA Polymerase II (RNA Pol II) is effected by binding of TBP to the TATA box upstream of the transcription start site, which results in the recruitment of TFIIIB, TFIIA, Pol II, TFIIF and TFIIH. This complex, named the pre-initiation complex starts transcription, and then moves into the next step, elongation, assisted by various factors and proteins (reviewed in [17]). Termination generally occurs when Pol II reaches the end of a gene, where CPSF (cleavage and polyadenylation specificity factor) and CSTF (cleavage stimulation factor) recognize the poly-A signal in the mRNA. CPSF and CSTF then recruit other proteins to effect cleavage and polyadenylation. The result is an RNA molecule containing a “tail” of adenine bases at its end [18]. Since our studies are focussed on RNA Pol II mediated transcriptional regulation, in this literature review, we will devote more attention to the transcriptional process in play for synthesis of mRNA from sequences of protein coding genes.

1.3 Factors controlling Transcription

As eukaryotic transcription is an elaborate multi-step process, several factors work in tandem to ensure regulated but efficient transcriptional output. These involve cis-regulatory elements, trans-regulatory elements, chromatin architecture, and modifications. Without the concerted action of these factors, the stimulation of expression or silencing of genes cannot occur. Hence, in the context of transcriptional regulation, it is imperative to gain knowledge on these essential regulatory aspects.

1.3.1 *Cis*-regulatory elements

Cis-regulatory elements can be defined as non-coding stretches of DNA which regulate the transcription of neighboring genes. They are bound by transcription factors and generally are found in the vicinity of genes whose expression they modulate. In eukaryotes, such elements include promoters, enhancers, and silencers.

1.3.1.1 Promoters

Promoters are short stretches of DNA, located just upstream of the transcription start site (TSS) (Figure 1.2). The TSS is marked as the +1 site, while promoters are found ranging between 100 to 1000 nucleotides upstream of this site (-100 to -1000) [19]. Promoters form the region where the RNA Polymerase complex binds to initiate the transcription of genes. In eukaryotes, the promoter comprises of three different elements: the core promoter, the proximal promoter, and the distal promoter.

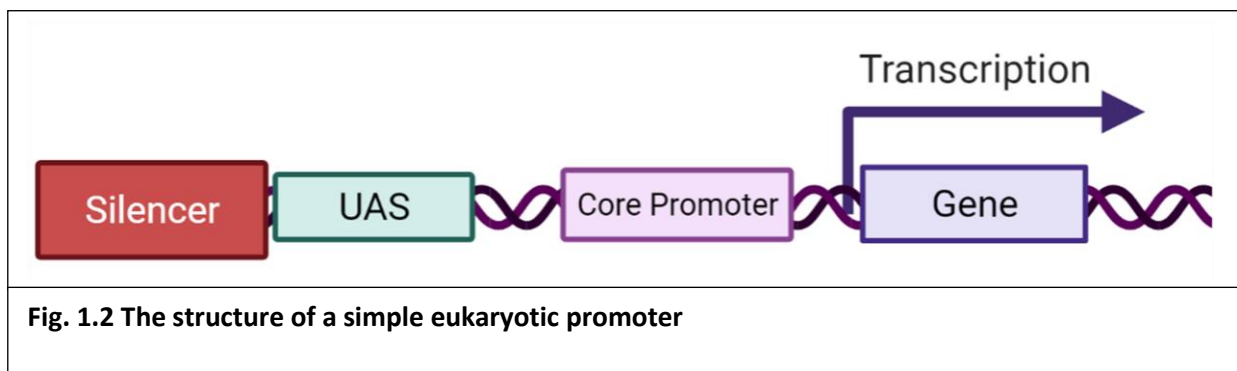


Fig. 1.2 The structure of a simple eukaryotic promoter

Core Promoter

The traditional view about core promoters is that it serves as the region where RNA Pol II and proteins of the transcriptional machinery bind. The binding of these factors results in the initiation of transcription, but this occurs at a much-subdued level. Though the definition of the transcription start site is ambiguous even now, it is widely accepted that the core promoter lies between the -50 to + 50 region of the transcription start site [20]. Based on several studies carried out on different sequences of core promoters, these have now been classified as “focused” or “sharp” core promoters and “dispersed” or “broad” promoters [21, 22]. Sharp core promoters have a single, distinct TSS with the essential regulatory elements located near the TSS. These promoters are normally associated with histone H3 Lys 4 trimethylation (H3K4me3) and H3 Lys 27 acetylation (H3K27ac), hallmarks of active transcription. Similar to the Pribnow box found in prokaryotes, about 48% of these promoters contain the Inr (Initiator) elements, which are bound by the TFIID for transcription initiation [23]. If a promoter is devoid of a TATA box, the Inr motif is complemented by another element, named the downstream promoter element (DPE). For the ~25% of promoters having the more well-known TATA box, the sequence 5'-TATAAA-3' serves as the recognition site for the TATA-binding protein (TBP) of the TFIID complex [24]. Dispersed promoters are normally associated with the expression of housekeeping genes, and are comprised of several transcription start sites located close to each other [25]. In case of key proteins involved in development, the promoters resemble those of constitutively expressing

genes but contain both the activating H3K4me3 and the repressive H3K27me3 histone modifications. This is probably to ensure their rapid activation in cells where they are needed and keep them silenced in other cell types.

Proximal Promoter

The proximal promoter region perhaps now serves as the most important genomic loci for the regulation of transcription of genes. The proximal promoter region is defined as about 50 bases downstream of the transcription start site. It was first identified in the Lis lab while studying the effect of heat shock on the transcription of *Hsp* genes in *Drosophila melanogaster*. Lis and colleagues defined this as the downstream region just adjacent to the promoter, where transcriptionally engaged RNA Pol II would “pause” briefly after it had synthesized about 20-60 nucleotides in the nascent RNA [26-28]. Later, detailed studies, including high throughput sequencing techniques have shown that the pausing of RNA Pol II in this region serves as an important regulatory event and is necessary for the rapid induction of several genes (Reviewed in [29-31]).

1.3.1.2 Distal promoter elements

Often, it has been observed that regulatory sequences occur not just adjacent to their targets, but may also be found kilobases away. Such elements termed as distal regulatory elements include enhancers, silencers, and insulators [32]. Gene regulation cannot be fully understood unless we have an idea of the functionality of these elements in transcription, and hence, it is essential we touch upon them, albeit briefly.

Enhancers

The term enhancer was coined when an SV40 DNA was shown to be able to activate the expression of the β -globin gene located a significant distance away [33]. Thus, the location of enhancer elements ranges from just a few kilobases, and may even extend to a megabase from their targets [34]. They are normally located within intergenic regions, but in some cases have also been found within gene bodies [35, 36]. Often, enhancers have also been seen to be located at the 3' region of genes. Though the definition of an enhancer element is still blurry, in essence, they consist of several transcription factor binding sites located in close proximity to each other. Using techniques like 3C, 4C, 5C and Hi-C, scientists have proposed that during transcriptional upregulation, enhancers are brought into proximity to their targets by the looping of DNA. Studies have also shown that the formation of the pre-initiation complex (PIC) may not necessarily occur at the core promoter, but at a distal enhancer

element, which would allow for more tightly regulated gene expression (reviewed in [20]). Enhancers are most associated with the histone H3 lysine 4 monomethylation (H3K4me1) and H3K27 acetylation (H3K27ac) activation marks [37]. When gene silencing is required, these enhancers are repressed by the Polycomb repressive complex-mediated H3K27 trimethylation or binding of other repressive factors. ChIP-Seq and DNase-seq have led to the identification of “super” enhancers, regions on DNA that are hubs for the binding of strong transcriptional activators, which generally occupy extensive regions on the genome [38]. Super enhancers have been found to be associated with pluripotent genes in mouse embryonic stem cells, or with oncogenes in several cancers [39, 40].

Silencers

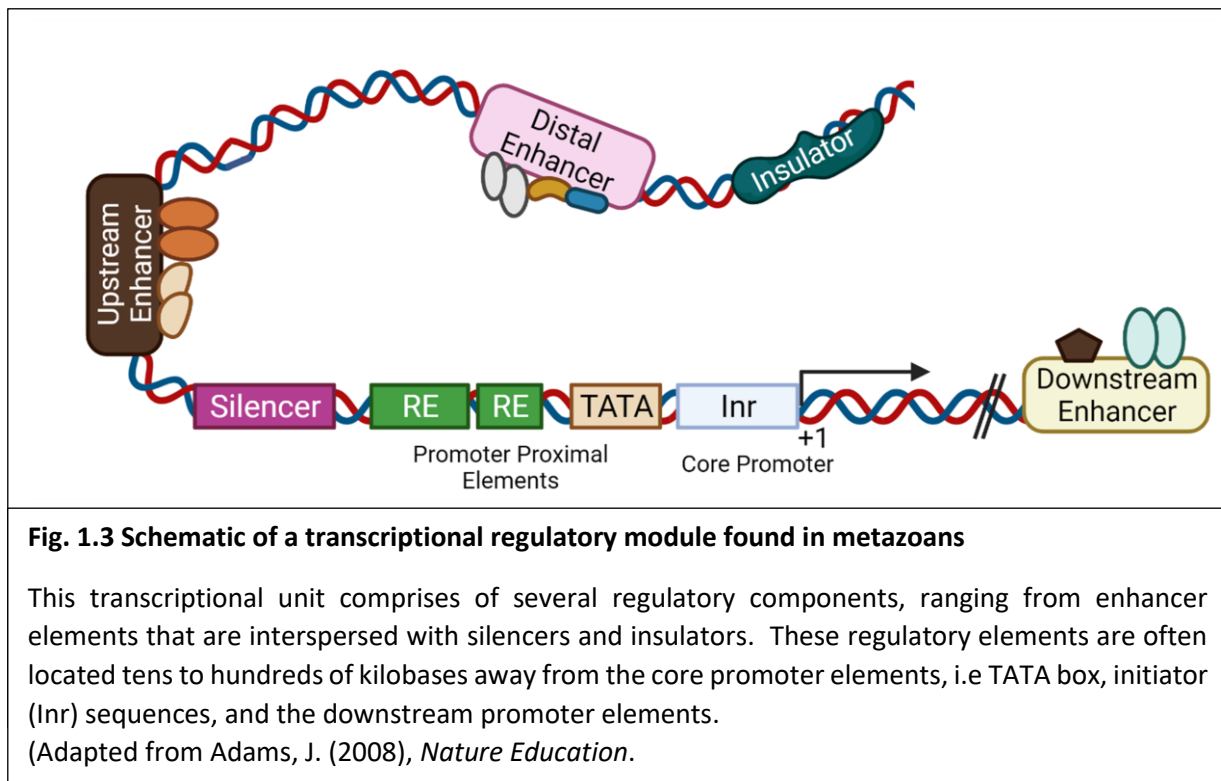
Like enhancers activate transcription from a core promoter through the enhanced binding of activating factors, silencer elements function to reduce the activity of a promoter, thereby repressing gene expression. Silencers are known to recruit transcriptional repressors which can modify chromatin or inhibit transcriptional activators. The role of silencers is to tightly regulate gene expression by preventing the expression of genes depending on the cellular state. Let us take the case of T-cell differentiation. T cells are of two subtypes, CD4 and CD8. When immature, both of these surface markers are expressed on the cells. However, upon maturation, when the lineage gets specified, the expression of one of the markers is silenced [41]. It has been observed that on several occasions, there is an overlap between silencers and enhancers. Some regions have been shown to act as both silencers and enhancers in different contexts or based on cell type. For example, a DNA element containing sites for binding of activators in one cell type can also recruit repressors in other cells [42-44]. Studies on silencers, though not many, have reported the presence of the PRC2 complex in these regions, which is one potential method of heterochromatinization. However, further reports seem to indicate that the enrichment of this complex specifically in the context of silencers may not be statistically significant [45-47]. It has also been reported that silencers often mediate gene repression by “anti-looping”, i.e. preventing the enhancers from reaching into close proximity with their targets [45, 48].

Insulators

These are DNA sequences that prevent interactions between genomic loci. Insulators function either by blocking enhancers or preventing heterochromatinization [49]. Enhancer blocking only can take place when the insulator element is located between the promoter and the enhancer. The enhancer is however free to regulate gene expression on the side not flanked by the insulator [50]. A number of blocking insulators have been reported and characterized in *Drosophila*, which include su(Hw), BEAF, and Zw5 [51-53]. In vertebrates, CTCF acts as an insulator by the formation of closed-loop domains that are unable to interact with each other [54]. Insulator elements are also known to prevent

the repression of genes by blocking the unregulated advancement of nearby heterochromatin. For example, in the yeast *Saccharomyces cerevisiae*, silencer proteins recruited to the *HM* locus act to repress gene expression. The spread of these silencing proteins beyond the gene boundary is carried out by barrier elements that prevent flanking genes to be repressed [55].

A representation of all these elements in metazoans is shown in Figure 1.3



1.3.2 *Trans*-regulatory elements

Cis-regulatory elements are generally involved in regulating the transcription of adjacent genes. On the other hand, *trans*-acting factors can affect the transcription and levels of expression of genes situated at a significant distance through intramolecular interactions. These elements can affect the expression of several genes located on either allele and are not bound by the genomic distance between the regulator and regulated bodies [56]. It has been often noticed that *cis* and *trans* elements act opposite to each other, thus helping to maintain a balance in the expression levels of their targets. On the basis of their targets, *trans*-acting regulatory elements can be divided into factors that act on DNA sequences or on mRNA.

DNA binding elements

These form the most common of all regulatory elements and affect gene expression by binding to consensus sequences located on the genome. The most common and widely known example of such elements are the components of the pre-initiation complex. The first step in transcription initiation involves the binding of the TBP to the consensus TATA box on the promoter of genes. This is followed by the binding of TFIIB and TFIIA, resulting in the recruitment of the RNA Pol II. All the modules of the initiation complex which bind to DNA, including the subunits of RNA Polymerases (I, II, and III) can be classified as *trans*-acting factors [57, 58]. However, when we traditionally think of *trans*-acting factors, we tend to restrict our classification to traditional transcription factors, which either bind to DNA or other proteins/transcription factors for (generally) the induction of transcription. Transcription factors are proteins involved in almost all cellular processes, such as cell speciation, differentiation, control of multiple cellular pathways and even directing immune responses [59-61]. It is thus not surprising that mutations in transcription factors or their binding sites lead to the development of several diseases [9]. Transcription factors often have their sequences conserved across species, indicating that in eukaryotes, several pathways are networks are similarly regulated and evolutionarily similar [62]. However, to regulate the expression of several tens of thousands of genes, until now, studies have been able to identify and characterize only about 1600 proteins that can behave as transcription factors [63]. This critical piece of data indicates that transcription factors are able to regulate the expression of multiple genes in different cells, possibly depending on the context and external stimuli, and observation that has been seen across different studies. Transcription factors, while binding to DNA, can recognize a short stretch of nucleotides on the DNA for their binding. This sequence is termed as a consensus site or a motif, and it is mostly observed that one gene will have multiple sequences that can recruit several transcription factors, which act co-operatively [64]. Several databases exist which profile and describe consensus sites for factor binding, such as TRANSFAC, JASPAR, HT-SELEX and Uniprobe[65-68].

Interestingly, transcription factors have also been seen to possess characteristic sequences that enable them to bind to DNA, thus facilitating their identification as well as classification. Earlier studies conducted using Electrophoretic mobility shift assays (EMSA), DNase I protection (footprinting) assays, Methylation interference assays, UV crosslinking, and Southwestern blotting have led to the identification that DNA binding transcription factors possess helix-turn-helix motifs, homeodomains, zinc finger motifs, leucine zipper motifs, and helix-loop-helix motifs (Reviewed in [58]). With the advent of modern techniques such as ChIP-seq, it has become easier to locate transcription factor binding sites across the genome and enabled us to identify other structural features of such proteins which enables their classification into superfamilies such as Forkhead boxes, THAP fingers, CxxC

motifs, AT-hooks, and T-boxes [63]. Transcription factors can function as activators or repressors. However, this mechanism is not always effected directly, but by the recruitment of accessory factors/proteins, which act as co-activators or co-repressors. Often, these accessory factors do not function alone and are instead found as part of large, multisubunit complexes which multiple regulatory modules. For example, the interferon- β -enhanceosome which contains ATF-2/c-Jun, IRF-3/IRF-7, and NF κ B binds to the *IFN- β* gene promoter and recruits the GCN5 and p300 acetyltransferases along with the SWI/SNF chromatin remodelers for activating transcription [69]. In the opposite fashion, the localization of the NCoR1/SMRT complex causes the recruitment of deacetylases such as HDAC3, HDAC4, HDAC5, or HDAC7 (in a context-dependent manner) or repressor proteins like SMRT/HDAC1-associated repressor protein (SHARP), which can bind RNA [70]. Transcription factors have a cell-type or tissue-specific expression, which enables one factor to regulate the expression of multiple pathways. For example, SOX2, OLIG1, and OCT7 express in the cerebral cortex, while GATA4 and TBX20 have high expression in cardiac tissue.

Transcription factors are the most commonly implicated reasons for the development of human diseases. Lambert et al reported the classification of about 1600 human transcription factors and identified diseases associated with their misregulation. Their genome-wide association studies (GWAS) revealed that human diseases ranging from multiple sclerosis, ischemic stroke, cancers, ulcerative colitis, hyperglycemia as well as celiac disease resulted because of improper functioning of transcription factors [63]. A comprehensive review implicating various transcription factors in several of the diseases covered above as well as in developmental and neurological disorders has been published previously [9].

mRNA binding and processing elements

Transcription of mRNA finally leads to the synthesis of proteins, which is why this review would be incomplete if we did not focus on elements that regulate the stability of mRNA, an important step that would determine the downstream expression of various classes of proteins. mRNA processing involves the marking of 5' ends of transcripts, when they are about 20-25 bases long, with a methylated guanosine "cap". Capping regulates the nuclear export of the mRNA, prevents its degradation by exonucleases, and promotes translation and excision of introns[71-73]. What is interesting is that the capping mechanism is brought close to the RNA Pol II only after it is phosphorylated by the TFIIF[74], as the recruitment of the capping enzymes RNA triphosphatase, RNA guanylyltransferase, and RNA-(guanine-7) methyltransferase require the phosphorylation at Ser5 of the CTD (C-terminal domain) which is carried out by the Kin28 kinase module of the TFIIF complex. Similarly, at their 3' end, most mRNAs are processed by cleavage and polyadenylation. This processing ensures the removal of the RNA Pol II only after the full-length transcript has been synthesized [18, 75]. Several reports have

suggested that the Ser2 phosphorylation on the CTD of RNA Pol II is necessary for the binding of poly-A factors and cleavage. Evidence is now emerging that shows 3' processing factors linked to the transcription machinery at the 5' ends of genes whose interactions enable efficient cleavage and polyadenylation (Reviewed in [76]).

Another method of transcriptional control involves the splicing of genes. A typical mammalian gene has about eight intronic sequences, spans about 30 kb, and possesses about nine exons that are only about 150-200 bases in length [77, 78]. About one-third of all human genes undergo alternative splicing, catalyzed by a spliceosome complex comprising five UsnRNPs and their associated proteins so that multiple products can be obtained from the same genomic sequences [79, 80]. These recognize the splice sites and branching points while other RNA binding elements, such as hnRNPs and SR proteins, recognize silencer elements as well as exonic and intronic splicing enhancers which together modulate splicing. Splicing is important as it regulates how the information stored in a DNA sequence is utilized, cutting and pasting together different sequences from a template together for the formation of a functional protein. Thus, splicing and its regulation play an essential role in the regulation of gene expression [80-83].

Control of gene transcription would not be complete if we did not touch upon small non-coding RNAs (sncRNA) and long non-coding RNAs (lncRNA) as research over the last few decades show that these elements perform significant functions in the regulation of gene expression. Small non-coding RNAs are of three types: short interfering RNAs (siRNAs), microRNAs (miRNAs), and PIWI-interacting RNAs (piRNAs). The siRNAs and miRNAs mostly are between 20 and 30 nucleotides long and originate from double-stranded DNA. The role of both these elements has been widely described in mRNA degradation, thus leading to the repression of transcription. siRNAs arise by the action of an enzyme called Dicer, wherein the siRNAs are bound to a complex termed RNA-induced silencing complex (RISC). After the double-stranded RNA molecule separates to form a single-stranded RNA, it then binds to its mRNA target, which is then cleaved by the action of the Ago2 protein (Reviewed in [84]). The miRNAs are transcribed as pri-miRNAs from their genomic sequences and pre-miRNAs. These are then processed by the endonuclease Dicer, forming a mature miRNA duplex. After strand separation, the mature miRNA is loaded onto members of the Ago protein family which mediates mRNA cleavage via the action of the RISC [85, 86]. Interestingly, it has also been seen that in some cases, miRNAs can mediate the upregulation of gene expression. For example, in serum starvation conditions, Ago2 and FXR1 can associate with AU-rich elements located at the 3' UTR and activate translation [87]. This same study also reported several miRNAs, including *let-7* to be capable of activating translation during conditions of cell cycle arrest. miRNA-mediated upregulation of expression has also been observed in oocytes and at the 5' UTR of mRNAs coding for ribosomal proteins under conditions of amino acid

starvation [88, 89]. Distinct from siRNAs and miRNAs, piRNAs are so named as they associate with PIWI-proteins, which are expressed primarily in germline cells and are necessary for spermatogenesis [90]. piRNAs are formed from single strand transcripts from genomic loci called piRNA clusters, rather than double-strand precursors. piRNAs can also form from 3' UTRs of several protein-coding genes. piRNA formation is catalyzed by RNA Polymerase II and the Rhino-Deadlock-Cutoff (RDC) complex, wherein TBP-related factor 2 (TRF2) and the Moonshiner (Moon) protein mediate transcription start. pre-piRNAs are processed by the endonuclease Zucchini and loaded onto PIWI proteins. Through the effect of the Ago3 and Aub proteins, mature piRNAs are formed which recognize and cleave target RNAs [91]. piRNAs are also involved in transposon silencing and DNA hypermethylation with their misregulation being implicated in several diseases [92, 93]. For example, several piRNAs, such as piR-020498, piR-823 and piR-Hep1 are seen to be upregulated in hepatocellular carcinomas [94, 95]. Analyses of other cancer tissues have revealed that the increased expression of piRNAs is prevalent across several cancers, such as piR-823 and piR-1245 in colorectal cancers, piR-651 in gastric cancers, piR-823 in oesophageal cancers and piR-10506469 in gall bladder cancers [96-100]. It is also interesting to note that while in most cases, piRNAs help in cancer progression by silencing tumor suppressor genes, some piRNAs, such as piR-017061 and piR-823 can also modulate attenuation of tumor progression in pancreatic and gastric cancers respectively [98, 101].

RNAs that are transcribed but not translated into proteins and are over 200 nucleotides long are termed as long noncoding RNAs (lncRNAs). About 16,000 possible lncRNAs have been identified so far *in silico*, but estimates suggest that over 100,000 such elements may be present in the genome [102, 103]. lncRNAs are generally transcribed by RNA Pol II, but maybe transcribed by other RNA Polymerases as well, and may arise from intergenic regions or from sense and antisense transcripts. RNA capture long sequencing techniques show that lncRNAs are similar in size to mRNAs, but have lesser and larger exons [104]. Most lncRNAs are transcribed by phosphorylation independent RNA Pol II, and termination of transcription is independent of polyadenylation signals, leading to accumulation of the lncRNAs on the chromatin, where they can be degraded by the RNA exosomes [105]. Often, lncRNAs may contain U1 snRNA binding sites, which recruit the U1 snRNP to transcriptionally active RNA Pol II, resulting in their tethering to the chromatin and preventing their degradation [106]. lncRNAs can be found both in the nucleus and the cytosol, facilitating their localization by the recruitment of specific factors. lncRNAs mostly regulate transcription by interacting with proteins and facilitating or preventing their binding to specific DNA sequences. Often, lncRNAs can have a direct effect on the transcription of genes, such as the CCAT1 lncRNA at the *Myc* locus [107, 108]. lncRNAs can interact with and enable the recruitment of chromatin modifiers to promoters to activate or inhibit transcription. They often do so by the formation of RNA-DNA hybrids which get recognized by transcription activators or suppressors, such as the *TARID* lncRNA mediated formation of R-loops

drives demethylation of the *TCF21* promoter [109, 110]. R loops also can form in *trans*, such as in the activation of auxin-regulated genes in *Arabidopsis thaliana* by *APOLO*. Here, the repressive chromatin state maintained by LHP1 is removed by its sequestration, leading to target gene activation [111]. lncRNAs can also modulate gene expression by acting as “decoys” of chromatin modifiers. For example, pluripotency in human embryonic stem cells is maintained by *lncPRESS1*, which sequesters the deacetylase SIRT6 from the promoters of target genes, thus maintaining H3K9Ac and H3K56Ac activation marks [112]. lncRNAs have also been implicated in the silencing of gene expression, such as the *XIST* lncRNA, which is responsible for the inactivation of the X chromosome in female mammals, or the lncRNA COOLAIR, which in cool climatic conditions is localized to the *FLOWERING* locus for PRC2 dependent gene silencing in *A. thaliana* [113, 114]. Similar to other regulatory elements, the misregulation of lncRNA development and functions have been implicated in the development of various human diseases such as cancers, metabolic disorders, and neurological conditions (Reviewed in [115])

Now that we have some idea about the different regulatory elements involved in the complicated and tightly regulated eukaryotic transcriptional process, let us look at how this multi-stage phenomenon is carried out faithfully.

1.4 Steps in Eukaryotic Transcription

Eukaryotic transcription is catalyzed by RNA Pol I, Pol II, or Pol III. RNA Polymerase I is responsible for the transcription of rRNAs while RNA Pol III transcribes the genes coding for small RNAs, tRNAs, and 5s rRNA. However, our studies are limited to transcription mediated by RNA Pol II, which facilitates the synthesis of mRNA of protein-coding genes. We will thus, examine in a few points, the process of transcription by RNA Pol II. For ease of understanding, the steps leading to the initial initiation and elongation phases are described pictorially in Figure 1.4.

1.4.1 Initiation

Genes in eukaryotes are present in compact chromosomes, where DNA strands are wrapped around histone proteins. Each such structure, named as a nucleosome, comprises of about 147 bases of DNA which are wrapped around a histone octamer core containing two copies each of histones H2A, H2B, H3, and H4 [116]. The closed structure is often associated with histone H3K9 methylation by SUV39H1, and H3K27 methylation by the PRC2 [117-119]. A repressive protein, HP1, also remains bound to these

methylation marks for maintenance of the heterochromatin structure. Before transcription is initiated, several ATP-driven chromatin remodeling complexes function to alter the structure of nucleosomes for the facilitation of transcription. Four such families of complexes have been described: SWI/SNF, INO80/SWR1, ISWI, and CHD.

The SWI/SNF complex, generally considered as a transcription stimulator, functions to slide or eject nucleosomes [120]. It also contains the RSC (Remodels Structure of Chromatin) complex, which has now been reported to be found at the promoters of several hundreds of genes [121, 122]. Hyperacetylation of nucleosomes at the promoter results in retention of these bromodomain-containing complexes. Here, it is interesting to note that RSC also functions at RNA Pol I and Pol III bound promoters, and is an essential factor for survival in *Saccharomyces* [123, 124]. In case of the INO80/SWR1 family, the remodeling is achieved not by displacement, but by alteration of the composition of the nucleosome. Here, ATP is used for the replacement of histone H2A with H2A.Z, which destabilizes the nucleosomal structure and thereby effects transcription [125, 126]. Along with transcription regulation, this complex is also involved in the repair of DNA [127, 128]. Contrary to the SWI/SNF and INO80/SWR1 families, the Imitation SWItch (ISWI) family functions to repress transcription possibly by cooperation with other repressive complexes and maintenance of a deacetylated state [129, 130]. Of all the remodeling complexes, the role of the CHD (Chromatin organization modifier, Helicase, and DNA-binding domains) family is the least studied. These proteins have a chromodomain at their N-terminus, which can bind to DNA, RNA as well as methylated histone H3 [131]. In *Drosophila* and yeast, these complexes have been shown to possess ATPase activity. In humans, CHD3 and CHD4 form a repressive complex known as the NURD (nucleosome remodeling and histone deacetylase) complex. The NURD complex contains HDAC1, HDAC2, RbAp48, RbAp46, MTA1, MTA2, MTA3, and MBD3, which together effect histone deacetylation along with DNA methylation for transcription repression.

The first step of transcription initiation involves the binding of TATA-binding protein (TBP) to the TATA box upstream of the transcription start site [132]. TBP is a component of the TFIID (Transcription Factor II D) complex, which also comprises of TBP-associated factors or TAFs. TBP, as part of the TFIID also gets recruited to TATA-less promoters (which occur in majority and are found upstream of housekeeping genes) mediated by the Inr elements [133]. Once the TFIID complex is recruited, it is followed by the TFIIB and the TFIIA. These complexes function to stabilize the interaction between the DNA and TFIID, along with countering the action of NC2 (Negative Cofactor 2) and Mot1 (Modifier of Transcription 1), which may cause the removal of TBP. TFIIA is normally seen to bind upstream to TBP, while TFIIB binding occurs downstream. Following this, RNA Pol II, TFIIF, and the Mediator are also recruited. Next, TFIIE associates with this complex and then recruits the TFIIH module, which

possesses kinase and helicase properties [134]. Once the mega-complex, termed as the PIC (Pre-Initiation Complex), is assembled on the DNA, the -1 nucleosome is then removed. Often partial PICs, which are devoid of TFIIF and RNA Pol II are assembled at the promoters of genes [20]. These remain bound to the DNA without loss of nucleosomes and are thought to serve as scaffold structures for rapid transcription induction. The XPB subunit of the TFIIF possesses ATPase and translocase properties and enables the DNA to “open up”, thereby enabling its entry into the RNA Pol II active site. This open DNA is now kept stabilized by TFIIF, TFIIE, and TFIIIB. For RNA Pol II to escape the promoter, it now needs to be phosphorylated. In yeast, the Kin28 (KINase 28) subunit of TFIIF, or its mammalian homolog, Cdk7, possess kinase activity [135]. This subunit now phosphorylates the heptad repeats of the CTD of RNA Pol II at its Ser5 and Ser7 positions [136]. This modification enables the transition from transcription initiation to elongation.

1.4.2 Elongation

Once the CTD of RNA Pol II is phosphorylated on its Ser5 and Ser7 positions, RNA Pol II elongation begins. However, it has been observed across multiple instances that elongating RNA Pol II, after traveling the first 20-60 nucleotides, suddenly pauses in its tracks [27, 137-140]. Most studies have attributed RNA Pol II promoter-proximal pausing to be mediated by the action of two proteins: DSIF (DRB Sensitivity Inducing Factor) and NELF (Negative Elongation factor). Several other factors have also been described to contribute to the pausing of Pol II, such as Gdown1, KAP1/TRIM28, FACT, GAGA factor, RNA polymerase II-associated factor 1 complex (PAF1C), and as per a recent report, even the TFIID (Reviewed in [17]). Multiple studies that have tried to answer as to why RNA Pol II pausing occurs seem to postulate it to be related to developmental control, with pausing serving as an event for a rapid response to external stimuli, inflammation, stress signals, and epigenetic changes. RNA Pol II is released from its paused state into productive elongation by the action of P-TEFb (Positive Transcription Elongation Factor-b), which is comprised of the kinase CDK9 and the associated CyclinT in *Drosophila* or CyclinT1/T2 in humans [141]. P-TEFb is normally found in an inactive complex, named the 7sk-snRNP complex, from where it is extracted out and recruited by multiple factors, such as BRD4, MYC, NF- κ B, G-Actin, p53, ER- α , MED23, and ZMYND8 [142-144]. The CDK9 module of this complex can now phosphorylate the CTD of RNA Pol II at its Ser2 position, which transitions the polymerase into productive elongation. In its active form, the P-TEFb is found associated with the Super Elongation Complex (SEC), initially thought to be a gigantic megadalton complex that facilitated elongation, but has been shown since to be dynamic in nature depending on cellular contexts, genes, and the stage of transcription [145, 146]. For bridging the transition between the initiation and elongation phases, the TFIID forms stable interactions with multiple components of the SEC, as indicated by previous studies from our group [147].

1.4.3 Termination

Termination of transcription does not occur at a specific site. RNA Pol II termination can occur proximal to the 3' end of genes, or in some cases, several kilobases away. For termination to occur, the pre-mRNA must undergo 3' end processing and have an intact polyadenylated tail. mRNA polyadenylation and cleavage occur downstream of a hexanucleotide sequence, AAUAAA, and upstream of a U/GU-rich element. The CPSF (cleavage and polyadenylation specificity factor) recognizes the AAUAAA motif, while the CstF (Cleavage stimulation factor) identifies the U/GU-rich region [18]. Multiple factors are involved with 3' end processing. For example, the CTD of the RNA Pol II, in its phosphorylated form at its Ser2 position, facilitates the recruitment of polyadenylation factors. Additionally, The Paf1 complex (Paf1C), identified initially as an elongation factor, also plays a role in mRNA 3'-end formation [148]. As RNA Pol II traverses the polyadenylation site (PAS), cleavage and polyadenylation (CPA) factors get associated with the Pol II machinery, thereby slowing down elongation. *In vitro* experiments have shown that the interaction of CPSF and CstF with Pol II as it passes a PAS leads to its slowing down and gradual dissociation from the DNA. Outside of the elongation complex, the nascent RNA can also hybridize with the open DNA, leading to the formation of thermodynamically stable R-loops. As Pol II encounters this bulky structure, it tends to backtrack, whence the nascent RNA formed downstream of the loop can wrap around the Pol II, leading to torsional stress and thereby, termination. Alternatively, if after traversing the PAS, the Pol II still continues synthesis of the nascent transcript, the exonuclease Xrn2 is recruited which degrades this formed transcript, eventually catching up with the Pol II, acting as a signal for its release from the DNA [149, 150].

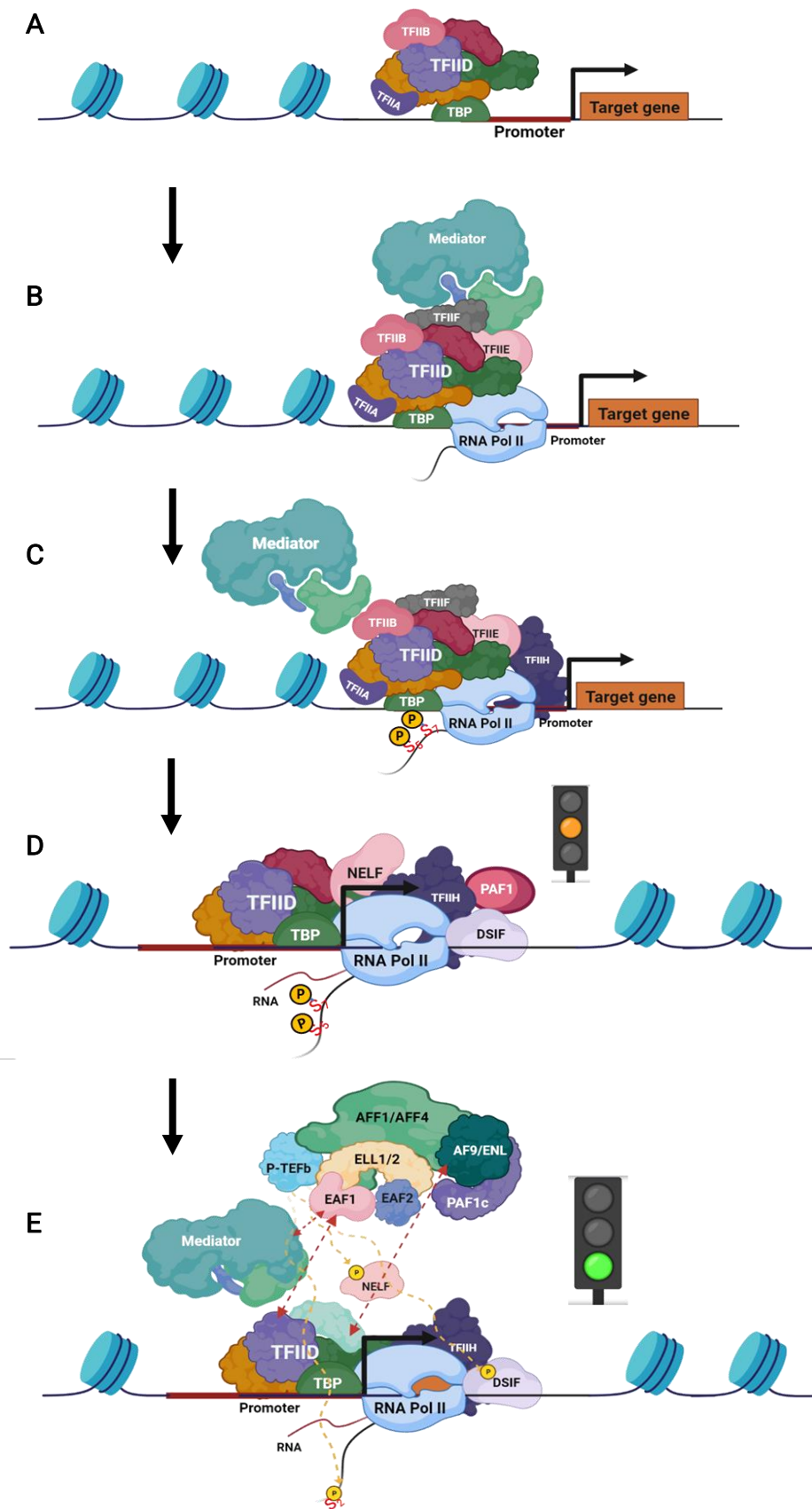


Fig. 1.4 An overview of the initial steps of transcription in eukaryotes

A. Binding of TBP to the TATA box results in the recruitment of the TFIID complex to the promoter region. This is followed by the arrival of the TFIIB and TFIIA complexes, which stabilize the interaction between TFIID and the DNA.

B. The Mediator complex gets recruited, along with the TFIIF complex and RNA Polymerase II. Following this, TFIIH gets associated with this complex

C. The TFIIH enables recruitment of the TFIIF complex, which possesses kinase and helicase activity. This is called the pre-initiation complex (PIC). Now, the CDK7 subunit of the TFIIF module phosphorylates the C-terminal domain (CTD) of RNA Polymerase II enabling the complex to move from initiation to the initial elongation phase.

D. After the synthesis of 20-60 nucleotides, the elongating RNA Pol II suddenly pauses in its tracks, particularly due to the action of two proteins, DSIF and NELF. Other proteins like PAF1 or KAP1 may also contribute to this promoter proximal pausing of RNA Pol II.

E. P-TEFb, the principal complex responsible for driving RNA Pol II into elongation is recruited by the action of several proteins, such as BRD4, Myc, or ZMYND8, among others. The CDK9 kinase module of the P-TEFb phosphorylates the negatively acting DSIF and NELF, by which their effect ceases. CDK9 also phosphorylates the CTD of RNA Pol II at its Ser2 position, a hallmark of transcriptionally engaged and elongating Pol II. The Mediator complex as well as the TFIID enables recruitment of the Super Elongation Complex (SEC) containing ELL as well as other factors necessary for transcriptional elongation. By the concerted effect of all these proteins, RNA Pol II proceeds into productive elongation.

(Adapted from Basu et al., (2020), *BBA-Gene Regul. Mech*)

1.5 The Super Elongation Complex in Transcription Regulation

Eukaryotic transcription, as described previously, consists of three major steps: initiation, elongation, and termination. Initiation is regulated by the formation of the PIC, which results in RNA Pol II traversing about 20 to 60 nucleotides downstream of the transcription start site, synthesizing a short nascent RNA strand. Upon reaching this condition, RNA Pol II pauses due to the action of two proteins, DSIF and NELF [29, 30]. RNA Pol II stays in this poised state until it receives signals for rapidly inducing transcription. RNA Pol II is made up of multiple subunits, the largest of them being Rpb1. In mammals, Rpb1 has a C-terminal domain (CTD) which contains 52 repeats of a heptad series of amino acids, in the sequence: Tyr-Ser-Pro-Thr-Ser-Pro-Ser (YSPTSPS). For RNA Pol II to initiate transcription, it is phosphorylated at its Ser5 and Ser7 positions by the Cdk7 kinase module of the TFIIF complex. Paused Pol II, poised to enter the productive elongation phase can only do so when CDK9 of the P-TEFb complex phosphorylates Ser2 of the CTD as well as DSIF and NELF, which either causes them to dissociate or get converted into a positive elongation factor [151, 152]. This P-TEFb, which is recruited to the RNA Pol II by multiple factors, is present as part of a giant complex, which facilitates Pol II-mediated transcription elongation, and is called the Super Elongation Complex (SEC). As the basis of

our studies is an RNA Pol II elongation factor, named ELL (Eleven Nineteen Lysine Rich Leukemia), which is an essential part of the SEC, we shall briefly be describing the role of the SEC and its components in transcription control.

The super elongation complex was proposed to contain the ELL family members ELL1, ELL2, and ELL3, AF4/FMR2 family member 1 (AFF1; also known as AF4), AFF4, eleven-nineteen leukemia (ENL) and ALL1-fused gene from chromosome 9 (AF9), and the RNA Pol II elongation factor P-TEFb[145]. However, subsequent studies have proven the presence of dynamic interactions between SEC components in a context-dependent manner. For example, AF4 and P-TEFb can form a complex, whilst higher-order complexes may be formed by AF9-AF4-P-TEFb or AF9-AFF4-P-TEFb. Similarly, the ELL-EAF1/2 complex was found to associate with AF4-P-TEFb to form an AF4-P-TEFb-ELL-EAF1/2 complex [146]. On the contrary, purification of AF9 interacting proteins failed to show the presence of ELL, even though both were thought to reside in the same complex, the SEC [146]. The SEC can activate genes with or without paused RNA Pol II. For example, upon HIV infection, the presence of the SEC is required for activation of the HIV-1-LTR [153, 154]. In case of genes where RNA Pol II is not paused, the SEC functions in rapid transcription activation as well. In mouse embryonic stem cells, the inactive *Cyp26a1* gene can be rapidly activated by the presence of retinoic acid (RA). The inactive *Cyp26a1* gene does not contain paused RNA Pol II, but upon treatment with RA, Pol II and the SEC are recruited, which activates transcription faster than that observed in RA-dependent genes where poised Pol II is present [155].

SEC recruitment may occur in several ways. For example, specific factors that bind to DNA could recruit the SEC. In the context of RA treatment, the RA receptor has been seen to bind to the SEC, which may result in its localization to the necessary genomic loci [145, 156]. Upon HIV infection, the Tat protein of the HIV serves as the recruiter of the SEC for transactivation [153]. As SEC components EAF1 and EAF2 interact with the Med26 subunit of the Mediator complex, these interactions may also serve as the bridge for channelizing the SEC to these regions [157]. Additionally, the PAF1 complex, through its interaction with AF9 or ENL, and the PIC component, TFIID, via its association with EAF1 and AF9, may also serve as methods by which the SEC reaches the necessary genomic loci [147, 158].

1.6 ELL (Eleven Nineteen Lysine Rich Leukemia)

Our work focuses on the detailed understanding of the RNA Polymerase II elongation factor, ELL. Hence, It is imperative that we gain some insight into what is known about ELL before detailed studies are conducted on it to understand its functioning. We have already mentioned previously that the presence of a giant complex, aptly named the super elongation complex (SEC) is important for the

elongation of RNA Pol II on gene bodies. What is interesting about the composition of the SEC is that several of its components, such as AF9, AF4, ENL, and of course, ELL, are known to be fusion partners of the Mixed Lineage Leukemia 1 (MLL1) protein [159, 160]. MLL1 has been shown to be a histone H3 lysine 4 (H3K4) methyltransferase. During embryonic growth, MLL1 is found localized at the genomic loci of several developmentally regulated genes, for example, the *Hox* cluster. During development, it is necessary that these genes express for an indicated period of time, and then get silenced [161, 162]. However, because of the presence of MLL1, gene silencing is hampered, leading to a condition named mixed lineage leukemia, which bears characteristic features of both Acute Lymphoblastic Leukemia (ALL) and Acute Myeloid Leukemia (AML) [163, 164].

Discovery of ELL and identification of its basic functions

By the end of 1994, six proteins had been found which could fuse to MLL1 and lead to leukemia. These were AF4 in t(4;11) (q21;q23), ENL in t(11;19) (q23;p13.3), AF9 in t(9;11) (p22;q23), AF6 in t(6;11) (q27;q23), AFlp in t(1;11) (p32;q23), and AFX in t(X;11)(q13;q23) leukemia. In August 1994, Janet Rowley's group from the University of Chicago Medical Center noted the presence of a novel gene that could fuse to MLL in a t(11;19) (q23;p13.1) translocation. This translocation was already quite commonly observed in *de novo* and therapy-related acute myeloid leukemia in infants, children, and adults and distinct from the t(11;19) (q23;p13.3) translocation, which had already been named ENL (Eleven nineteen leukemia). Experiments indicated that this gene was similar to the PARP protein, was hydrophilic and basic in nature, with a lysine-rich region at its C-terminus. Keeping these factors in mind, this gene was named Eleven Nineteen Lysine-Rich Leukemia (ELL, in short) [165]. A few short months later, researchers at the University of Tokyo reported the presence of the same protein from three different adult patient samples, all suffering from AML, and confirmed the similar properties it had to ELL. They however named this protein MEN (Myeloid Eleven Nineteen). Their experiments confirmed that MEN was expressed in multiple tissues and that its coding sequence was almost fully similar to that of ELL reported some months ago [166]. In a later study, this same group reported that MEN (or ELL) was localized to the nucleus in the K562 myelogenous leukemic cell line [167]. In the midst of this, mechanistic studies were carried out on ELL in the Conaway lab, which showed that upon addition of RNA polymerase II, TBP, TFIIB, TFIIE, TFIIIF, TFIIH, and a DNA sequence of the adenovirus 2 major late (AdML) promoter, the ELL protein, purified from a rat liver, could stimulate the rate of transcription. ELL, in the presence of initiation factors, was thus found to be an RNA Polymerase II elongation factor [168]. Subsequent studies by the same group showed that ELL could also inhibit initiation by the sequestration of RNA Pol II as seen in an *in vitro* assay. The inhibitory effect, distinct from its elongation property, resulted due to the effect of the first 50 amino acids, which also overlapped with its elongation domain located just adjacent to the inhibitory region [169].

Furthermore, detailed analyses of this dual mode of regulation by the purification and study of ELL and associated proteins from rat livers showed that ELL, a ~ 80 kDa protein, was found associated with three other polypeptides, with sizes ranging from 20, 30, and 45 kDa, which were named EAP20, EAP30, EAP45. This complex, named the Holo-ELL complex, was found to be unable to inhibit transcription, thereby leading to the speculation that the interaction of either of these three new ELL-associated proteins with the N-terminus of ELL led to the sequestration of RNA Pol II [170]. Another RNA Pol II elongation factor, seen to have about 66% identity to ELL was subsequently identified and named ELL2. ELL and ELL2 shared three regions of high similarity: at the N-terminus (7-352 amino acids), the middle lysine-rich region (443-474 amino acids), and between the C terminal 516-640 amino acids. As both these similar proteins were found to carry out the same functions, it was speculated that these proteins acted in tandem in different tissues and genes [171].

ELL has roles in multiple cellular processes

Once ELL had been structurally characterized (Figure 1.5), several studies elegantly deciphered mechanistic details into its functions. ELL had been described to increase the rate of RNA Pol II-mediated transcription by suppressing promoter-proximal pausing. The Hatakeyama group reported the presence of interaction between ELL and p53 via a yeast two-hybrid assay. The interaction mediated by the elongation domain of ELL and the C terminus of p53 resulted in the inhibition of both p53 as well as ELL. By inhibiting ELL activity, p53 could also inhibit RNA Pol II-mediated transcription. Similarly, upon genotoxic stress, the p53-dependent induction of the p21 gene was prevented by ELL, which saved the cells from apoptosis [172]. Additionally, this observation forms the basis for possible p53 inactivation by the MLL-ELL fusion protein, which leads to leukemia. ELL was also found to exist in *Drosophila*, encoding a 912 amino acid protein, and localized to the nucleus in follicular and germ cells. *Drosophila* ELL (dELL) was seen to be mainly expressed in the early embryonic stages, indicating that its role was essential in the early stages of development [173]. Detailed analysis of dELL by the Shilatifard group showed that dELL was indeed present in *Drosophila*, where its expression was heightened during development. dELL was found exclusively in the nuclei of *Drosophila* embryos, where it localized within the polytene chromosomes and was found to be exclusively associated with the hyperphosphorylated, elongating form of RNA Pol II [174]. The role of dELL was seen to be essential, as ELL functional mutants were lethal along with the development of segmentation defects, and their functions could not be rescued by the introduction of other elongation factors *in vivo* [175]. Domain analysis of dELL showed that distinct regions on the protein are responsible for different functions. While the N-terminal region was necessary for ELL recruitment and RNA Pol II binding, leading to transcriptional elongation, the middle region was responsible for dELL localization to gene loci upon rapid transcription induction. The C-terminus region was seen to be important for the

regulation of developmental processes[176]. In *Xenopus*, two homologs of ELL have been identified, named xELL1 and xELL2. Here, a deletion fragment of ELL with the elongation domain removed was seen to inhibit the Wnt pathway by binding to the Dishevelled protein in the cytoplasm. Further analyses revealed that the occluding homology domain present at the C-terminus was responsible for this inhibitory effect, thus showing that ELL, which was seen as a nuclear protein, could also affect signaling pathways in the cytoplasm [177]. Interest piqued in the potential role of ELL in regulating other cellular processes, and in 2005, Marc Lombes' group reported the interaction of the N-terminus of the mineralocorticoid receptor (MR) with ELL. In this case, ELL was shown to specifically act as an MR activator, with the opposite effect on the glucocorticoid receptor. The transactivation potential of ELL was lost in the MLL-ELL fusion, indicating that this first report of an interaction between a nuclear receptor and a transcription elongation factor is dependent on the elongation domain of this protein [178]. A possible mechanism of MLL-ELL mediated leukemogenesis was elucidated by Zhou et al., who showed that ELL could induce the expression of the thrombospondin-1 (*TSP-1*) gene, a known anti-angiogenic factor. This ability of ELL was also lost in the MLL-ELL fusion protein, speculating that this could be one of the ways in which ELL can act to prevent cancer[179]. Incidentally, our studies show, as will be presented later, that the depletion of ELL most often leads to cancer, suggesting a possible property of this protein as an anti-tumor factor. While ELL2 was studied to be involved in the development of B cells in the bone marrow and their maturation in the peripheral organs, the role of ELL as an anti-cancer protein was reported by yet another study, which demonstrated ELL as a HIF-1 α binding partner and inhibitor, the interaction between which led to inhibition cell proliferation and colony formation in PC3 (prostatic adenocarcinoma) cells [180-183]. Interestingly, a detailed mechanistic study also showed that ELL possesses E3 ligase-like activity, by virtue of which it can degrade the oncoprotein, c-Myc, resulting in reduced tumor growth [184].

ELL and its role in RNA Pol II-mediated transcription

Several studies have also elucidated in detail the role of ELL in RNA Pol II-mediated transcriptional regulation. For example, while studying the recruitment of the lysine acetyltransferase p300 to the promoter-proximal regions in T cells, the Gardner group identified several genes that rapidly recruited p300 as well as RNA Pol II to the gene promoters upon mitogen stimulation. This mediated the translocation of factors like P-TEFb, BRD4, and ELL for rapid transcription. After termination, RNA Pol II (in its unphosphorylated form) and p300 remained bound to the promoters for rapid re-recruitment of these elongation factors upon further stimuli [185]. In another study, the same group reported that ELL which was demonstrated priorly, to have had a role in transcription elongation, is also essential for the recruitment of the RNA Pol II, and regulates the process of transcription initiation as well as its entry into the pausing site. Afterward, upon receipt of the stimulatory signal, ELL now gets

incorporated into the super elongation complex and assists in the release of the paused RNA Pol II into the productive elongation stage [186]. dELL, the *Drosophila* analog of ELL had already been seen to perform a similar function at the *Hsp70* loci upon heat shock for the rapid induction of heat shock response genes [187]. In humans, HTLV-1 (Human T-Lymphotropic Virus type 1 or Human T-Cell Leukaemia Virus type 1) is responsible for the development of adult T-cell leukemia. It possesses a protein, Tax, that mediates the transcription of viral RNA. Like most viral transcription, HTLV-1 hijacks the cellular transcription machinery for the transcription of its own genes. In this case, the Tax protein induces ELL incorporation into complexes consisting of p300 and RNA Pol II, thereby inducing gene transcription [188]. As part of the super elongation complex (SEC), ELL along with P-TEFb was found to regulate gene expression in undifferentiated embryonic stem cells, enabling rapid transcription and effecting differentiation upon receiving the appropriate signals [155]. A proposed mechanism for this event involved the Mediator subunit MED26, which was found to interact first with the TFIID, and then with the SEC containing ELL and EAF1, for driving the RNA Pol II into productive elongation [157]. Incidentally, ELL was shown to be part of not only the SEC but a smaller complex as well, which was aptly named the Little Elongation Complex (LEC). The LEC was found to have four members: ELL, EAF1 KIAA0947, and NARG2, and was first identified as to be highly enriched at small nuclear RNA (snRNA) genes whose transcription was mediated by RNA Pol II, in spite of the absence of the P-TEFb complex [189]. Like the SEC was involved in the transcription of mRNA, the LEC specifically affected the transcription of snRNA. While no other studies could elaborate much on the roles of this complex, an elegant study published in 2014 identified a unique function of the LEC. The study showed ELL, as part of the Little Elongation Complex, to be able to interact with the Cdk7 subunit of the TFIIH complex. As the TFIIH had already been reported to play a role in DNA damage repair, this study presented this complex as necessary in transcription coupled DNA repair response. Upon successful repair of the damaged DNA lesion, the ELL containing LEC was then seen to play an essential part in the resumption of transcription [190].

ELL associated factors: EAF1 and EAF2

Work on ELL and its associated partners resulted in the identification of two proteins that could stably associate with ELL, which were named EAF1 and EAF2 (ELL-associated factors 1 and 2). EAF1 was found to interact with ELL1 and ELL2, forming stable complexes that could remain associated even under 1.2 M salt concentrations. The expression of ELL was seen to increase EAF1 expression. Additionally, EAF1 colocalized with ELL in a speckled pattern in the nucleus, but upon overexpression of MLL-ELL, it was seen to be delocalized [191]. The same group also identified EAF2, which had 74% of its amino acid sequence conserved compared to EAF1. EAF2 could also co-localize with ELL in the nucleus, and its expression was seen to be directly linked to developmental regulation in mice. Unlike EAF1, EAF2 was

seen to bind to the C terminus of ELL and could not interact with the MLL-ELL fusion protein [192, 193]. EAF1 and EAF2 were distinct from EAP20, EAP30, and EAP45 and hence it would be interesting to study their role in transcription to understand their interplay with ELL. Kong et al., by *in vitro* assays showed that EAF1, by virtue of its interaction with the N-terminus of ELL, could enhance Pol II-mediated transcription elongation from the AdML promoter. In similar assays, the group also showed that EAF2 could also positively regulate ELL-mediated RNA Pol II elongation activity [194]. While most studies were focused on the effect of cellular processes regulated by the ELL-EAF1 complex, several reports also elucidated the interplay between EAF2 and ELL. Primarily, EAF2 was found to act as an activator of transcription. However, the localization of EAF2 to nuclear speckles within the cell, enhancement of its potential to activate transcription, as well as the stability of EAF2 itself was seen to be dependent on the presence of ELL [195]. Subsequent studies by the same group revealed that EAF2 could act as a tumor suppressor gene and cause apoptosis upon interaction with ELL. The authors stressed on the point that in several cancers, including those of the prostate or leukemia, the EAF2 expression is either significantly decreased or ELL-EAF2 interaction is hindered by different mechanisms, which impedes apoptotic pathways and thus causes uncontrolled growth [196]. This hypothesis gained weight by studies in EAF2 knockout mice which showed them to have higher rates of occurrence of various cancers such as carcinomas of the lung and liver, prostate cancer as well as B cell lymphomas [197]. As it had been observed before that the ELL could stabilize EAF2, but no mechanistic insight was gained into this mode of regulation, an elegant study described how EAF2 levels within the cell were regulated by proteasomal degradation, and ELL acted as a “protector” of EAF2 by binding to it and preventing its interaction with the E3 ubiquitin ligase, Siah2 [198]. Thus EAF1 and EAF2, inspite of having significant similarities in their composition, have been seen to regulate widely different pathways within the cell, albeit in conjunction with the same protein, ELL. A recent study has also described that both these proteins, i.e EAF1/2 inhibit the formation of the SEC by preventing AFF1 interaction with ELL. The authors show that this negative mode of regulation can occur in the context of Tat protein-dependent HIV-1 transcription [199]. Though all studies until now had described EAF1 and EAF2 as positive regulators of ELL elongation activity, the reporting of such new information lends credence to the fact that inspite of being identified about two decades ago, the detailed interplay between ELL and its binding partners, EAF1 and EAF2 have not yet been fully understood. Thus detailed studies are the need of the hour to elucidate mechanistically how these two similar proteins act along with ELL to regulate such diverse processes within the cell.

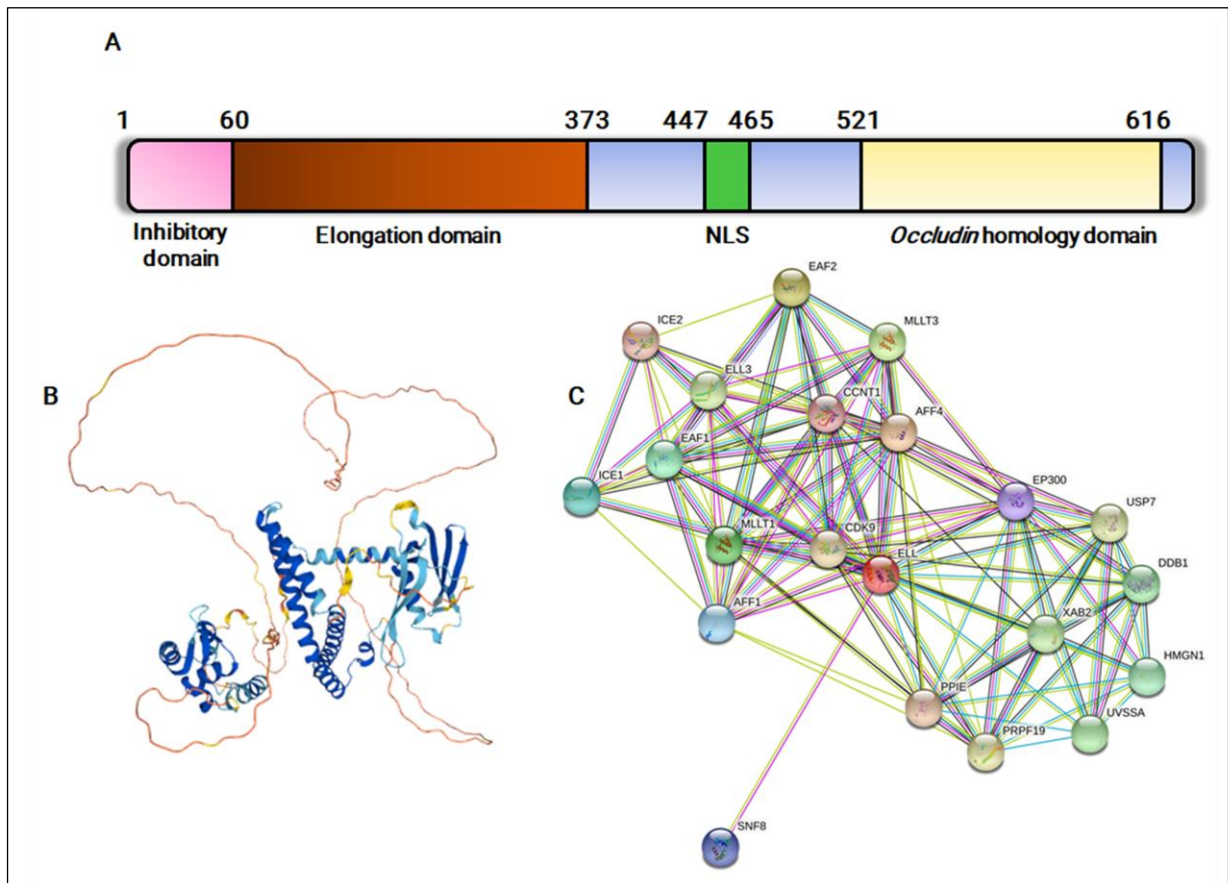


Fig. 1.5: Structure and functions of the Eleven Nineteen Lysine Rich Leukemia (ELL) protein

A. Functional assays show that ELL has several structural domains, such as the N terminal inhibitory domain, the elongation domain and an occluding homology domain at its C-terminus

B. Structure prediction of ELL shows that this protein is made up mostly of α -helices, as well as β -sheets along with several unstructured regions. As the full structure of ELL has not been solved experimentally, the 3D conformation of ELL was predicted using the Alphafold software *in silico*.

C. The predicted interactome of ELL indicates that it interacts strongly with proteins involved in the regulation of transcription as well as the cell cycle. This network was obtained using the STRING database.

1.7 DBC1 (Deleted in Breast Cancer 1)

DBC1 goes by many names. It is also known as CCAR2 (Cell Cycle and Apoptosis Regulator 2), KIAA1967, or p30 DBC. It was observed by Michael Wigler's lab at Cold Spring Harbor Laboratory, USA, that a particular genetic locus, at the human chromosome 8p21, was frequently deleted in several breast cancers [200]. Mapping of the genomic loci indicated that this region contained six genes: *DBC1*, *DBC2*, *TNFRSF10A*, *TNFRSF10B*, *EGR3*, and *KIAA0273*. Over the years, studies were conducted

to understand the functional implication of this newly identified protein, DBC1, in cancer progression and cell survival. However, pan-cancer studies indicated that DBC1 could not function as a universal tumor suppressor gene as it was neither seen to be lost across major cancer types, and even if present, its expression remained mostly unchanged [200, 201]. Thus, the name DBC1 appeared to be a misnomer, and based on subsequent detailed studies, this protein started being referred to as CCAR2, as it was found to play major roles in the regulation of apoptotic processes [202].

DBC1- An interactor of multiple deacetylases

After the discovery of DBC1 in 2002, not much significant progress was made in understanding the different roles this protein might play in cellular processes until 2008 when in the same issue of Nature, two groups reported that DBC1 could function as a negative regulator of the histone deacetylase, SIRT1. This inhibition of SIRT1 activity, mediated by the binding of the N-terminal Nudix hydrolase domain of DBC1 to the C-terminal deacetylase core of SIRT1, resulted in the acetylation of p53, thus inducing apoptosis upon genotoxic stress [203-205]. The discovery of this role of DBC1 suggested significant implications for this protein in the regulation of cell survival as well as metabolic processes, as SIRT1 is a known modulator of several metabolic pathways [206-208]. Interestingly, the very next year it was reported that DBC1 could inhibit the activity of Suv39H1, a methyltransferase enzyme that catalyzes heterochromatinization [209]. Additionally, it is important to note that SIRT1 also accentuates the activity of Suv39H1. Thus, by a two-pronged approach of inhibiting both Suv39H1 as well as SIRT1, DBC1 was implicated in chromatin remodeling leading to enhanced gene expression and cell proliferation [210, 211]. DBC1 was also seen to be an inhibitor of the deacetylase activity of HDAC3 via binding through its LZ motif. Through this interaction, DBC1 was able to maintain the acetylated and active state of the transcription factor MEF2D [212]. Subsequent studies led to the identification and characterization of several functional domains on the 923 amino acids containing DBC1 protein, such as the S1-like RNA binding domain (55-111 amino acids), a Leucine Zipper (LZ) motif between amino acids 243-264, which mediates interactions between histone deacetylases and a nuclear localization signal (202-219 amino acids). Furthermore, amino acids 339-462 constituted a Nudix hydrolase domain for binding to NAD⁺, an EF-hand (amino acids 704-748) for calcium ion binding, and a coiled-coil domain at the extreme C terminus (amino acids 794-918) which mediates protein-protein interactions (reviewed in [213]).

Transcriptional regulation by DBC1

In particular reference to the regulation of transcription, DBC1 not only modulates the activity of deacetylases or chromatin-modifying enzymes, but several other reports have touched upon its role in affecting other cellular processes. For example, DBC1 has been seen to interact with both the

androgen receptor (AR) as well as the estrogen receptors ER α and ER β through its N-terminus [214-216]. However, it regulates the expression of these proteins in different ways. DBC1 mediates chromatin binding and thus enhances transcription of the AR, while in the case of ER α and ER β , DBC1 binding directly to these two proteins stabilizes them and results in increased proliferation as well as survival of breast cancer cells [217]. In a contrasting report, DBC1 was identified as a potential tumor suppressor, wherein as a member of the nuclear receptor co-regulator (NRC) interacting factor 1 (NIF-1) complex it mediated activation of the retinoic acid receptor causing stimulation of its target genes, thus preventing the growth of cancer cells [218]. The first finding of DBC1 association with other proteins mediated through its C-terminus was in 2013, when the Chini group identified that DBC1 binding with the Rev-Erba protein receptor was able to protect it from polyubiquitination and proteasome-mediated degradation. Upon serum shock, DBC1 levels were seen to be decreased, thus resulting in Rev-Erba degradation and subsequent upregulation of its downstream genes [219]. This report implicated DBC1 in the regulation of the circadian rhythm as well. In a recent study, a complex containing ZNF326 (a.k.a ZIRD) and DBC1 was identified which was named the DBIRD complex [220]. The DBIRD complex was seen to interact directly with RNA Pol II and could regulate alternative splicing at exons. Depletion of DBIRD resulted in reduced transcriptional elongation, indicating that the DBC1 containing complex functions to regulate elongation as well as alternative splicing.

DBC1 and cellular metabolism

The regulation of cellular metabolism by DBC1 is mainly due to its interaction and modulation of the functions of SIRT1. When treated with a high-fat diet, the DBC1-SIRT1 interaction is markedly increased. In DBC1 knockout mice, SIRT1 levels are high, protecting the mice against steatohepatitis and inflammation [221]. The depletion of DBC1 causes high-fat-diet fed mice to be more obese but are sensitive to insulin, do not show pronounced atherosclerosis, and have longer life spans [222]. This is probably due to the activation of stearoyl-coenzyme A desaturase 1 (SCD1) by SIRT1, which causes unsaturated fatty acid accumulation [223]. Since lipid metabolism also involves the deacetylation of Peroxisome Proliferator-Activated Receptor (PPAR γ) by SIRT1, the genetic removal of DBC1 from mice results in them developing glucose intolerance with age. As DBC1 regulates the expression and activity of Rev-Erba and SIRT1, which in turn are implicated in the regulation of Phosphoenolpyruvate Carboxykinase (PEPCK), depletion of DBC1 has been shown to result in higher PEPCK protein expression leading to enhanced gluconeogenesis [224].

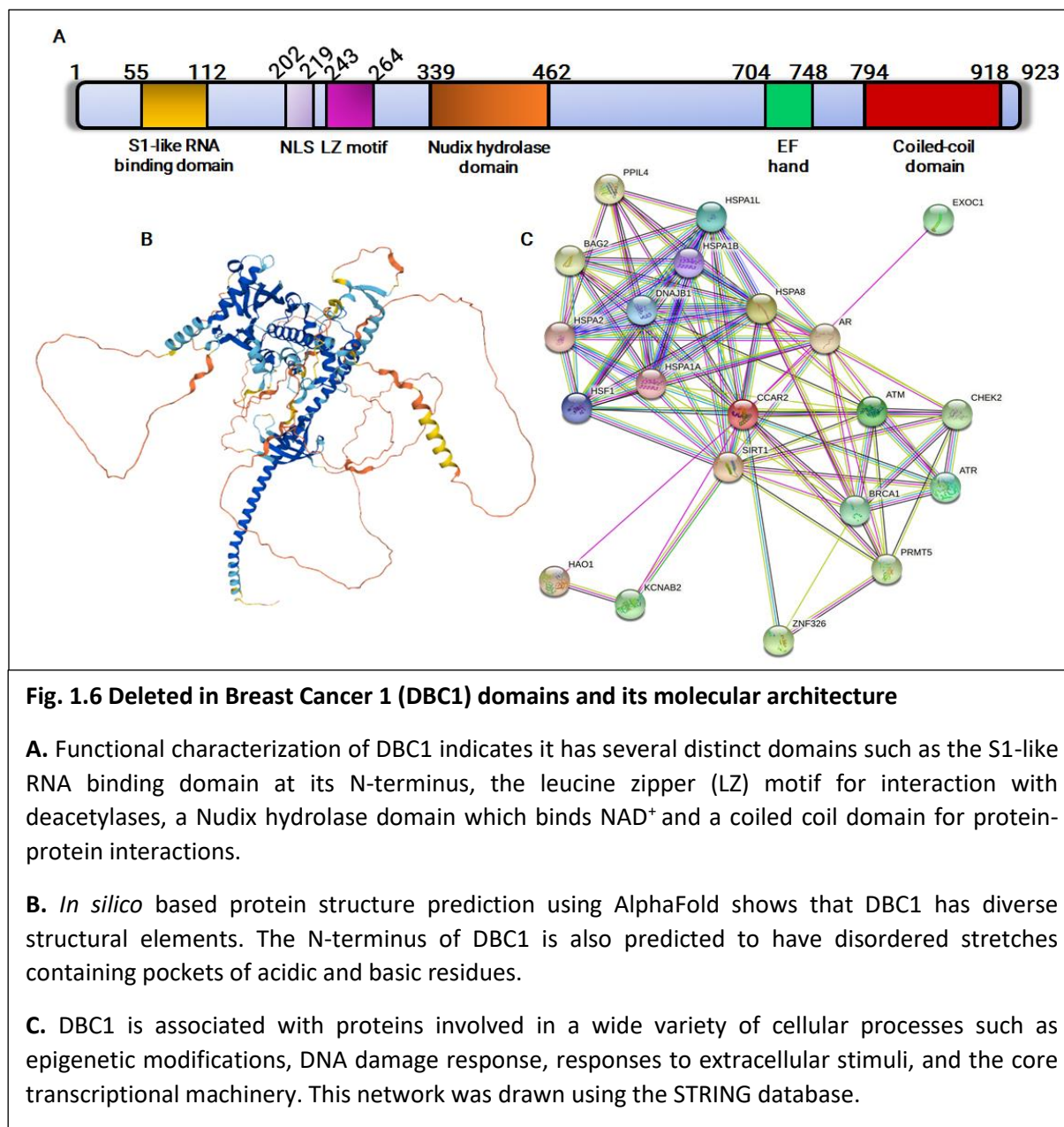
DBC1 is a versatile protein that regulates multiple cellular processes

The versatile DBC1 has also been implicated in the regulation of immune cells. For example, the CD40/BAFFR signaling pathway results in the activation of DBC1, leading to the suppression of RelB

and P52 proteins from the NF- κ B pathway, causing reduced formation and proliferation of B cells [225]. Additionally, IKK α -mediated phosphorylation of DBC1 results in the suppression of NF- κ B, which can also cause a reduction in B cell numbers [226]. A characteristic protein expressed by regulatory T cells (Tregs) is Forkhead box P3 (FOXP3), which is acetylated in its active form, and degraded by ubiquitination upon deacetylation in the presence of SIRT1 [227, 228]. DBC1 can interact with FOXP3 and negatively regulate its expression and function, possibly by the Caspase8 mediated signaling pathway [229].

A summary of DBC1 functions would not be complete if its role in the DNA damage response is not mentioned. Upon DNA damage, the KRAB domain associated protein 1 (KAP1) gets activated by ATM and CHK2-mediated phosphorylation on Ser824 and Ser473 respectively, which facilitates heterochromatinization by the recruitment of enzymes such as deacetylases and nucleosome remodelers [230, 231]. DBC1, being indispensable for the activation of CHK2, promotes the phosphorylation of KAP1, thus promoting heterochromatin relaxation and DNA repair [232]. A recent article has also suggested a possible role of DBC1 in DNA repair as a result of its interaction with poly-ADP ribose polymerase (PARP1) [233]. PARP1 has been extensively studied for its functions in the DNA damage response (reviewed in [234]). The report demonstrated that DBC1 binding to PARP1 inhibited its function in the DNA repair processes. In young cells, the presence of sufficient concentrations of NAD⁺ resulted in its binding to the Nudix-hydrolase domain of DBC1, which in turn caused its dissociation from PARP1, leading to efficient DNA repair. The study showed that with old age, as NAD⁺ declined, the DBC1-SIRT1 association was enhanced, leading to defects in DNA repair, with a possible promotion of tumorigenesis.

Thus, a multifaceted protein like DBC1 appears to have enormous implications in various cellular processes and thus is a prime candidate for studying and gaining mechanistic insights into the regulation of various processes within the cell. A pictorial representation of the DBC1 structure and its known interactions are represented in Figure 1.6.



Chapter 2

Rationale of the Study

2. Rationale of the work

Transcription regulates almost all cellular processes and is essential for the maintenance of a homeotic environment within the cell. In eukaryotes, transcription is carried out by three RNA Polymerases, which are named RNA Polymerase I, II, and III[235]. While RNA Polymerase I is restricted to the transcription and synthesis of the three largest rRNAs, namely, 28S, 18S, and 5.8S, where S denotes the Svedberg unit, RNA Polymerase III mediates the transcription of the smallest rRNA, the 5s rRNA, along with other transfer RNAs (tRNAs)[16, 236, 237]. RNA Polymerase III also transcribes RNAs involved in splicing and protein transport, such as small nucleolar RNAs (snRNAs) and small conditional RNAs (scRNAs). The bulk of the transcriptional burden within the cell falls on RNA Polymerase II, which effects the transcription of mRNAs leading to the synthesis of functional proteins, the building blocks of all components that perform cellular functions[151, 238].

As RNA Polymerase II mediates transcription of the most abundant and yet, critically the most important cellular components, it is essential to study the processes by which these components, ie: proteins, form. The mammalian genome can code for approximately 20,000 to 25,000 protein coding genes, but all of these are not expressed at the same time[239]. Rather, two types of genes are mostly found in the cell: housekeeping genes, and inducible genes. Housekeeping genes are constitutively expressed and are mostly essential for the maintenance of regular cellular functions. Examples include genes responsible for the synthesis of RNA, such as *AARF*, *ATF2*, *POLR2A*, as well as other cellular proteins such as *PPID*, *CCNK*, *ARPC2*, *TUBB*, and *EFNA3*, among numerous others[240]. Inducible genes are only expressed upon detection of certain extracellular or intracellular stimuli. Thus, the mechanism of turning on or turning off certain genes, which begins from the stage of neonatal development, and continues all throughout the life span of an organism, needs to be tightly regulated in order for cellular processes to be maintained[63, 241].

Regulation of gene expression has been one of the most well-studied phenomena in living organisms, directly affecting protein synthesis and thus cellular functions, thus having implications in the development of various diseases. Up until about two decades ago, the study of transcriptional regulation in eukaryotes was mostly focused on the steps preceding elongation, i.e. pre-initiation and initiation. The chronology of recruitment of different factors/complexes, and the roles they play has been a topic of interest for several years. Elongation was merely thought of as the simple addition of nucleotides to the growing mRNA. However, studies from several groups elucidated the importance of transcriptional control at the stage of elongation itself, emphasizing an occurrence termed RNA Pol II “pausing”[29, 30]. Pausing of RNA Pol II was seen to occur after it had synthesized a transcript

approximately 20-60 nucleotides in length, after which it underwent stalling, awaiting an external signal/stimulus. This phenomenon was found to be of utmost importance for the regulation of the expression of several genes, particularly those involved in development. Subsequent studies demonstrated the presence of several elongation factors in a megadalton complex, termed as the Super Elongation Complex (SEC)[145]. Though several studies prove convincingly that the components of the SEC are not rigid, and the complex itself is highly dynamic, several member proteins have been isolated, studied, and characterized[146]. These include AF9, ENL, AFF1/AFF4, ELL, and P-TEFb (a heterodimer of CyclinT1 and the kinase CDK9). Interestingly, the majority of these SEC components are found to form stable fusions with the Mixed Lineage Leukemia 1 (MLL1) protein, which methylates histone H3 at lysine 4 (H3K4), resulting in transcriptional activation[159]. However, aberrant transcriptional upregulation of several developmentally regulated genes by MLL fusion proteins is often the principal cause for the development of mixed lineage leukemias, which are extremely aggressive and have a poor prognosis. For example, the MLL fusion partners AF4 and ENL give rise to acute lymphoid leukemia (ALL), whereas the fusion partners AF9, ELL, and AF10 give rise to acute myeloid leukemia[160]. As MLL1 associates with key proteins involved in transcriptional elongation, aberrant transcriptional output mediated by these proteins in conjunction with MLL1 may be a significant cause for the development of such cancers. Hence, elucidation of such mechanisms may prove to be instrumental in developing diagnostics, identifying markers, and formulating therapeutics for the treatment of such neoplasms.

Of all the SEC components, ELL (Eleven nineteen lysine-rich leukemia) is of particular interest as it is the only SEC component that possesses the potential to stimulate RNA Pol II-mediated transcriptional elongation by itself. Several studies have shown that ELL alone or in association with ELL-associated factors 1 and 2 (EAF1/2) are able to stimulate RNA Pol II-mediated transcription *in vitro*[146, 168]. ELL also forms a complex known as the Little Elongation Complex (LEC) along with EAF1, KIAA0947, and NARG2 for the transcription of snRNA genes as well as the resumption of transcription after transcription-coupled DNA repair (TCR)[190]. ELL has also been seen to be necessary for pre-initiation complex formation, RNA Pol II entry into the pause site, and its subsequent release into productive elongation[186]. Thus, ELL serves as a versatile protein with several different key functions within the cell. However, studies on how ELL functions are regulated had not been studied, and thus was an exciting avenue for our exploration. In trying to understand how functions of ELL were regulated, we sought to identify ELL interacting proteins. What piqued our interest was the strong presence of the versatile Deleted in Breast Cancer 1 (DBC1) protein among ELL interactors. DBC1 has been well characterized as necessary for the action of p53, a negative regulator of the deacetylases SIRT1 and HDAC3, and interacting with HDAC5 and SIRT7[212, 219, 242]. It inhibits the histone methyltransferase Suv39H1 and is involved in RNA processing as part of the DBIRD complex[209, 220]. DBC1 has also

been seen to interact with BRCA1 during UV-induced DNA damage, activating p53, and causing apoptosis[243]. In the liver, DBC1 depletion leads to an increase in PEPCK and thus increased gluconeogenesis[224]. DBC1 also regulates adipocyte differentiation and lipid accumulation within cells[244]. However, inspite of its versatility and implication in various cellular processes, not much attention has been paid to DBC1 regulating functions of key transcription-associated factors other than p53. Hence, in this study, our primary aim was to elucidate the interplay between the key transcriptional elongation factor ELL and DBC1 and understand the molecular processes behind ELL regulation in order to understand potential mechanisms by which it could function in affecting leukemia and possibly other diseases in humans.

Chapter 3

Materials and Methods

3. Materials and Methods

3.1 Experimental Models

For performing all our experiments, we have used mammalian, insect, and bacterial cells. Mammalian cells used for our assays were HEK-293T (ATCC CRL-3216™) and HCT116 (ATCC CCL-247™). These were cultured and maintained in Dulbecco's Modified Eagle Media (Hi-glucose), containing 10% Fetal Bovine Serum (Gibco, USA) and supplemented with 100 U/ml of Penicillin-Streptomycin (Life Technologies, USA) at 37°C and 5% CO₂ in a humidified incubator. For baculoviral reconstitution assays, Sf9 cells were used. These were cultured in Grace's Insect Media (Himedia, India), supplemented with 10% FBS and 7 µg/ml Gentamycin, and kept in a BOD incubator at 26°C. For the purpose of cloning, we utilized *E. coli* DH5α cells, grown in Luria Bertani broth (Himedia, India), with appropriate selection markers as per the plasmid transformed. Protein expression for purification and use in *in vitro* experiments was done in *E. coli* BL21(DE3) (New England Biolabs, USA). All bacterial cells were grown at 37°C with constant shaking at 160-200 RPM.

3.2 Cell Culture and Transfection

For ectopic expression of proteins, appropriate plasmids were transfected into HEK-293T and HCT-116. Lipofectamine2000 (Invitrogen, USA) was used as the transfection reagent of choice for both cell lines and used as per the manufacturer's protocol. Briefly, seeded cells were grown to a confluence of ~60-70% in appropriate cell culture grade plates/dishes (BD Biosciences, USA). Plasmids to be transfected were added to Opti-MEM reduced serum media (Life Technologies, USA). A separate mix was also made with Opti-MEM and Lipofectamine2000 maintaining the ratio of 3:1, i.e: 3 µl Lipofectamine2000 for 1 µg of DNA transfected. Cells were harvested ~48 hours post-transfection (unless otherwise mentioned) and processed for downstream applications. For the preparation of baculoviruses, Sf9 insect cells were transfected using Cellfectin II reagent (ThermoFisher, USA) as instructed by the manufacturer. Sf9 cells were kept in Grace's Insect media without FBS or antibiotics. Baculoviral bacmid DNA was mixed in 100 µl Grace's insect media. For each mix, 8µl Cellfectin II was used. The mix was kept at room temperature for 30 minutes prior to adding to the cells. Post transfection, the media was replaced with complete media (containing FBS and antibiotics) for the generation of baculovirus.

3.3 Generation of plasmid constructs for expression in mammalian, insect, and bacterial expression systems

For our assays in mammalian and insect cells, as well as purification of proteins in bacterial expression systems, a wide variety of vectors were generated in different expression plasmid vectors. These plasmid constructs are listed in Table 1. For cloning of genes of interest in appropriate expression vectors, the plasmids containing the ORFs of the genes were obtained from Open Biosystems, USA. These include the ORFs of ELL, EAF1, EAF2, DBC1, and CDK9. Other expression plasmids, such as those of HDACs 1-8, SIRT1-7, p300, CBP, PCAF, GCN5, MOF, ubiquitin, and Siah1 were purchased from Addgene. For expression in mammalian cells, the majority of the genes were cloned in different epitope-tagged pcDNA5/FRT/TO plasmid vectors. *In vitro* studies were performed with purified proteins obtained after cloning in pET-GST vector (for GST epitope tag), and 6xHis-pET 11d vector (for N-terminal His tag). For use in insect cells, bacmids were generated by cloning genes of interest in pFASTBAC vector containing different epitope tags as mentioned. For cloning the gene of choice into the desired expression vector, the entire ORF was amplified by PCR, generating compatible restriction sites as available in the destination vector. After purification of the PCR products using the Qiaquick PCR purification kit (Qiagen, Germany), the products were digested with desired restriction enzymes as mentioned, obtained from NEB, USA. Similarly, the vector of choice was also digested, and the products were purified after resolving them in an agarose gel. Gel purification was carried out using Qiaquick Gel Extraction Kit, following the recommended protocol. The vector and inserts were ligated at room temperature for 3 hours to overnight in the presence of T4 DNA ligase (New England Biolabs, USA). Ligated DNA was transformed into chemically competent *E. coli* DH5 α cells and colonies obtained screened for the clone of choice by plasmid purification, followed by restriction digestion. The details of all clones generated in this study, as well as other clones that have been created and used earlier in previous studies, are summarized in Table 1 along with details of their cloning information.

Table 1: List of plasmid constructs

Name of Plasmid	Description	Source
M10	FLAG-HA pcDNA5-FRT-TO vector	[146]
M13	AF9 cloned into FLAG-HA pcDNA5-FRT-TO vector	[146]
M15	ELL cloned into FLAG-HA pcDNA5-FRT-TO vector	[146]
M24	AFF1 cloned into FLAG-HA pcDNA5-FRT-TO vector	[146]

M56	pET-GST vector	GE Amersham
M61	ELL cloned into pET-GST vector	[147]
M63	ELL (1-521 aa) cloned into pET-GST vector	[245]
M65	ELL (1-421 aa) cloned into pET-GST vector	[245]
M67	ELL (1-321 aa) cloned into pET-GST vector	[245]
M69	ELL (1-221 aa) cloned into pET-GST vector	[245]
M72	ELL (1-71 aa) cloned into pET-GST vector	[245]
M237	EAF1 cloned into FLAG-HA pCDNA5-FRT-TO vector	[146]
M238	EAF2 cloned into FLAG-HA pCDNA5-FRT-TO vector	[146]
M250	CDK9 cloned into FLAG-HA pCDNA5-FRT-TO vector	[146]
M296	psPAX2 lentivirus packaging plasmid	Addgene plasmid#12260
M297	pMD2.G lentivirus envelope plasmid	Addgene plasmid#12259
M298	Lentiviral pLKO.1 vector containing scramble sequence (as a control for shRNA knockdown)	[147]
M336	DBC1 cloned into FLAG-HA pCDNA5-FRT-TO vector	[245]
M378	FLAG-PCAF in pcl vector	Addgene Plasmid #8941
M379	HA-CBP in pRc/RSV vector	Addgene Plasmid #16701
M380	FLAG-HDAC1 in pcDNA 3.1 vector	Addgene Plasmid #13820
M381	FLAG-SIRT1 in pECE vector	Addgene Plasmid #13820
M434	EAF1 cloned into pET-GST vector	[147]
M435	EAF1 cloned into 6xHis-pET-11d vector	This study
M442	VP16-activation domain (AD) cloned into pM-DNA-BD vector	[245]
M455	EAF1 cloned into FLAG pCDNA5-FRT-TO vector	[147]
M458	EAF2 cloned into FLAG pCDNA5-FRT-TO vector	[147]
M464	DBC1 cloned into FLAG pCDNA5-FRT-TO vector	[245]
M465	DBC1 cloned into HA pCDNA5-FRT-TO vector	[245]
M496	pEGFP-N2 vector	Clontech (Takara)
M498	ELL cloned into pEGFP-N2 vector	[245]
M499	EAF1 cloned into pEGFP-N2 vector	This study
M532	P300 cloned into myc-pCDNA5-FRT-TO vector	[246]

M580	AFF4 cloned into FLAG-HA pCDNA5-FRT-TO vector	[246]
M594	Myc-tag DBC1 in pcDNA.3	Addgene plasmid #35096
M686	P300 HAT domain cloned into 6xHis-pET-11d vector	[246]
M742	ELL(1-500aa) fragment cloned into FLAG pCDNA5-FRT-TO vector	[245]
M743	ELL(1-373aa) fragment cloned into FLAG pCDNA5-FRT-TO vector	[245]
M744	ELL(45-621aa) fragment cloned into FLAG pCDNA5-FRT-TO vector	[245]
M745	ELL(374-621aa) fragment cloned into FLAG pCDNA5-FRT-TO vector	[245]
M746	ELL(45-373aa) fragment cloned into FLAG pCDNA5-FRT-TO vector	[245]
M747	ELL cloned into FLAG pCDNA5-FRT-TO vector	[245]
M748	ELL cloned into Myc- pCDNA5-FRT-TO vector	[245]
M758	DBC1(1-794aa) fragment cloned into FLAG-HA pCDNA5-FRT-TO vector	[245]
M759	DBC1(1-704aa) fragment cloned into FLAG-HA pCDNA5-FRT-TO vector	[245]
M760	DBC1(1-340aa) fragment cloned into FLAG-HA pCDNA5-FRT-TO vector	[245]
M761	DBC1(1-243aa) fragment cloned into FLAG-HA pCDNA5-FRT-TO vector	[245]
M762	DBC1(112-923aa) fragment cloned into FLAG-HA pCDNA5-FRT-TO vector	[245]
M773	MOF cloned into His- pCDNA5-FRT-TO vector	[246]
M787	HDAC4 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #30485
M788	HDAC7 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13824
M789	HDAC8 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13825
M790	HDAC2 (mouse) cloned into FLAG pCDNA3.1 vector	Addgene plasmid #68117
M791	HDAC6 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13823
M792	HDAC3 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13819
M793	SIRT2 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13813
M794	FLAG-GCN5 in pAdEasy vector	Addgene plasmid #14106

M796	HDAC5 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13822
M800	SIRT7 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13818
M801	SIRT5 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13816
M802	SIRT4 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13815
M804	SIRT6 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13817
M806	SIRT3 cloned into FLAG pCDNA3.1 vector	Addgene plasmid #13814
M912	Ubiquitin cloned into EGFP-C1 vector	Addgene plasmid#11928
M987	ELL (K5A) point mutant cloned into FLAG-HA-pcDNA5-FRT-TO vector	[245]
M993	ELL(61-621aa) fragment cloned into pET-GST vector	[245]
M1005	HDAC3 cloned into HA pCDNA5-FRT-TO vector	This study
M1006	ELL (K5A, K29A) point mutant cloned into FLAG-HA-pcDNA5-FRT-TO vector	[245]
M1013	Siah1 cloned into FLAG-HA-pcDNA5-FRT-TO vector	Addgene plasmid #45569
M1028	TRIM28 cloned into pKH vector	Addgene plasmid #45569
M1156	ELL cloned into 6xHis-pET-11d-GFP vector	This study
M1256	EAF1 cloned into 6xHis-pET-11d-RFP vector	This study

3.4 Generation of plasmids by site-directed mutagenesis (SDM)

In our studies, we generated mutants of ELL and TRIM28 harboring single amino acid substitutions for different functional assays. For example, we cloned the ELL K5A, and K5A, K29A mutants for identification of the key acetylation as well as ubiquitylation residues on ELL. To generate the above mentioned mutants for our studies, we used the Quikchange II Site-Directed Mutagenesis Kit (Agilent Technologies, USA). Amplicons were generated potentially harboring the modified sequences by PCR amplification using Phusion High Fidelity DNA polymerase (Thermo Scientific, USA). The amplified products were subjected to DpnI mediated digestion to cleave methylated DNA. The non-methylated PCR products were then transformed into chemically competent *E. coli* XL1-Blue cells and selected using appropriate selection markers. The obtained plasmids, after purification, were sent for Sanger sequencing for the identification of positive clones. Clones obtained were used in our transfection experiments.

3.5 Bacmid and baculovirus preparation for Sf9 experiments

The Bac to Bac[®] expression system (Invitrogen, USA) was used to generate baculoviruses for protein purification as well as reconstitution assays. This system uses site-specific transposition for the expression of high levels of recombinant proteins. For this, a gene of choice is first cloned into desired pFastBac vectors which contain baculovirus-specific promoters. The cloned plasmid is transformed into the modified *E. coli* DH10Bac cells, which already harbors two vectors: a bacmid (bMON14272), containing a lacZ peptide with the transposition site for the gene of choice, and the helper pMON7124 plasmid, which encodes the transposase. For our assays, we cloned the DBC1, ELL, EAF1, EAF2, and CDK9 genes into FLAG, His, and untagged pFastBac vectors respectively. These were then transformed into *E. coli* DH10Bac by electroporation, and selected with Kanamycin, Tetracycline, Gentamycin, IPTG, and X-gal. The use of IPTG and X-gal is necessary as the successful insertion of the gene of choice within the transposition cassette disrupts the lacZ peptide and thus would cause the formation of white colonies of the Luria Agar plate. After incubation of the transformed bacteria for 48 hours at 37°C, plates were checked for the presence of white colonies. Blue colonies, which were formed due to non-disruption of the transposase were discarded, and the white colonies were inoculated with the mentioned antibiotics for DNA purification. After the recombinant DNA bacmids were isolated, these were then transfected into Sf9 cells using Cellfectin II as per the protocol mentioned earlier. 72 hours post-transfection, when cells started showing signs of stress, thus indicating the formation of viral particles, the supernatant was collected and marked as P₀ stage baculovirus. As these are a low titer viral stock, Sf9 cells in 6 well plates were infected with these P₀ stage baculovirus and incubated in a 26°C incubator for 48 hours. After 48 hours, cells showed signs of stress indicating the production of a higher titer virus. This supernatant was then collected and marked as P₁ virus and stored at -80°C. Further amplification of viral stock was done by again infecting cells with the P₁ virus, incubating at 26°C in an incubator for another 48 hours, after which the supernatant was collected, named as P₂, and stored protected from light at 4°C. For all our experiments involving expression studies in Sf9 cells, we have used the P2 viral stock of different proteins in different combinations as was required.

3.6 Generation of Stable Overexpressing Cell Lines

In our studies, we have generated two stable expressing cell lines: one for ELL and the other for DBC1. For this, at first, the ORFs of both ELL and DBC1 were cloned into the FLAG-HA-tagged-pcDNA5/FRT/TO vector. Details of these clones have already been provided in Table 1. This vector is an inducible expression vector for generating stable cells using the Flp-In T-Rex System (Invitrogen, USA). In this

system, the Flp-In™ T-REx™-293 Cell line, containing a single FRT site, serves as the destination for the overexpressed gene fragment. 500 ng each of the two plasmids, one being the FLAG-HA-tagged-pcDNA5/FRT/TO vector, along with pOG44 (encoding the Flp recombinase) were transfected into the cell line using the protocol mentioned previously. The next day, cells were passaged and plated into a 100mm cell culture plate. On attaining ~70-80% confluency, Hygromycin was added to the media at 200 µg/ml, and cells were maintained at 37°C and 5% CO₂ in a humidified incubator. Colonies of resistant cells developed after about ~1 week. These were then picked up carefully by scratching with a tip, and cultured in a 24-well plate. When these also grew, multiple colonies of cells were then transferred to a 6 well plate, where the expression of the target proteins was induced by the addition of 0.5 µg/ml Doxycycline. Induced cells were maintained for 48 hours at 37°C and 5% CO₂, after which they were harvested in PBS, and lysed in RIPA buffer (50mM Tris-Cl (pH 8), 150mM NaCl, 1% Triton X-100, 0.1% SDS, 4 µl/ml PMSF and 0.7 µl/ml β-Mercaptoethanol). The expression of DBC1 and ELL were checked by immunoblotting against the FLAG tag.

3.7 Generation of Stable Knockdown Cell Lines

Since our studies involve the interplay between several different factors, we prepared cell lines with several different factors knocked down. We depleted DBC1, HDAC3, EAF1, EAF2, and TRIM28 within 293T as well as HCT116 cells. In order to do this, we first designed several shRNAs against the target proteins. The sequences of the shRNAs used are given in Table 2. The shRNAs were so generated so that they could be cloned into the pLKO.1 vector obtained from Addgene, USA. The p300 knockdown shRNAs were obtained from the source as mentioned [247]. Following the protocol available from the Addgene database, cloned plasmids containing the shRNAs of interest were transfected into HEK-293T cells along with the packaging plasmid psPAX2 and the VSV-G envelope plasmid pMD2.G (both purchased from Addgene, USA) in the ratio 4:3:1. After removing and replenishing the media after 24 hours, cells were incubated at 37°C and 5% CO₂ for a further 48 hours, following which the supernatant containing the lentiviral particles were collected and stored aseptically at -80°C. For stably knocking down target factors in HEK-293T as well as HCT116 cells, a similar protocol was followed. 400 µl of viral supernatant was used to transduce cells in the presence of 8 µg/ml Hexadimethrine bromide a.k.a Polybrene (Sigma, USA). 16 hours post-transduction, the polybrene was removed and cells were subjected to selection in the presence of 3 µg/ml puromycin (Thermo Fisher, USA) after another 24-hour gap. Cells resistant to puromycin were grown till they attained confluency, following which they were passaged in puromycin-containing media at least another two times to ensure the selection of a pure population of stable cells. After this, the knockdown of specific factors was checked at the protein

level by SDS-PAGE followed by immunoblotting, keeping β -Actin as a control. The mRNA levels of the target factors were also determined using qRT-PCR, using the expression of 18s rRNA as an internal control. Values obtained were normalized to protein/RNA levels in control scramble knockdown cells to calculate the knockdown efficiency.

Table 2: Oligonucleotides used for generating shRNA constructs

Target	Upper oligo	Lower oligo
DBC1 sh#1	CCGGGCCAAAGGAAAGGATCTCTTTCTCG AGAAAGAGATCCTTTCCTTTGGCTTTTGTG	AATTCAAAAGCCAAAGGAAAGGATCTCT TTCTCGAGAAAGAGATCCTTTCCTTTGGC
DBC1 sh#2	CCGGGCATTGATTTGAGCGGCTGTACTCG AGTACAGCCGCTCAAATCAATGCTTTTTG	AATTCAAAAGCATTGATTTGAGCGGCTGT ACTCGAGTACAGCCGCTCAAATCAATGC
HDAC3 sh#1	CCGGAAGATGCTGAACCATGCACCTCTCG AGAGGTGCATGGTTCAGCATCTTTTTT	AATTAAAAAAGATGCTGAACCATGCACCT CTCGAGAGGTGCATGGTTCAGCATCTT
HDAC3 sh#2	CCGGCAAGAGTCTTAATGCCTTCAACTCG AGTTGAAGGCATTAAGACTCTTGTTTT	AATTAAAAACAAGAGTCTTAATGCCTTCAA CTCGAGTTGAAGGCATTAAGACTCTTG
P300 sh#1	Wang et al., (2011) <i>Science</i>	
P300 sh#2	Wang et al., (2011) <i>Science</i>	

3.8 Mass Spectrometry Analysis of DBC1 associated proteins

The study that showed the identification of different ELL-associated partners through a stable overexpressing cell line, which led to the putative identification of DBC1 as an ELL interacting protein, has been published elsewhere [146]. In order to perform a reciprocal mass spectrometry analysis to test if ELL could be identified as a DBC1-associated partner, nuclear protein extract was obtained from FLAG-HA-DBC1 stable expressing cells. After immunoprecipitation using anti-FLAG affinity beads (Sigma, USA) and stringent washing with high salt buffer, the obtained protein extract was resolved on a 4-12% NuPAGE Bis-Tris protein gel (ThermoFisher, USA). Following this, the gel was stained using Coomassie Brilliant Blue G-250 (Sigma, USA). The target protein bands were excised carefully with a scalpel and destained with 20% methanol. The cysteine residues were reduced and then alkylated with Iodoacetamide (IAA). The proteins were digested overnight with Trypsin and resolved nanocapillary reverse phase column using a 1% acetic acid/acetonitrile gradient at a 400 nl/min rate of flow. The resolved peptides were introduced into a linear ion-trap mass spectrometer (LTQ XL,

ThermoFisher, USA). MS/MS spectra were collected and proteins were identified by comparing the data against the Human IPI database (v 3.41) using the X1Tandem/Trans-Proteomic Pipeline (TPP) software suite.

3.9 Nuclear Extract Preparation from Plated Cells

For the purpose of identifying ELL and DBC1 associated proteins by mass spectrometry, as well as for subsequent immunoprecipitation analyses and western blotting, we have used protein extracts from the nuclear fraction of HEK-293T cells. In order to obtain these extracts, cells were harvested in cold PBS and centrifuged at 800 $\times g$ for 5 minutes. The packed cell volume (PCV) was estimated and two volumes of the PCV (2x PCV) of Nuclear Extraction Buffer 1 (NE1; 10mM Tris-HCl pH7.3, 1.5mM MgCl₂, 10mM NaCl, and 0.7 μ l/ml β -mercaptoethanol) added to the cells after discarding the PBS. As NE1 is hypotonic, cells were allowed to swell up in ice for about 15 minutes. After this, they were taken up in a 2ml sterile syringe with a 23 gauge needle pre-washed with NE1, and forcefully expelled out rapidly into a micro-centrifuge tube. Cells were aspirated again and released out at least 7 more times in order to lyse them. The lysate was then centrifuged at 6000 $\times g$ for 5 minutes in a refrigerated centrifuge. Following this, the supernatant containing the cytoplasmic extract was collected and discarded. The pellet containing the nuclei was used in the subsequent steps. First, the volume of the pellet was estimated and marked as Nuclear Pellet Volume (NPV). This was resuspended in 0.5 volumes of NE Buffer 2 (20mM Tris-Cl pH 7.3, 1.5mM MgCl₂, 20mM NaCl, 0.2mM EDTA, 25% Glycerol, 0.7 μ l/ml β -Mercaptoethanol and Protease inhibitor Cocktails (PIC) (Roche, USA)). An equal volume of NE Buffer 3 (20mM Tris-Cl pH 7.3, 1.5mM MgCl₂, 1.2M NaCl, 0.2mM EDTA, 25% Glycerol, 0.7 μ l/ml β -Mercaptoethanol and PIC) was added to the resuspended cells following which they were immediately vortexed. The cell pellet-buffer mixture was vortexed after every 3 minutes for about 45 minutes. In the interim, the homogenates were maintained on ice. After intermittent vortexing for about 15 times, the lysate was subjected to centrifugation at 12000 $\times g$ at 4°C for 20 minutes. The supernatant thus obtained constitutes the nuclear extract, which was either used immediately or frozen and stored at -80°C for future use.

3.10 Immunoprecipitation Analyses

A major part of our studies involved the identification of binding partners of ELL, DBC1, HDACs1-8, EAF1, EAF2, and TRIM28 for our detailed mechanistic analyses. For this, we used co-immunoprecipitation analyses to identify target factor-associated proteins. When experiments were

conducted using ectopically expressed proteins, these were first transfected in HEK-293T cells using the method described previously. Then, the cells were harvested in PBS, and lysed in BC₁₅₀ buffer (20mM Tris-Cl, pH 8.0, 150mM KCl, 2mM EDTA, 20% Glycerol, 0.1% NP40, 4µl/ml PMSF, 0.7 µl/ml β-Mercaptoethanol) for 1.5 hours at 4°C. The collected supernatant, after centrifugation at the highest speed for 20 minutes, was subjected to affinity-based immunoprecipitation using either α-FLAG or α-HA tag agarose beads (Sigma, USA) as per the plasmid constructs used in the experimental setup. For the purpose of purification of proteins from mammalian cells, cell lysis, as well as immunoprecipitation, was performed in BC₁₀₀₀ buffer (20mM Tris-Cl, pH 8.0, 1M KCl, 2mM EDTA, 20% Glycerol, 0.1% NP40, 4µl/ml PMSF, 0.7 µl/ml β-Mercaptoethanol). Similarly, for studying post-translational modifications of the target proteins, we performed the immunoprecipitation step in R₁₀₀₀buffer (50mM Tris-Cl (pH 8), 1M NaCl, 1% Triton X-100, 0.1% SDS, 4 µl/ml PMSF and 0.7 µl/ml β-Mercaptoethanol). After incubation of the cell extracts with the agarose beads overnight, the tubes were centrifuged briefly at 3000 RPM, and the beads containing bound proteins were washed thrice with wash buffer, and finally eluted by boiled in 1X SDS loading dye at 95°C for 10 minutes. In case of purifying active proteins, the bound proteins were eluted from the beads by incubation with 3X FLAG-peptide (Sigma, USA). Subsequent assays were carried out as necessary.

For the purpose of immunoprecipitating endogenous proteins, the protocol followed was a bit different. In this case, first, whole-cell lysate extraction/ nuclear extract prep was carried out. After this, approximately 500µg protein lysate was subjected to pre-clearing using magnetic Protein-G Agarose Beads (BioRad, USA). Simultaneously, in a separate micro-centrifuge tube Protein-G Agarose Beads were blocked for 1 hour with 1% BSA, following which they were washed thrice with BC₁₀₀ buffer (20mM Tris-Cl, pH 8.0, 100mM KCl, 2mM EDTA, 20% Glycerol, 0.1% NP40). After this, the beads were then incubated with 2µg primary antibody. The pre-cleared lysate was then added to the antibody-conjugated beads, and incubated overnight. The next day, the beads were centrifuged at 3000 RPM for 1 minute, and the supernatant containing unbound proteins was discarded. The beads were then washed thrice with BC₁₀₀ buffer + 0.1% NP40 and bound proteins eluted in 1X SDS loading buffer by heating at 95°C for 10 minutes. Immunoprecipitated proteins were subsequently separated by SDS-PAGE and detected by western blotting.

3.11 Luciferase based Reporter Gene assay

The Luciferase assay is a simple system to measure gene expression at the transcriptional level. In this assay, a reporter gene is used to track the effect of a target gene on its transcriptional regulation. In most cases, the luciferase gene from the firefly, *Photinus pyralis* is used as the reporter. This gene

yields a protein, which oxidizes D-luciferin in the presence of ATP, oxygen, and Mg^{2+} . The resultant fluorescence can be measured and quantified. In our assays, we have used an in-house generated cell line named G-293T. G-293T is a HEK-293T-based cell line, where the firefly luciferase gene is cloned downstream of a GAL4 activation sequence, and this entire sequence is chromosomally integrated into the genome. Of note, this cell line does not contain any activators of the GAL4 reporter, and thus, the readout obtained is due to the effect of the transfected plasmids themselves. In our assays, we have tried to understand the effect of DBC1 and ELL on transcriptional regulation. For this, we have transfected G-293T cells with plasmids expressing DBC1 and ELL in different combinations along with a GAL4-VP16 activator. 48 hours after transfection, cells were harvested and processed as per the process recommended in the Dual-Glo[®] Luciferase Assay System (Promega, USA). Cells were lysed in Passive Lysis Buffer (PLB), following which the lysate was mixed with Luciferase Assay Reagent II (LAR II). After incubation, luciferase activity was measured using the GloMax[®] 20/20 Luminometer (Promega, USA).

3.12 Cycloheximide Chase Assay

Analysis of protein levels at different time points provides insight into their regulation and how different stimuli affect the turnover rate of different proteins. For example, inhibition of a regulatory pathway, such as proteasome-mediated degradation, and the measurement of steady-state levels of proteins herewith, may help in studying how these pathways affect protein levels. However, as the use of these assays does not provide a dynamic measurement of protein turnover rates, in this context, a cycloheximide chase assay provides an easy, convenient and reliable way to study protein regulation. Cycloheximide (CHX) is a chemical compound isolated from the bacterium *Streptomyces griseus*, which inhibits translation by impaired translocation of the ribosomes. Thus, the addition of this compound prevents new proteins from being synthesized, whereas the regular degradation processes continue, enabling a detailed study into this process. In our assays, we have tried to understand the degradation kinetics of ELL and also delved deeper into its regulation by understanding its interplay with EAF1/2. In case of endogenous proteins, $\sim 1 \times 10^6$ HEK-293T cells were seeded into wells of a 6 well plate, and cycloheximide was added at a concentration of 100 $\mu\text{g/ml}$. Cells were incubated for indicated time points and then harvested in ice-cold PBS. For studying the turnover rate of exogenously expressed proteins, 0.6×10^6 HEK-293T cells were seeded into each well of a 6 well plate and transfected with indicated plasmids as necessary in the experiment. 24 hours post-transfection, as before, 100 $\mu\text{g/ml}$ cycloheximide was added to each well, and cells were incubated for pre-determined time points. After harvesting the cells in PBS, they were lysed using RIPA buffer

(along with PICs, PMSF, and β -Mercaptoethanol), and protein levels were measured by SDS-PAGE followed by western blotting with respective antibodies.

3.13 Purification of recombinant proteins

Several of our assays involved the use of purified recombinant proteins for use in *in vitro* experiments. For this, we have expressed and purified several proteins from bacterial as well as mammalian cells. As an example, for our direct interaction analyses, we purified FLAG-DBC1, FLAG-HDAC3, FLAG-AFF1, FLAG-AF9, FLAG-ELL, FLAG-CDK9, and FLAG-HA-EAF2 from HEK-293T cells. 48 hours after transfection of the indicated mammalian expression plasmids, cells were lysed in BC₃₀₀ buffer (20mM Tris-Cl, pH 8.0, 300mM KCl, 2mM EDTA, 20% Glycerol, 0.1% NP40, 4 μ l/ml PMSF, 0.7 μ l/ml β -Mercaptoethanol) for 1.5 hours at 4°C. Following this, the lysate was incubated with pre-cleared anti-FLAG M2 magnetic beads (Sigma, USA) at 4°C for 16 hours. After incubation, the beads were washed thrice with wash buffer (BC₃₀₀ buffer + 0.1% NP40). For elution of the purified proteins when required, the FLAG beads-bound proteins were incubated with 3X FLAG peptide for 90 minutes.

Several of the proteins used in our analyses have been purified via bacterial expression systems as well. GST-tagging of proteins was done by cloning the necessary ORFs of ELL, EAF1, and TRIM28 in the pET-GST vector. For generating His tagged proteins, we cloned the p300 HAT domain, EAF1, and ELL in 6xHis pET-11d vector or its derivatives. The plasmids were transformed into *E. coli* BL21(DE3) cells and a single colony was grown overnight in LB broth. From this overnight culture, a secondary culture was inoculated and grown at 37°C with continuous shaking at 250 RPM till the OD reached 0.4. Subsequently, cells were induced with 1 mM IPTG (Goldbio, USA), and grown for a further 4 hours. After harvesting, cells were resuspended in 0.1 culture volumes of lysis buffer (BC300: 20mM Tris-Cl, pH 8.0, 300mM KCl, 2mM EDTA, 20% Glycerol, 0.1% NP40, 4 μ l/ml PMSF, 0.7 μ l/ml β -Mercaptoethanol) and lysed by sonication for 10 cycles, at 60% amplitude, with alternating 30 second on/off phases. For purification of His-tagged proteins, the buffer used comprised 50 mM Na₂HPO₄, 300 mM NaCl, 10 mM Imidazole, 20% Glycerol, PICs, and PMSF. The supernatant was incubated with either anti-GST agarose beads (Thermo Scientific, USA) or Ni-NTA agarose beads (Genetix Biotech Asia, India) for 4 hours at 4°C. Bead bound proteins were subsequently washed thrice with GST wash buffer (BC₃₀₀ buffer + 0.1% NP40) or His-wash buffer (50 mM Na₂HPO₄, 300 mM NaCl, 20 mM Imidazole, 20% Glycerol, 0.1 % NP40). When elution was necessary, GST-tagged proteins were eluted using GST-elution buffer (30mM reduced L-Glutathione (Sigma, USA) in 100mM Tris-Cl, pH 8) by incubating for 2 hours at 4°C. For His-tagged proteins, elution was carried out in His-elution buffer (50 mM Na₂HPO₄, 300 mM NaCl, 250 mM Imidazole, 20% Glycerol, 0.1 % NP40).

In case of GST-TRIM28, we consistently observed the protein localizing to the inclusion bodies in the bacterial cells. In the event of such an occurrence, we used a protocol originally used for isolating histones, tailored to suit our needs. Briefly, the harvested pellet after induction was resuspended in 0.02 culture volumes of Wash/Lysis (WL) buffer (50 mM Sodium phosphate pH 8.0, 100 mM NaCl, 5 mM Imidazole, 10% Glycerol). The solution was then sonicated using 30-second cycles at 60% amplitude for 5 minutes (70-80% power, 90% duty cycle using 20-second bursts for 2 minutes) or until the solution was no longer viscous. After centrifugation at 18000g for 15 minutes, the supernatant was discarded and the pellet containing the inclusion bodies was resuspended in 1.5 ml of WL buffer and 500 μ l of glass beads (Himedia, India) by vigorous vortexing. After the addition of 9 ml WLT buffer (WL buffer + 1% Triton-X), the solution was vortexed again and the pellet was washed with WL buffer after centrifugation. The pellet was then resuspended in 250 μ l DMSO and 12 ml freshly prepared D500 buffer (6M Urea, 500 mM NaCl, 50 mM Tris-Cl pH 8.0, 10 mM β -Mercaptoethanol, and 10% Glycerol), and rotated at 4°C for at least 1 hour. The supernatant collected after centrifugation contained the protein lysate, which was incubated with pre-washed GST agarose beads. Bound proteins were washed once with D500 buffer and twice with D1000 buffer (6M Urea, 1M NaCl, 50 mM Tris-Cl pH 8.0, 10 mM β -Mercaptoethanol, and 10% Glycerol) to remove unbound proteins and contaminants. The bound protein was then eluted and subsequently dialyzed at decreasing salt concentrations. Finally, it was concentrated using a centricon (Millipore, USA).

3.14 Western Blotting

For a major part of our experiments, in order to detect the presence of proteins or their expression levels, we have employed western blot analyses to achieve our aim. Western blotting involves first, the boiling of the protein lysate in a reducing buffer (Laemmli buffer, 5X: 312.5 mM Tris, 250 mM SDS, 50% glycerol, 25% β -Mercaptoethanol, and Bromophenol blue as an indicator), and then resolving proteins according to their size by sodium dodecyl sulphate- polyacrylamide gel electrophoresis (SDS-PAGE). Once the proteins were resolved, they were transferred onto a nitrocellulose membrane (BioRad, USA). The membrane containing the immobilized proteins was then blocked for 1 hour at room temperature in 5% non-fat dry milk (Himedia, India), dissolved in PBST (Phosphate buffered saline + 0.1% Tween-20). Specific portions of the membrane corresponding to the molecular mass of the proteins of choice were then incubated overnight in primary antibody diluted in PBST. A list of antibodies used has been provided in Table 3. The following day, membranes were washed thrice with PBST and then incubated with either rabbit or mouse HRP conjugated secondary antibodies (depending on the source of the primary antibodies used) for 1 hour at room temperature. After three

more washes with PBST, membranes were incubated in appropriate volumes of ECL western blotting substrate (BioRad, USA), and the chemiluminescent signal was detected by exposure to X-Ray films or in a chemi-documentation system (Azure Biosystems, USA or iBright Imaging System, Thermo Fisher, USA). ImageJ was used to quantify the chemiluminescent signal on the blots in all our experiments.

Table 3. Antibodies used for Immunoblots and Immunoprecipitation Assays

Antibody	Source	Catalogue No.
Acetylated Lysine	Cell Signaling Technology	9441S
AF9	Bethyl Laboratories	A300-595A
AFF1	Abcam	ab196788
AFF4	Abcam	ab103586
CCNT1	SantaCruz Biotechnology	sc-10750
CDK9	SantaCruz Biotechnology Cell Signaling Technology	sc-13130 2316S
DBC1	Bethyl Laboratories Cell Signaling Technology	A300-433A 5693S
EAF1	SantaCruz Biotechnology	sc-373832
EAF2	Cell Signaling Technology	14159S
ELL	Bethyl Laboratories Cell Signaling Technology	A301-644A 14468S
FLAG epitope	Sigma	F7425
GAL4	SantaCruz Biotechnology	sc-510
GFP	SantaCruz Biotechnology BioBharati Life Science	sc-8334 BB-AB0065
GST tag	SantaCruz Biotechnology	sc-53909
HA epitope	Raised from clone 12CA5	-
HDAC3	Cell Signaling Technology	85057S
His epitope	SantaCruz Biotechnology	sc-8036

HRP conjugated secondary antibody (α -Rabbit)	BioRad Laboratories	1706515
HRP conjugated secondary antibody (α -Mouse)	Cell Signaling Technology	7076S
Myc epitope	Cell Signaling Technology	sc-40
Normal Mouse IgG	Cell Signaling Technology	5415S
Normal Rabbit IgG	Cell Signaling Technology	2729S
P300	Cell Signaling Technology	86377S
Phospho Rpb1 CTD (Ser2)	Cell Signaling Technology	13499S
Phospho Rpb1 CTD (Ser5)	Cell Signaling Technology	13523S
Rpb1 CTD (4H8 clone)	Cell Signaling Technology	629S
TAF1	Cell Signaling Technology	12781S
TAF5	Bethyl Laboratories	A303-686A
TAF6	Bethyl Laboratories	A301-276A
TAF7	Generated in house	-
TBP	Cell Signaling Technology	8515S
TRIM28	Cell Signaling Technology	4124S
Ubiquitin	Cell Signaling Technology	3933S
β -Actin	SantaCruz Biotechnology Biobharati Life Science	sc-47778 BB-AB0024

3.15 *In vitro* interaction assays

In our studies, often we have used *in vitro* interaction assays to demonstrate a direct interaction between two or more proteins, as well as to study competitive interactions between a group of proteins. Our experimental setup was such that in most cases, GST-tagged proteins were kept immobilized on GST-agarose beads. These include GST-ELL and its deletion mutants, GST-EAF1 and its

deletion mutants, as well as GST-TRIM28. These served as the bait proteins. The prey proteins were then added to the reaction mix, either individually or in specific combinations as per the experimental set-up. The proteins were incubated in a binding buffer (20mM Tris-Cl, pH 8.0, 100mM KCl, 2mM EDTA, 20% Glycerol, 0.1% NP40, 20 ng/μl BSA) overnight at 4°C. The next day, the supernatant was carefully removed and the beads were washed thrice in wash buffer (BC₁₀₀ + 0.1% NP40). Bound proteins were then eluted by boiling the beads in 1X Laemmli buffer for 10 minutes at 95°C. These were then resolved by SDS-PAGE and detected by western blotting using factor/tag-specific antibodies.

3.16 *In vitro* Acetylation Assay

The ability of p300 to serve as the *bonafide* acetyltransferase for ELL, and the identification of the region on ELL acetylated by p300 were determined by *in vitro* acetylation assays. First, GST-tagged ELL and its deletion mutants were immobilized on GST-agarose beads. Following this, an equal amount of ELL or its domains were incubated with specific concentrations of the p300 HAT domain, in presence of a HAT buffer (75mM Tris-Cl (pH 8.0), 1.25mM EDTA, 12.5mM DTT, 0.25% Tween20, 25% glycerol). Acetyl Co-A served as the source for acetyl groups in the experimental setup. The tubes containing the reaction mix were kept at 30°C and tapped intermittently to ensure uniform mixing. After 2 hours, the reaction was stopped by adding Laemmli buffer and boiling the mixture for 10 minutes at 95°C. The proteins resolved by SDS-PAGE were incubated with pan-acetylated lysine-specific antibody to detect acetylation signals.

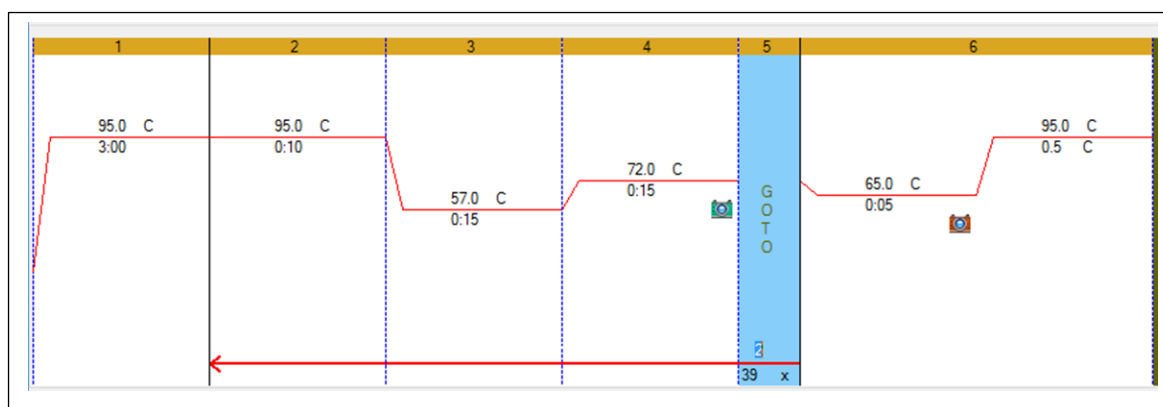
3.17 *In vitro* Deacetylation Assay

To confirm HDAC3 as the deacetylase for ELL, we first purified HDAC3 after its overexpression in 293T cells. Subsequently, an *in vitro* acetylation assay was carried out by the method described above, using GST-ELL and p300 HAT, but kept for 1 hour at 30°C. Then, the GST beads containing the immobilized and acetylated ELL were washed thrice with BC₃₀₀ + 0.1% NP40, so that other proteins were removed. This acetylated ELL was then used as the substrate for the deacetylation assay. In this, the acetylated ELL was incubated with HDAC3 in HDAC buffer (50mM Tris [pH 8.0], 4mM MgCl₂, 0.2mM DTT), and rotated on a rotor for 3 hours at 37°C. The reaction was stopped by adding 1X Laemmli buffer and boiling the samples at 95°C for 10 minutes. The samples were then resolved by SDS-PAGE and probed by acetylated lysine antibody to check for the deacetylation of ELL in the presence of HDAC3.

3.18 RNA Extraction and Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

In several of our experiments, we have sought to understand how the overexpression or depletion of a certain target protein affects the expression of another protein. For example, we have examined if the overexpression of DBC1 or HDAC3 affects ELL protein or RNA levels. Similarly, we have also tested for the basal level expression of several target genes of ELL upon DBC1 knockdown, and how the mRNA levels of these genes are affected upon induction under specific conditions. For studying the mRNA levels of target genes, first, cells were grown in wells of a 6 well plate till they reached confluency. $\sim 1 \times 10^6$ cells were harvested from each well in ice-cold PBS and washed twice. RNA extraction was done using TRIzol reagent (Thermo Fisher Scientific, USA) as per the manufacturer's protocol. Briefly, 800 μ l TRIzol was added to each cell sample and mixed well by pipetting. The samples were left at room temperature for 5 minutes for dissociation of the nucleoprotein complexes. After this, 160 μ l chloroform (Sigma, USA) was added to each tube, and the mixture was mixed by vigorous shaking. The tubes were then left to stand for about 5 minutes at room temperature and then centrifuged at 12000 $\times g$ for 15 minutes at 4°C. The upper transparent aqueous phase was then collected carefully in a new microcentrifuge tube without disturbing the lower layers. Now, 400 μ l 2-propanol (Sigma, USA) was added to each tube and again mixed by shaking. This mixture was again subjected to centrifugation at 12000 $\times g$ for 15 minutes at 4°C. After this step, RNA could be seen as a small pellet sticking to the lower part of the tube. This RNA pellet was washed with 75% Ethanol (Merck, USA), and the solution centrifuged at 7500 $\times g$ for 10 minutes at 4°C. The supernatant was subsequently discarded and the RNA pellet was air-dried for 10-15 minutes at room temperature. After drying, the pellet was dissolved in UltraPure™ DNase/RNase-Free Distilled Water (Thermo Fisher Scientific, USA) and stored at -80°C for long-term use.

1 μ g of each RNA sample was used as a template to synthesize cDNA using the Verso cDNA Synthesis Kit (Thermo Fisher Scientific, USA). RNA was mixed with 5X RT Buffer, anchored Oligo dT, 20 mM dNTPs, RT Enhancer, and RT enzyme, mixed well, and incubated at 45°C for 1 hour. The reaction was terminated by heating at 95°C for 2 minutes, and the cDNA obtained was diluted 25 times and then used as a template for the qRT-PCR reaction. In the qRT-PCR reaction mix, 2X iTaq Universal SYBR Green Supermix (BioRad, USA) was used as the reporter dye, along with 3 μ M of each primer and the diluted cDNA. The cycling was carried out in a CFX96 Touch Real-Time PCR Detection System (BioRad, USA) as per the following protocol:



The expression level of the target genes was calculated after normalizing with 18s rRNA levels as specified. The primers used in the qRT-PCR analyses are given in Table 4.

Table 4. Primers used for qRT-PCR analyses for mRNA expression of target genes

Gene	Forward Sequence	Reverse Sequence
AFF1	GTCATCCAAGTCAGCTTAC	CATGTTCAAAATGGACTGG
AF9	GAGAGAAAGACACATTCTGC	GGATGTTCCAGATGTTTCC
APOE	GCGGATGGAGGAGATG	CTCGAACCAGCTCTTG
ATF2	GTACCAGGCCCATTTCTCTTC	GAACGAGTGGGACTGCAGCTG
ATF6B	CATCGTTCCCGCTCCTATGCCT	CTTCTTTCTCCGGGACTGGCAG
BCL6	CCACACAGGAGAGAAACCTTAC	GCAGGTCAGTGGCTGACAC
CAMK2A	GAAGTGCTGCGGAAGGACCC	CAGGCGGTGCTGGTCCTCATC
CAMK2D	TATCAACAATGGGGACTTTG	GTAGAATCGGTGAAAATCC
CAMK2G	GATGAGGATCAGCACAAGCTG	GGTCAGCATCTGGTTGATCAAG
CCND1	TCTAAGATGAAGGAGACCAT	GGAAGTGTTCAATGAAATCG
CCNE2	GCTGGCCTATGTGACTTACC	GGTAGACTTCGATGGGCCC
CDKN1A	GGACAGCAGAGGAAGACCATG	CTGTCATGCTGGTCTGCC
CDKN1B	CGACGATTCTTCTACTCAA	TTACGTTTGACGTCTTCTG
CDKN2C	GGATTGGAAGGACTGCG	GCTTGAAACTCCAGCAAAGTC

CDK6	GGAGTGTGGCTGCATATTTG	CGATATCTGTTACAACTTC
CDK9	GCATCATGGCAGAGATGTG	GTTGTCCACGTTTGGCC
CNN3	CTACAGAACCTGTCATTC	ATCTCTGGGGTAGTCATC
CREB1	GATTCACAGGAGTCAGTGGATAG	GCACTCCTGGTGCATCAGAAG
CREB5	GAAGGCTGCATTGACTCACCAC	GTGCATGTGATGGTGTGGCGTC
DBC1	GAGGAGTTTGCAGGAGC	GTAGCCACACCAGTTGG
DCAF4	AAACCTCTACTTCACCAACC	CTGGGTGACTATTGACGAA
DDB2	TACCCCTTATGAATTGAGG	CAGATGAGAATGTGGTAAC
EAH1	CAGCATCTATAGACACTTCC	GCACATATTCACCAGTGTC
EAH2	GCATCTCCAATAGATGATATCG	GTATCAGAAGTAGAGGATTTGC
ELL	GACCAACACCAACTACAGCCAGG	GTAATCGGCGATGAGCCTCTTG
GLP1R	GCATCCTTCATCCTGCGAGCAT	GCATGAGCAGAAACACCAGGCG
GLUT1	GAAGTGTACCCACAGCCCTTC	GGGCCACAGGTCCTTGTTGC
HDAC3	CATGACGGTGTCTTCCAC	GGCACGTTTACAGACAGTAGTAG
KIF1A	AGAAGGTGAGCAAAATCAG	CGGAATGAAATCTGTCTTC
KIF1B	AATCAGAGTGACTTTTCGTC	ATTCGACTGATTTCTTCTGG
MDM2	TTGGATCAGGATTCAGTTTC	GAGAGTTCTTGTCTTCTTC
MED12	CAGGATGAACAACGCGAG	GAGTGCCTCATGCATAAGC
MED15	GACCTTTAAACACACCTGTG	GGCTCTTCATCTTACTCAGG
MEF2D	GACCTGAACAGTGCTAACG	GTGATGACTCGCAGGTC
MYO10	CTTCAGGATGAGGCCATC	GCCAGCTGTACAGGTTG
NAV2	GTAATTAACACACCTGAGC	TCTTCTTGAGTCACTCTC
NEDD9	CCCATCCAGATACCAAAGGAC	GTAGGCGGGATGTCATACAC
NFKB2	GAGAACGGAGACACACC	CAGAAAGCTCACCACAC

PEPCK	GGTACCTCTGCCACCACCAATC	CACCAGGTGGAAGAGGCTGGTC
PCNA	GGAGGAAGCTGTTACCATAGAG	CCTCGATCTTGGGAGCCAAG
PRKCA	CAGGAAGGGCACAGAAGAACTG	GAACGGGGGTTTGTCAAGCAG
PRKCB	GAATTGGAAGGTTCAGTGAGCC	CCTGGATATAGCCTTGATGGTC
PTPRF	CAATGCCAGCTTCCTGGATG	ATGATGGTGGAAATTGTGCTC
SIAH1	CTGCTTTGACTATGTGTTACCGC	ACTGAATTAGCCACTTTCTCCAT
SKI	CCAAGTACTCGGCCAGATC	GCACGGAATCTACGGCTCC
TAF1	GATGAAGAAGGAGACAATCC	GCTATCCTCCAAACCATG
TAF3	CAGCGTGGTGACTGAG	GTCATGATTCCAACACAGG
TAF5	CTGGGACGTCCTGAATG	CCTAAGTGAACAGACTGTATC
TAF6	CTTGCTGAAGGTTCTG	GAAGTGGTGGTGGAAAC
TAF7	CCAGGAATGTCTGGTCAC	GTCCTGTAGCTGTCTCTC
TAF9	CAGTTACTAGCAGACCAAG	GCAGGAATTGAAGCTTTTAC
TAF12	ATAGTACTGCAGTGGTAAAG	TCTCGATAAAATCATCAGCA
VAMP2	GAGACTGCAGCAGACCCAGG	AGTGCATCTGCACGGTCGTC
VDAC1	CAAGTATCAGATTGACCCTG	GTCAGTTTAATACCTGGC
VPS29	GCATATAATGCCTTGAAAC	CACATCATCTCCAATTAGC
WEE1	GCTGTCCGCTTCTAGAAAG	CGAGATGTTCTATTACTCTGGG
18s rRNA	GTAACCCGTTGAACCCATT	CCATCCAATCGGTAGTAGCG

3.19 Chromatin Immunoprecipitation Assay

Recruitment of different factors such as ELL, RNA Pol II, CDK9, and TBP was tested at the transcription start site (TSS) of different target genes under different conditions by chromatin immunoprecipitation assays. In this assay, cells were first cross-linked by the addition of 1% formaldehyde and incubated for 10 minutes with gentle swirling. The cells were then washed with ice-cold PBS and the reaction

was quenched by adding 125 mM glycine (in PBS) and again incubating for 5 minutes at room temperature. The cells were washed again with PBS and then collected in a micro-centrifuge tube. The cells were lysed in nuclear extraction buffer (20mM Tris (pH 7.5), 300mM NaCl, 0.5% NP40, 1% Triton X-100, 2mM EDTA and protease inhibitor cocktail (PIC)). The pellet obtained after centrifugation, containing the nuclei were then resuspended in shearing buffer (50mM Tris (pH 8.1), 1% SDS, 10mM EDTA, and PICs) and sonicated using a UCD-200 Bioruptor (Diagenode Inc, USA), for 30 minutes with 30 second on/30 second off pulses. The sonicated chromatin was checked on an agarose gel to verify if the sonicated fragment lengths ranged between 200-500 bp. After this, the sonicated samples were pre-cleared with Protein A agarose beads for 1 hour at 4°C. The pre-cleared protein lysate was then diluted 10 times with ChIP dilution buffer (20mM Tris-Cl (pH 8.0), 1.1mM EDTA, 167mM NaCl, 0.01% SDS, 1.1% Triton X-100) and incubated overnight with 2 µg of each primary antibody at 4°C. The next day, the antibody-protein complexes were incubated with protein G agarose beads for 1 hour. The beads containing the antibody-bound protein complexes were washed first with Low Salt buffer (20mM Tris-Cl (pH 8.0), 150mM NaCl, 2mM EDTA, 0.1% SDS, 1% TritonX-100 and PIC), and then with High Salt Wash Buffer (20mM Tris-Cl (pH 8.0), 500mM NaCl, 2mM EDTA, 0.1% SDS, 1% TritonX-100 and PIC). The third wash was performed with LiCl wash buffer containing 20mM TrisCl(pH8.0), 1mM EDTA, 1% deoxycholate, 0.5M LiCl, and 1% NP40. Bead-bound DNA complexes were then eluted using elution buffer (1% SDS, 0.1M NaHCO₃) for 30 minutes. The protein-DNA crosslinks were reversed by incubating the samples at 65°C for 4 hours. Following Proteinase K treatment, the DNA was then purified by the QIAquick PCR purification kit (Qiagen, Germany). The purified DNA was subsequently used as a template for the qRT-PCR assays. The primers used for analyses of factor enrichment at the TSS of the target genes are mentioned in Table 5.

Table 5. Primers used for qRT-PCR for testing factor enrichment after chromatin immunoprecipitation assays at the TSS of target genes

Gene TSS	Forward Sequence	Reverse Sequence
BCL6	CCTCTTATGTCACCGAATAGC	GGTTCCGGTTCGAGGCTC
CCND1	CGGGCTTTGATCTTTGCTTA	CTGCTGCTCGCTGCTACT
GLUT1	GGTCCTATAAACGCTACGGTCC	CGTTTGGTCTCCTGCTCCCTG
KIF1B	TCCTCCGCAGCCGTCAC	CATCACGCAGCAGCCCG
NEDD9	GAAGCACAGAGAGGAAATGA	CTAGAGGCAGCTGTAAGC

PTPRF	TCCAGCTTCGGCTCCGGC	TCCACCTGGTCCCGCTCC
SKI	CAGAAGACGCTGGAGCAGTTCC	CTTGGCGCTCTCCTTCTTGTAG

3.20 Cell Proliferation Assay

To test how the knockdown of DBC1 and EAF1/2 results in phenotypic changes in HEK-293T as well as HCT-116 cells, we employed cell proliferation assays. ~ 6×10^4 cells of each cell line were counted and seeded into each well of a 24 well plate and left to grow at 37°C and 5% CO₂ in a humidified incubator. On the indicated days, cells were harvested and collected in microcentrifuge tubes. They were then stained with Trypan Blue and counted in a Neubauer Haemocytometer. The total cell number was determined as per the formula:

$$\text{No. of cells/ml} = (\text{Total live cell count of all quadrants}/4) \times \text{Dilution factor} \times 10^4$$

Counting of cells was done on indicated days for up to 7 days. The data is represented as the fold increase in the number of cells with respect to the cells present on Day 0. Three replicates were performed for each experiment.

3.21 Clonogenic Assay

The clonogenic or colony formation assay tests how cells are able to survive under different conditions. We have tested in detail how the depletion of EAF1/2 affects the survival of both HEK-293T cells as well as the HCT-116 cancer cell line by this method. Here, we seeded 1.5×10^4 cells into each well of a 6 well plate and allowed them to grow and proliferate for 7-14 days. After the pre-determined time period of the experiment, the media was removed and cells fixed in a Fixing Solution (Methanol: Acetic Acid in a 3: 1 ratio) for 10 minutes at room temperature. The cells were then washed and stained in 0.5% Crystal Violet solution prepared in methanol, for about 15 minutes. The excessive staining solution was then removed by multiple gentle washes with water. Stained colonies which were then visible were photographed and counted.

3.22 Analysis of High Throughput NGS Data and Gene Ontology Analyses

Understanding the physiological significance of the DBC1-ELL axis was very important to determine processes regulated by the expression of these proteins. Our principal aim was to understand how the expression of DBC1 affects the levels of ELL target genes. For this, we conducted an extensive literature survey and identified two individual studies identifying genes regulated by DBC1 and ELL, which were publicly available in the GEO dataset database (GSE35480 and GSE34104)[186, 220]. From these two datasets, genes that were most significantly affected were identified and sorted into two lists of the top 10,000 genes. Overlapping of these two datasets yielded 2987 genes co-regulated by this axis. These genes were initially sorted as per their functions by the Protein Annotation Through Evolutionary Relationship (PANTHER) database [248]. Detailed gene ontology analyses, functional classification, pathways affected, and prediction of human diseases that could be caused by perturbations of these genes was done by using the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 [249].

3.23 Peripheral Blood Mononuclear Cell Isolation

Through our extensive studies, we identified that several pathways in humans were affected by the DBC1-ELL axis, and a wide majority (n= 750, approximately) of the genes were predicted to be involved in the regulation of metabolic processes. Of these, about 75 genes regulated glucose metabolism, and the perturbation of which could possibly lead to Type 2 Diabetes. To verify this hypothesis, we sought to perform gene expression studies on peripheral blood mononuclear cells (PBMCs) isolated from treatment naïve Type 2 Diabetes patients. Upon getting the necessary ethical approvals, ~ 5ml of whole blood was collected from each treatment naïve diabetic patient, after getting their informed consent. The blood was then immediately transferred to BD Vacutainer Ethylenediaminetetraacetic Acid (EDTA) Tubes (BD Biosciences, USA) and processed within 1 hour. HiSep LSM1077 (Himedia, India) was used to isolate PBMCs and used as per the manufacturer's protocol. The solution was added slowly into sterile 15 ml centrifuge tubes. The whole blood collected in the EDTA vials was gently layered onto the HiSep layer. Care was taken to maintain that the blood: HiSep ratio remained equal. The tubes were then centrifuged at 500xg for 30 minutes at room temperature. Care was taken that no brakes were used during the deceleration of the centrifuge. The interface of the plasma and HiSep showed the presence of a buffy coat, containing the mononuclear cells. These were carefully collected and washed thrice with 1X PBS. The cells were then divided equally into two tubes. One was lysed

using RIPA for detecting proteins by SDS-PAGE and western blotting. RNA was extracted from the remaining cells by using TRIzol for studying mRNA levels of target genes by qRT-PCR.

Chapter 4

Results

Part 1

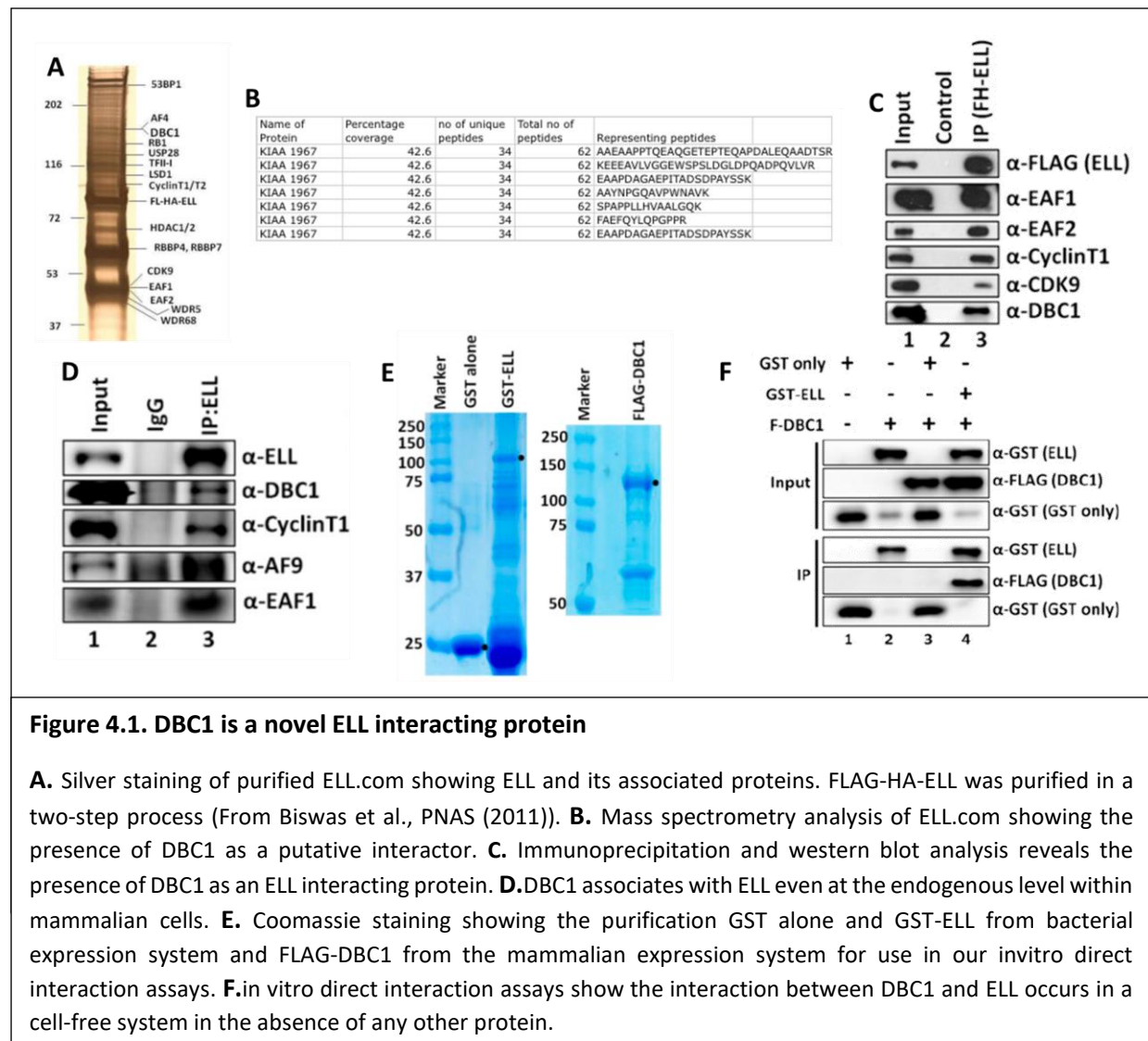
**Interplay between DBC1, HDAC3, p300,
and Siah1 for the maintenance of ELL
stability and functions**

4. Results

4.1 Identification of DBC1 as a novel interactor of ELL

The aim of this study was to understand the functional regulation of the RNA Polymerase II elongation factor, ELL, and thereby study its implications in the context of various human diseases or under different stimuli. For this, we have generated a stable cell line that expresses ELL as FLAG-HA-tagged. Using this cell line, we have purified ELL and its interacting proteins, separated them by SDS-PAGE, and then identified them by mass spectrometry. As these proteins remain associated with ELL even after two rigorous purification steps at high salt concentrations, it may safely be assumed that their interactions are quite specific in nature. In our earlier study, we have observed the presence of several of these interacting proteins using the ProteoSilver™ Silver Stain Kit (Sigma, USA). We reported these observations in detail and also demonstrated that ELL interacts with several components of the Super Elongation Complex (SEC) ([146] and Fig. 4.1A). A few initial studies have shown that other than its role as an RNA Pol II elongation factor, ELL plays a role in the development of *Drosophila* embryos, can act as an E3 ligase for c-Myc, and plays a role in the restart of transcription after DNA repair [173, 174, 184, 190]. As not much is known about the interaction of ELL with components outside the SEC, careful analysis of our ELL.com mass spectrometry data revealed the possible presence of a versatile protein, DBC1, hitherto unknown as an interactor of ELL (Fig. 4.1B). As this was an intriguing observation, we immunoprecipitated FLAG-HA-ELL from the ELL stable cell line using anti-FLAG M2 magnetic beads. Subsequent Western blotting revealed the presence of several known interactors of ELL, such as EAF1, EAF2, CyclinT1, and CDK9, along with the clear presence of DBC1 as an ELL-associated protein (Fig. 4.1C). As these observations were generated from ectopically expressed proteins, we were curious to know and most interested in nullifying the bias that may be created with the overexpression of proteins in different assays. Hence, we decided to explore the interaction between DBC1 and ELL in an endogenous context. For this, we immunoprecipitated endogenous ELL from HEK-293T cells using an ELL-specific antibody, which was conjugated subsequently to protein-G agarose beads to facilitate easy washing out of unbound proteins. Western blot analysis of proteins that could interact with endogenous ELL within HEK-293T cells showed the presence of DBC1 as a stable interactor of ELL along with other known ELL-associated proteins such as EAF1, AF9, and CyclinT1 (Fig. 4.1D, lane 3). Next, we also wished to know whether this interaction between ELL and DBC1 is direct, or if any other proteins form an intermediary between them for their interaction. For this, we first purified GST alone and GST-ELL from expression in a bacterial expression system and FLAG-DBC1 after overexpression in the mammalian cell system. The proteins obtained were quite pure and had minimal contaminants (Fig. 4.1E). Using these proteins, we decided to set up an *in vitro* direct interaction assay. For our convenience, ELL was chosen to be the bait protein and remained bound to the anti-GST agarose

beads, whilst free FLAG-DBC1 was used as the prey. Incubation of these two proteins, followed by extensive washing and then SDS-PAGE followed by immunoblotting revealed that DBC1 is indeed able to interact directly and strongly with ELL even in a cell-free system (Fig. 4.1F, lane 4 of the DBC1 blot in the IP panel). Our results thus point towards the identification of the DBC1 protein as a novel interactor of the RNA polymerase II elongation factor, ELL.



4.2 DBC1 associates with ELL as well as other human SEC components

Our initial experiments led to the identification of DBC1 as a novel interactor of ELL. However, we had not explored the strength of ELL association with DBC1. For this, we first generated FLAG-HA-DBC1 expressing stable cell lines in a HEK-293T background using the Flp-In-TREx expression system. After selection with G418, colonies of cells were grown and their expression checked to identify DBC1 expressing cell lines. We tested cells grown from twelve different colonies, and all of them were found

to have a robust expression of FLAG-HA-DBC1 (Fig. 4.2A). We then induced expression of FLAG-HA-DBC1 from one of these stable lines and immunoprecipitated FLAG-HA-DBC1 by a two-step purification protocol. This stringency was required to ensure the presence of only strong stable interactors of DBC1 in our prep. The first round of purification was carried out using anti-FLAG M2 beads, and the purified protein eluate was then subjected to the second round of immunoprecipitation, this time using anti-HA agarose beads. Subsequently, this two-step purified eluate was resolved on an SDS-PAGE gel and stained by silver staining. Visualization of the eluted proteins in our DBC1.com prep indicated the presence of several previously known as well as novel interactors of DBC1 (Fig. 4.2B). Our next aim was to try and identify some of these proteins that showed a stable and strong association with DBC1. For this, we resolved the proteins on by SDS-PAGE and proceeded to identify certain target interactors by western blotting. Our results, as shown in Fig. 4.2C clearly show that DBC1 interacts quite strongly with ELL. However, when we tested for the presence of other known super elongation complex components in our DBC1.com prep, we were pleasantly surprised to find that along with ELL, DBC1 could also associate with several other SEC components as shown, such as EAF1, EAF2, AFF1, and CDK9. AF9 was also seen to interact with DBC1, albeit the strength of this interaction was quite weak. Subsequently, our next experimental step was to address whether, in a co-immunoprecipitation reaction, we would see DBC1 interaction with ELL. Hence, we co-expressed FLAG-DBC1 with GFP-ELL and after immunoprecipitation with anti-FLAG agarose beads, found that indeed, within cells, DBC1 interacts with ELL (Fig. 4.2D, lane 4 of ELL plot in the IP panel). Next, we also wanted to test the association between these two proteins in a heterologous expression system. In this case, we chose to express our target factors in Sf9 insect cells. Sf9 cells are a common tool to test protein-protein interactions as the expression of any protein in them is possible by the use of baculoviruses. Additionally, the absence of our target proteins in Sf9 eliminates the possibility of non-specific interactions as well. Keeping these above-mentioned factors in mind, we first decided to address the interaction between DBC1 and ELL in these cells. So, we infected the cells with FLAG-DBC1 and His-ELL expressing baculoviruses and performed immunoprecipitation with anti-FLAG beads. Not surprisingly, in this system as well, the interaction between DBC1 and ELL was observed, even at a high salt concentration (300 mM), indicating that these two factors interact stably within mammalian cells (Fig. 4.2E, lane 4 of ELL blot in the IP panel). After having this observation, we now turned our attention to the other SEC factors that were seen to interact with DBC1 previously from the DBC1.com prep (Fig. 4.2C). Hence, we also prepared baculoviruses that could express non-tagged EAF2, CDK9, and AF9 for the purpose of understanding whether they would interact with DBC1 in this heterologous system as well. Our results clearly show that all three chosen proteins, i.e EAF2, CDK9, and AF9 interact stably and strongly with DBC1 (Fig.

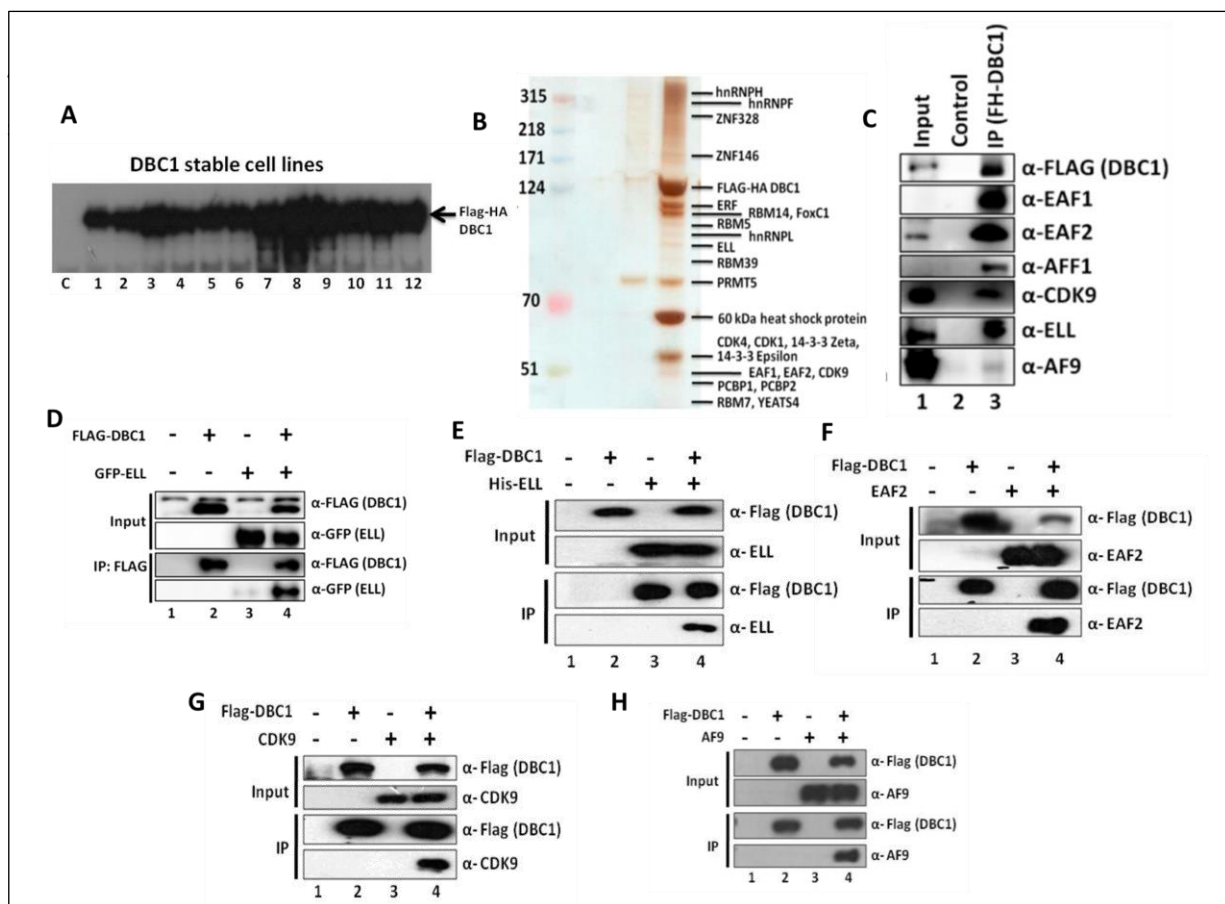


Fig 4.2 DBC1 forms stable associations with several SEC components

A. Western blots showing the expression of FLAG-HA-DBC1 in generated stable cell lines. **B.** Resolution of DBC1.com prep and staining by silver stain shows the presence of several novel proteins as interactors of DBC1. **C.** Immunoprecipitation and western blotting confirms that ELL interacts with DBC1 within mammalian cells. Additionally, several SEC components are also seen to associate with DBC1. **D.** Co-immunoprecipitation and subsequent immunoblotting analyses show that ELL interacts with DBC1 within mammalian cells. **E.** DBC1 and ELL were also seen to associate with each other under high salt conditions when tested in a heterologous expression based system (Sf9 cells). Interestingly, several other SEC components are also seen to interact with DBC1, such as **F.** EAF2, **G.** CDK9, and **H.** AF9

4.3 DBC1 and ELL interaction occurs through specific structural domains

We have observed that a direct stable interaction exists between DBC1 and ELL within mammalian cells. This interaction is seen not only in the case of ectopically expressed proteins but for endogenous proteins as well. In order to understand the implications of the interaction between these two proteins, we decided to generate deletion mutants of both ELL and DBC1, comprising of different structural domains that have already been reported and studied. First, we wanted to address the detailed interaction between different domains of DBC1 with ELL. Co-expression of these deletion fragments of DBC1 (Fig. 4.3A, Top panel) with full-length ELL yielded some interesting results. Deletion

of the first 111 amino acids from the N-terminus region of DBC1, which consists of the S1-RNA like binding domain completely abolished ELL interaction with DBC1 (Fig. 4.3A, Compare ELL interaction intensity in lane 4 vs lane 9 in the IP panel). Additionally, the interaction between DBC1 (1-794 aa) and ELL was found to be quite strong, however, further deletion of 90 amino acids renders the association between DBC1 and ELL to almost negligible levels (See the interaction intensity in lane 5 vs lane 6). A DBC1 domain containing the amino acids 1-340 (containing the LZ domain) was seen to interact quite strongly with ELL, but the deletion of this region almost completely abolished ELL-DBC1 interaction (Compare ELL interaction between lanes 7 and 8). Thus, we can conclude that the interaction between DBC1 and ELL is dependent on various structural regions of DBC1. The presence of the S1-RNA like binding domain and LZ motif is essential for its interaction with ELL, whilst the 90 amino acid stretch between amino acids 705 to 794 inhibits its association with ELL. Next, we also wanted to identify the region of ELL that is necessary for its interaction with DBC1. For this, we made deletion fragments of ELL as indicated (Fig. 4.3B, Top panel). Co-immunoprecipitation of these ELL fragments with ectopically expressed DBC1 shows distinct interactions with specific regions of ELL. All deletion mutants of ELL were seen to interact with DBC1, barring one fragment, which expressed the region between amino acids 374-621 (Fig. 4.3B, Compare DBC1 interaction across lanes in the IP panel, noting the loss of interaction with ELL in Lane 8). The common feature in all ELL deletion domains is the presence of its elongation domain (Amino acids 60-373), which is missing in the fragment ELL (374-621). Incidentally, it is this fragment itself that loses its interaction with the DBC1. Thus, our results conclusively prove that while ELL requires multiple regions of DBC1 to interact, the elongation domain of ELL is key for its association with DBC1 (Fig. 4.3B).

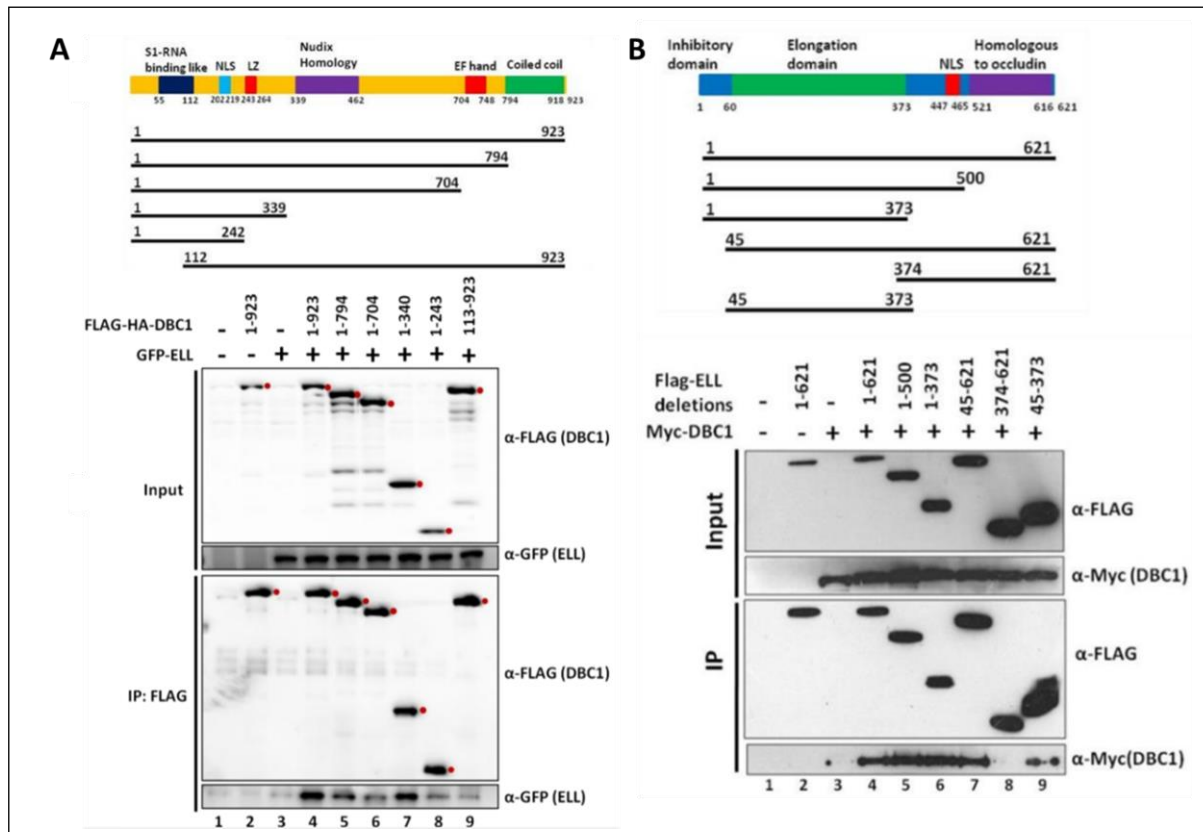


Fig 4.3 DBC1 and ELL interact with each other through specific structural domains

A. Western blots of immunoprecipitation assays for identifying the minimal region of DBC1 necessary for interacting with ELL shows that the N-terminal S1-RNA like binding domain and the region between amino acids 705-794 bearing the EF hand are the two domains important for this interaction **B.** A similar assay carried out, designed to identify the region of ELL necessary for interacting with DBC1 indicates that the elongation domain of ELL is essential for its association with DBC1.

4.4 ELL and DBC1 have an additive effect on RNA Pol II-mediated transcription

Our experiments show that DBC1 is a novel interactor of ELL, and the interaction depends on the elongation domain of ELL. ELL is known to be a bona fide RNA Pol II elongation factor. Keeping these two factors into account, it is not difficult to hypothesize that the interaction between DBC1 and ELL could have implications for the regulation of transcription. To test this, we have taken the advantage of the Luciferase reporter assay system. Our reporter system consists of a stable cell line containing 5 GAL DNA binding sequences inserted upstream of the luciferase reporter gene [144]. When the target construct is introduced into the cells, if the expressed gene is a transcriptional activator, the Luciferase gene is expressed, which causes oxidation of its substrate luciferin, leading to the release of light, which is then quantified. First, we transfected increasing concentrations of DBC1 to test what effect it would have on RNA Pol II-mediated transcription. In the presence of the synthetic activator GAL4-

VP16, DBC1 was able to modestly increase the rate of transcription (Fig. 4.4A, Luciferase reading represented by the bar graph and blots indicate the expression of transfected DBC1). However, this observation also raised the possibility that DBC1 by itself may not have a stimulatory effect on transcription, but may affect the expression of the GAL4-VP16 activator, leading to increased readout. To test this, we transfected increasing amounts of DBC1 plasmid along with a fixed concentration of GAL4-VP16 and tested its expression levels by western blotting. As shown in Fig. 4.4B, the increased DBC1 did not have any effect on the expression of GAL4-VP16, whose levels remain constant, indicating that: a) DBC1 does not have any effect on the expression of the GAL4-VP16 activator, and b) The stimulatory effect on transcription seen in the presence of DBC1 (as shown in Fig. 4.4A) is due to the effect of DBC1 alone. Next, we also wanted to test the effect of ELL overexpression on the transcription of the luciferase reporter gene. For this, we again introduced increasing concentrations of ELL plasmid within our stable cell line and measured the luciferase activity. ELL has already been established as an RNA Pol II elongation factor, and hence, its presence should cause enhanced transcriptional output. Our experiments also revealed the same trend, wherein ELL was able to strongly stimulate transcription of the Luciferase gene (Fig. 4.4C, Bar graph at the bottom part and the expression of ELL in the top panel). AS DBC1 and ELL have been seen to interact with each other, we were curious to know what effect these two proteins together would have on transcriptional output. To understand this, in our first experiment, we co-transfected increasing concentrations of DBC1 along with a fixed amount of ELL and tested the effect of this treatment by the luciferase assay. We found that the stimulation of transcription is enhanced when compared to DBC1 or ELL alone, but the cumulative effect is not synergistic, but rather additive in nature (Fig. 4.4D). Finally, we also checked the effect of transfecting a constant concentration of DBC1 with increasing ELL on the luciferase gene expression. In this experiment, we observed that the cumulative effect of this co-transfection gave a synergistic output and the increase in activity of the luciferase gene was manifold when compared to all our previous setups (Fig. 4.4E). Thus, we can conclude that both DBC1 and ELL are capable of stimulating RNA Pol II-mediated transcription within mammalian cells.

While performing our Luciferase assays, for each experiment we tested the expression of the proteins transfected by western blotting. This essentially serves as our internal control, ensuring that the transfected plasmids have expressed as desired. When we were performing experiments to test the effect of ELL and DBC1 together on the expression of the luciferase gene, we had two setups: one with increasing ELL and constant DBC1 (Fig. 4.4E) and the other with a fixed concentration of ELL and increasing DBC1 (Fig. 4.4D). Western blotting assays of these two experiments yielded some interesting results. When constant DBC1 was transfected along with increasing ELL, the western blots revealed that with increasing ELL expression, the protein levels of DBC1 also increased linearly (Fig. 4.4E, Compare DBC1 expression in lanes 3-7). This observation, though unexpected, can still be

explained by the logic that since ELL is a Pol II elongation factor, it may have stimulated transcription of the DBC1 plasmid itself. The more astonishing result was obtained when we were testing the effect of a fixed concentration of ELL in the presence of increasing DBC1. Here, much to our surprise, we observed a significantly enhanced ELL protein expression in the presence of DBC1 (Fig. 4.4D, Compare ELL expression in lane 4 to that in lanes 5-8). This observation is unexpected, as there is no report of DBC1 acting as a transcriptional stimulator, and thus forms the basis of asking the most pertinent question: how does the presence of DBC1 cause an increase in expression of the ELL protein?

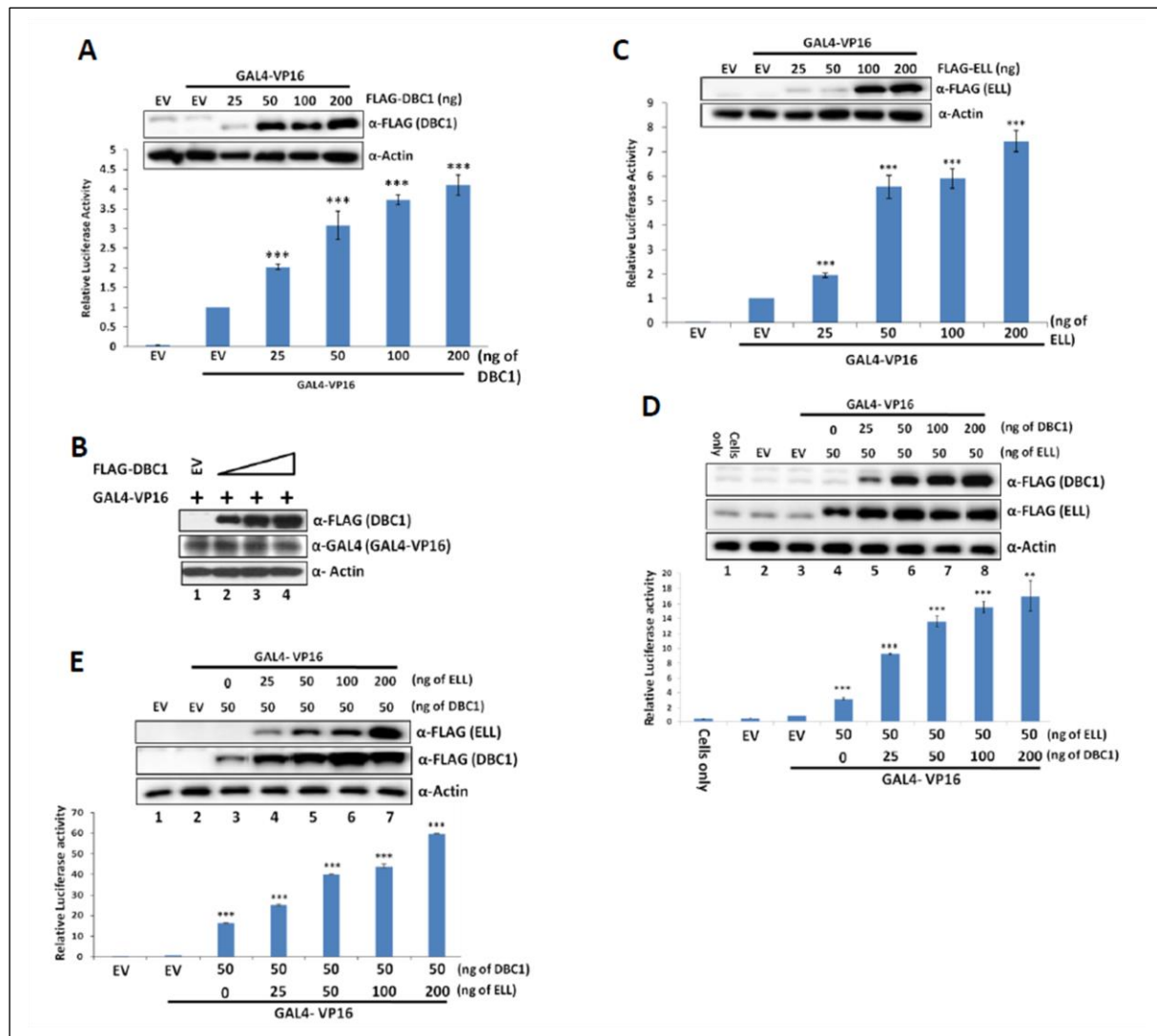


Fig 4.4 DBC1 and ELL together can stimulate RNA Pol II mediated transcription

A. A luciferase assay shows that DBC1 can enhance transcription of the luciferase reporter gene within mammalian cells. The immunoblots (top panel) indicate the amount of DBC1 transfected. **B.** Immunoblots show that the introduction of DBC1 plasmid by transfection has no effect on the synthetic GAL4-VP16 activator. **C.** ELL stimulates RNA pol II mediated transcription as denoted by the enhanced expression of the luciferase gene. **D.** The combined effect of increasing DBC1 and constant ELL has a rather additive effect on luciferase activity. However, in the presence of increased DBC1, the ELL expression also increases as seen in the immunoblots **E.** DBC1 and ELL synergize to stimulate transcription when constant DBC1 is co-transfected along with increasing ELL within mammalian cells.

4.5 DBC1 stabilizes ELL protein expression within mammalian cells

Luciferase reporter assays conducted to understand the effect of DBC1 on eukaryotic transcription yielded some interesting results for our study. While the co-transfection of DBC1 and ELL has an additive effect on gene transcription, when we introduced increasing concentrations of DBC1 with a fixed concentration of ELL, we were surprised to see that the ELL protein levels seemed to increase linearly with increasing DBC1 (Fig. 4.5D, ELL blot, lanes 4 vs lanes 5-8). This prompted us to ask the question whether DBC1 has any role in the regulation of expression of ELL within cells. To understand this, first, we transfected a minimal concentration of ELL along with increasing amounts of DBC1 within HEK-293T cells and tested if any changes were observed in ELL expression. We see, in our assay, that the presence of DBC1 greatly increases ELL levels (Fig. 4.5A, Compare ELL expression in lane 1 to lanes 2-5 and quantitation in the bar graph below). Thus, the DBC1 protein seems to have a stimulatory effect on ELL expression. However, it could be possible that since DBC1 has also been seen to have a positive effect on transcription, the increased ELL levels could be due to enhanced transcription of ELL mRNA. To test this, we used the transfected cells obtained from our earlier experiment (Shown in Fig. 4.5A), extracted RNA, and set up qRT-PCR assays from the synthesized cDNA. Our results indicate that enhanced transcript levels of DBC1 are seen upon transfection of the DBC1 plasmid (Fig. 4.5B, upper panel) indicating that our assay conditions are optimized. However, in this experimental setup, the ELL mRNA remained largely unchanged upon ectopic DBC1 expression when compared to the ELL-only control (Fig. 4.5B, Lower panel, See ELL mRNA levels across the different labeled lanes). Thus, it becomes very clear that the effect of DBC1 seen on ELL is at the protein level, and DBC1 does not have any significant effect on the transcription of ELL mRNA. As our earlier experiments were performed under overexpression conditions, we were also curious to check whether this effect of DBC1 on its target, ELL, would be prevalent under endogenous conditions. To test this hypothesis, we transfected a DBC1 expressing plasmid in HEK-293T cells and checked for the levels of ELL protein by western blotting. Incidentally, we found that the presence of DBC1 was able to increase the expression of endogenous ELL protein as well (Fig. 4.5C, lane 1 vs lanes 2-3 in the ELL blot and quantification of ELL signal relative to Actin in the bar graph below). Thus, the effect of DBC1 on ELL protein is not limited to its ectopic expression. However, a concern arose that it could be highly possible that this effect of DBC1 is universal and not specific to ELL, or any other protein. Thus, we checked if DBC1 would have any effect on the expression of a control (GFP expressing) plasmid within mammalian cells. Our results suggest, however, that the stabilizing effect of DBC1 is specific for ELL only as no changes were seen in the expression of GFP whether in the presence or absence of DBC1 (Fig. 4.5D). We also determined by our experiments that the stimulatory effect of DBC1 on ELL protein is not cell type-specific, as the transfection of DBC1 in a colorectal carcinoma cell line, HCT116, also showed a significant rise in ELL

levels upon DBC1 overexpression (Fig. 4.5E). Thus, our experiments confirm that DBC1 stabilizes ELL at the protein level across all cell types.

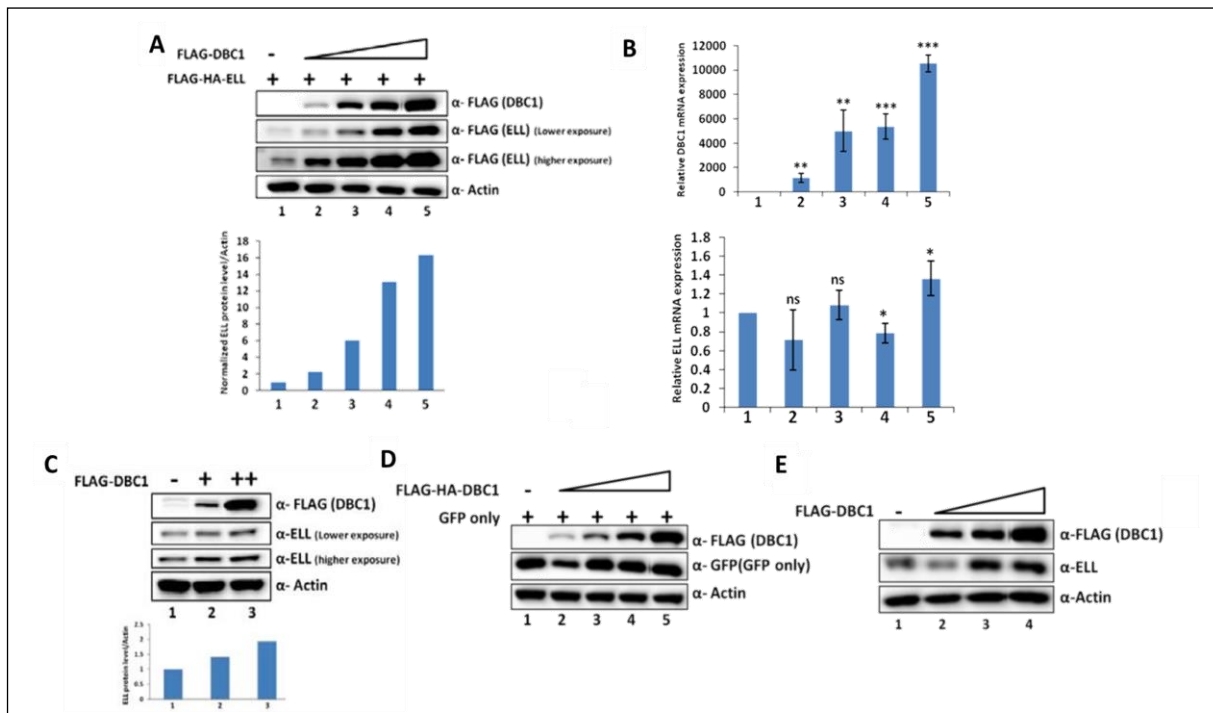


Fig 4.5 DBC1 stabilizes ELL protein within mammalian cells

A. Immunoblots show that in the presence of DBC1, ectopically expressed ELL levels are greatly enhanced (Top panel). The bar graph at the bottom denotes the ELL signal relative to Actin in the marked lanes. **B.** The stabilizing effect of DBC1 is not due to an effect on ELL transcription as seen by qRT-PCR assays. The numbers denote lane numbers of the samples shown in Fig. 4.5A **C.** DBC1 can also stabilize endogenous ELL levels as shown by this western blot assay. **D.** Immunoblots show the stabilization effect of DBC1 is specific to ELL as the overexpression of DBC1 does not have any effect on the control green fluorescence protein. **E.** Endogenous ELL protein is also stabilized in the presence of DBC1 in HCT116 cells indicating that this effect of DBC1 is not cell type specific.

The qRT-PCR data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, **denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

4.6 ELL interacts with HDAC3 within mammalian cells

DBC1 has been identified by our detailed studies as a novel interactor of ELL. In addition to interacting with ELL for the stimulation of transcription, DBC1 also provides stability to ELL protein. Keeping these two observations in mind we wondered what could be the implications of the DBC1-ELL interaction within cells. Looking back at published literature, we came across multiple reports where DBC1 has been shown to maintain p53 levels within cancer cells [203, 204]. p53, in an active form, stays in an acetylated state. Deacetylation by SIRT1 causes p53 to be destabilized and degraded by the E3 ubiquitin E3 ligase, Mdm2. DBC1 maintains p53 levels by preventing the activity of SIRT1, thus

ensuring an anti-tumor response. Indeed, several cancer cell lines have been seen to have low or nil expression of DBC1, thus cementing the role of DBC1 as a tumor suppressor. As such a detailed analysis of DBC1 function has been done previously, we wondered whether this same model could be extrapolated to our observations. We thus hypothesized that like p53, ELL could also interact with SIRT1 which would cause a loss in its function, and DBC1 would act to prevent this from taking effect. Hence, we first decided to test if ELL associates with SIRT1 within mammalian cells. For this, we co-expressed ELL (as FLAG-HA tagged) and SIRT1 (as FLAG-tagged) and performed an immunoprecipitation assay with anti-HA agarose beads. However, western blot analysis indicated that in spite of a significant amount of ELL being immunoprecipitated in our setup, SIRT1 could not be captured as an ELL-associated protein (Fig. 4.6A). Hence, we delved back into available literature to identify other proteins that may be regulated by DBC1 and found a report that showed the negative regulation of the lysine deacetylase, HDAC3, by DBC1 [212]. With this information in mind, we now decided to test if ELL would interact with any HDAC within mammalian cells and transfected ELL along with FLAG-tagged HDACs 1-8 in HEK-293T cells. Subsequent immunoprecipitation of the HDACs with anti-FLAG agarose beads showed that ELL could interact stably with HDACs 1, 3, and 6 within cells (Fig. 4.6B, see lanes 4,6 and 8 of the ELL blot in the IP panel). A similar assay performed to test the association of HDACs with DBC1 also showed that within cells, DBC1 could associate most strongly with HDACs 1 and 3 (Fig. 4.6C, lanes 4 and 6 of the DBC1 blot in the IP panel). Both proteins were seen to interact with other HDACs as well, but we chose to focus on HDACs 1 and 3 as their interaction strength was the most. Taking both proteins into consideration, both DBC1 and ELL were seen to interact with HDAC3 with the highest affinity. Thus, we decided to proceed with HDAC3 for our further experiments. First, we tested if ELL would interact with endogenous HDAC3. Upon transfection and immunoprecipitation of FLAG-ELL from a stably expressing cell line, western blots indicated that HDAC3 indeed was present as an interactor of ELL (Fig. 4.6D, lane 3). Interestingly, a reciprocal experiment, where FLAG-HDAC3 was immunoprecipitated from a stably expressing cell line also showed the presence of both endogenous DBC1 as well as ELL as interactors of HDAC3 (Fig. 4.6E). Thus, our experiments identified HDAC3, among all lysine deacetylases, as a stable interactor of ELL as well as DBC1 within mammalian cells.

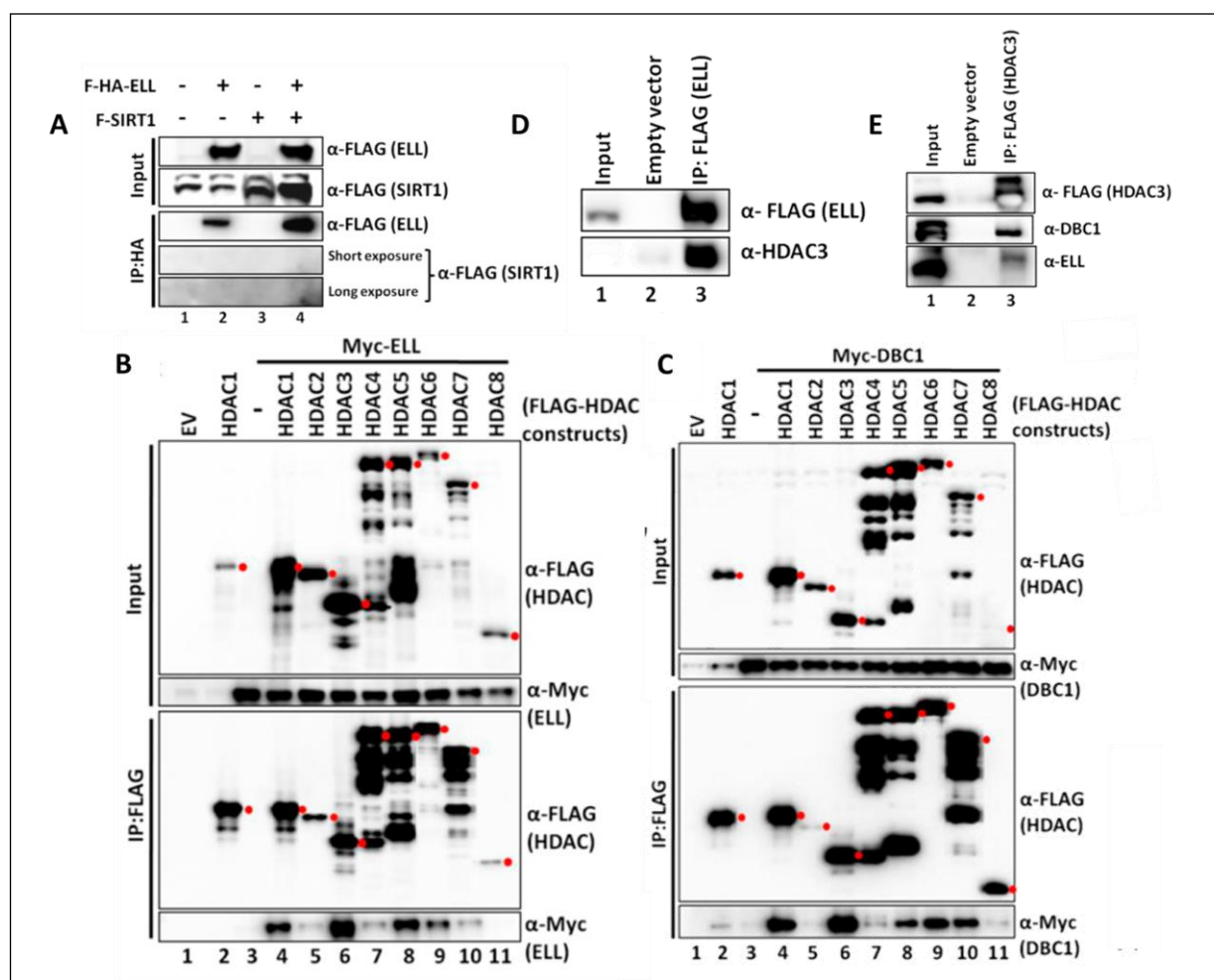


Fig 4.6 A novel interaction is seen between ELL and HDAC3 within mammalian cells

A. An immunoprecipitation assay shows that ELL does not interact with the deacetylase, SIRT1 within mammalian cells. **B.** ELL interacts with HDACs 1, 3 and 6 as tested by immunoprecipitation followed by western blotting assays at physiological salt concentrations. **C.** Immunoblots show the strong interaction between ectopically expressed DBC1 with HDACs 1 and 3 within mammalian cells **D.** ELL associates with endogenous HDAC3 as seen in this immunoprecipitation assay. **E.** A reverse experiment confirms that endogenous DBC1 and ELL do indeed interact with HDAC3 within cells.

4.7 ELL is degraded specifically by HDAC3 within cells

Our screening assays have shown that both DBC1 and ELL interact strongly with HDACs 1 and 3. As the strength of interaction was maximum for both ELL as well as DBC1, and previous reports have stated an interplay between HDAC3 and DBC1, we chose to focus on this lysine deacetylase for our detailed studies. Subsequent immunoprecipitation studies showed ELL interacting with endogenous HDAC3 and *vice versa*. These observations prompted us to think about what could be the underlying implication of this interaction. We then noted a previous study elucidating a mechanism by which the deacetylase SIRT1 causes degradation of p53 which is inhibited by DBC1 [203]. Thus, we then wanted

to understand if HDAC3 could cause destabilization of ELL, much like SIRT1 and p53. Hence, we co-expressed ELL along with increasing concentrations of HDAC3 and tested ELL levels by western blotting. We observed that the presence of HDAC3 almost depleted all the ectopically expressed ELL in HEK-293T cells (Fig. 4.7A, Compare ELL expression in lane 2 vs lanes 3-6). Hence, from this observation, it appears that the expression of HDAC3 is detrimental to ELL in mammalian cells. However, this effect of HDAC3 could also be a result of adversely affecting ELL transcription, rather than affecting ELL protein. We tested this hypothesis by a very simple assay. Using the setup described in the earlier experiment, we transfected cells in the indicated combinations as mentioned previously (Fig. 4.7A). these cells were then harvested and RNA extracted for performing qRT-PCR for both HDAC3 and ELL. Our results from this simple experiment show that HDAC3 mRNA levels increase linearly as more HDAC3 is transfected, indicating that our assay conditions are perfect (Fig. 4.7B, Top panel). However, there is no decrease in ELL mRNA levels when high concentrations of HDAC3 are co-expressed (Fig. 4.7B, Bottom panel). In fact, at the highest concentrations of HDAC3, where the ELL protein amount is seen to be the least, the ELL mRNA levels appear to be the highest in the entire assay (ELL mRNA in lanes 5 and 6). Thus, it is confirmed that this effect of HDAC3 is not on ELL transcription, but by some currently unknown mechanism that affects ELL protein levels. One concern that we faced like previously we had had with DBC1, is that this observation could be cell type-specific and limited to HEK-293T only. To address this, we carried out essentially the same experiment as shown in Fig. 4.7A, but this time performed the transfection in HCT116 cancer cells. as expected, the results obtained in this experiment also indicated that ELL was indeed degraded when HDAC3 was overexpressed, thereby confirming that this effect of HDAC3 on ELL is universal in nature and not restricted to any specific cell type. However, our earlier experiments have indicated that ELL does interact with several HDACs, and hence this HDCA3 mediated LL regulation could also be carried out by some other lysine deacetylases. To test this, we simply co-transfected ELL with all the HDACs we had in our repertoire and tested which of these enzymes have an effect, if any, on ELL protein. Our experimental results clearly show that of all the HDACs we have tested in our assay, only HDAC3 was able to significantly decrease ectopically expressed ELL expression under our assay conditions (Fig. 4.7D, Compare ELL expression in lane 1 vs lane 4). It is interesting to note that with the overexpression of HDACs 1 and 6, ELL did show a slightly decreased expression, but this effect was not as significant as that of HDAC3 (See lanes 2 and 7 in the ELL blot). As DBC1 is also known to interact and modulate the effect of SIRT, we also decided to test if any SIRT would be able to modulate ELL expression. However, when we co-transfected all human SIRTs (i.e. SIRT1-7) along with ELL within HEK-293T cells, we did not find ELL expression to be affected in our immunoblots (Fig. 4.7E). Thus, our results indicate that none of the SIRTs have the property to modulate ELL expression, and within the HDAC family, only HDAC3 specifically can regulate the levels of ELL within mammalian cells.

A method of testing the stability and turnover of proteins within the cell is by exposure to cycloheximide (CHX), which inhibits protein translation. As a result, new protein synthesis in the cell is stopped, and only the available pool is now affected by different degradation pathways. Thus, this assay, termed the cycloheximide chase assay, is a useful tool to test the turnover rate of any protein within cells. Initially, we wanted to get an idea about the degradation kinetics of our target protein, ELL, in mammalian cells. Hence, we treated the cells with cycloheximide for different time points, harvested the cells with cold PBS, and obtained the protein extract from these cells for testing by immunoblotting. Our results indicate that ELL is a highly stable protein, and does not get degraded easily, an observation that can be clearly inferred as its levels remain static even after 24 hours of cycloheximide treatment (Fig. 4.7F). Incidentally, upon testing, we also found that DBC1, which can interact stably and directly with ELL, is also highly stable as its expression also remained unperturbed after a 24-hour chase assay. Thus, it is aptly clear that both DBC1 and ELL are highly stable proteins whose expression is ubiquitous within cells. Now, we wanted to test if the introduction of HDAC3 would have any effect on ELL turnover, thus possibly giving us an idea as to how HDAC3 regulates ELL. Hence, we then tested ELL degradation kinetics in the absence as well as the presence of HDAC3. As measured by western blots, we find that ELL indeed remains stable for the entire duration of our CHX chase (Fig. 4.7G, See ELL expression in lanes 1-3). However, as HDAC3 is overexpressed, ELL protein decreases by about 25% within a short span of 4 hours, and a further 4-hour treatment shows only about 30% of the ELL remaining when we compare it to its abundance at the start of the assay (ELL blot, lanes 4-6, and the representation of the quantification of the ELL signal across treatments as indicated in the line graph on the right). Hence, HDAC3 does affect ELL's turnover rate and greatly increases its degradation kinetics within cells. However, until now, one caveat of our experiments with respect to HDAC3-mediated ELL regulation was that all assays were performed under overexpression conditions. Thus, to validate our observations in an endogenous context, we decided to generate stable cell lines where HDAC3 had been knocked down. Using HDAC3 shRNA sequences from earlier published reports, we were able to use two shRNA constructs for stably knocking down HDAC3 within HEK-293T cells (Fig. 4.7H). Next, we wanted to examine how this depletion of HDAC3 affects endogenous ELL levels. Western blot analysis from whole-cell protein extracts clearly shows that upon HDAC3 knockdown, the expression of endogenous ELL is markedly increased (Fig. 4.7I, ELL blot, compare lane 1 vs lanes 2 and 3). To eliminate the possibility that this increase in ELL is not due to any effect on ELL transcription, we performed qRT-PCR assays for testing the mRNA levels of ELL from RNA extracted from control scramble as well as HDAC3 knockdown cells. The results obtained in this assay clearly indicate that ELL mRNA remains unchanged upon HDAC3 knockdown (Fig. 4.7J). Thus, our results clearly show that HDAC3 can degrade ELL protein within mammalian cells without affecting its

rate of transcription. This effect is specific for HDAC3 as ELL is impervious to the overexpression of any other lysine deacetylase within the cells.

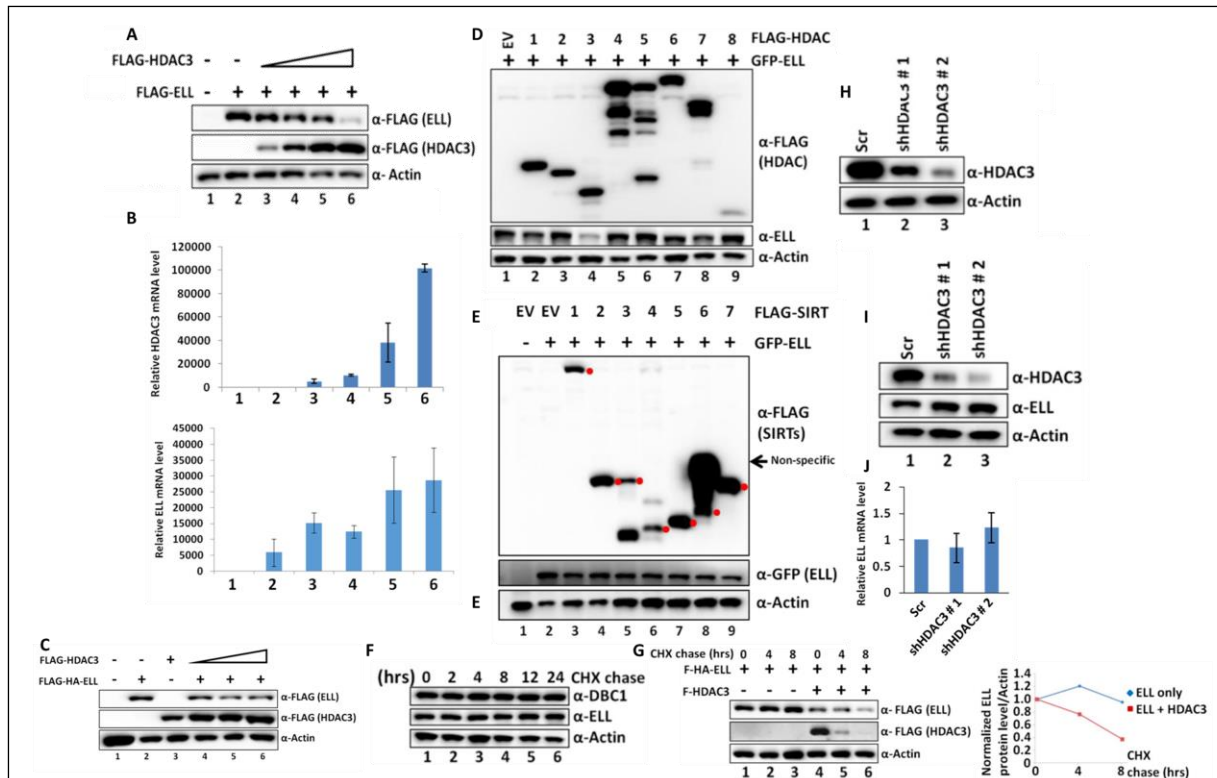


Fig 4.7 HDAC3 causes degradation of ELL within mammalian cells

A. Western blots showing that the overexpression of HDAC3 causes a linear decrease in ectopically expressed ELL levels within HEK-293T cells **B.** qRT-PCR assays indicate that the effect of HDAC3 on ELL expression is not at the transcript level as ELL mRNA does not decrease with HDAC3 overexpression. The numbers denote the template used for performing the reaction, corresponding to the lane numbers of the samples loaded in the previous experiment **C.** ELL degradation by HDAC3 is not specific to HEK-293T cells as the same results are obtained upon transfection of the indicated plasmids in HCT116 cancer cells **D.** Immunoblots show that ELL is specifically degraded by HDAC3 and remains unaffected by the overexpression of other HDACs. **E.** Western blot analysis to determine whether the presence of any SIRT1s have any effect on the ELL levels show that ELL protein remains unchanged if co-expressed with any sirtuin. **F.** A cycloheximide chase assay carried out in HEK-293T cells indicates that ELL and DBC1 are highly stable proteins that are maintained at constant levels even after 24 hours of CHX treatment **G.** ELL, by itself, is highly stable, but the overexpression of HDAC3 greatly increases its degradation rate as shown in this CHX chase assay. The graph on the right indicates the changes in ELL levels over time. **H.** Immunoblots showing the knockdown of endogenous HDAC3 in HEK-293T cells by two different shRNA constructs. **I.** ELL levels are enhanced in a cell lines where HDAC3 has been stably knocked down by shRNAs as indicated in these western blots **J.** qRT-PCR analysis indicate that the increase in endogenous ELL levels upon HDAC3 knockdown are not because of an increase in ELL transcription, but is an effect at the protein level.

The qRT-PCR data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, **denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

4.8 DBC1 acts oppositely to HDAC3 for stabilization of ELL

Our experiments started with the identification of a novel protein interacting with ELL, named DBC1. DBC1 was also seen to be capable of stimulating RNA Pol II-mediated transcription by a reporter gene assay. Incidentally, the overexpression of DBC1 caused an increase in ELL protein levels without affecting its transcription, but another novel interactor of ELL, i.e. HDAC3 was seen to cause destabilization of ELL protein as well as increase the kinetics of its degradation within mammalian cells. Our observations thus point towards the regulation of an essential protein by the effect of two other factors. We had already seen that HDAC3 was capable of causing ELL degradation within mammalian cells. Taking this forward, we were curious to know whether DBC1, which had shown the capability to increase ELL levels, could reverse HDAC3 mediated ELL degradation. We thus first expressed ELL alone and with HDAC3 within mammalian cells. Western blots of these samples indicated that in this case as well, HDAC3 caused ELL degradation. Next, in the presence of both ELL and HDAC3, we introduced increasing concentrations of DBC1 and tested ELL protein levels by western blotting. As the results in Fig. 4.8A show, the presence of DBC1 rescues ELL expression which was diminished by HDAC3 overexpression (Compare lane 3 vs lanes 4-6 of the ELL blot, and quantification of the ELL protein with respect to Actin as shown by the bar graph below). Now, we had seen previously that HDAC3 drastically increases ELL degradation kinetics (Fig. 4.7G) and thus wanted to check if, in the presence of DBC1, ELL gets stabilized. Hence, we performed another CHX chase assay wherein cells transfected with plasmids in multiple indicated combinations were treated with cycloheximide (Fig. 4.8B). Protein levels were measured by immunoblotting against the respective antibodies. The results obtained by us clearly show that ELL remains stable even after 8 hours of CHX treatment (Lanes 1-3). However, in the presence of HDAC3, ELL is rapidly degraded to about half of its original levels over the course of this treatment (See lanes 4-6 and quantification represented by the graph below). Finally though, when DBC1 is also transfected along with ELL and HDAC3, the ELL protein levels again remain mostly constant with very minor fluctuations (Lanes 7-9 of the ELL blot). Thus, these experiments make it very clear that ELL degradation by HDAC3 can be reversed by DBC1, the presence of which also decreases the rate of ELL turnover within cells.

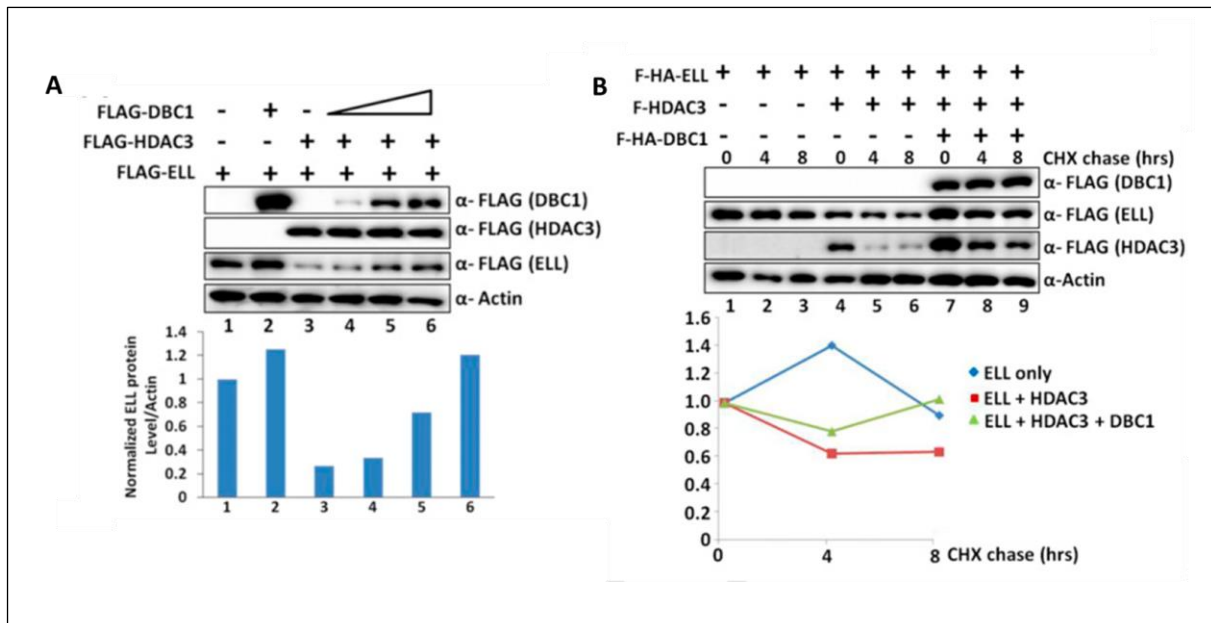


Fig 4.8 DBC1 protects ELL from being degraded by HDAC3

A. Western blots show that HDAC3 mediated ELL degradation can be rescued by the overexpression of DBC1 within HEK-293T cells. The bar graph at the bottom shows ELL levels relative to actin in the respective lanes as labelled **B.** Cycloheximide chase assays confirm that ELL degradation rate, which is enhanced in the presence of HDAC3, is markedly reduced when DBC1 is overexpressed. The line graph at the bottom panel shows the ELL levels relative to Actin at the indicated time points

4.9 HDAC3 and DBC1 compete with each other for binding to ELL

Among all the deacetylases we have tested, we found only HDAC3 as capable of degrading cellular ELL, which otherwise is a very stable protein. This stability of ELL is derived from its interaction with DBC1, which has also been seen to be able to rescue ELL from HDAC3-mediated destabilization. We wanted to delve deeper into this intricate network of regulation. In order to do so, and to understand the mechanics of ELL-HDAC3 interaction, we turned to detailed interaction analysis between these two proteins using different deletion fragments. Using the same deletion mutants of ELL that we had used in our earlier experiments (Fig. 4.3B), we set up a co-immunoprecipitation assay of these fragments with sub-optimal concentrations of HDAC3. Our results obtained in this experiment show that the first 44 amino acids of ELL are important for its interaction with HDAC3 (Fig. 4.9A, compare HDAC3 signal in the IP blot for lanes 4 vs 7). Additionally, we find that the deletion of amino acids 1-373 from the N terminus of ELL completely abolishes its interaction with HDAC3 (Fig. 4.9A, Compare HDAC3 signal in lane 4 vs lane 8 of the IP panel). Looking back at our earlier analyses of the interaction of ELL deletion fragments with DBC1 (Fig. 4.3), we find a similar pattern of interaction as that observed for HDAC3. Both these two proteins lose their interaction with ELL if its elongation domain, located between amino acids 60-373, is deleted. Thus, it is not very difficult to fathom a plausible scenario for the mechanism of interaction of these two factors with ELL. Taking into consideration the interaction

patterns of DBC1 and HDAC3 with ELL, it appears that both these proteins depend on the elongation domain for their association with ELL protein. Since these two proteins have diametrically opposite effects on ELL, there possibly exists a competition between them for binding to the ELL elongation domain. To test our hypothesis, we performed an *in vivo* displacement assay. Here, we expressed a fixed concentration of ELL with increasing amounts of DBC1 within HEK-293T cells. Subsequently, the ELL was immunoprecipitated using anti-HA tag agarose beads, and the association of endogenous HDAC3 was tested in the presence and absence of DBC1. Western blot analysis of the immunoprecipitated samples clearly shows that as DBC1 has been transfected in incremental quantities, the association between ELL and DBC1 was seen to increase (Fig. 4.9B, Compare lane 2 vs lane 4 of the DBC1 blot). With an enhanced DBC1 interaction with ELL, the concomitant HDAC3 -ELL association decreased significantly (See the HDAC3 signals in lane 1 vs lanes 2-4 in the HDAC3 blot of the IP panel). These observations make it quite clear that DBC1 and ELL do compete with each other for binding to ELL within cells. However, it was also imperative that to directly verify this hypothesis, we perform an experiment to test this observation under *in vitro* conditions. For this, we first purified HDAC3 by plasmid-mediated overexpression in HEK-293T cells. From the cell lysate, FLAG-HDAC3 was affinity purified using antibody-coated beads and eluted under high salt conditions to ensure its purity (Fig. 4.9C). ELL, as GST tagged, had already been purified earlier (Fig. 4.1E). To standardize our assay conditions, and also to determine whether HDAC3 and ELL do interact directly in a cell-free environment, we first set up a direct interaction assay between these two proteins. Keeping GST-ELL as the bait, immobilized on anti-GST beads, we incubated it with HDAC3 in a neutral interaction buffer. After overnight incubation with gentle agitation, the beads were rigorously washed, and bound proteins were identified by western blotting. As the results in Fig. 4.9D indicate, HDAC3 was found to interact with ELL directly in a cell-free system (Lanes 4 and 5 of the HDAC3 blot). Thus, HDAC3, like DBC1, has the potential to directly associate with ELL. But the critical question remained if DBC1 could compete with HDAC3 for directly binding to ELL. This hypothesis was also tested by an *in vitro* competitive binding assay. In this assay, we used the immobilized GST-ELL as bait and added HDAC3 into the reaction mix. Subsequently, purified DBC1 was introduced in incremental amounts and the association between ELL and HDAC3 was tested by western blotting. Once again, we observed a direct interaction between ELL-HDAC3 and ELL-DBC1 in our assay (Fig. 4.9E, lane 5 of the HDAC3 blot and lane 6 of the DBC1 blot). However, our experimental results indicate that when the DBC1 concentration in the reaction is increased, the HDAC3-ELL association is dramatically perturbed and diminishes significantly (Compare lane 5 vs lanes 7 and 8 of the HDAC3 blot in the IP panel). Thus, our experiments, taken together, clearly indicate that both DBC1 and HDAC3 bind to similar regions of ELL, and DBC1 mediated ELL stabilization arises from its ability to competitively prevent HDAC3-ELL association. ELL within the cell is destabilized in the presence of HDAC3. However, DBC1 prevents

HDAC3 from interacting with ELL and thus maintains cellular ELL levels as illustrated in the cartoon (Fig. 4.9F).

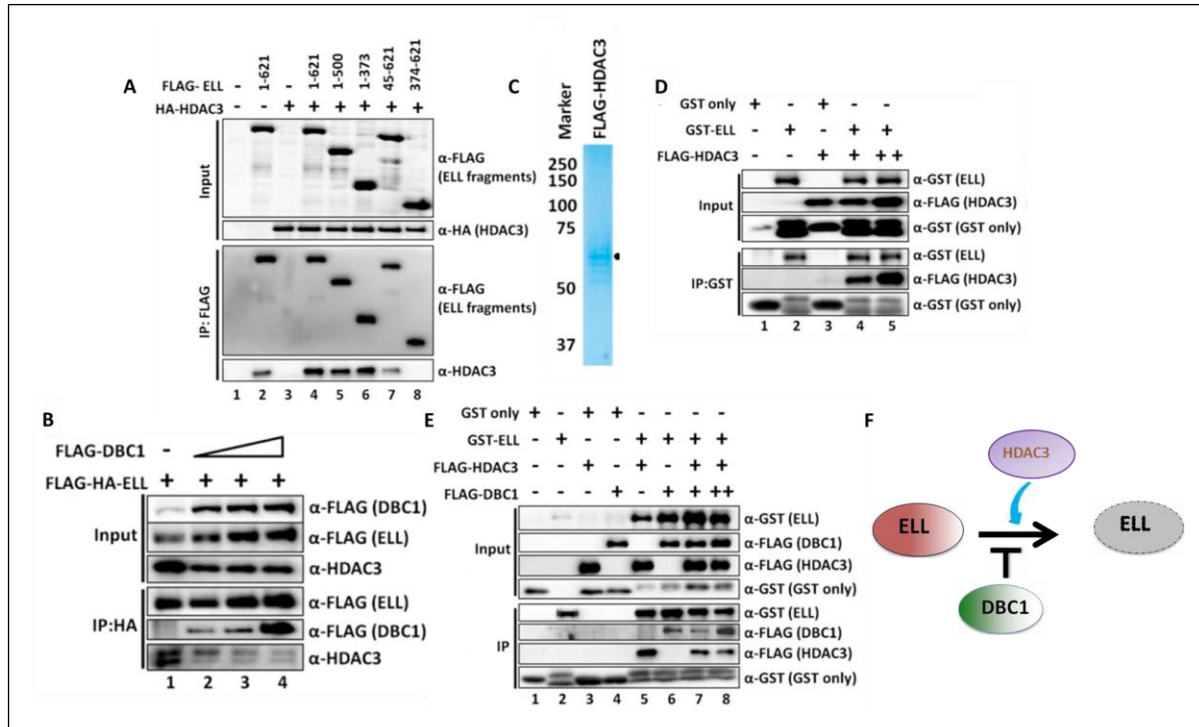


Fig 4.9 DBC1 and ELL competitively bind to the same region of ELL

A. Immunoprecipitation assays between HDAC3 and different deletion mutants of ELL indicate that the N terminus and elongation domain of ELL is necessary for stable HDAC3 association. **B.** Overexpression of DBC1 along with ELL within HEK-293T cells concomitantly reduces ELL-HDAC3 association as shown by these immunoblots. **C.** Coomassie staining of a SDS-PAGE gel showing the purification of HDAC3 after its overexpression in mammalian cells for use in our *in vitro* interaction assays. The purification of GST only, GST-ELL and FLAG-DBC1 has already been described previously. **D.** Immunoblots of *in vitro* direct interaction assays show that purified HDAC3 interacts with ELL directly in a cell free system under physiological salt concentrations. **E.** An *in vitro* competitive binding assay shows that when the three proteins, ELL, HDAC3 and DBC1 are put together in a cell free system, the presence of DBC1 impedes HDAC3 interaction with DBC1. **F.** Cartoon diagram showing the interplay between DBC1 and HDAC3 for regulation of ELL levels. The presence of HDAC3 destabilizes ELL, but if DBC1 is in the vicinity, it prevents ELL-HDAC3 interaction and thus maintains ELL levels within the cell.

4.10 The lysine acetyltransferase p300 acetylates ELL

Detailed studies performed previously show that DBC1 is a negative regulator of several histone deacetylases, including HDAC3 and SIRT1. It has also been reported in studies that deacetylation often increases the propensity of proteins to undergo degradation by different means. Incidentally, we have also seen that when HDAC3 is present, ELL gets destabilized significantly and its rate of degradation is highly enhanced. Thus, ELL could be a substrate for the acetylation and deacetylation cycle. To explore

this possibility in detail, we first sought to identify if ELL would be acetylated within mammalian cells. for this, we co-expressed ELL along with lysine acetyltransferases of different families, such as PCAF, GCN5, p300, CBP, and MOF. We then immunoprecipitated the ELL under very high salt concentrations (~ 1M KCl) to ensure that no interactors or non-specific proteins were associated with it in the immunoprecipitate. Probing this ELL with a pan-acetyl lysine antibody showed the presence of acetylation signals in the presence of p300, CBP, and MOF (Fig. 4.10A, lanes 4-6 of the acetyl-lysine blot). However, of all these acetyltransferases, the most robust acetylation signal was observed in the presence of p300, indicating that this protein predominantly acetylates ELL within mammalian cells. Next, to re-confirm this observation, we also co-transfected increasing concentrations of p300 along with ELL in HEK-293T cells. This ELL was also subjected to affinity-based purification at high salt concentrations, followed by blotting with a pan-acetyl-lysine antibody. The results obtained indicate that ELL is robustly acetylated within cells upon co-expression of p300 (Fig. 4.10B, compare acetylation signal in lane 2 vs lanes 4-6). However, we also needed to confirm p300 as a *bona fide* acetyltransferase of ELL, which could acetylate ELL when tested under purified reaction conditions. Hence, we purified GST-EV and GST-ELL as substrates and the HAT domain of p300 from transformation and expression in bacterial cells (Fig. 4.10C). The decision to purify the HAT domain and not the full-length p300 was taken intentionally as the high molecular weight p300 cannot be purified efficiently from cells, and thus studies have widely used the described domain for their reactions. Now, using these purified proteins after checking their purity and concentration by Coomassie staining, we set up an *in vitro* acetylation assay to determine if p300 could acetylate ELL in the absence of the cellular environment. One of the properties of p300 is its ability to acetylate itself, and the presence of this autoacetylation signal in our assay indicates that the purified p300 is enzymatically active. In presence of this p300 HAT domain, we were able to observe significant acetylation of the GST-ELL substrate, while GST-ELL by itself, in the absence of p300, showed no such signal (Fig. 4.10D, See lane 3 vs lanes 4-6). This observation clearly establishes p300 as capable of acetylating ELL. However, a caveat of our performed experiments is perhaps the use of overexpression systems to identify cellular mechanisms. To overcome this limitation in this context, we decided to test how endogenous ELL acetylation is perturbed by p300 depletion. For this, we generate two cell lines where p300 was stably knocked down by two different shRNAs. From these two cell lines, we immunoprecipitated endogenous ELL with antibodies and then examined how the removal of p300 affects its acetylation levels. We observed that, in both p300 knockdown cell lines, acetylation levels of endogenous ELL were severely hampered when compared to control scramble knockdown cells (Fig. 4.10E, Compare acetylation signal in lane 1 vs lanes 2 and 3. The bar graph given in the bottom panel serves as a quantification of the acetylation signal normalized with respect to the

immunoprecipitated ELL level). Thus, from our detailed experiments, we conclude that ELL is post-translationally acetylated within mammalian cells and p300 serves as the acetyltransferase for ELL.

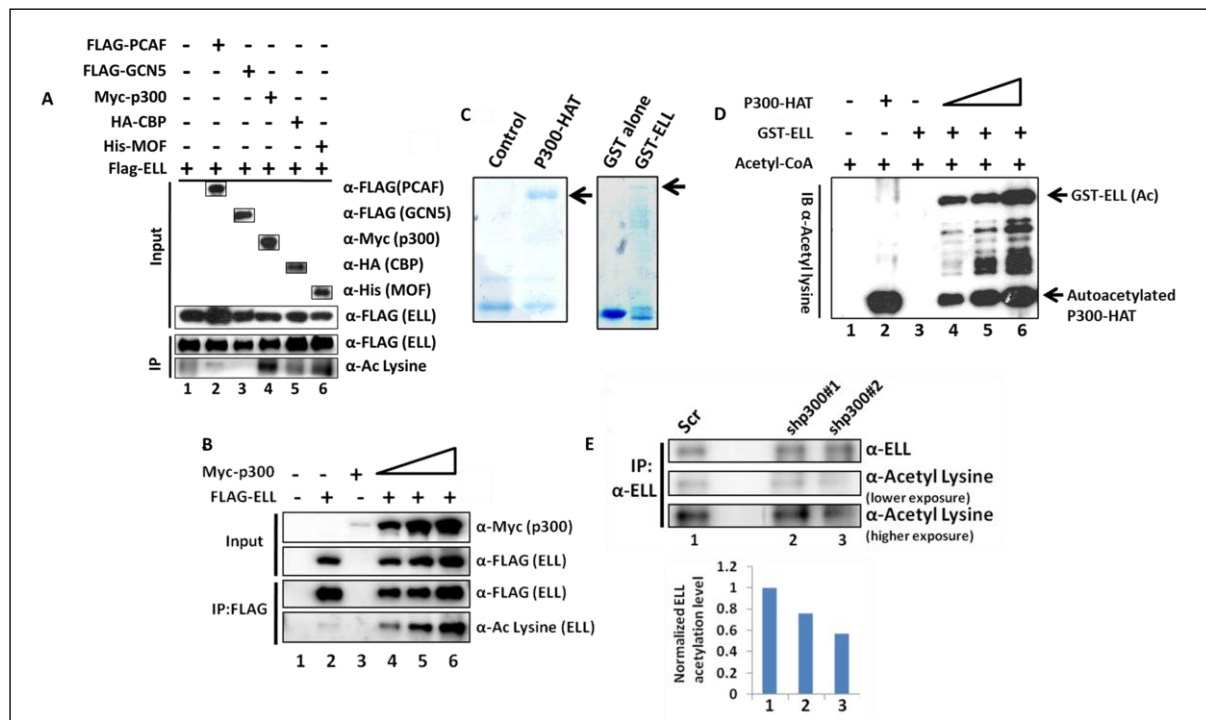


Fig 4.10 p300 acetylates ELL

A. Co-transfection of ELL with different lysine acetyltransferases followed by western blotting indicates that ELL is acetylated by several HATs within cells, but the most robust acetylation occurs in presence of p300 **B.** Co-expression of increasing concentrations of p300 along with ELL results in a significantly enhanced levels of acetylated ELL protein. **C.** Coomassie staining of a SDS-PAGE gel showing the purification of His-p300 HAT domain, GST-EV and GST-ELL after their overexpression in *E. coli* BL21(DE3) cells for use in our *in vitro* acetylation assays. **D.** Immunoblots of an *in vitro* acetylation assay show that in presence of purified p300 HAT domain, ELL gets acetylated even in a cell free system. **E.** Knockdown of p300 by two different shRNAs in HEK-293T cells causes the acetylation levels of endogenous ELL to be significantly reduced. The bar graph in the lower panel represents the quantification of the acetylation signal of ELL normalized to the ELL protein immunoprecipitated using specific antibodies.

4.11 ELL is stabilized upon p300 mediated acetylation

DBC1, a novel interactor of ELL has been observed to stabilize ELL and protect it from HDAC3 mediated degradation. It does so by binding to ELL and thus preventing HDAC3 interaction, which would otherwise channelize it towards degradation. Co-expression of ELL with increasing concentrations of DBC1 showed that ELL protein levels are enhanced when DBC1 is present (Fig. 4.5A). While testing if p300 acetylates ELL within mammalian cells, we had performed a co-transfection of ELL with p300. Interestingly, in this experiment, we observed that ELL protein signals were markedly enhanced when p300 was present, even though the same amount of ELL had been transfected in all wells (Fig. 4.10B,

ELL blot in the Input panel). This mirrors the results we have observed for DBC1 previously and begged us to ask the question whether p300 mediated acetylation could stabilize/increase ELL protein. In this section, we tested this hypothesis in multiple ways. First, we overexpressed a minimal concentration of ELL along with increasing concentrations of p300. In this experiment, we also observed that when p300 is present, the ELL protein is significantly increased as indicated in Fig. 4.11A (Compare ELL level in lane 2 vs lanes 4-7). A quantification of ELL proteins relative to the housekeeping gene, Actin is given in the panel below. As this experiment, like previously, was performed by ectopically expressing both plasmids, we next focused our attention on the effect of p300 overexpression on endogenous ELL. After transfection of the indicated amounts of p300 in HEK-293T cells, we performed immunoblots for testing if there were any changes in endogenous ELL protein. As the results in Fig. 4.11B show, endogenous ELL protein is significantly elevated when p300 is overexpressed in mammalian cells. Indeed, quantification of ELL protein with respect to Actin shows a six-fold enhancement when p300 is present (Fig. 4.11B, bottom panel). P300, however, has been reported to positively regulate transcription directly, or by epigenetic modifications. Thus it could be possible that the increase in ELL is due to a positive effect on its transcription. To check this hypothesis, we purified RNA from the cells where p300 had been transfected (As shown in Fig. 4.11B), and measured ELL mRNA by qRT-PCR. The results obtained clearly indicate that though the immunoblots show a dramatic increase in ELL protein, the mRNA levels are not adversely affected, and remained unchanged in spite of overexpression of high concentrations of p300 (Fig. 4.11C, The numbers correspond to the concentration of p300 transfected as shown in the previous figure). It is worth mentioning, though, that the caveat of our studies being based on overexpression systems arises again, for which we hypothesized, that if ELL stability is indeed dependent on p300, then the depletion of p300 should cause ELL to be diminished within the cell. In order to test this, we generated two cell lines in a HEK-293T background where p300 had been stably knocked down by the lentivirus-mediated introduction of shRNA. A whole-cell extract of protein samples from control scramble knockdown cells and the two putative p300 depleted cells was then subjected to SDS-PAGE followed by western blotting with factor-specific antibodies. Our assay results clearly show that endogenous p300 has been knocked down in these two cell lines (Fig. 4.11D, p300 blot). As we hypothesized ELL to be maintained by p300, we saw that the knockdown of p300 causes endogenous ELL to be severely diminished (Fig. 4.11D, ELL blot). Again, since p300 is known to affect transcription, the depletion of this important factor may have perturbed the transcriptional machinery, causing ELL to be reduced. Hence, to confirm that the effect of p300 on ELL is independent of its effect on RNA Pol II-mediated transcription, we also checked the levels of ELL mRNA in control scramble shRNA, as well as p300 knockdown cells. Our expectations were that since ELL is putatively regulated at the protein level by p300, upon its depletion, at most ELL mRNA would remain unchanged. However, the results obtained surprised us when we saw that in cell lines where

p300 had been stably knocked down, ELL transcript levels were significantly increased (Fig. 4.11 E)! This observation, which was quite opposite to our expectations baffled us but confirmed that the ELL-p300 axis functions at the protein level, leaving the transcription machinery undisturbed. Since all our experiments with respect to p300 and ELL have, until now been performed in HEK-293T, we also wanted to understand whether this mode of regulation was universal or specific to a particular cell type. With this in mind, we also generated stable p300 knockdown cells in a HCT116 background and performed western blots to find out if ELL was affected in a similar way in these cells. indeed, our results show that when p300 is stably depleted in HCT116 cells, endogenous ELL expression decreases significantly in these cell lines too. Thus, p300 serves to act as a factor essential for regulating ELL protein within cells. Backtracking to our earlier data, we find that ELL, by itself, is a mightily stable protein and remains unaffected by even 24 hours of cycloheximide treatment. However, the overexpression of HDAC3 does cause ELL to degrade within 8 hours of treatment. Thus, if p300 indeed contributes to ELL stability, it should also have some effect on its degradation rate. To find this out, we performed cycloheximide chase assays in scramble control and p300 knockdown cell lines for up to 12 hours. Within that time point, we found that while, like before ELL was unperturbed in the control cells, the depletion of p300 caused ELL to be degraded by about half within the duration of the experiment (Fig. 4.11G and line graph in the bottom panel for quantification). Thus, ELL, which is itself a stable protein and remains fairly constant within cells, is readily destabilized by the depletion of p300. Our results thus indicate that the stability of the RNA Pol II elongation factor, ELL, is highly dependent on the lysine acetyltransferase, p300, which regulates it at the protein level.

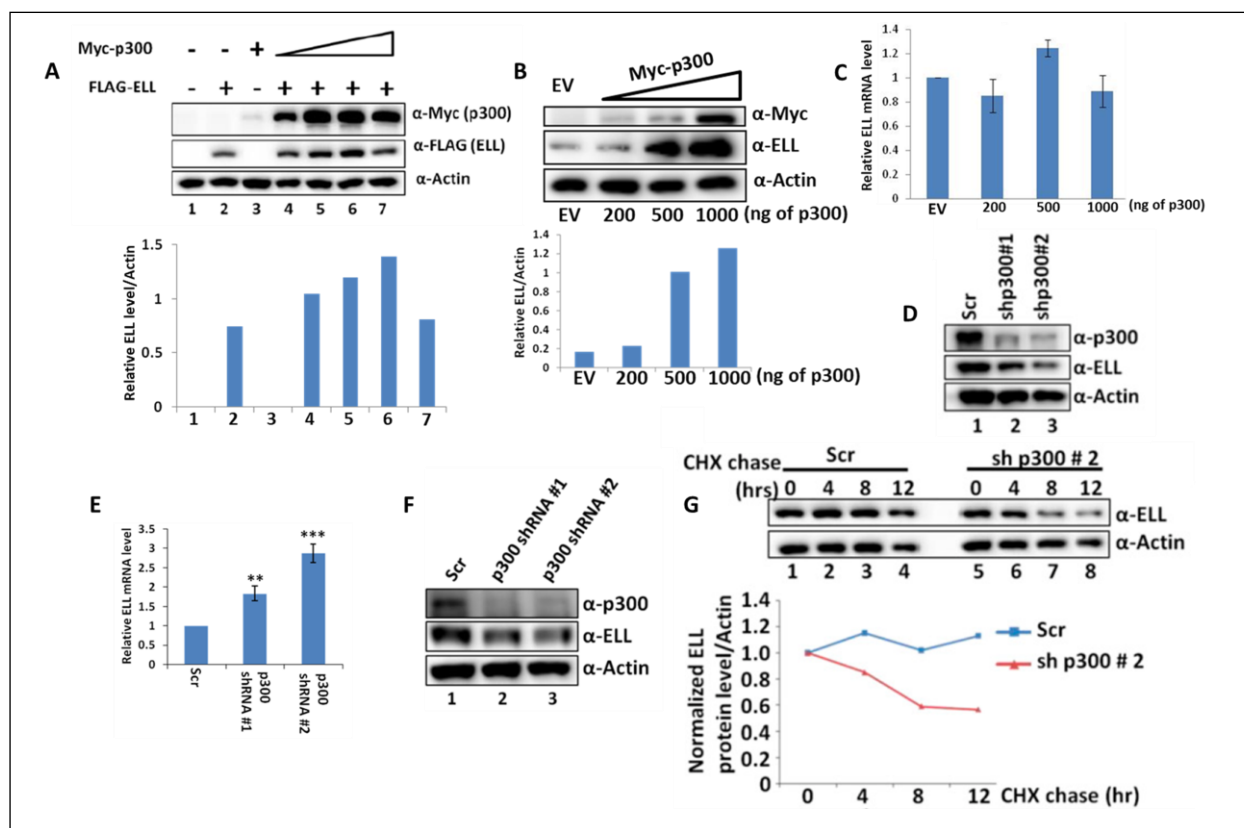


Fig 4.11 ELL is stabilized by p300 within cells

A. Western blots show that ectopically expressed ELL levels are greatly increased when p300 is co-expressed within mammalian cells. The lower panel represents the quantification of the ELL protein signal with respect to actin. **B.** Overexpression of p300 also significantly enhances endogenous ELL protein as seen in these immunoblots. The quantification of the ELL protein signal relative to Actin is represented by the bar graph in the lower panel. **C.** qRT-PCR analyses from cells where p300 had been overexpressed show no changes in the mRNA levels of ELL even though its protein concentrations are highly increased. **D.** Immunoblots from whole cell extracts of control scramble as well as p300 knockdown cell lines verify that p300 has indeed been diminished in these cell lines, and, additionally, there is also a drastic reduction in ELL protein levels. **E.** Knockdown of p300 by two different shRNAs in HEK-293T cells reduces endogenous ELL protein expression but has no effect on the transcription of ELL mRNA as indicated by these qRT-PCR assays. In fact, ELL transcripts were found to be significantly enhanced in these cell lines. **F.** To address a cell type specific bias in our observations, p300 was also stably knocked down in HT116 cells as shown in these immunoblots. Here also, the endogenous ELL protein was seen to significantly decreased. **G.** Stable knockdown of p300 greatly enhances the rate of ELL degradation as indicated by this cycloheximide chase assay. The line graph at the bottom represents the change in ELL protein upon CHX treatment with time.

The qRT-PCR data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, **denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

4.12 HDAC3 can act as the deacetylase for ELL

We have observed p300 mediated acetylation being able to stabilize ELL and increase its protein levels within cells. However, depletion of endogenous p300 renders ELL susceptible to a faster rate of degradation. If the presence of an acetyltransferase provides protein stability, then, when a deacetylase is present, the protein should be rendered unstable and get degraded. This is the exact pattern we observe for p300 and HDAC3 with respect to ELL. Hence, if p300 is the acetyltransferase for ELL, then it is highly likely that HDAC3 is its deacetylase. We chose to test this hypothesis first within cells. hence, we co-expressed ELL along with sub-optimal concentrations of HDAC3, so that the ELL would not be degraded in the presence of the HDAC3. We then immunoprecipitated the ELL under high salt conditions and tested for its acetylation status using a pan acetyl-lysine-specific antibody. The results obtained in our experiment show convincingly that acetylation of ELL within cells decreases significantly with the overexpression of HDAC3 (Fig. 4.12A, Compare the acetylation signal between lanes 1 and 3-6. The quantification of the acetylation signal relative to the immunoprecipitated ELL levels are represented in the bar graph below). It could, however, be possible that the decrease in ELL acetylation is because of diminished p300 expression upon HDAC3 overexpression. However, western blotting to detect endogenous p300 levels clearly indicate that it remains constant across samples (p300 blot in the input panel). Thus, it could be affirmed quite convincingly that HDAC3 causes ELL to get deacetylated within mammalian cells. In spite of this, however, we were looking to get some direct evidence of our hypothesis, particularly when no other possible interfering factors would be present in the milieu. For this, we decided to test if ELL could be deacetylated by HDAC3 *in vitro*. This experiment needed a two-step protocol. In the first step, we incubated ELL with the HAT domain of p300 for a period of time, following which the bead-bound, now acetylated ELL was washed rigorously to remove all associated contaminants. This acetylated ELL was then used as a substrate for deacetylation by incubating with increasing concentrations of HDAC3 in a reaction buffer. As the immunoblots in Fig. 4.12B show clearly, as more and more HDAC3 was used in the reaction mix, the greater was the deacetylation of ELL (See the Acetyl-lysine blots, comparing lane 1 vs lanes 2-4). Thus, HDAC3 could successfully cause deacetylation of its substrate, ELL, under *in vitro* conditions. These results prove that HDAC3 indeed can deacetylate ELL both *in vitro* as well as within cells. We also intended to confirm that HDAC3 does indeed affect endogenous ELL under natural circumstances. We tested this by using the HDAC3 stable knockdown cells at our disposal. We immunoprecipitated the endogenous ELL from the protein extract of these cells and tested its acetylation status using the acetylated lysine antibody. We had earlier observed that the depletion of HDAC3 enhanced endogenous ELL protein, but the ELL acetylation signals in these cell lines were also seen to be significantly enhanced when compared to control scramble cells. Thus, it is quite evident from our

results that the presence of HDAC3 causes deacetylation of ELL which subsequently leads to its degradation.

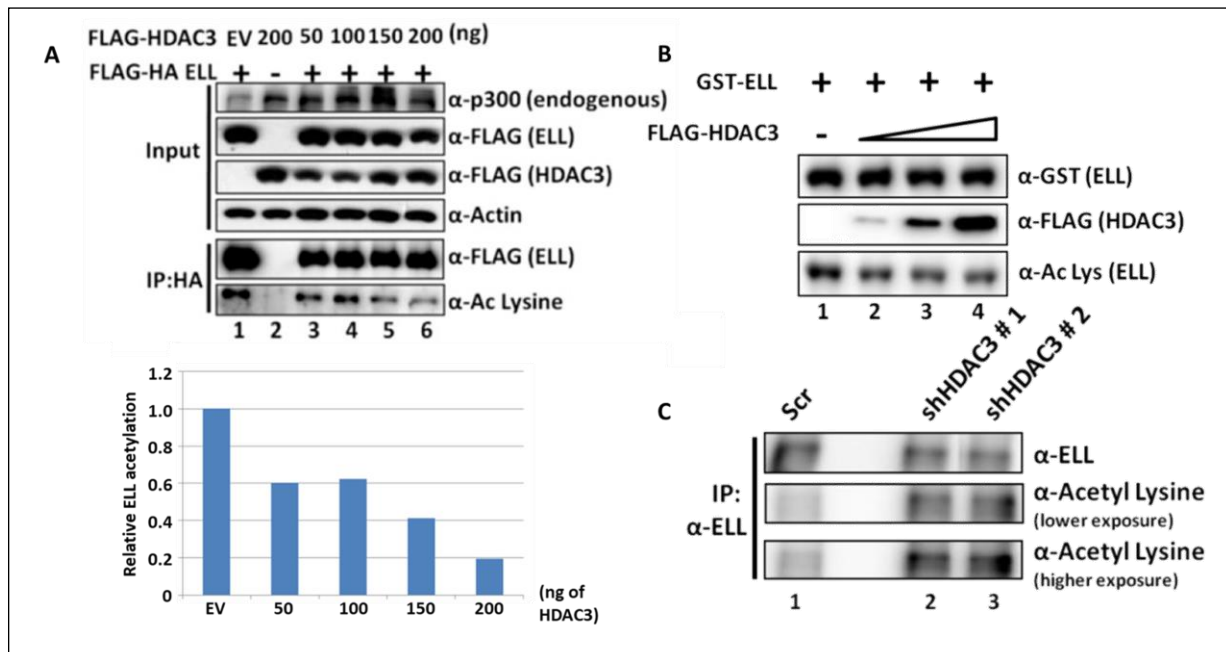


Fig 4.12 HDAC3 acts as the deacetylating enzyme for ELL

A. Western blots show that ectopically expressing HDAC3 in HEK-293T cells diminishes co-expressed ELL acetylation levels when present in sub optimal concentrations. The bar graph below represents the relative ELL acetylation signal with respect to ELL protein immunoprecipitated. **B.** An *in vitro* deacetylation assay with priorly acetylated ELL shows that when purified HDAC3 is present, ELL deacetylation occurs. **C.** Immunoprecipitation of endogenous ELL from control scramble and HDAC3 knockdown cells, followed by probing with a acetylation specific antibody shows ELL acetylation to be increased when HDAC3 has been depleted.

4.13 DBC1 and p300 act in opposition to HDAC3 for the maintenance of ELL acetylation levels

The lysine acetyltransferase p300 acetylates ELL and provides it stability. This acetylation of ELL provides it stability and enables it to be expressed ubiquitously within cells. In the presence of HDAC3, ELL gets deacetylated and then degraded, by mechanisms hitherto unknown. However, it is interesting to note that both p300, as well as the human DBC1, can stabilize/increase ELL within the cell. As both of these two dissimilar proteins have the same output, and DBC1 has been previously described as having the potential to protect the acetylation status of p53, we wondered if the same would hold true for ELL. In simple terms, we were curious to know what would happen to ELL acetylation in the presence of DBC1. To get an answer to this question, we co-expressed both ELL as well as DBC1 in combinations as indicated in Fig. 4.13A. The overexpressed ELL was then immunoprecipitated using FLAG-tag specific agarose beads under high salt conditions, and the acetylation status of this ELL was tested in the presence and absence of DBC1. The results we obtained clearly show that the acetylation

signal of ELL is greatly enhanced in when DBC1 is co-expressed, confirming that indeed DBC1 acts to protect ELL acetylation (Fig. 4.13A, Compare ELL acetylation signal in lane 1 vs lanes 3 and 4. The bar graph in the right panel also represents a quantification of the ELL acetylation signal in the indicated lanes of the immunoblots). However, if we dig deeper into this mode of regulation, could we see DBC1 being capable of rescuing HDAC3 mediated deacetylation of ELL? As we have already observed the DBC1 preventing ELL-HDAC3 association, thus this seems to be an entirely plausible scenario. We tested this by co-expressing ELL and HDAC3 in HEK-293T cells, with the possible rescue of ELL acetylation being carried out by transfection of increasing concentrations of DBC1. The input blots of this assay clearly showed that ELL levels were enhanced in presence of DBC1, depleted when HDAC3 was expressed, and somewhat rescued upon DBC1 co-expression (Fig. 4.13B, ELL Input blot). Upon performing immunoprecipitation of this ELL, and probing with specific antibodies, we find that ELL acetylation, which increased when only DBC1 was expressed, diminished in the presence of HDAC3 (See lanes 1 vs 2 vs 3 of the acetyl-lysine blot in the IP panel). However, and expectantly, the acetylation levels were restored when DBC1 was co-expressed along with ELL and HDAC3 (Mark the increase in acetylation signals in lanes 4-6, and the quantification indicated by the bar graphs on the right). This is a clear indication that DBC1 is able to maintain acetylation levels of ELL within cells, as well as render ELL stable by maintaining and enhancing its acetylation levels. As DBC1 can prevent HDAC3 mediated ELL deacetylation, it is expected that p300, the acetyltransferase for ELL, would also be able to do so. We also tested this by essentially utilizing the same setup as the previous experiment, but substituted p300 in place of ELL. In this experiment as well, the input blots showed ELL to be enhanced upon p300 overexpression and decreased in the presence of HDAC3. The immunoprecipitated ELL also displayed a similar trend as seen previously. A robust acetylation signal was observed when p300 was overexpressed, which diminished significantly in the presence of HDAC3. However, upon co-expression of p300 as well as HDAC3, the ELL acetylation was again seen to be restored significantly (Fig. 4.13D, Acetylation signal in the IP panel). None of our observations are important though, if what we see under overexpression conditions cannot be reproduced in the context of endogenous proteins. Hence, we decided to test endogenous ELL acetylation upon overexpression of an empty plasmid vector, and DBC1. After transfection of these two plasmids, we pulled down endogenous ELL with antibodies and tested for the effect of DBC1 expression on HDAC3 association with ELL as well as its acetylation level. Our results clearly show the presence of DBC1, and that it interacts with endogenous ELL (Fig. 4.13E, DBC1 blot, compare lane 3 vs lane 4). As the ELL-DBC1 interaction increased, we observed an impaired association between endogenous ELL and HDAC3 (Lane 3 vs lane 4 of the HDAC3 blot). Following decreased HDAC3 interaction and increased DBC1 association, there was a concomitant enhancement in the ELL acetylation signal, reflecting the opposite role of these two proteins in regulating this modification. Our studies, until now, thus show

an interplay of interactions between four different proteins, i.e. ELL, DBC1, HDAC3, and p300, where p300 mediated ELL acetylation renders it stable within mammalian cells. HDAC3 acts as the deacetylase of ELL, leading it to its degradation. However, the novel interacting partner of ELL, known as DBC1, prevents this from happening and maintains ELL acetylated state by competitively binding to ELL and inhibiting HDAC3 association. A cartoon diagram representing this mechanism is shown in Fig. 4.13E.

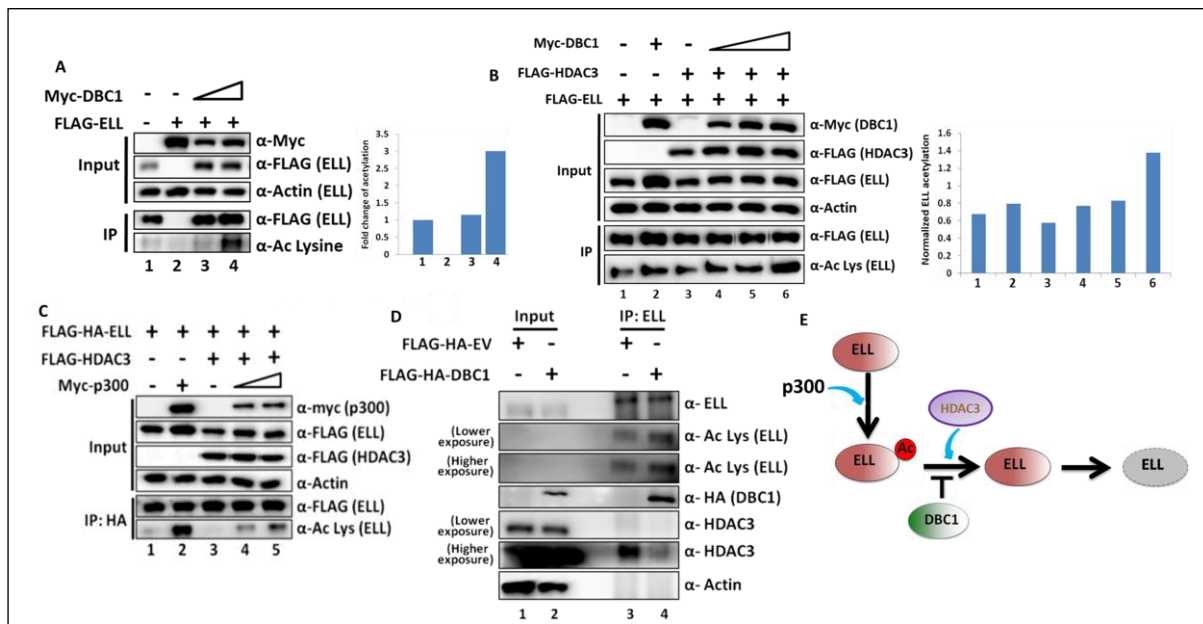


Fig 4.13 DBC1 and p300 prevent HDAC3 action and maintain acetylated ELL within cells

A. Western blots show that ectopically expressing DBC1 levels significantly enhances the acetylation of ELL in HEK-293T cells. The panel on the right represents the quantification of the ELL acetylation signal with respect to immunoprecipitated ELL. **B.** Overexpression of DBC1 significantly enhances ELL acetylation as well as prevents its deacetylation that would erstwhile be carried out by HDAC3 within mammalian cells. The quantification of the ELL acetylation signal relative to immunoprecipitated ELL is represented by the bar graph on the right. Numbers indicate the lane numbers as shown in the immunoblots. **C.** Like DBC1, overexpressing p300 also rescues ELL deacetylation by HDAC3 and thus contributes to higher numbers of acetylated ELL species being present in the cell. **D.** Transfection of DBC1 within mammalian cells dramatically decreases the association of endogenous ELL with HDAC3 as well as leads to ELL acetylation being increased. **E.** A cartoon diagram to summarize the findings until this part. The diagram shows a novel acetylation of ELL by the acetyltransferase p300, which can then be removed by the action of HDAC3. HDAC3 mediated ELL deacetylation leads to its degradation, which is prevented by DBC1 by competitively binding to ELL and thus maintaining its acetylated state.

4.14 ELL is subjected to proteasome-mediated degradation

ELL is an acetylated protein within the cell, a reaction which is catalyzed by the lysine acetyltransferase p300. The acetylated population of ELL is maintained by DBC1, which prevents HDAC3-ELL association through competitive binding. Because of this association, HDAC3 is unable to deacetylate ELL, which

would subsequently lead to its degradation. HDAC3 not only has the capability to destabilize ELL, but can also increase the rate of its degradation. As we have observed ELL levels to decrease upon HDAC3 overexpression, we were curious to know how this occurs within the cell. Protein degradation within the cell occurs by proteasomal degradation or lysosomal proteolysis. Of these, the most common method of protein level control is proteasome-mediated degradation. To determine if HDAC3 mediated ELL degradation occurs through this pathway, we first co-expressed ELL with HDAC3. 40 hours post-transfection, cells were then treated with increasing concentrations of MG132, which is a proteasome degradation inhibitor, for another 8 hours. Following this, ELL protein levels were tested by immunoblotting. As the results in Fig 4.14A indicate, ELL is significantly reduced upon overexpression of HDAC3. However, upon treatment with MG132, even though HDAC3 is present, ELL protein was significantly rescued (See the change in ELL protein in lane 2 vs lane 3, and compare the rescue in lanes 4-6. The quantification of ELL in the different lanes is denoted by the bar graph in the lower panel which clearly shows a significant increase when MG132 is administered). Degradation by the proteasome involves tagging a protein with ubiquitin molecules, a phenomenon termed as polyubiquitination/ polyubiquitylation. This ubiquitin conjugated protein is then targeted to the proteasome. As we had already observed HDAC3 mediated ELL degradation being rescued by MG132, we were curious to know whether, in the presence of HDAC3, ELL would be polyubiquitylated. Hence, after transfection of cells with EL and HDAC3, followed by MG132 treatment, the ELL was then subsequently immunoprecipitated and then its ubiquitination status was checked by using an anti-ubiquitin antibody. If ELL is polyubiquitylated, and since it is already being expressed as FLAG epitope-tagged, then probing with an anti-FLAG antibody should display characteristic polyubiquitination signals. This is exactly what we have observed in our assays as well. When ELL alone is transfected, it does not show long smears of ubiquitination. However, co-expression of HDAC3 does indeed result in the formation of polyubiquitylation “trails” in our western blots (Fig. 4.14B, lane 3 of ELL blot in the IP panel). With the addition of MG132, more ubiquitylated ELL species accumulate within the cells, as their degradation is inhibited, and hence, in this case, the ubiquitination signals are much more robust (lane 4). Probing with an anti-FLAG antibody would not be enough to prove our point, as such “smears” in western blots also appear due to post-translational modifications such as sumoylation. To confirm whether HDAC3 does indeed cause ELL ubiquitylation, we also probed these samples with an anti-ubiquitin antibody, observing the exact same results as our anti-FLAG immunoblot (Anti ubiquitin blot, compare the signals in lanes 1 vs 3 and 3 vs 4 in the IP panel). Thus, ELL does indeed get ubiquitylated within the cell in presence of overexpressed HDAC3. However, we had also observed that the presence of an acetyltransferase renders ELL stable, and that HDAC3, which degrades ELL, can also cause its deacetylation. We thus wished to understand if both these events are independent or form two parts of the same process. Hence, we co-expressed ELL along with sub-optimal

concentrations of HDAC3 for the purpose of testing both the acetylation as well as the ubiquitylation of ELL under these circumstances. As seen previously, we did observe robust ELL polyubiquitylation upon co-expression of HDAC3. Additionally, when this immunoprecipitated ELL was probed with an anti-acetyl lysine antibody, we also have observed that the acetylation of ELL diminishes (Fig. 4.14C, See the increase in ubiquitylation signals in lane 1 vs lanes 3-6, while the acetylation decreases in these same lanes). Reconciling these two separate observations thus confirms that the deacetylation of ELL by HDAC3 leads to its polyubiquitylation and subsequent degradation by the proteasome.

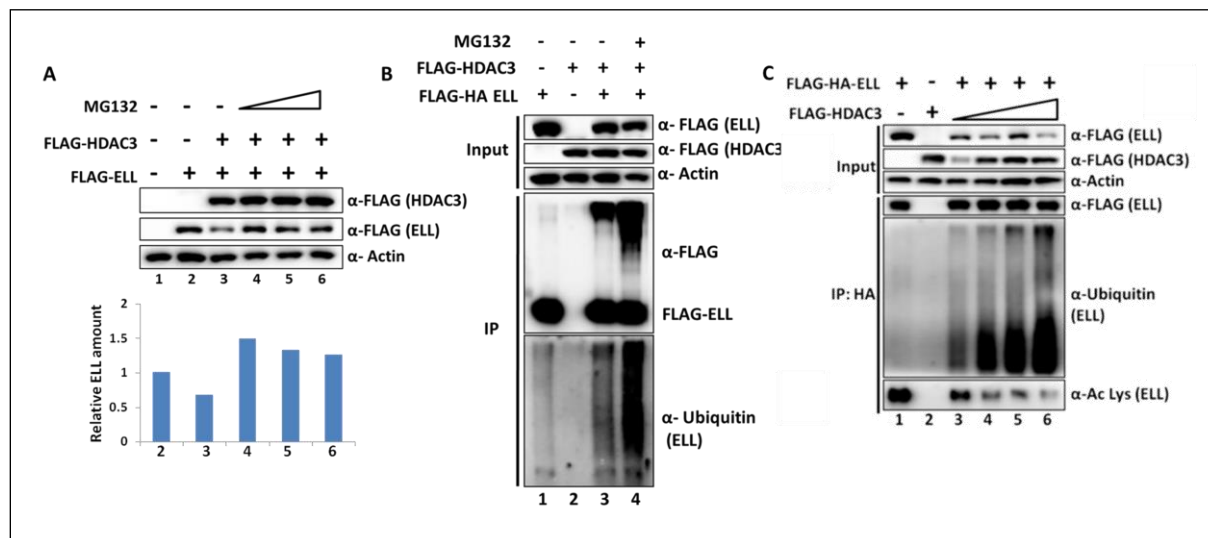


Fig 4.14 ELL is degraded by the proteasome within mammalian cells

A. Western blots show that HDAC3 mediated ELL degradation can be rescued by the treatment of cells with MG132, indicating that the degradation of ELL occurs via the proteasome. The panel on the bottom represents the quantification of the ELL protein signal as observed in the different marked lanes with respect to Actin. **B.** Overexpression of HDAC3 significantly enhances ELL polyubiquitylation as shown by this western blot assay. The treatment of cells with MG132 further increases the polyubiquitylation signal indicating that ELL indeed is polyubiquitylated in presence of HDAC3. **C.** The overexpression of HDAC3 causes ELL degradation by deacetylating it, which in turn increases its polyubiquitination as well, as seen in these immunoblots.

4.15 DBC1 and p300 decrease ELL polyubiquitination by increasing its acetylation

ELL, the RNA Pol II elongation factor, is a mostly stable protein within cells. So much so, that a 24-hour treatment of cells with cycloheximide, an inhibitor of translation, also has no effect on ELL levels. However, in the presence of excessive HDAC3, ELL is targeted towards degradation by the proteasome. HDAC3 mediated ELL degradation first involves deacetylation of ELL, which renders ELL protein less functional and unstable. We have observed earlier that by maintaining the acetylation of ELL within cells, DBC1 and p300 act in opposition to HDAC3. P300 keeps ELL acetylated and DBC1

maintains this state by not allowing HDAC3-ELL association. Thus, it is also expected that these two proteins, which cause ELL to remain acetylated, would result in its decreased ubiquitylation. We intended next, to perform some experiments to test this hypothesis. But first, we focused on another point, where we wanted to test if by bypassing the overexpression of HDAC3, and only upon the transfection of ubiquitin, would a) ELL get degraded, and b) if DBC1 or p300 could prevent this degradation. We first focused our attention on the interplay between DBC1 and ubiquitin. Overexpression of a ubiquitin-containing construct resulted in a significant decrease in ELL protein levels as seen by western blotting (Fig. 4.15A, compare lane 3 vs lane 5 in the ELL blot). However, as observed in the case of HDAC3, upon ectopically overexpressing increasing concentrations of DBC1, we observed the ubiquitin-mediated degradation of ELL to be significantly rescued (Compare ELL expression in lane 5 vs lanes 6-8). Next, we repeated this experiment using essentially the same setup but replaced DBC1 with a plasmid expressing p300. In this experiment as well, ELL was significantly destabilized in the presence of ubiquitin, but the overexpression of p300 caused significant restoration of ELL protein as indicated in Fig. 4.15B (ELL blot, lane 5 vs lanes 6-8). We know from our earlier experiments that in the presence of p300 and DBC1, ELL gets and stays acetylated. Putting these two observations together, we can conclude that even if HDAC3 is not overexpressed, the degradation of ELL by ubiquitin is possibly prevented because the presence of these two proteins maintains ELL in an acetylated state. Thus, we may presume that ELL stability would be dictated by its post-translationally modified state. With this theory in mind, we next wanted to test the relevance of this hypothesis. So, first, we tested the interplay between HDAC3 and DBC1 in affecting ELL modifications. In our transfection-based assay, we overexpressed ELL, and ELL along with HDAC3, and then all three plasmids, ELL, HDAC3, and DBC1 were transfected. The input blots obtained reproduced observations we had seen earlier, wherein DBC1 was able to rescue HDAC3 mediated ELL destabilization, possibly because of its ability to competitively bind to ELL and displace HDAC3 (Fig. 4.15C, Compare ELL protein in the input panel). We had also observed HDAC3 increasing ELL deacetylation and causing its increased ubiquitination (Fig. 4.14C). Contrary to this, when DBC1 is also overexpressed, we found the acetylated ELL species to increase significantly, and along with this, there was a sharp decrease in its polyubiquitylation (Fig. 4.15C, See the acetylation and ubiquitylation signals in lanes 1-3, and note the opposite trend between acetylation and ubiquitylation signals). Next, we checked if this same observation would be reproduced in presence of p300 as well, since p300 has the ability to acetylate ELL directly. Once again, as observed previously, HDAC3 mediated ELL degradation decreased dramatically when p300 was overexpressed (Fig. 4.15D, ELL blot in the input panel, lane 2 vs lanes 3-5). When we tested the acetylation status of ELL across these samples, we found, as expected, that ELL was deacetylated when only HDAC3 was present, but the overexpression of p300 rescued the ELL deacetylation, causing the acetylated ELL species to increase within the cell. Interestingly, even though

HDAC3 was present, the overexpression of p300, which resulted in increased ELL acetylation, also displayed a concomitant loss in polyubiquitylation (Note the acetylation and corresponding ubiquitylation signals in lane 2 vs lanes 3-5). Indeed, in the higher concentrations of expressed p300, the polyubiquitylated ELL species was almost fully rescued in our assay (Lane 6 of the ubiquitin blot, and compare its signal to lane 1). As noted previously in this section itself, the overexpression of both DBC1 and p300 could rescue ELL from ubiquitin-mediated degradation. For testing the ubiquitination as well as acetylation status of ELL in these samples, we performed immunoprecipitation of ELL followed by western blotting with specific antibodies. We noted a significant polyubiquitination signal in the presence of ubiquitin, but this rapidly was lost when DBC1 was also overexpressed. The overexpression of DBC1 also resulted in a concomitant increase in the acetylation of ELL (Fig. 4.15E, Note the exactly opposite trends of the acetylation and ubiquitin signals in lane 1,2 and 3 in the IP panel). A similar observation was also seen when the ubiquitylation state of ELL was probed in presence of ubiquitin and p300. As p300 is the acetyltransferase of ELL, it is indeed expected that the acetylated species of ELL would dramatically increase upon overexpression of p300. Additionally, we observe, that as more ELL moieties are present in an acetylated state, their overall ubiquitin levels decrease even though ubiquitin has been overexpressed in our assay (Fig. 4.15F, IP panel). Our results thus indicate two different post-translational modifications of the same protein, ELL, leading to drastically different results. While in the presence of DBC1 and p300, ELL stays acetylated, leading to its stability and maintenance of a steady-state, the overexpression of HDAC3 results in its deacetylation, which finally causes its increased polyubiquitylation and subsequent degradation. Thus, an interplay between three proteins, i.e. DBC1, p300, and HDAC3 occurs for the regulation of levels and function of another protein, ELL, within the cell.

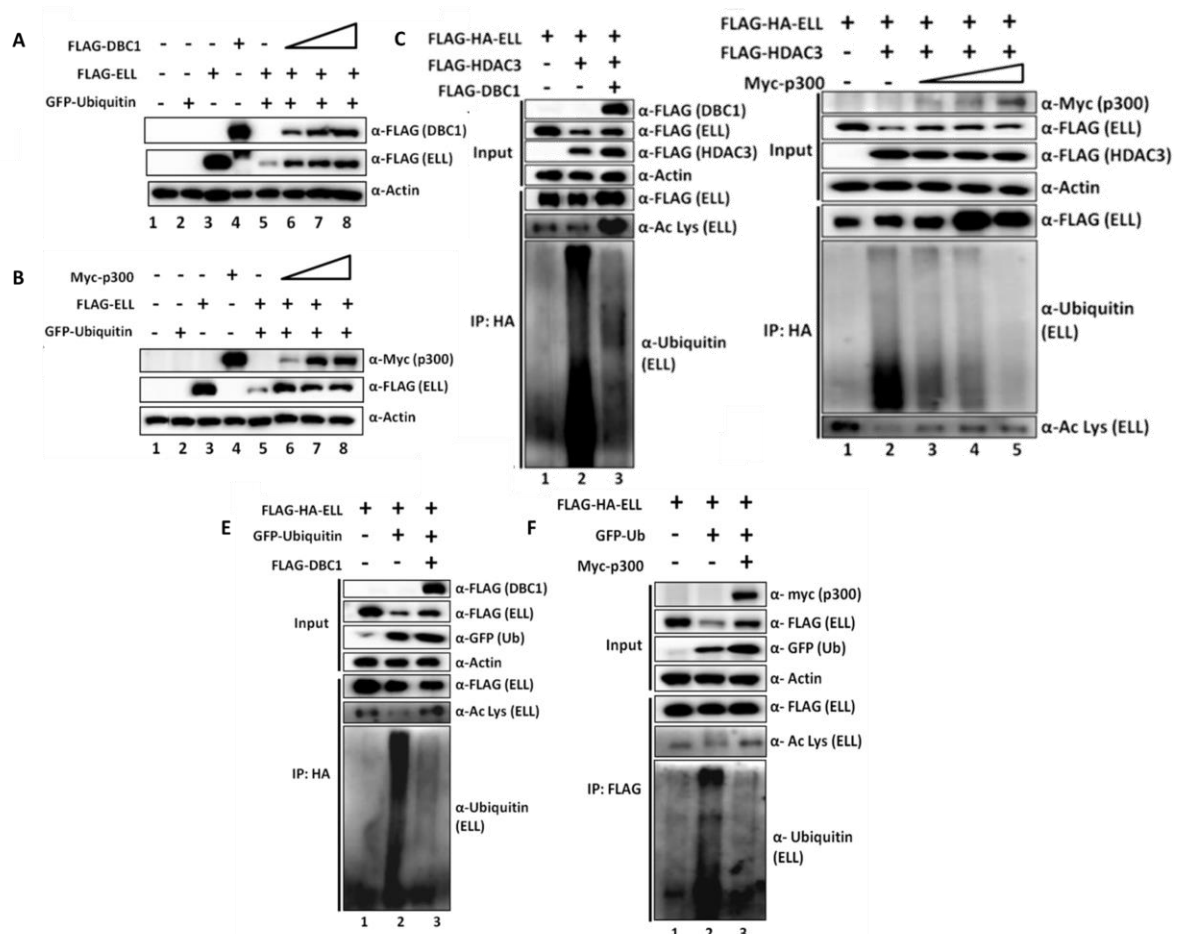


Fig 4.15 DBC1 and p300 increase ELL acetylation and decrease its ubiquitylation within cells

A. Western blots show that ubiquitin mediated ELL degradation can be rescued by the overexpression of DBC1. **B.** Overexpression of p300 significantly reduces ELL degradation in the presence of ubiquitin as shown by this western blot assay. **C.** The overexpression of DBC1 causes increase in acetylated ELL species which are less prone to ubiquitylation as indicated by this experiment. **D.** Introduction of p300 in the presence of ELL and ubiquitin results in enhanced ELL acetylation, causing decreased polyubiquitylation. **E.** ELL polyubiquitylation in the presence of ubiquitin decreases when DBC1 is overexpressed because of maintenance of acetylated ELL within the cell. **F.** Immunoblots show that ELL acetylation increases with a decrease in its polyubiquitylation when p300 is expressed alongside ubiquitin in HEK-293T cells.

4.16 Two lysine residues at the N-terminus of ELL, K5, and K29 are targets for ubiquitylation

Upon understanding that the acetylation of ELL results in its decreased polyubiquitylation, we were curious to gain some mechanistic insight into this regulation. For this, we sought to identify amino acids on ELL which would be the targets for conjugation with ubiquitin. Using the reagents at our disposal, we first tried to identify the deletion fragment of ELL (which we had used earlier for our interaction analysis) which would be sensitive to ubiquitin-mediated degradation. For this, we

overexpressed each deletion fragment along with the full-length ELL and compared the changes in their expression upon transfection of ubiquitin. As the western blot assay in Fig. 4.16A shows, the full-length ELL (1-621), was found to be highly sensitive to ubiquitin-mediated degradation (Compare ELL expression in lane 1 vs lane 2 and Actin as loading control). The subsequent challenge of the other deletion mutants of ELL with ubiquitin showed that whereas the ELL (1-500) and ELL (1-373) deletion mutants were degraded significantly upon ubiquitin overexpression, the fragments in which the first 44 amino acids of ELL were deleted (namely the 45-373 aa and 45-621 aa fragments) were fully resistant to the presence of ubiquitin (Compare ELL expression in lane 3 vs 4, lane 5 vs 6, and mark the difference of these two fragments compared to lane 7 vs 8 and lane 11 vs 12). This analysis indicates beyond doubt that the target residues for ubiquitination on ELL lie within the first 44 amino acids at its N-terminus. Hence, next, to identify possible amino acid targets for ubiquitination, we turned our attention to some *in silico* analysis. Here, we obtained the amino acid sequences of ELL from different species and aligned them using Clustal Omega, in order to identify amino acids that would be conserved across species. It is well known that the tagging of ubiquitin occurs mostly on lysine residues on proteins, and our analyses showed that only three such residues occurred within the first 60 amino acids on human ELL (Fig. 4.16B). Alignment of the amino acid sequence of human ELL with ELL sequences for different species, such as *Mus musculus*, *Rattus norvegicus*, and *Danio rerio*, we identified two lysine residues that were conserved across all species, located at position 5 (K5) and 29 (K29). To understand which of these residues would be the target for ubiquitin-mediated degradation, we generated point mutants of these amino acids, replacing the lysines with alanines (K → A). These mutants were then co-transfected with increasing concentrations of ubiquitin within HEK-293T cells. Western blot assays performed for changes in ELL expression upon ubiquitin co-transfection yielded some interesting results. The expression of wild-type ELL, as expected, was seen to decrease considerably when ubiquitin was co-expressed (Fig. 4.16C, Top panel, Compare ELL in lane 2 vs lanes 4-6). Challenging the ELL (K5A) mutant with ubiquitin showed that this mutant was somewhat less sensitive to ubiquitin, but however, at higher concentrations, partial degradation could be observed when ubiquitin was overexpressed (Fig. 4.16C, middle panel). Next, when we co-expressed the ELL (K5A, K29A) double mutant along with increasing concentrations of ubiquitin, we observed by immunoblotting that even at the highest levels of ubiquitin overexpression, the ELL protein remained unaffected (Fig. 4.16C, Bottom panel, Observe no changes in ELL signal in lane 2 vs lanes 4-6). This experiment, in totality, shows that both these lysine residues at the N terminus of ELL are indeed targets of ubiquitin conjugation. However, we also wanted to see whether this double mutant of ELL indeed would be resistant to polyubiquitylation. Hence, we performed two sets of parallel transfections. In the first set, ELL (WT) was co-expressed with increasing concentrations of ubiquitin, while the second set contained ELL (K5A, K29A) mutant along with ubiquitin as well. Western blot

analysis showed that whereas the wild-type ELL was significantly destabilized by ubiquitin overexpression, the ELL mutant protein remained largely unaffected, as seen previously as well (Fig. 4.16D, ELL blot in the input panel). Immunoprecipitation of this ELL and probing with an anti-ubiquitin antibody, for visualization of polyubiquitination “trails”, showed that while the wild type ELL was robustly polyubiquitylated upon overexpression of ubiquitin, the K5A, K29A mutant showed almost no difference in ubiquitination across all the samples, irrespective of the absence or presence of ubiquitin (Fig. 4.16D, Ubiquitin blot in the IP panel). As overexpression of HDAC3 also has been seen to increase ELL polyubiquitylation, we also tested the sensitivity of this double mutant to HDAC3 dependent ubiquitylation. Our results showed that not only was the double mutant significantly more resistant to HDAC3 mediated degradation (Fig. 4.16E, ELL blot in the input panel), in the presence of HDAC3, the polyubiquitination of this mutant was also dramatically low when compared to the ELL wild type construct. Thus, taken together, we can indeed conclude that the ELL is targeted for proteasomal degradation by conjugating two lysine residues at its N terminus, i.e K5 and K29 with ubiquitin.

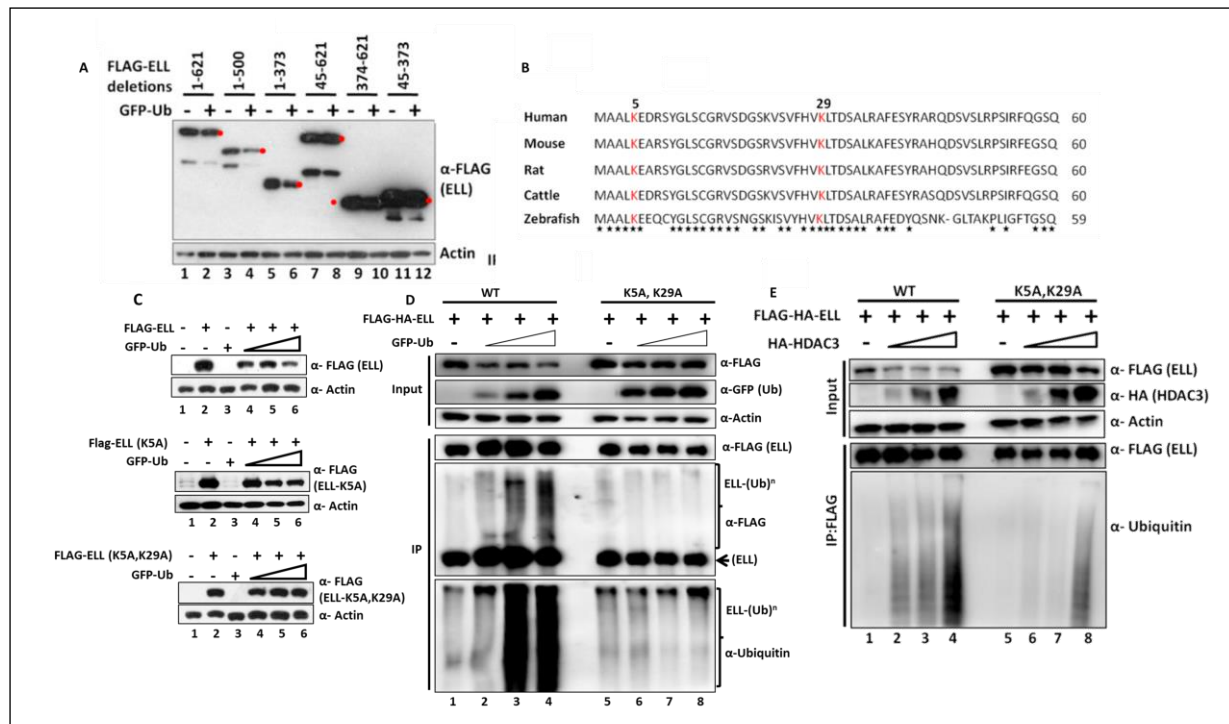


Fig 4.16 ELL is ubiquitylated at two lysine residues at its extreme N terminus

A. Western blots show that ELL deletion fragments containing the first 44 amino acids are sensitive to ubiquitin mediated degradation within mammalian cells. **B.** Multiple sequence alignment of amino acid sequences of ELL protein across different species shows the presence of two conserved lysine residues at the 5th and 29th position. **C.** The overexpression of ubiquitin causes degradation in the ELL wild type, but a mutant with the lysines modified at position 5 and 29 remains unaffected. **D.** Upon overexpression of ubiquitin, ELL (WT) gets heavily ubiquitylated, but the ELL (K5A, K29A) point mutant does not. **E.** In a similar assay, while the ELL (WT) is susceptible to HDAC3 mediated ubiquitination, the state of the ELL (K5A, K29A) mutant remains unchanged.

4.17 The lysines K5 and K29 are also the predominant targets of p300 mediated ELL acetylation

We have understood from our earlier observations that there is an interconnection between the acetylation and ubiquitylation of ELL. Thus, our experiments conclude that the acetylated ELL is stable and cannot be ubiquitylated, whilst, upon deacetylation, the same ELL now becomes prone to proteasomal degradation. Thus, there exists a possibility that the acetylation, as well as ubiquitylation, occurs on the same residues of ELL. To test this hypothesis, we first decided to screen different deletion mutants of ELL for their potential to get acetylated. The mutants were cloned in such a way that 100 amino acids were serially deleted from the C terminus of each fragment. One mutant was devoid of the N terminal 60 amino acids. The expression and purification of the different ELL deletion fragments are shown in Fig. 4.17A. As p300 has been established as the acetyltransferase of ELL, we performed *in vitro* acetylation assays with the different ELL domains as the substrate and p300 HAT as the enzyme. Our results from this *in vitro* assay indicate that all fragments containing the N terminal amino acids were heavily acetylated, in amounts comparable to the full length (Fig. 4.17B, Compare acetylation signal in lanes 3 to 7). However, the deletion of the N terminal 60 amino acids caused this acetylation signal to be significantly decreased when compared to the full-length ELL (Lane 9). Thus, it is safe to conclude that ELL acetylation mediated by p300 occurs predominantly at its N terminus. Now, we have established earlier that only two lysine residues, at K5 and K29, are evolutionarily conserved across species. Thus, given this trend of our results, we expect that these lysines will also be the targets for p300 mediated acetylation. In order to test this, we co-expressed ELL (WT) and ELL (K5A, K29A) along with increasing concentrations of p300 within mammalian cells. Immunoprecipitation of this ELL and subsequent blotting with a pan acetyl lysine antibody showed a minimal increase in acetylation signals in the K5A, K29A mutant as compared to the wild type, which displayed robust acetylation (Fig. 4.17 C, Note the change in acetylation signal between lanes 1 vs 2-3, and compare with lane 4 vs lanes 5-6 in the IP panel). To exclude the possibility of interference of other cellular factors, and to confirm our observations, we also purified FLAG-ELL (WT) and FLAG-ELL (K5A, K29A) mutant by overexpression from HEK-293T cells. These bead-immobilized proteins were incubated with the p300 HAT domain for an *in vitro* acetylation reaction. Here also, we observed, that similar to our results obtained within cells, while the wild type ELL was adequately acetylated because of the presence of the p300 HAT domain, the K5A, K29A mutant showed almost no acetylation at all. These results indicate that the same two lysine residues, present at the 5th and 29th positions of the ELL protein, are the predominant targets of p300 mediated acetylation. While because of our *in vitro* experiment, we cannot rule out the possibility that other acetylation targets may exist on the ELL protein, but the most important ones, particularly involved in the regulation of ELL stability, are found

at its extreme N terminus. Since DBC1 maintains ELL acetylation state and renders it stable, we also wanted to confirm if indeed the acetylation of these two lysine residues is protected by DBC1. Hence, we co-expressed the ELL wild type and the ELL K5A, K29A mutant along with increasing concentrations of DBC1. These samples were subjected to immunoprecipitation with a pan acetyl lysine antibody after immunoprecipitation. The wild-type ELL, like before, demonstrated a significant enhancement in its acetylated species (Fig. 4.17E, lane 2 vs lanes 4-6 of the acetyl-lysine blot in the IP panel). However, even upon overexpression of a significantly high concentration of ELL, the acetylation levels of the K5A, K29A double mutant did not change significantly. Incidentally, a luciferase reporter assay performed with ELL wild type and the K5A, K29A mutant shows that the wild type ELL is able to positively modulate transcription more significantly than the mutant. This observation confirms the importance of these two lysine residues and demonstrates the impact of their differential modification in the regulation of transcription. Thus, our experimental results show that it is these two lysine residues, present at the 5th and 29th position of ELL, that are the targets for acetylation by p300, and when they get deacetylated in the presence of HDAC3, are conjugated with ubiquitin for proteasomal degradation.

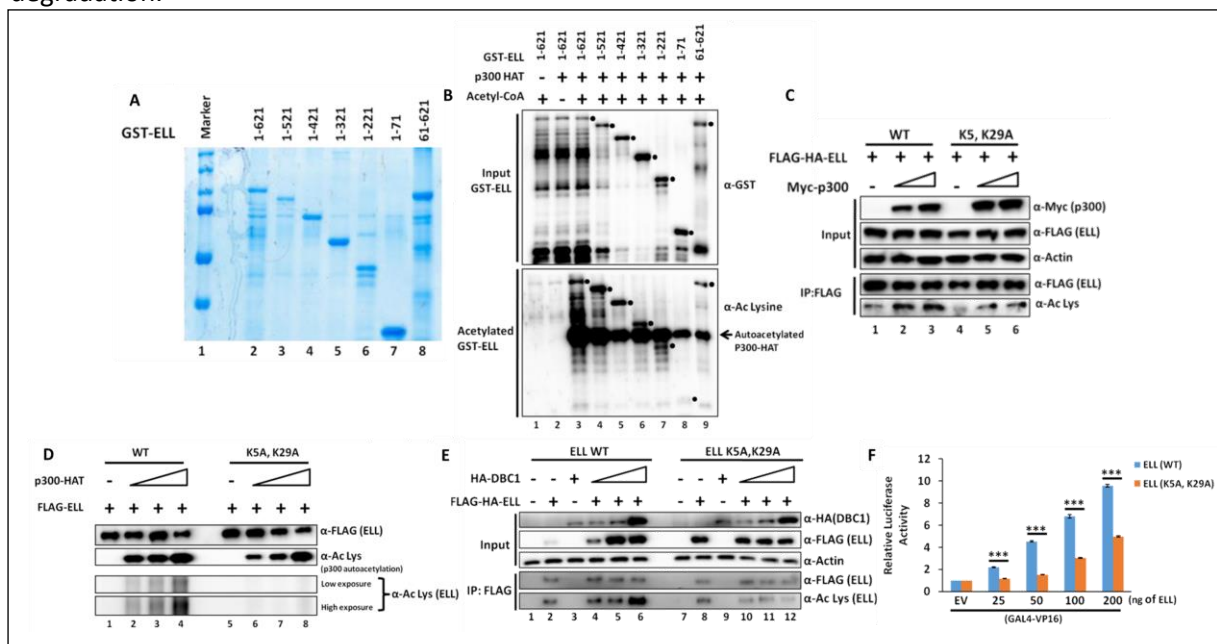


Fig 4.17 ELL K5 and K29 are the predominant targets for p300 mediated acetylation

A. Coomassie staining shows the expression and purification of ELL full length and its different (serially deleted) domains for use as substrates in our in vitro acetylation assay. **B.** In vitro acetylation assays show that the presence of the first 60 amino acids of ELL results in its robust acetylation in the presence of p300. Deletion of this region causes a significant reduction in ELL acetylation. **C.** The ELL K5A, K29A mutant is acetylation impaired even when expressed in cells along with p300, as indicated in these immunoblots. **D.** An in vitro acetylation reaction performed with ELL wild type and its K5A, K29A mutant in the presence of p300 HAT domain shows that while the wild type ELL is significantly acetylated, the double mutant does not show any acetylation at all. **E.** Lysines 5 and 29 of ELL are the targets for DBC1 mediated maintenance of ELL acetylation as shown by the immunoblots in this experiment. **F.** Luciferase assays indicate that the wild type ELL is much more potent in stimulating transcription than its K5A, K29A mutant counterpart.

4.18 Siah1 acts as the E3 ligase for ELL

Our detailed studies have demonstrated the presence of an intricate regulatory network within cells comprising of multiple proteins for the regulation of the levels and function of the RNA Pol II elongation factor, ELL. We have found that ELL is acetylated at the N terminus, which provides it stability and causes it to be functionally active. The ELL acetylation is maintained by DBC1, in the absence of which, deacetylation occurs, which leads to ELL degradation. The key element missing in our studies is the identification of the E3 ligase for ELL. One such E3 ligase, named Siah1, is often found associated with members of the super elongation complex. We noted a recent study that has also described Siah1 as the E3 ligase for ELL2. However, this study failed to notice any discernable effect of Siah1 on ELL levels. Based on our observations, we reasoned that the cause of this observation might have been the natively acetylated status of ELL within cells. Thus, we hypothesized that Siah1 mediated ELL degradation would only occur when ELL was deacetylated, i.e. in the presence of HDAC3. We tested our theory by performing a transfection-based assay in HEK-293T cells. Here, ELL was transfected alone, or in combination with increasing concentrations of Siah1. The results obtained showed that, as reported previously, Siah1 overexpression could not alter ELL protein levels at all (Fig 4.18A, Note similar ELL expression in lane 2 vs lanes 3-5). Next, we transfected HDAC3 along with ELL, and, as seen in our earlier assays, there was a significant destabilization of ELL protein (Compare ELL in lane 2 vs lane 6, and quantitation in the lower panel). Upon maintaining this fixed amount of HDAC3, and the introduction of increasing concentrations of Siah1, which had earlier failed to show any effect on ELL protein, we observed a sharp decline in the levels of ectopically expressed ELL (See ELL expression in lane 2 vs lanes 7-9, and quantitation of ELL relative to Actin in the lower panel). This result indicates that the presence of HDAC3 is indeed essential for the destabilization of ELL by the E3 ligase Siah1. However, since it has already been reported that Siah1 by itself can cause the degradation of several factors involved in transcription, such as AFF1 and ELL2, we wanted to confirm that this effect of Siah1 on ELL is not because of impaired transcription due to the overexpression of Siah1. Hence, in this setup itself, we tested the mRNA levels of Siah1 and ELL. As presented in Fig. 4.18B, our assays showed that ELL mRNA levels remained constant, or strikingly, increased significantly upon transfection of cells with Siah1 (See mRNA levels in lane 2 vs lanes 3-9). This solidifies our theory that Siah1 mediated ELL depletion occurs at the protein level and ELL transcription is not impaired upon Siah1 overexpression. As a positive control for our qRT-PCR assay, we also tested the mRNA levels of Siah1 and did observe linearly increasing mRNA levels coinciding with the amounts transfected (Bottom panel of Fig. 4.18B). Thus, our results clearly show the role of Siah1 as the E3 ligase of ELL, which can cause its degradation only when DAC3 is present.

Quickly recapitulating what we have been able to understand so far from our experiments, we see that the RNA Pol II elongation factor, ELL, has two key lysine residues at its N terminus, at positions 5 and 29. These lysines are targets for p300 mediated acetylation, which makes ELL protein stable and functionally able to stimulate transcription. If these critical lysines are deacetylated by HDAC3, it paves the way for conjugation of ubiquitin at these residues itself by the E3 ligase, Siah1, which finally causes ELL degradation by the proteasome. This deacetylation of ELL is prevented by DBC1, which competitively binds and inhibits the ELL-HDAC3 association, thereby preventing its degradation. A cartoon diagram representing this regulation is shown in Fig. 4.18C.

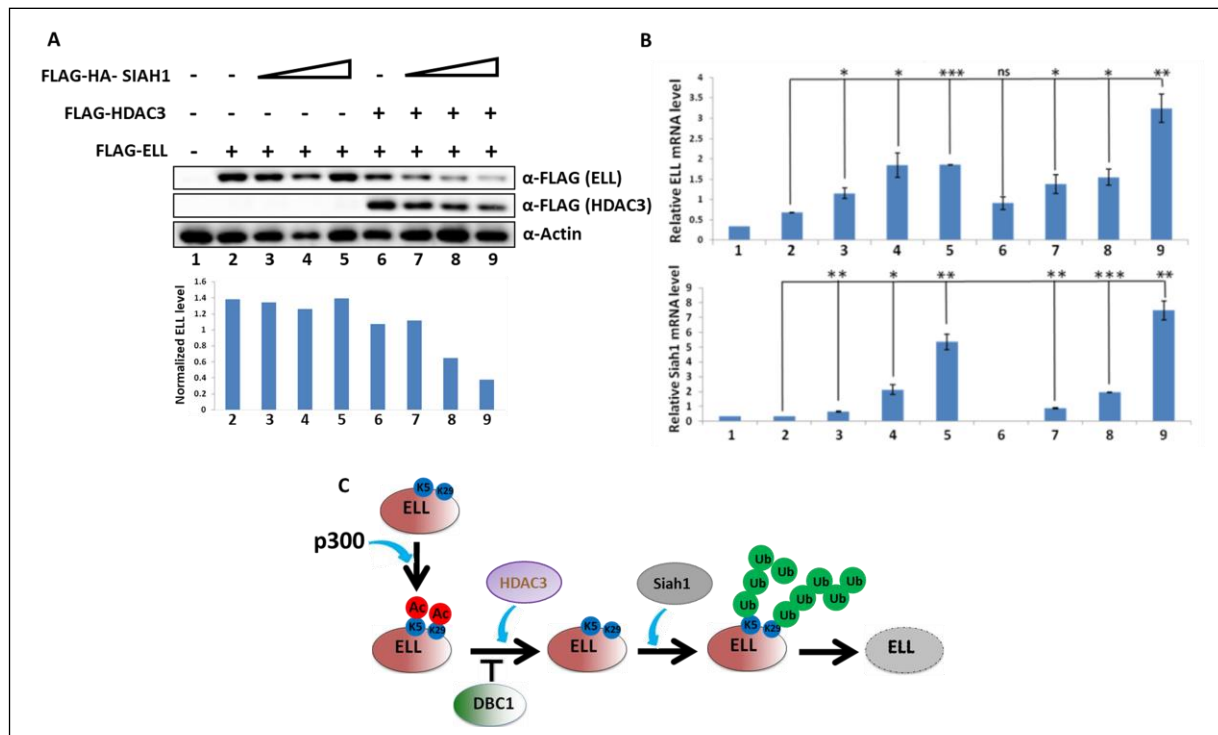


Fig 4.18 ELL is degraded by the action of the E3 ubiquitin ligase, Siah1

A. A western blot based transfection assay shows Siah1 alone does not have any effect on ELL expression. However, co-transfection of HDAC3 along with Siah1 results in rapid degradation of ELL protein. The bar graph in the bottom panel shows the quantification of ELL protein relative to Actin. **B.** qRT-PCR assays indicate that the effect of Siah1 on degrading ELL does not occur at the mRNA level but is restricted to ELL protein expression as ELL mRNA remains unaffected by Siah1 overexpression. As a control for our assay, Siah1 mRNA was seen to increase linearly with increasing concentrations of Siah1 plasmid transfected. **C.** Cartoon diagram showing the multivalent mechanism of ELL regulation. ELL is acetylated by p300, which renders it stable and functionally active. However, HDAC3 has the potential to deacetylate ELL and cause its degradation via the E3 ligase, Siah1. In this milieu, DBC1 acts the protector of ELL by preventing ELL-HDAC3 interaction, thus keeping ELL stable and active within cells.

The qRT-PCR data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, ** denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

4.19 DBC1 and ELL co-regulate diverse sets of genes

Our studies have resulted in the identification of a novel interacting protein of the transcription elongation factor, ELL. This novel interacting protein, DBC1, functions to stabilize and maintain the acetylated form of ELL within cells, thus maintaining the cellular pool of active ELL, and resulting in proper expression of its target genes. Hence, it can be proposed that this DBC1-ELL axis functions to maintain the expression of a specific subset of genes within the cell. In order to thus identify the physiological function of this regulatory network, it is essential to identify genes that are co-regulated by these two proteins. For this purpose, we took the help of published literature and identified two independent studies, one which had identified genes impacted by DBC1 knockdown, and the other which studied genes hampered by the depletion of ELL, both carried out in HEK-293T cells (GSE34104 for ELL knockdown and GSE35480 for DBC1 knockdown from the GEO Database maintained by NCBI)[186, 220]. From each of these two datasets, we obtained the top 10,000 genes that were adversely affected due to the knockdown of ELL or DBC1. These two lists were then cross-referenced to identify the genes that are most affected by the depletion of DBC1 as well as ELL within the cell. Our analysis of the two datasets showed a total of 2,987 genes to be co-regulated by DBC1 and ELL within the cell (Fig. 4.19A). However, as these were *in silico* analyses, we would need to validate them using wet-lab techniques. Hence, we generated cell lines where DBC1 would be stably knocked down. Analysis of the DBC1 levels in these cell lines showed that they were significantly lower than in the control scramble cell line, indicating that the knockdown experiment was successful (Fig. 4.19B, DBC1 panel). Testing the ELL protein level in these cell lines, we also noticed that due to the knockdown of DBC1, the ELL protein expression was also drastically reduced (Note the ELL panel in Fig. 4.19B). This observation also solidifies our claim that DBC1 indeed regulates ELL stability within cells. To rule out this effect as specific to a single cell type, we also generated stable cell lines with DBC1 knocked down in a HCT116 background. Western blot assay for target proteins in these cell lines also showed a significant decrease in DBC1, which in turn also caused the ELL protein levels to be reduced (Fig. 4.19C). Next, we sought to validate the results of our *in-silico* analysis of genes putatively co-regulated by the ELL-DBC1 axis. qRT-PCR assays using cDNA from the control scramble as well as DBC1 knockdown cell lines showed a significant decrease in DBC1 mRNA in the DBC1 knockdown cell lines. However, this effect on DBC1 mRNA did not result in any changes in ELL transcription, proving that the DBC1 mediated ELL regulation occurs at the protein level (Fig. 4.19D). Subsequently, upon checking the effect of DBC1 depletion on ELL target genes, we found that all the genes we tested had diminished transcriptional output upon DBC1 knockdown. This effect was specific as several genes, which are not targets of either of these proteins, such as *NEDD9*, *PTPRF*, *MYO10*, *NFKB2* and *NAV2* were seen to be unaffected by the loss of both DBC1 and ELL. Our qRT-PCR analyses also revealed that

several of these ELL target genes were reported to be involved in the regulation of cell growth and the cell cycle, such as *CCND1*, *CCNE2*, *CDK6*, *MED12*, *SKI*, and *PCNA*. Hence, we wanted to test if there actually was a phenotypic output of loss of DBC1 mediated ELL depletion. Thus, we performed cell proliferation assays in control scramble and DBC1 knockdown cells by seeding equal numbers of cells and monitoring their proliferation rate by counting on different days of growth. Our analyses revealed that upon DBC1 knockdown, the proliferation rate of cells decreases modestly, and this may be due to a result of impaired transcription of genes involved (Fig. 4.19E). One of the questions that could arise from our qRT-PCR analyses is whether ELL itself is responsible for the regulation of expression of these tested genes, or other factors are involved in this process. To answer this question we performed a rescue experiment, wherein we transfected DBC1 knockdown cells, where ELL levels were found to be low, with either an empty vector or a plasmid expressing FLAG-ELL. After extraction of RNA from these cells, we compared the expression of the tested target genes in control scramble, DBC1 knockdown, empty vector-transfected, and ELL overexpressed cell lines by qRT-PCR. Our data reveals that for all the genes we have tested, such as *SKI*, *MDM2*, *KIF1B*, *MED12* and *BCL6* we observe decreased expression upon DBC1 knockdown, that remained mostly unchanged upon transfection with the empty vector plasmid. However, ectopically expressing ELL containing plasmid in the DBC1 knockdown cell line significantly enhanced the mRNA levels of these target genes as compared to the empty vector-transfected samples (Fig. 4.19F). In several cases, the induction in gene expression upon ELL transfection was higher than their expression levels in control scramble cells as well, indicating that the hampered transcription of these genes occurred due to the deficiency of ELL itself, and not because of any other factor. Thus, we can conclude that the DBC1-ELL axis is involved in the regulation of the expression of a wide variety of target genes within cells.

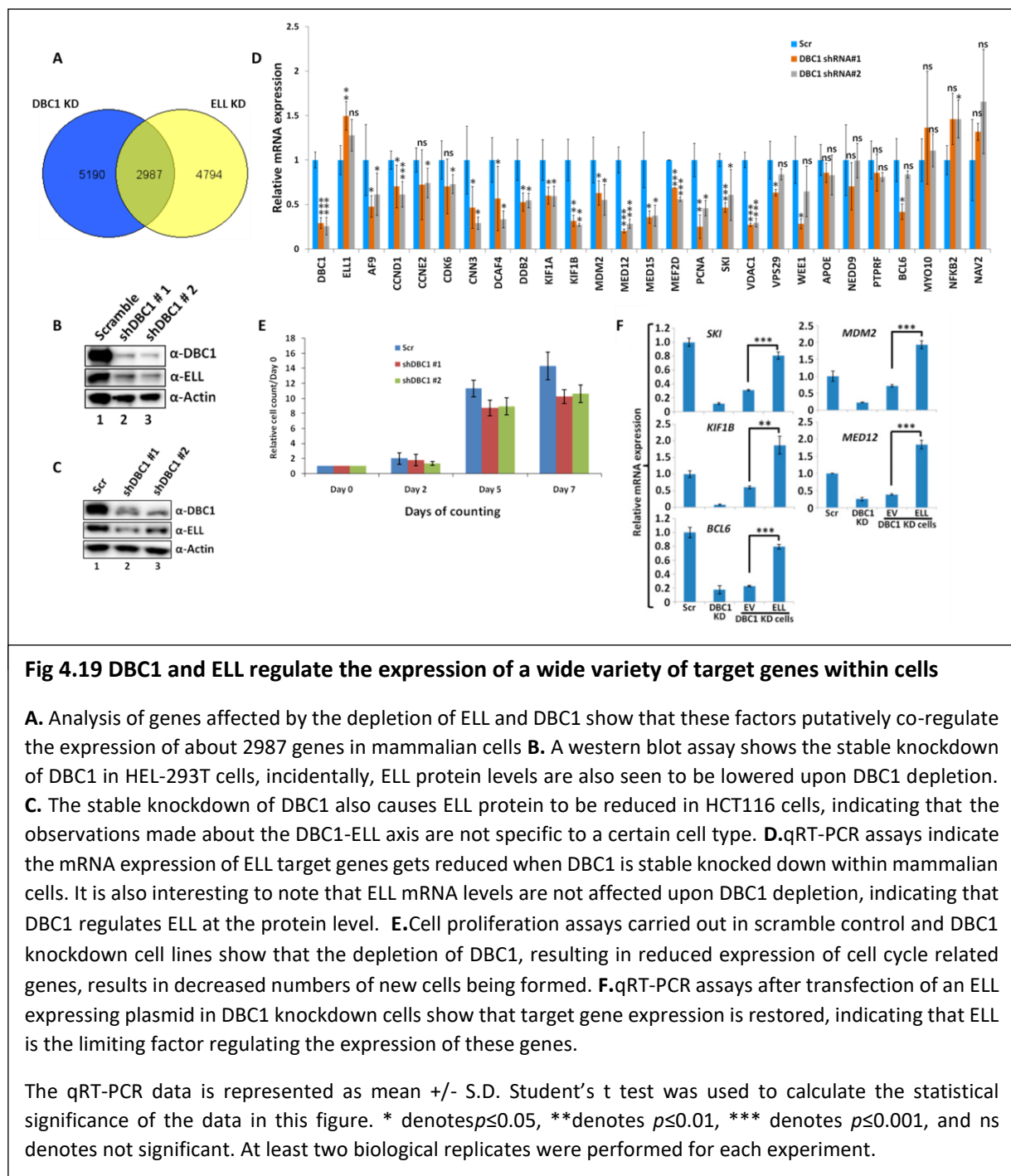


Fig 4.19 DBC1 and ELL regulate the expression of a wide variety of target genes within cells

A. Analysis of genes affected by the depletion of ELL and DBC1 show that these factors putatively co-regulate the expression of about 2987 genes in mammalian cells **B.** A western blot assay shows the stable knockdown of DBC1 in HEL-293T cells, incidentally, ELL protein levels are also seen to be lowered upon DBC1 depletion. **C.** The stable knockdown of DBC1 also causes ELL protein to be reduced in HCT116 cells, indicating that the observations made about the DBC1-ELL axis are not specific to a certain cell type. **D.** qRT-PCR assays indicate the mRNA expression of ELL target genes gets reduced when DBC1 is stable knocked down within mammalian cells. It is also interesting to note that ELL mRNA levels are not affected upon DBC1 depletion, indicating that DBC1 regulates ELL at the protein level. **E.** Cell proliferation assays carried out in scramble control and DBC1 knockdown cell lines show that the depletion of DBC1, resulting in reduced expression of cell cycle related genes, results in decreased numbers of new cells being formed. **F.** qRT-PCR assays after transfection of an ELL expressing plasmid in DBC1 knockdown cells show that target gene expression is restored, indicating that ELL is the limiting factor regulating the expression of these genes.

The qRT-PCR data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, **denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

4.20 Depletion of DBC1 and ELL results in the inefficient formation of the transcription complex

While trying to understand the physiological significance of the DBC1-ELL axis, qRT-PCR based assays have shown that the knockdown of DBC1, which results in lowered expression of ELL, causes decreased transcription of a diverse set of target genes. These genes range in function from cell growth, proliferation, cell cycle control, apoptosis, as well as proteins involved in transcription. In order to understand how this regulation occurs, we performed chromatin immunoprecipitation (ChIP) analyses for the recruitment of various factors in the vicinity of the transcription start site (TSS) of

genes whose expression was adversely affected by the depletion of DBC1. The knockdown of DBC1 causes a loss in cellular ELL, thus it was expected that its recruitment would be hampered at the TSS of the target genes. Our results show that for all the target genes we have tested, ELL recruitment was indeed decreased in DBC1 knockdown cells (Fig. 4.20A, Topmost panel). An earlier study has shown that the ELL functions to recruit SEC components at the transcription start site, and we have also observed similar results while testing for the presence of CDK9 (as a representative of the SEC) on these genes. These results prove that the loss of ELL causes decreased SEC recruitment at target genes leading to their impaired transcription. Interestingly, when we performed ChIP analyses for the presence of RNA Pol II as well as its phosphorylated forms, we noticed that upon DBC1 knockdown, the RNA Pol II recruitment at the TSS also underwent significant reduction for all target genes (Fig. 4.20A). RNA Polymerase II phosphorylated at its Ser 2 and Ser 5 residues is a mark of Pol II actively engaged in transcription elongation. Reduction in the presence of these two forms of RNA Pol II also indicates that upon knockdown of DBC1, and subsequent depletion of ELL, transcription is hindered at the target genes additionally due to the loss of Pol II. Thus, the depletion of DBC1 impairs transcription due to reduced recruitment of SEC components as well as RNA Pol II. This effect, that we observe on DBC1 knockdown is not a general phenomenon, but is specific for only DBC1-ELL co-regulated genes, as a parallel analysis of factor recruitment at genes not regulated by ELL, viz. *NEDD9* and *PTPRF* show no adverse effect of the reduction of DBC1 or ELL on the presence of ELL, CDK9, or even RNA Pol II at these genes (Fig. 4.20B). It is probably due to this reason itself that the expression of these genes does not change when ELL levels in the cell are reduced as shown in Fig. 4.19D.

As DBC1 was seen to interact with several members of the transcription machinery, we also wanted to check if the loss of DBC1 causes any changes to the components of the pre-initiation complex. To test this, we also performed chromatin immunoprecipitation for TBP (a part of the TFIID complex) recruitment at the TSS of the target genes. Our assays revealed that upon knockdown of DBC1, the presence of TBP was also significantly diminished for all target genes tested (Fig. 4.20A, Bottom most panel). Thus, our results enable us to conclude that when DBC1 is depleted, a loss of the pre-initiation complex as well as the functional SEC, and RNA Pol II, at the transcriptional start site of target genes results in subdued transcription. To gain some more insight into DBC1 mediated regulation of TFIID or SEC recruitment, we decided to perform some more experiments. First we focused on DBC1 regulating TFIID assembly. For this, first, we decided to test if the protein levels of the TFIID complex components would be affected by DBC1. Hence, we simply checked, by western blot analysis, if TFIID components were affected upon DBC1 knockdown. Our analyses revealed that almost all TFIID complex members, such as TAF1, TAF5, TAF6, and TAF7 displayed significantly decreased protein levels when DBC1 was depleted (Fig. 4.20C). Interestingly, however, we found that TBP remained unaffected by the knockdown of DBC1. To gain some mechanistic insight into the regulation of the TFIID assembly by

DBC1, we also performed qRT-PCR assays to assess the effect of DBC1 depletion on the mRNA levels of the TAF subunits. Our analyses show that although the mRNA levels of most of the TAFs remain unperturbed, the transcript levels of TAF3 and TAF6, in particular, are significantly decreased (Fig. 4.20D). Using TAF3 as an example, we tested if its decreased transcriptional output was due to reduced ELL in DBC1 knockdown cells by performing a complementation experiment wherein an ELL expressing plasmid was transfected in DBC1 knockdown cells. The results obtained indicated that the overexpression of ELL could rescue the diminished TAF3 transcripts that were erstwhile reduced upon DBC1 knockdown (Fig. 4.20E). These datasets prove indeed that the recruitment and assembly of the TFIID is indeed dependent on DBC1, as well as ELL. Since TAF6 is known to form a scaffold upon which the other TFIID members rest for associating with each other, its decreased expression upon loss of DBC1 and ELL would effectively inhibit the formation of the intact functional TFIID complex. Recapitulating to our initial observations, wherein we had noticed the ability of DBC1 to associate with several components of the super elongation complex, such as AFF1, AF9, EAF1, EAF2, and CDK9, we also wondered if DBC1 would also have the ability to regulate the expression and functions of these proteins. A western blot based assay to detect the protein levels of these SEC members show that with the stable knockdown of DBC1, the expression of several key components of this essential complex, such as AF4 (also known as AFF1), AFF1, and CDK9 are also significantly reduced (Fig. 4.20F). Like TAF6 forms the scaffold for the TFIID complex, AFF1 and AFF4 have also been described to act as the scaffold for the SEC. Thus, their decreased protein levels on DBC1 knockdown may be the reason for decreased transcriptional output in DBC1 depleted cells, due to inefficient formation of the SEC. Strikingly, our assays showed that the expression of the ELL interacting protein, EAF1, was enhanced, indicating that there may exist a reciprocal relationship between DBC1 and EAF1 (Fig. 4.20F, EAF1 blot). Interestingly, when we tested if DBC1 affected the expression of these key SEC components at the transcript level, we found that the mRNA levels of none of the genes tested were perturbed (Fig. 4.20G). This observation, similar to the one we have had for the ELL-DBC1 interplay, suggests that DBC1 may be central to regulating eukaryotic transcription by modulating the stability and functions of key components of this tightly regulated machinery.

Our extensive studies thus lead us to propose a model of eukaryotic transcription regulation, wherein ELL, a key transcription elongation factor, upon acetylation at its N-terminus by p300, becomes active and causes efficient RNA Pol II mediated transcription of its target genes. On the other hand, if ELL is deacetylated by HDAC3, it then becomes a target for Siah1 mediated ubiquitination and proteasomal degradation, which significantly impairs the expression of diverse sets of genes within the cell. DBC1, on the other hand, helps to maintain ELL protein within the cell by inhibiting its association with HDAC3, and thus preventing its degradation. A cartoon diagram describing this mode of eukaryotic transcriptional regulation is shown in Fig. 4.20H.

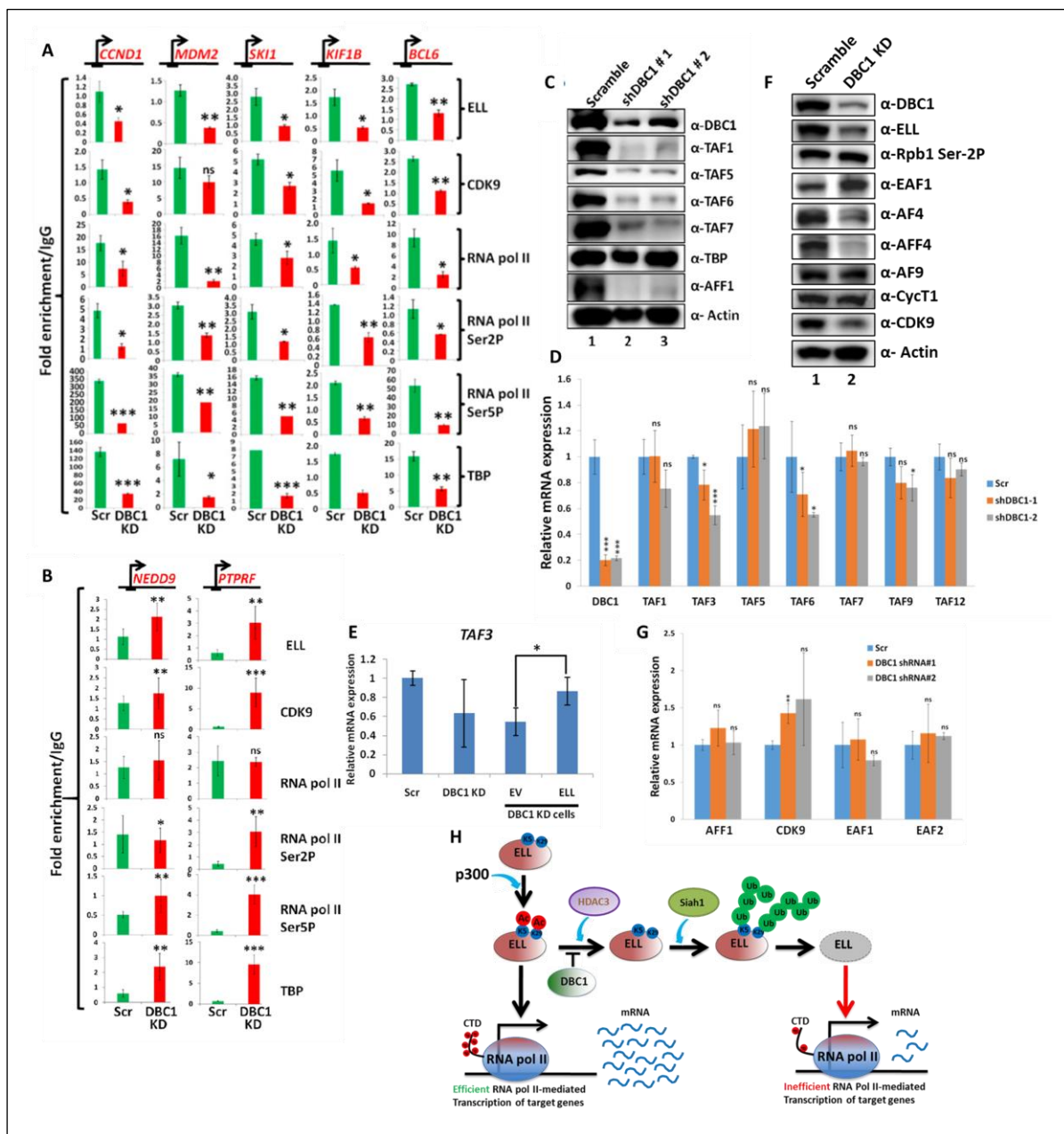


Fig 4.20 DBC1 and ELL regulate the formation of a functional transcription complex

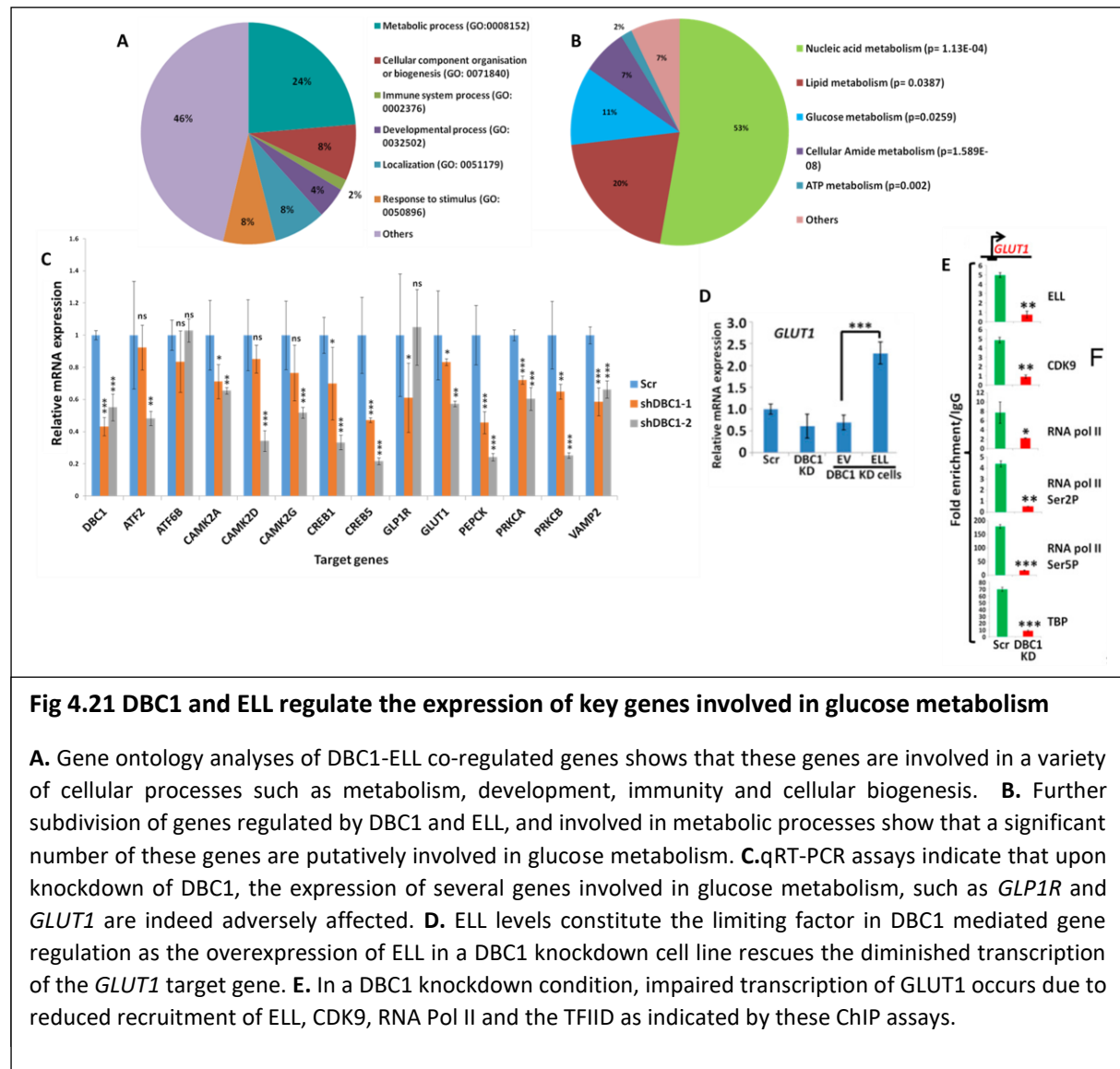
A. Chromatin immunoprecipitation assays of several key components of the eukaryotic transcription machinery show the knockdown of DBC1 results in the reduced recruitment of ELL, CDK9, RNA Pol II and TBP (of the TFIID complex) at the transcription start site of ELL target genes. **B.** DBC1 depletion results in reduced presence of the transcriptional machinery only on ELL target genes as the transcription of non-target genes are unaffected by the reduced DBC1 expression. **C.** Immunoblots showing that the stable knockdown of DBC1 results in reduced protein levels of several components of the TFIID complex **D.** qRT-PCR assays indicate that DBC1 knockdown affects transcript levels of TAF3 and TAF6 of the TFIID complex. **E.** Overexpression of ELL within DBC1 knockdown cells results in the rescue of TAF3 transcription which was seen to be significantly decreased upon DBC1 depletion. **F.** The stable knockdown of DBC1 causes reduced expression of several key SEC components such as AFF1, AFF4, and CDK9 indicating that DBC1 may play an important role in regulating the expression of these proteins. **G.** DBC1 knockdown does not affect the transcript levels of SEC components as seen in this qRT-PCR based assay, indicating that the regulation of these components by DBC1 occurs at the protein level. **H.** The proposed model for the described mechanism of transcription regulation wherein ELL, acetylated by p300, enables efficient transcription of its target genes. This acetylated form of ELL is maintained by the action of DBC1 protein. Upon HDAC3 mediated deacetylation, ELL is subsequently tagged for degradation by the proteasome.

The qRT-PCR data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, ** denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

4.21 The ELL-DBC1 axis regulates key genes involved in glucose metabolism

Detailed studies performed by us have resulted in gaining deep mechanistic insights into the regulation of eukaryotic transcription by modulating the stability and functions of the key RNA Pol II elongation factor, ELL, by the action of several different proteins. After deciphering the players in this intricate network, we wondered if there would be any physiological relevance of this tightly regulated system. To understand this, we went back to the 2,987 genes that we had found to be co-regulated by DBC1 and ELL within cells. Using these genes as a template, we performed gene ontology analyses using the DAVID software, with the aim being to try and classify genes into functions they would regulate within the cell. The results obtained showed that of the 2,987 genes under study, about 717 (~24%) were involved in metabolic processes. Other categories of processes regulated by these common target genes included the organization of cellular components, development, and regulation of the immune system (about 240 genes in each category) (Fig. 4.21A). As DBC1 has already been reported to be involved in the regulation of different metabolic processes, we chose to focus on this category for further analysis. So, in the next step, taking these 717 genes seen to be involved in metabolic processes, we further sub-classified them according to the processes they affected. The results we obtained turned out to be highly interesting. Detailed classification of these genes showed that while half of them would be involved in nucleic acid metabolism (including transcription and replication), about 20% of the genes (~ 143 genes) were putatively involved in lipid metabolism (Fig. 4.21B). As the role of DBC1 in lipid metabolism has been widely discussed, we decided not to focus our attention on this category. Interestingly, about 78 of these genes (amounting to 11% of the total) were seen to be involved in glucose metabolism pathways. As no study had yet described a role for DBC1 or ELL in this process, we chose to focus our attention on these target genes. Hence, in the first step, we tested, by qRT-PCR analysis, whether the knockdown of DBC1 would indeed adversely affect the transcript levels of these target genes. Our studies showed that all the metabolic genes we targeted were indeed regulated by DBC1, as its depletion resulted in a diminished transcriptional output of these genes (Fig. 4.21C). We further tested if DBC1 by itself affected the transcription of these genes, or if ELL would be the rate-limiting factor in this case as well, by performing a complementation assay in DBC1 knockdown cells. These cells, when transfected with an ELL expressing plasmid, showed markedly enhanced rescue of transcription of *GLUT1*, a glucose transporter, indicating that in these cellular processes as well, ELL regulated by DBC1 is the key player

(Fig. 4.21D). Indeed, chromatin immunoprecipitation analyses of recruitment of target factors at the TSS of *GLUT1* showed that upon DBC1 knockdown, the presence of ELL, CDK9, RNA Pol II and TBP decreases significantly, resulting in a markedly lower rate of transcription (Fig. 4.21E). Thus, it is clear from our studies that the DBC1-ELL axis affects the expression of several sets of genes within the cells, and is also involved in the regulation of metabolic processes such as glucose and lipid metabolism.



4.22 Dysregulation of the ELL-DBC1 axis is a potential cause of Type 2 Diabetes

The ELL-DBC1 axis is involved in the regulation of a wide variety of target genes. Our analyses have revealed that these genes include ones regulating transcription, cell growth and proliferation, immunity, and metabolism. Detailed analyses show that several of the ELL target genes are involved in glucose metabolism, with genes related to glucose uptake and transport being specifically affected (Fig. 4.21C). Hence, we decided to take a closer look at our list of ELL-DBC1 co-regulated genes. Here, we now classified the genes according to the disease they may potentially cause when their misregulation occurs. These analyses showed that while the majority of these genes may play a part in cancer development, about 12% of these genes have the potential to cause Type 2 Diabetes if their expression is not modulated properly (Fig. 4.22A). With this new information in our hands, we scoured the available literature to see if there would have been any reports linking any of our studied proteins with Type 2 Diabetes. In this endeavor, we chanced upon a solitary report which stated that when type 2 diabetes occurred, the activity of HDAC3 was seen to be specifically increased[250]. This got us wondering if the enhanced HDAC3 would somehow be responsible for affecting ELL, which in turn would result in the impaired expression of genes involved in glucose homeostasis. We chose to test this hypothesis directly from samples obtained from diabetic patients. Hence, we created a cohort of 30 treatment naïve type 2 diabetes patients and isolated peripheral blood mononuclear cells (PBMCs) from their blood after obtaining informed consent. As a control group, PBMCs were also collected from 30 healthy, age-matched individuals. After obtaining protein extracts from the cells, they were then subjected to western blotting to check for differences, if any, between the protein levels of DBC1, ELL and HDAC3 in healthy and type 2 diabetic samples. Our extensive experiments showed that the protein levels of both DBC1 and ELL were significantly reduced in the diabetic patient samples when compared to the healthy controls. However, the HDAC3 protein levels were the same across all samples tested (Fig. 4.22B). Taking our analyses one step further, we also tested for the mRNA expression of *ELL*, and other genes involved in glucose homeostasis such as *GLUT*, *GLP1R*, and *CREB5*. The results obtained from our qRT-PCR assays show that though ELL transcription remains unaffected, the mRNA levels of *GLUT*, *GLP1R* and *CREB5* are all significantly reduced in the treatment naïve patient samples. Thus, it is quite clear that the DBC1-ELL axis plays an important role in the pathogenesis of type 2 diabetes. From our understanding, we believe that as HDAC3 remains unchanged and DBC1 is seen to be low in the diseased state, the decreased DBC1 results in enhanced susceptibility of ELL to undergo HDAC3 mediated deacetylation, which subsequently leads to its degradation. Decreased ELL results in inefficient transcription of its target genes, resulting in a misregulation in the maintenance of glucose homeostasis in the body. Due to the decreased uptake of glucose-mediated by GLUT1, or

storage by GLP1R, the glucose concentration in the peripheral system may increase rapidly, thus leading to the development of type 2 diabetes. In healthy controls, maintenance of ELL levels leads to regulated transcription of its target genes, causing the glucose balance to be maintained in the body. A model depicting this mode of regulation is shown in Fig. 4.23.

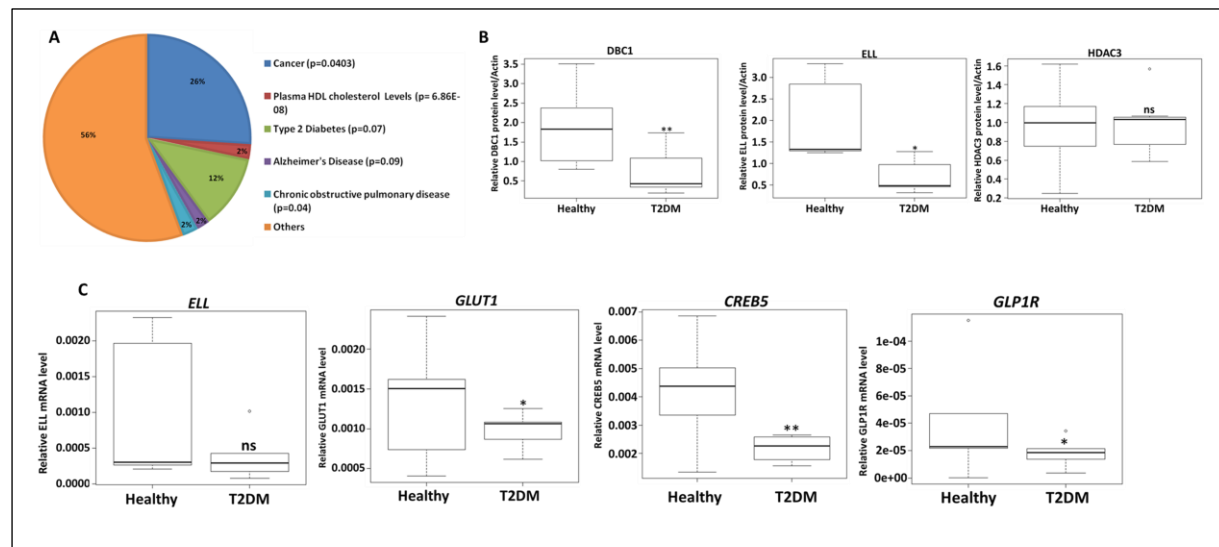


Fig 4.22 Dysregulation in the DBC1-ELL axis potentially causes Type 2 Diabetes

A. Analyses of genes co-regulated by DBC1 and ELL and prediction of diseases they could be associated with shows that misregulation of these genes may cause cancer, type 2 diabetes, Alzheimer's disease or COPD. **B.** Quantification of the levels of different proteins in healthy control vs treatment naïve type 2 diabetic PBMCs (n=30 for each) show that the expression of DBC1 and ELL are markedly reduced upon the development of diabetes. HDAC3 levels, however, mostly remains unchanged across samples. **C.** Transcription of genes involved in the maintenance of glucose homeostasis is significantly reduced in type 2 diabetic patients when compared to healthy controls. ELL transcript levels remain unchanged confirming that the DBC1-ELL axis functions at the protein level. For the qRT-PCR assays, n=25 for each set.

The qRT-PCR data is represented as mean +/- S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, ** denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant.

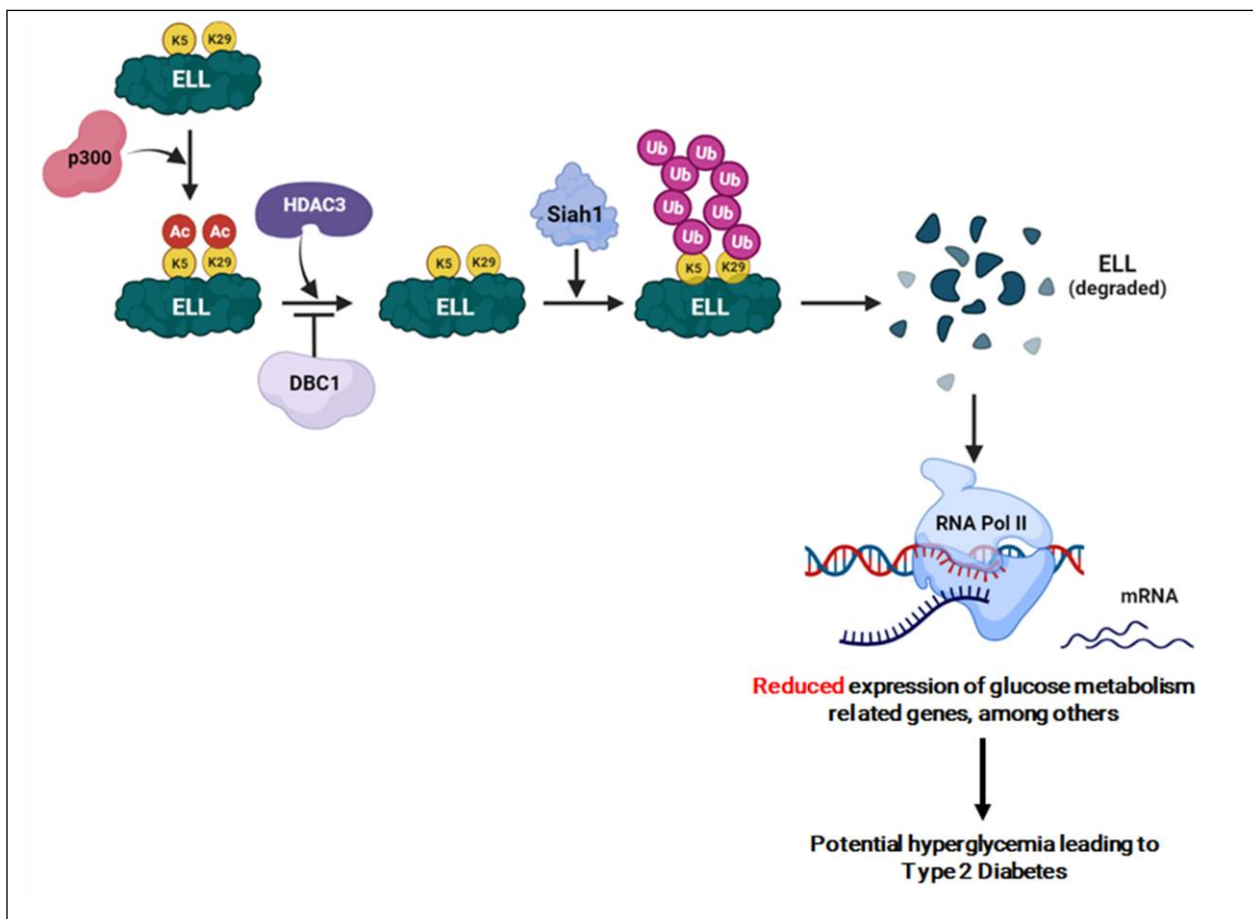


Fig 4.23 Model of functional regulation of ELL by the concerted action of multiple proteins

ELL possesses two key lysine residues at its N terminus, at positions K5 and K29. These lysines are acetylated by p300, which makes ELL functionally active and renders it stable by preventing its degradation by ubiquitylation. In the presence of HDAC3, these acetylation marks are removed, which enables the tagging of these lysine residues with polyubiquitin chains by the E3 ligase, Siah1. This leads to the degradation of ELL by the proteasome. The loss in cellular ELL causes reduced transcription of its target genes, which are involved in multiple cellular processes, including glucose metabolism. The decreased transcription of these genes leads to inefficient glucose transport and uptake, thus causing hyperglycemia and type 2 diabetes. DBC1 acts as a “protector” of ELL by preventing its interaction with HDAC3, thereby maintaining the cellular pool of ELL by keeping it in an acetylated state.

Chapter 5

Results

Part 2

**DBC1 and EAF1/EAF2 maintain ELL
stability and functions via a negative
feedback loop mechanism**

This section focuses on detailed studies on the interplay between ELL and EAF1/EAF2 and the effect of this association on the regulation of eukaryotic transcription. The identification of EAF1 as an ELL interacting factor in 2001, and similar observations with EAF2 heralded a new era in the studies of SEC mediated transcriptional regulation. Several reports have shown that the ELL.com, or the ELL-EAF1 containing complex is able to stimulate RNA Pol II-mediated transcriptional elongation on a DNA template as tested in an in vitro transcription assay. Furthermore, a role of EAF1 containing Little Elongation Complex (LEC) has been seen in the interaction with ELL for pausing during, and subsequent restart of transcription after, DNA damage repair. However, as detailed studies have not been performed on the interplay between ELL and EAF1/2, we tried to address the significance behind the strong association between these proteins.

5.1 The presence of EAF1 and EAF2 increases the level of ELL protein

The interactions between ELL and EAF1/EAF2 have been described earlier in detail, and thus, at the beginning of our studies, we decided to re-test these observations within our experimental settings. For this, we transfected both ELL and EAF1 or EAF2 expressing plasmids within HEK-293T cells. Surprisingly, with the overexpression of EAF1, we found that exogenously expressed ELL levels increased concomitantly (Fig. 5.1A). Similar observations were also seen with the co-expression of EAF2 along with ELL, wherein, exogenous ELL levels increased linearly with increasing expression of EAF2 (Fig. 5.1B). However, from these assays, it was not aptly clear whether the increase in ELL levels is because of an effect on ELL protein, or simply due to increased transcription. To test this, we overexpressed increasing concentrations of both EAF1 and EAF2 within 293T cells along with a fixed pre-determined amount of ELL. Cells were collected and RNA extracted for testing the effect of EAF1 and EAF2 on ELL mRNA levels. To our surprise, and as shown in Fig. 5.1C, increasing EAF1 amounts had a very minimal effect on EAF1 mRNA. The increase in ELL mRNA levels was seen only in presence of the highest concentrations of EAF1 used in our assay, and, the increase of ELL mRNA was minimal as compared to the significant increase observed for its protein levels as seen previously. We also had similar observations with EAF2, wherein, the overexpression of EAF2 had caused a drastic increase in ELL protein levels (Fig. 5.1B), however, in this case as well, with repeated experiments, the ELL mRNA levels remained largely unchanged (Fig. 5.1D). Our experiments were performed at least twice, and relative mRNA levels were calculated taking GAPDH as the housekeeping gene. Overall, our observations confirm that EAF1 and EAF2 do indeed affect ELL protein levels and do not have any significant effect on its transcription. Taken together, these serendipitous observations indicate that the human EAF1 and EAF2 are able to increase ELL levels within mammalian cells.

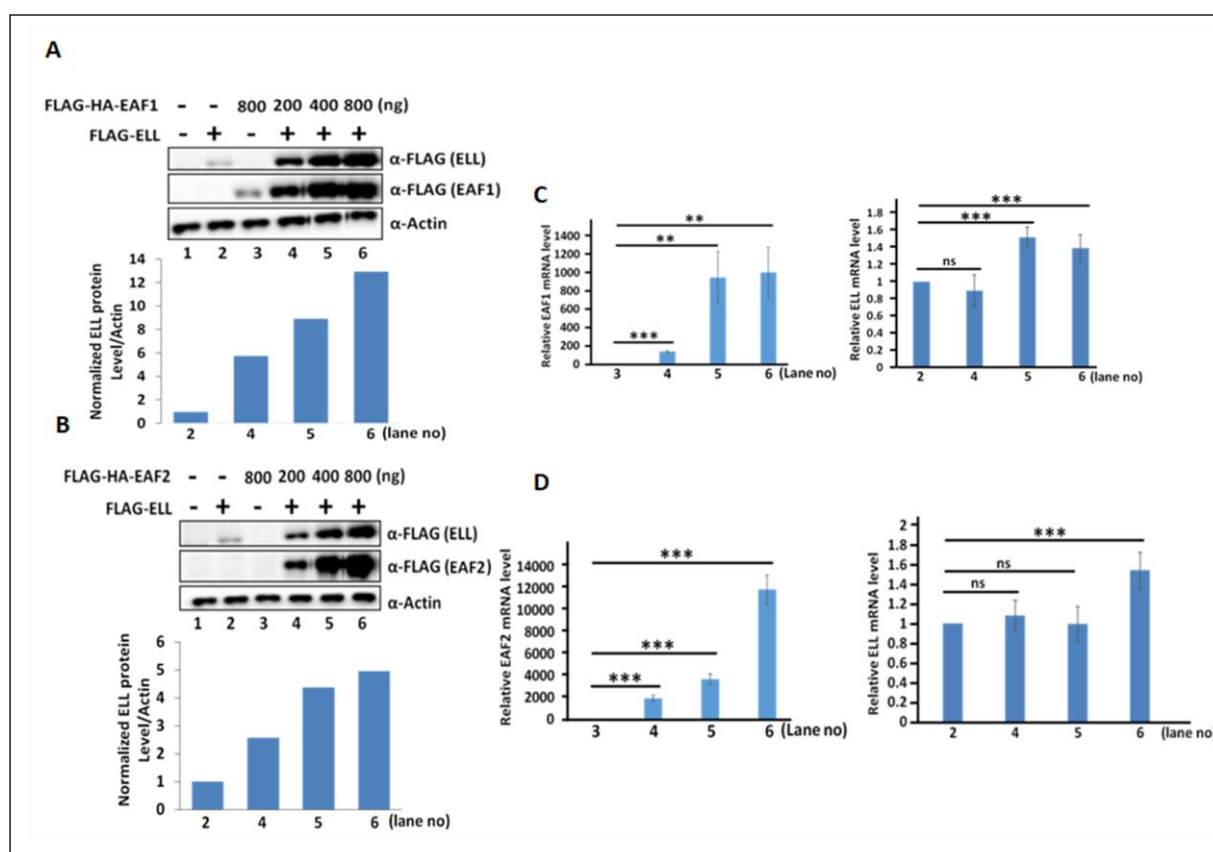


Fig 5.1 Overexpression of EAF1 and EAF2 increase ELL levels within cells

A. Ectopic ELL levels increase upon overexpressing EAF1 within mammalian cells. **B.** Similar observations are seen upon the overexpression of EAF2 as well. **C.** qRT-PCR data showing that overexpression of either EAF1 or **B.** EAF2 does not affect ELL mRNA levels.

In these experiments, the data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, ** denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

5.2 The stabilizing effect of EAF1 is limited only to ELL

We have already observed that EAF1 and EAF2 are able to increase ELL protein levels without affecting its transcription. This effect on ELL appears to be directly dependent on both EAF1 and EAF2 as both these proteins are known to associate directly with ELL. As EAF1/EAF2 are members of the Super Elongation Complex (SEC), we also wished to understand whether this effect of these two proteins is limited to ELL or other components of the SEC. As ELL interacts directly with EAF1/EAF2, we hypothesized that the interaction between EAF1/EAF2 and its target proteins would be necessary for their stabilizing effect. To test this hypothesis, we decided to identify the members of the SEC that could interact directly with EAF1 (as a representation of EAF1 and EAF2). For this, we overexpressed and purified the SEC components AFF1, AF9, ELL, and CDK9 from mammalian HEK-293T cells. EAF1

was expressed as GST tagged in *E. coli* BL21(DE3) cells and left immobilized on GST-agarose beads after affinity purification (Fig. 5.2A). These purified SEC proteins were then incubated with the immobilized GST-EAF1 with continuous agitation for 12 hours, at 4°C. Bound proteins were then washed with wash buffer, and EAF1-associated proteins were identified by SDS-PAGE, followed by western blotting. Our analyses from these experiments indicate that of all the SEC components tested, ELL, as expected, was found to associate with EAF1, directly (Fig. 5.2B, lane 9). Additionally, and much to our surprise, the kinase module of P-TEFb, namely CDK9, was seen to interact with EAF1 *in vitro* (Fig. 5.2B, lane 10). Based on this observation, we reasoned that the EAF1 would be able to stabilize only these two members of the SEC. We thus transfected human AFF1, AFF4, AF9, CDK9, and ELL along with EAF1 in HEK-293T cells. Analysis of the effect of EAF1 on the expression of these SEC components by western blotting, clearly indicates that of all the proteins tested, EAF1 had a significant effect in increasing the protein levels of both ELL as well as CDK9 (Fig. 5.2C, compare lane 9 vs lane 10 and lane 11 vs lane 12). The effect of EAF1 on CDK9 expression is modest but reproducible. In contrast, the overexpression of EAF1 resulted in a decrease in protein levels of AFF1, AFF4, and AF9. These observations, taken together, lend credibility to the observation that the interaction with EAF1 is the determining criteria for its stabilizing effect on different proteins.

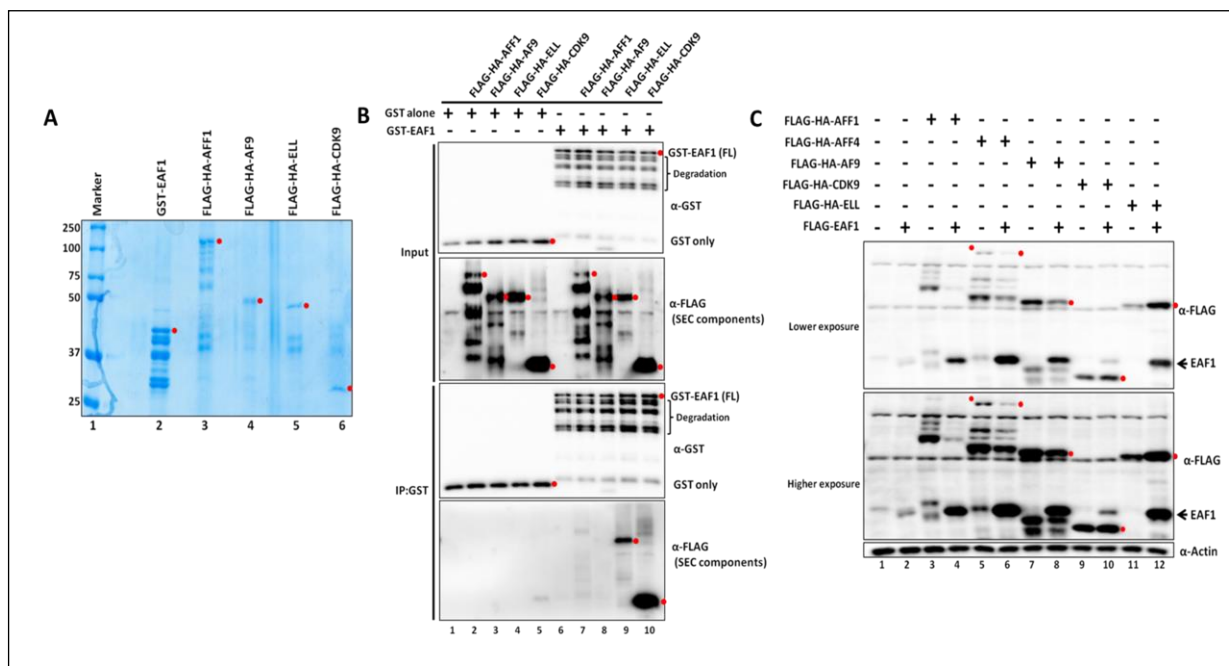


Fig. 5.2 EAF1 stabilizes only ELL and CDK9 in the SEC

A. Coomassie staining showing the purification of different proteins as used in our *in vitro* interaction assays. **B.** *In vitro* interaction assays showing the direct association of EAF1 with ELL and CDK9 components of the SEC. **C.** Western blot analysis showing the effect of EAF1 on the expression of different SEC components as indicated. EAF1 was able to stabilize only ELL and CDK9 components of the SEC.

5.3 EAF1 and EAF2 protect ELL from ubiquitin-mediated degradation

In the earlier chapter, we have shown that ELL is a target for ubiquitin-mediated degradation and that the presence of DBC1 can protect ELL from ubiquitination and subsequent degradation by maintaining its acetylated state. Similar to DBC1, the acetyl transferase p300, which acetylates ELL on its K5 and K29 N terminal residues, also prevents ELL ubiquitylation. As we have observed from our earlier experiments, both EAF1 and EAF2 have a stabilizing effect on ELL and increase its protein levels without affecting its mRNA transcription. We then wondered if both EAF1 and EAF2 can have such a similar effect on ELL, i.e: whether, in the presence of these two proteins, ELL degradation by ubiquitin can be rescued or prevented. To test this initially, we co-transfected ELL along with ubiquitin in HEK-293T cells. As seen in Fig. 5.3A, and consistent with our earlier observations, the presence of ubiquitin decreased ELL protein levels significantly. Subsequently, in this experimental setup, we also transfected increasing concentrations of both EAF1 and EAF2. In the presence of EAF1, ELL degradation mediated by ubiquitin was markedly reduced, and with an increasing amount of transfected EAF1, the levels of ELL increased even though ubiquitin was present in an equal amount in the experimental setup (Fig. 5.3A, compare lane 2 vs lane 3, and lane 3 vs lanes 4-6). The overexpression of EAF2 also had a similar effect, and ELL destabilization by ubiquitin was seen to be rescued significantly (Lane 3 vs lanes 7-9). When we observed that in the presence of EAF1 and EAF2, ELL is protected from ubiquitin-mediated degradation, we were then curious to know whether the presence of these two proteins would influence ELL turnover rate within cells by regulating its degradation kinetics. For this, we employed CHX chase assays to understand the kinetics of degradation of ELL under normal conditions and in the presence of ubiquitin. Subsequently, we introduced EAF1 and EAF2 in this setup and tested for the effect of these two proteins on the overall reaction by measuring protein levels by western blotting. As the results show in Fig. 5.3B, ELL itself is highly stable and does not undergo degradation even when treated with CHX for 8 hours (Lanes 1-3). However, in the presence of ubiquitin, ELL protein levels decrease rapidly, as seen within the first 4 hours of treatment. Now, with the introduction of EAF1, there is a dramatic effect, wherein, ELL levels are stabilized to such a great extent in its presence, that the presence of CHX does not cause any change to its steady-state levels. The overexpression of EAF2 has a more modest outcome, where ELL protein levels are significantly stabilized and it is affected very marginally by CHX treatment. The dynamic changes in the levels of ELL protein at different time points are visualized by the graph (Fig. 5.3B, bottom panel) wherein the rapid degradation of ELL occurs because of the presence of ubiquitin. This degradation rate, however, is almost nullified because of the presence of EAF1 and EAF2. We have already observed previously that the degradation of ELL occurs by the ubiquitin-proteasome pathway, i.e: it is necessary for ELL to be polyubiquitylated for it to undergo degradation. Upon this

observation of EAF1/2 being able to rescue ubiquitin-mediated destabilization, we were curious to test if the rescue of ubiquitin-mediated ELL degradation by EAF1 or EAF2 is a result of its decreased polyubiquitylation. For this, we again transfected HEK-293T cells with plasmids expressing FLAG-HA-ELL, GFP-Ubiquitin, and FLAG-EAF1/EAF2 in the indicated combinations similar to that of our previous experiment. In this setup, however, the ubiquitin amount was decreased slightly so that the complete degradation of all exogenously expressed ELL would not occur. In this experiment as well, as seen previously, ELL was destabilized in the presence of ubiquitin, and this destabilization could be rescued by the overexpression of both EAF1 and EAF2. The FLAG-HA-tagged ELL was subsequently immunoprecipitated using anti-HA-tagged agarose beads. The immunoprecipitation was done under high salt conditions in order to eliminate non-specific binding to the beads. After washing, the polyubiquitylation of ELL was tested by examining for the presence of characteristic “trails” of ubiquitin-tagged proteins using an anti-ubiquitin antibody. As shown in Fig. 5.3C, in the presence of overexpressed ubiquitin, characteristic polyubiquitination trails could be seen (Lane 2). However, upon overexpressing EAF1, almost no polyubiquitylated species of ELL could be observed in our experimental conditions (Compare lane 2 with lanes 3-5, ubiquitin blot) when probed with an anti-ubiquitin antibody. With the overexpression of EAF2 in this setup, moderate ELL polyubiquitylation was observed, which was, however, much lesser than the signal seen in the presence of ubiquitin only (Fig. 5.3C, compare lane 2 vs lanes 6-8). Interestingly, and lending credibility to our hypothesis, we observed that with increasing expression of EAF2, the ubiquitylation signal of ELL decreased, which suggests a clear role of this protein in protecting ELL.

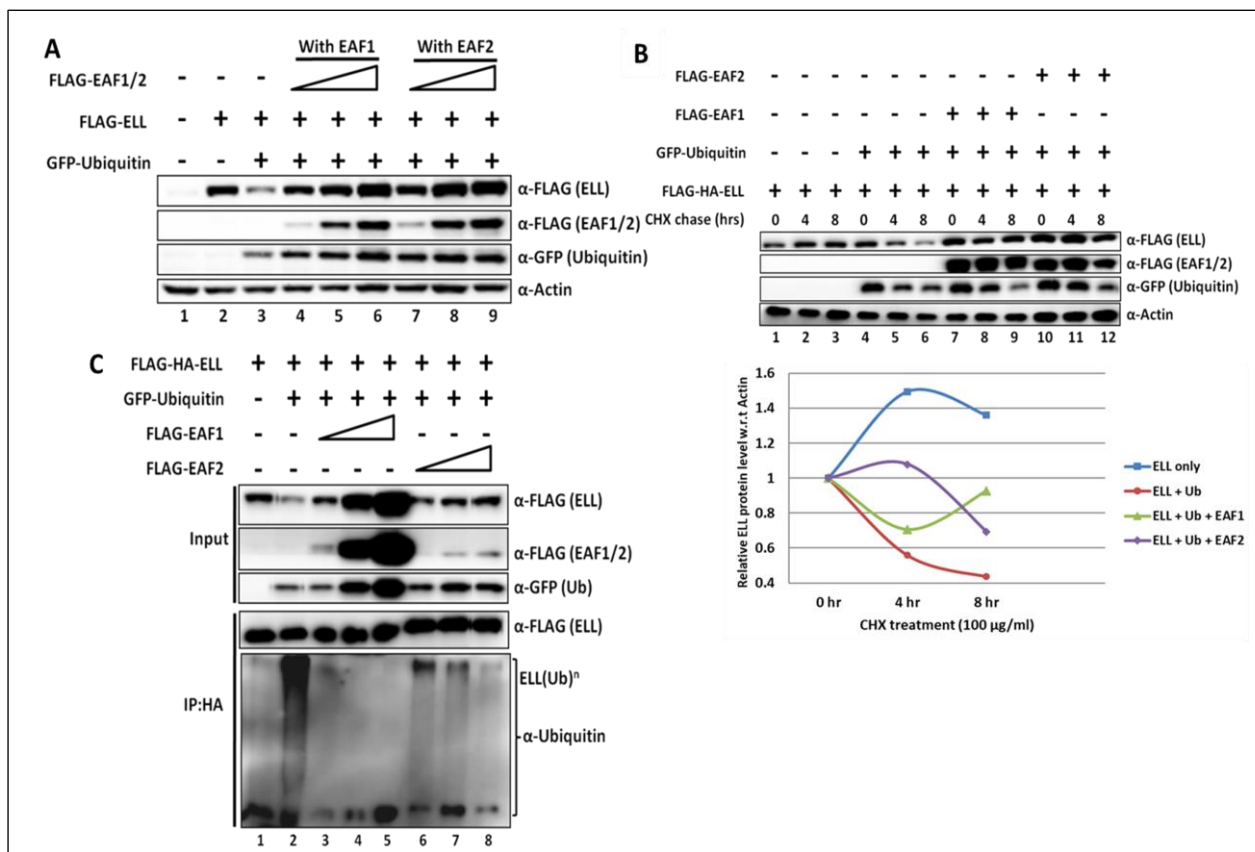


Fig 5.3 ELL is protected from ubiquitin mediated degradation by EAF1 and EAF2

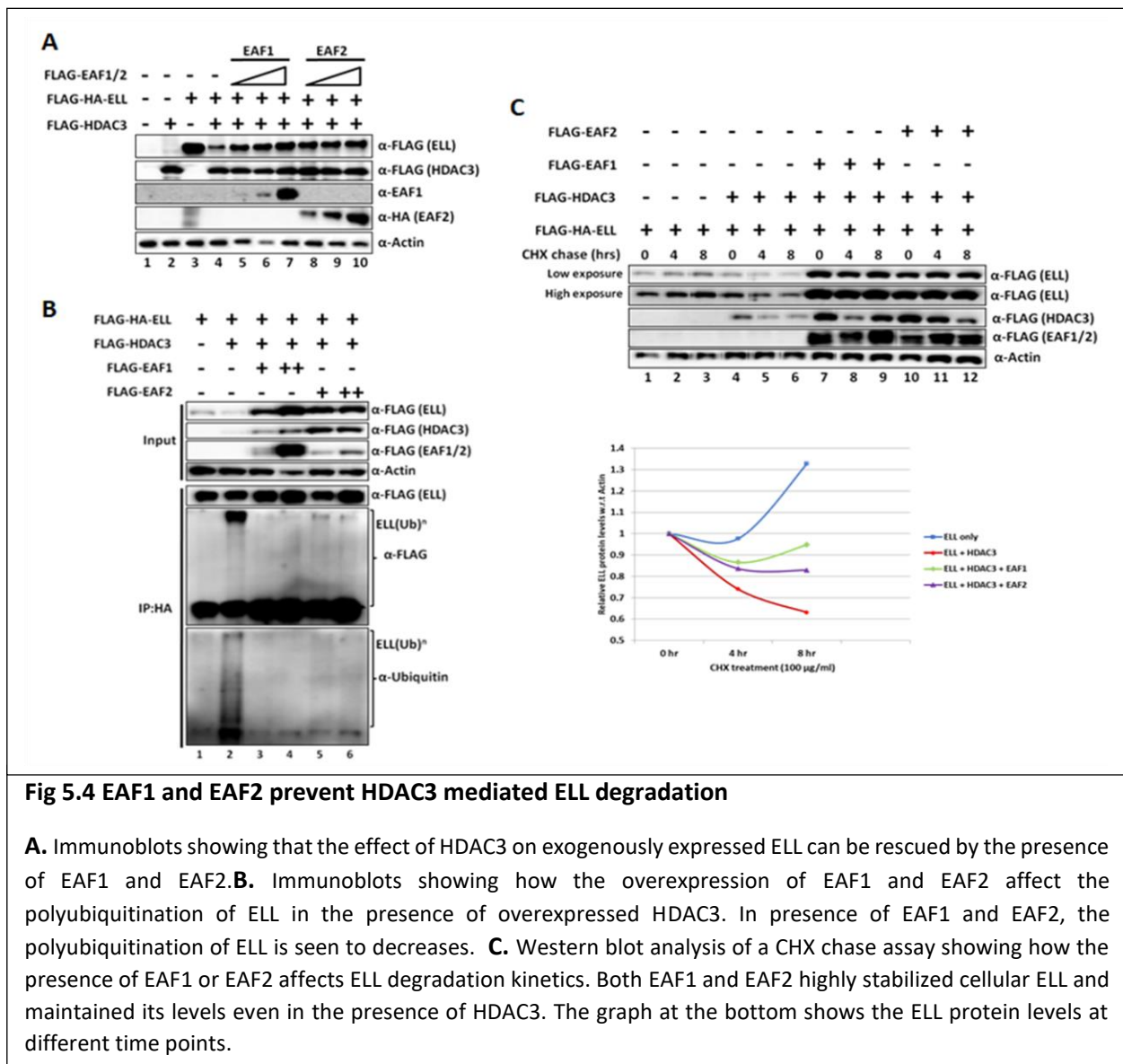
A. Immunoblots showing the effect of ubiquitin on exogenously expressed ELL and how the presence of EAF1 and EAF2 rescues ELL from ubiquitin mediated degradation. **B.** Western blot analysis of a CHX chase assay showing the effect of EAF1 or EAF2 on ELL degradation kinetics. Both EAF1 and EAF2 highly stabilized cellular ELL and maintained its levels even in the presence of ubiquitin. The bottom graph shows the ELL protein levels at different time points. **C.** Immunoblots showing how the overexpression of EAF1 and EAF2 affect the polyubiquitination of ELL in the presence of overexpressed ubiquitin. In presence of EAF1 and EAF2, the polyubiquitination of ELL decreases even if ubiquitin is overexpressed.

5.4 HDAC3-mediated ELL degradation is prevented by EAF1 and EAF2

We have previously observed that ELL is a highly stable protein, which exists predominantly in an acetylated state and is degraded by ubiquitination only upon deacetylation by HDAC3. Our experiments show that in order to maintain the cellular level of ELL, it is acetylated by p300, and this acetylated form of ELL is maintained by DBC1, which renders it functional and prevents its degradation. As mentioned above, we have also observed that EAF1 and EAF2 can protect ELL from ubiquitin-mediated degradation, and achieves this by simply preventing the polyubiquitination of ELL which is necessary for its degradation by the ubiquitin-proteasome system. This overall renders ELL stable and maintains its cellular levels. We now wondered if HDAC3 mediated ELL degradation could also be prevented in the presence of these two factors, i.e EAF1 and EAF2. For this, we transfected ELL

alone, ELL with HDAC3 and ELL and HDAC3 in combination of increasing concentrations of EAF1 and 2 as indicated in Figure 5.4A. We know from our previous experiments that ELL is significantly destabilized in the presence of HDAC3, and this observation was replicated in this experiment as well, wherein, in the presence of HDAC3, we observed degradation of about 60% of the cellular ELL levels (Fig. 5.4A, compare lane 3 vs lane 4, ELL panel). We now introduced increasing concentrations of both EAF1 and EAF2, individually, into the transfection mix. As shown in Fig. 5.4A, and similar to as seen previously in Fig. 5.5, when EAF1 is present, ELL degradation by HDAC3 is markedly rescued by about 2-3 folds (Compare lane 4 vs lane 5-7). A similar effect of EAF2 overexpression is seen wherein, in the presence of EAF2, HDAC3 mediated ELL degradation is rescued significantly (Compare lane 4 vs lanes 8-10). As we previously tested, in our earlier experiments we have seen that the presence of DBC1 and p300 is able to protect ELL from HDAC3 mediated deacetylation and thus results in the cellular pool of ELL being kept at a constant level. This is achieved by reducing ELL polyubiquitylation and thus preventing its degradation. In our subsequent experiments, when we observed that EAF1 and EAF2 were able to protect ELL from HDAC3 mediated degradation, we reasoned that the same mechanism might function in maintaining the cellular ELL pool. To examine this, we co-transfected ELL, HDAC3, and increasing concentrations of EAF1 and EAF2 within mammalian cells. After recapitulating our earlier results that the presence of both EAF1/2 are able to prevent ELL from HDAC3-mediated degradation and that in the presence of EAF1 and EAF2, the levels of ELL are markedly increased (Fig. 5.4B, Input panel), we immunoprecipitated the ELL using HA epitope tag-specific agarose beads and tested whether it was conjugated with ubiquitin moieties or not. As is shown in Fig. 5.4B, (IP panel, α -FLAG and α -ubiquitin blot), we could reproduce our earlier results that in the presence of HDAC3, ELL is polyubiquitylated (Lane 1 vs lane 2). However, when EAF1 and EAF2 are overexpressed in increasing concentrations, as the cellular levels of these two proteins increased, the polyubiquitylation “trails” decreased and were almost reduced to levels where there was no overexpression of HDAC3 at all (Compare lane 2 vs lanes 3-4 and 5-6, and match with lane 1). These observations suggest that like EAF1 and EAF2 prevent ubiquitin-mediated ELL degradation, these two proteins are able to protect ELL from HDAC3 mediated destabilization as well, simply by reducing ELL polyubiquitylation within the cell. Our previous observations have also indicated that ELL is a highly stable protein, maintained by the presence of DBC1 and p300 in the cell. We have observed that when DBC1 or p300 are overexpressed, the degradation rate of ELL is markedly reduced, an observation we have seen for ubiquitin-mediated ELL degradation as well. To verify that HDAC3 mediated ELL destabilization rate is reduced in the presence of EAF1 and EAF2, we performed cycloheximide chase assays for ELL in the presence of HDAC3, wherein like before, about half of the ELL present was seen to be degraded. However, as shown in Fig. 5.4C, the overexpression of EAF1/2 caused a minute reduction in the exogenously expressed ELL levels even after an 8-hour chase experiment (Compare lane 4-6 vs lanes

7-9 for EAF1 and lanes 10-12 for EAF2. These results indicate that indeed, these two proteins play an integral role in the stability and function of ELL within the cell by preventing its degradation by the ubiquitin-proteasome system.



5.5 EAF1 and EAF2 bind to similar regions as that of HDAC3 on ELL protein

Our earlier detailed analyses indicate that in spite of ELL being a stable protein within cells, its levels can be regulated by making it prone to degradation by the ubiquitin-proteasome system. As ELL predominantly exists in an acetylated stable state, its deacetylation by HDAC3 is a key prerequisite for its channelization to the proteasome by the E3 ligase, Siah1. In this milieu of proteins, EAF1 and EAF2

play an important role in regulating ELL levels and maintaining its stability by reducing HDAC3 and ubiquitin-mediated ELL polyubiquitylation (Fig. 5.4 and Fig. 5.3 respectively). Next, we were curious to know how these two proteins can achieve this phenomenon. Both EAF1 and EAF2 have been previously identified and described as stable interactors of ELL [191, 193] and detailed interactions between these proteins reported. Additionally, our experiments have shown a direct HDAC3-ELL association, including in the context of endogenous proteins. We have already established that the interaction between ELL and EAF1 is essential for ELL stabilization (Fig. 5.2). We thus hypothesized that, simply, all three proteins, namely EAF1, EAF2, and HDAC3 interacted with similar regions on the ELL protein. Now, to address whether the three proteins, EAF1, EAF2, and HDAC3 bind to similar regions on ELL, possibly located at its N terminus, we purified FLAG-HDAC3 and FLAG-HA-EAF2 after overexpressing them in mammalian HEK-293T cells. Additionally, we also purified His-EAF1 and GST-ELL (1-621 aa) as well as GST-ELL (61-621 aa) from overexpression in *E. coli* BL21(DE3) cells (Fig. 5.5A). The decision to use this deletion mutant of ELL (61-621 aa) was taken knowingly as the first 60 amino acids of ELL constitute its “inhibitory domain”, and we wished to minimize any off-target effect due to the presence of any amino acids from the mentioned region. As it has been already described previously by several studies that ELL interacts directly with EAF1, we wished to ensure that our assay system and conditions were sound for carrying out subsequent experiments. For this, our initial *in vitro* direct interaction assay involved only ELL and EAF1 where GST-ELL was kept immobilized on anti-GST agarose beads and free His-EAF1 was taken as the prey protein (Fig. 5.5B). As can be seen, in our purified assay system, His-EAF1 was seen to interact with GST-ELL alone and did not show any binding with GST alone, indicating that our assay conditions were standardized to prevent any non-specific interactions (Compare lane 3 vs lane 4). Next, taking our experiments one step further, we now immobilized GST alone (as control), GST-ELL (1-621), and GST-ELL (61-621) on anti-GST beads and tested the strength of their interaction with free His-EAF1. As shown in Fig. 5.5C, even though inadvertently higher concentrations of GST-ELL (61-621) were added in the reaction mix, and inspite of a high amount of His-EAF1 being present (Fig. 5.5C, Input panel), the deletion of the first 60 amino acids from the N-terminus of ELL reduced its interaction with ELL to a great extent (Compare lane 5 vs lane 6, EAF1 blot, IP panel, and quantitation of relative intensities as mentioned on the right using ImageJ software). Along similar lines, we then used this immobilized GST-ELL (1-621) and GST-ELL (61-621) as the bait for testing their interaction with free FLAG-HA-EAF2. EAF1 and EAF2 are similar proteins. Keeping this in mind, we expected that the EAF2 binding on ELL would also be restricted to its N-terminus. As hypothesized, our experiments showed that (i) EAF2 indeed binds directly to ELL (Fig. 5.5D, lane 5 of IP panel) and, (ii) as seen previously for EAF1, the deletion of the first 60 amino acids from the N-terminus of ELL drastically reduces the strength of EAF2 binding to ELL (Fig. 5.5D, compare lanes 5 and 6 for the EAF2 immunoblot in the IP panel) by about 80%. This exciting

observation prompted us to expect that HDAC3 binding to ELL would also possibly be restricted to its N-terminus indicating that these three proteins interact with similar regions of ELL. Our earlier observation that HDAC3 interacts directly with ELL (1-621) was replicated in our assay conditions (Fig. 5.5E, lane 5). Furthermore, we observed, that the N-terminal deletion mutant of ELL (61-621) had a four-fold decrease in interaction intensity as compared to the full-length ELL (Fig. 5.5E, lane 5 vs lane 6, and right panel for quantification). Our combined results thus indicated convincingly that the presence of the N terminus of ELL is essential for its stabilization by EAF1 or EAF2 and the binding of these three proteins, i.e EAF1, EAF2, or HDAC3 occurs within the first 60 amino acids of ELL. It is interesting to note that our earlier studies have shown that it is this N terminal region of ELL which harbors the two key lysine residues, namely, K5, and K29, which are the sites for its acetylation as well as ubiquitylation.

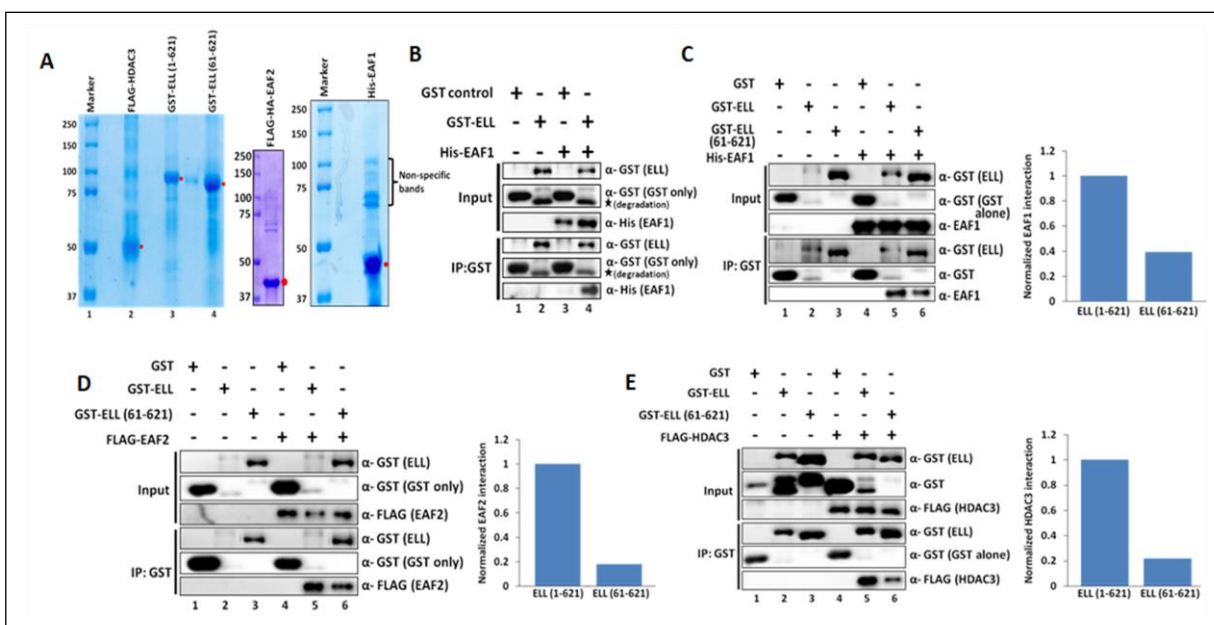


Fig 5.5 Interaction of EAF1, EAF2 and HDAC3 occur predominantly at the N-terminus of ELL

A. Purification of different proteins for use in our *in vitro* interaction assays as visualized by Coomassie staining after resolution on a SDS-PAGE gel. The FLAG-HDAC3 and FLAG-HA-EAF2 were purified from HEK-293T cells after overexpression. His-EAF1, GST-EV, GST-ELL (1-621) and GST-ELL (61-621) were all purified after growing them in *E. coli* BL21(DE3) and inducing protein expression with IPTG. **B.** Interaction between ELL and EAF1 is direct as visualized by our *in vitro* direct interaction assay **C.** Immunoblots showing that the extreme N-terminus of ELL is important for its interaction with EAF1 **D.** In vitro interaction assays also show that the interaction between ELL and EAF2 is dependent on the N-terminal 60 amino acids of ELL. **E.** Like EAF1 and EAF2, HDAC3 is also seen to associate directly with the N-terminus of ELL protein.

5.6 ELL-HDAC3 interaction is reduced in the presence of EAF1 and EAF2

Since we observed in our experiments that EAF1, EAF2, and HDAC3 predominantly bind within the N-terminal amino acids on ELL (Fig. 5.5), and our earlier experiments conclusively show that the presence of EAF1 and EAF2 deter HDAC3 mediated ELL degradation, we hypothesized that these two proteins competitively inhibit ELL-HDAC3 interaction due to their stable binding to ELL. We thus sought to verify this hypothesis by *in vitro* competitive binding assays using purified proteins. This system was chosen to conduct our experiments to nullify the effects of any other proteins that may influence the interaction interplay between our target factors. At first, we reproduced our observation that both EAF1 and HDAC3 bind directly to full-length ELL even in a cell-free assay system (Fig. 5.6A, lane 5 for HDAC3 interaction, and lane 6 for EAF1 interaction in the IP panel). Now, in this reaction mix, we added increasing concentrations of EAF1, keeping HDAC3 constant. When EAF1 was linearly increased in our reaction system, we observed the concomitant binding between ELL and HDAC3 decreased significantly (Fig. 5.6A, lane 5 vs lanes 7-8 in the HDAC3 blot of the IP panel). As seen earlier, the EAF1 and EAF2 have similar effects on ELL, and also counteract the destabilization effect of HDAC3 on ELL in a similar fashion. Thus, it would not be too prudent to assume that in the presence of EAF2, HDAC3 association with ELL would also be reduced. Hence, we also proceeded to test this hypothesis by means of our cell-free *in vitro* interaction and competitive binding assay system. Like EAF1, EAF2 also bound ELL directly (Fig. 5.6B, lane 6 of the FLAG immunoblot for EAF2 in the IP panel). The intensity of this interaction was also found to be quite strong, as seen previously. In the presence of HDAC3, we next introduced increasing amounts of EAF2, expecting that, similar to EAF1, the HDAC3 association would decrease if EAF2 was given. As seen in Fig. 5.6B, and along the same lines as we had observed for EAF1, we observed a linear dissociation between ELL and HDAC3 as the interaction between ELL-EAF2 grew stronger (Fig. 5.6B, lane 5 vs lanes 7-8 in the HDAC3 blot of the IP panel). These two separate observations clearly point to a similar mechanism of action of these two proteins wherein, when EAF1 or EAF2 are present in high concentrations, they negatively impact ELL-HDAC3 interaction due to the presence of competition for the same binding sites at the N-terminus of ELL. Next, we were also curious to understand whether these interactions and cross talk that we observe between these proteins in a cell-free system is replicated within cells as well. For this, we utilized a simple overexpression-based assay in HEK-293T cells. In our experiment, we transfected ELL alone, and ELL along with HDAC3. Then we immunoprecipitated the ELL and tested if the interaction between ELL and endogenous EAF1 is hampered by the simultaneous overexpression of HDAC3 within the system. The immunoblots presented in Fig. 5.6C indicate that ELL does interact with HDAC3 within the cell (Lane 4, HDAC3 blot). Additionally, and strikingly, the stable interaction between cellular ELL and EAF1

was disrupted when HDAC3 was also overexpressed along with ELL (Compare lane 2 vs lane 4 of the EAF1 blot in the IP panel). These observations indicate that in a context-dependent manner, these proteins interact with each other, possibly being regulated by their cellular concentrations. When EAF1 or EAF2 are high, or in closer proximity to ELL, they displace the HDAC3 associated with ELL. Similarly, an increased level of HDAC3 results in a concomitant decrease in the ELL-EAF1/EAF2 interaction, finally leading to the degradation of ELL by the ubiquitin-proteasome system. We next wondered as to what could be the implication of this enhanced ELL-EAF1 interaction which causes stabilized ELL levels within the cell. We know that ELL, when acetylated, is stable within cells. Additionally, this acetylated ELL appears to be transcriptionally active as well. Taking these two observations into account, we decided to examine the state of EAF1 associated ELL within cells. For this, we simply co-transfected both ELL and EAF1 in HEK-293T cells and subsequently performed an immunoprecipitation for the ELL protein. This immunoprecipitated ELL was probed with a pan-acetyl lysine antibody to ascertain its nature when EAF1 is present. As the immunoblots show in Fig. 5.6D, with increasing levels of transfected EAF1, levels of acetylated ELL also increased linearly (Compare lane 2 vs lanes 3-5 of the IP blot for ELL, and note the quantification of acetylation signal relative to the corresponding ELL pull down as mentioned in the bar graph on the right). It is safe to say that this enhanced acetylation is not due to any changes in levels of p300, the *bona fide* acetyltransferase of ELL, as it is clearly observed in our input blots that cellular p300 remains constant across all samples (Fig. 5.6D, p300 blot, Input panel). From these experiments, it is obvious that keeping ELL in the centre, EAF1/EAF2 and HDAC3 interact with it, possibly in a context dependent manner. When an excess amount of HDAC3 is present, ELL-EAF1/EAF2 interaction decreases, whilst the presence of EAF1/EAF2 reduces ELL-HDAC3 association, thereby maintaining acetylated stable ELL pools within the cell.

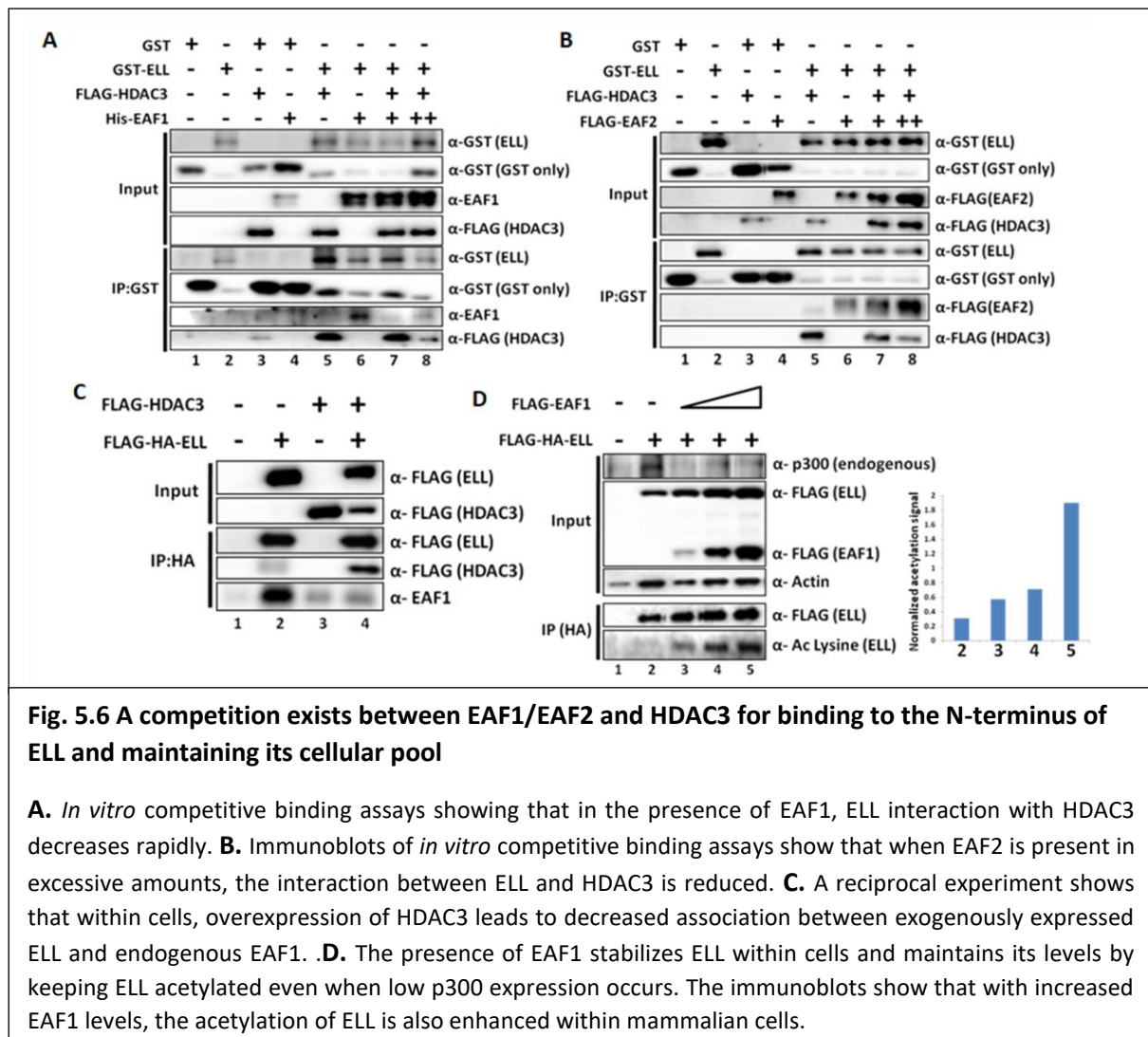


Fig. 5.6 A competition exists between EAF1/EAF2 and HDAC3 for binding to the N-terminus of ELL and maintaining its cellular pool

A. *In vitro* competitive binding assays showing that in the presence of EAF1, ELL interaction with HDAC3 decreases rapidly. **B.** Immunoblots of *in vitro* competitive binding assays show that when EAF2 is present in excessive amounts, the interaction between ELL and HDAC3 is reduced. **C.** A reciprocal experiment shows that within cells, overexpression of HDAC3 leads to decreased association between exogenously expressed ELL and endogenous EAF1. **D.** The presence of EAF1 stabilizes ELL within cells and maintains its levels by keeping ELL acetylated even when low p300 expression occurs. The immunoblots show that with increased EAF1 levels, the acetylation of ELL is also enhanced within mammalian cells.

5.7 An interplay exists between DBC1 and EAF1 for regulating each other's levels

In the previous chapter, we have described an elaborate mechanism involved in regulating ELL stability and functions by the DBC1 protein and the lysine acetyltransferase p300, which work to maintain acetylated ELL levels in the cell, preventing HDAC3 mediated deacetylation, leading to its ubiquitylation by the E3 ligase Siah1, and its subsequent degradation by the ubiquitin-proteasome system. Our observations described in this chapter show a mechanism of ELL regulation by EAF1 and EAF2 that appears to be similar to that of DBC1 and p300. Because of these sets of similar observations, we wondered whether DBC1 and EAF1 act upon ELL under different context-dependent circumstances in order to regulate its stability and functions. If such a situation exists, then it is also possible that there may be a link between EAF1 and DBC1 by virtue of which they can associate with

ELL under different scenarios, as the need may arise. To test this, we first co-expressed DBC1 along with increasing concentrations of EAF1 within HEK-293T cells. In this experiment, we observed, much to our surprise, that with the increase in ectopic EAF1 levels, the DBC1 expression was seen to decrease in a linear manner, suggesting that in the presence of EAF1, DBC1 levels decrease (Fig. 5.7A, Compare DBC1 expression in lane 1 vs lanes 2-4). A reciprocal experiment was also carried out in HEK-293T cells wherein this time, EAF1 was kept constant, whilst DBC1 was transfected in increasing amounts. In this experiment, we observed that the increasing DBC1 levels led to the depletion of ectopic EAF1 in a dose-dependent manner (Fig. 5.7A, Compare lane 5 vs lanes 6-8 for the EAF1 blot). The most striking observation that lends credence to our hypothesis that the cellular EAF1 and DBC1 function in tandem for regulating ELL, was the observation of how these two treatments led to changes in the endogenous ELL levels. Under both scenarios, i.e: overexpression of EAF1 or overexpression of DBC1, though the other factor decreased, in both cases, the expression of endogenous ELL was seen to increase. This critical observation indicates that the two proteins, EAF1, and DBC1 regulate ELL levels by two non-overlapping pathways, which follow similar mechanisms but appear to be unique depending on the cellular system. Next, we wanted to verify that our observations are relevant in an endogenous context and are not influenced due to any bias in using overexpression constructs. Hence, we first transfected increasing concentrations of DBC1 and tested for the effect it would have on endogenous EAF1. As we had expected, we observed endogenous EAF1 levels decreasing with increased DBC1 levels (Fig. 5.7B, top panel). Subsequently, we were curious to know whether this effect is at the level of mRNA regulation or protein stability. For this, we also tested the effect of DBC1 overexpression on endogenous EAF1 mRNA levels by qRT-PCR. From our assays, we found that though upon DBC1 overexpression, EAF1 protein levels are reduced, there was absolutely no trend in the change of EAF1 mRNA levels (Fig. 5.7B, bottom panel). From these results, we can safely conclude that DBC1 causes a decrease in EAF1 by virtue of its action on EAF1 protein, and has no effect on the transcription of EAF1 mRNA. Along similar lines, we also wanted to check if EAF1 overexpression would show any effect on endogenous DBC1 levels. Transfection of increasing concentrations of EAF1 in HEK-293T cells, followed by immunoblotting showed that, similar to our observations for ectopically expressed DBC1, the endogenous DBC1 was also seen to decrease linearly (Fig. 5.7C, top panel). qRT-PCR analysis of these samples to understand the effect of EAF1 on endogenous DBC1 transcription showed that transcript levels of DBC1 reduced drastically with overexpression of EAF1 (Fig. 5.7C, lower panel). These results taken together clearly indicate that both DBC1 and EAF1 regulate each other's levels in an antagonistic manner, wherein EAF1 causes decreased DBC1 transcription, on the other hand, the presence of DBC1 adversely affects EAF1 protein stability and has no effect on EAF1 mRNA.

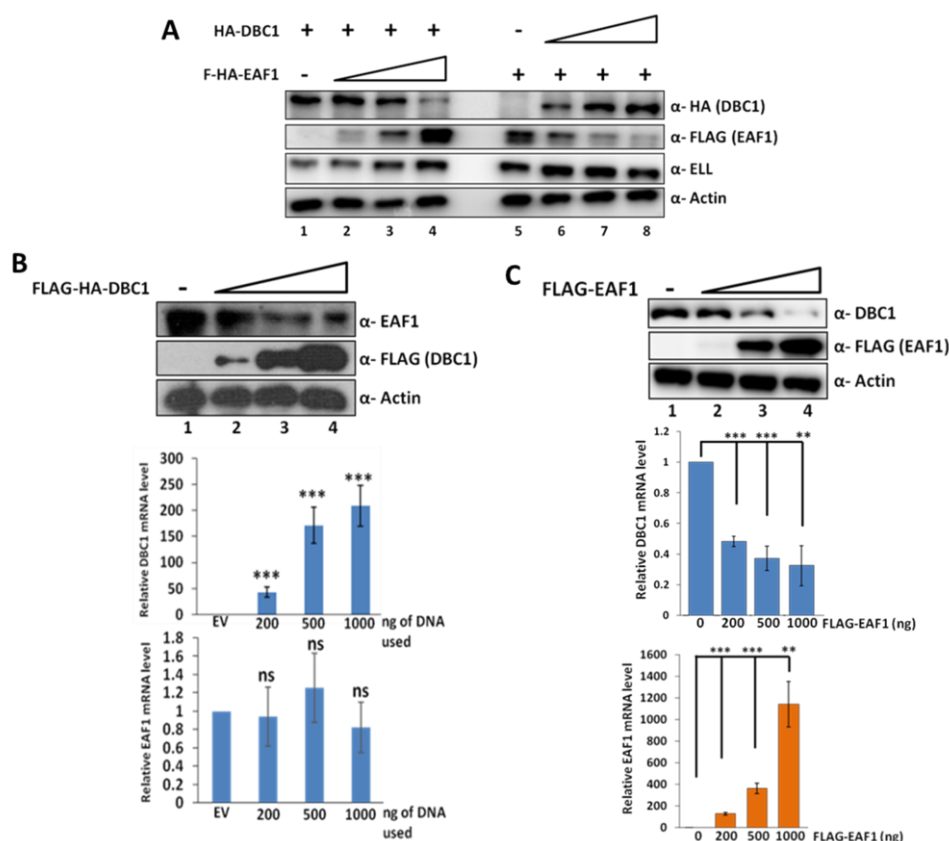


Fig. 5.7 DBC1 and EAF1 antagonize each other within the cell

A. Immunoblots showing that the overexpression of EAF1 decreases the expression of exogenous EAF1 (Lanes 1-4) while the overexpression of DBC1 decreases ectopically expressed EAF1 levels (lanes 5-8) when transfected within HEK-293T cells. **B.** Western blot analysis showing that the overexpression of DBC1 within mammalian cells is able to cause a significant decrease in endogenous EAF1 levels (Top panel) whilst not having any effect on the transcription of EAF1 mRNA (bottom panel). **C.** Overexpressing EAF1 in HEK-293T cells leads to a decrease in endogenous DBC1 levels as shown by western blotting (top panel). This decrease in DBC1 levels is due to reduced transcription of DBC1 mRNA as shown in the bottom panel.

The data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, ** denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

5.8 EAF1 is polyubiquitylated in the presence of DBC1

As we have observed that EAF1 protein decreases in the presence of DBC1, we sought to identify how EAF1 levels are regulated within the cell. One of the most common mechanisms of protein concentration control within the cell is by the ubiquitin-proteasome system, in which a protein is tagged with ubiquitin moieties and subsequently degraded. Keeping this in mind, we had the question whether EAF1 would be regulated by ubiquitination. To answer this, we co-expressed EAF1 along with ubiquitin. As the western blot analysis in Fig. 5.8A shows, with the overexpression of ubiquitin, the protein levels of EAF1 were seen to decrease significantly (Compare lane 2 vs lane 3 for the EAF1 blot in the Input panel). This indicates that the EAF1 protein is possibly regulated by ubiquitination within

the cell. Additionally, we treated the cells with increasing concentrations of MG132. MG132 is an inhibitor of the ubiquitin-proteasome system, and prevents the degradation of ubiquitin-tagged proteins. Hence, this treatment would serve two purposes: first, it would reinforce our observations that EAF1 is indeed degraded by the proteasome, and second, the presence of MG132 would greatly amplify the polyubiquitylation signal. In our experiments, we observed that the treatment of cells with MG132 could significantly rescue EAF1 degradation by ubiquitin (Fig. 5.8A, lane 2 vs lane 3 and lane 3 vs lanes 4-6 in the input panel). As the characteristic feature of proteasome-mediated degradation is the formation of “smears” of polyubiquitylation trails, we tested whether, in the presence of ubiquitin, we would be able to observe such “trails” for the EAF1 protein. Hence, when we immunoprecipitated EAF1 from these cell extracts and tested for polyubiquitylation signals, we were ecstatic to find that in the presence of ubiquitin, EAF1 undergoes ubiquitination, as shown in Fig. 5.8A (Ubiquitin blot in the IP panel). As we have seen previously that the presence of DBC1 causes EAF1 protein levels to decrease, we wondered if DBC1 mediated EAF1 degradation was due to ubiquitination as well. For this, we performed a simple assay. Here, we transfected DBC1 in HEK-293T cells and then subsequently treated the cells with increasing concentrations of MG132. Our assays indicate that in the presence of DBC1, endogenous EAF1 protein decreases, as seen previously. However, with the treatment of MG132, the EAF1 levels are rescued significantly and almost revert back to their original levels (Fig. 5.8B, Compare EAF1 in lane 1 vs 2, and then lane 2 vs lanes 3-5). These results clearly indicate that the DBC1 mediated destabilization of EAF1 occurs by the proteasomal degradation system.

Next, we also wanted to confirm whether the presence of DBC1 is enough to cause EAF1 polyubiquitylation. For this, we co-transfected EAF1 with sub-optimal levels of DBC1 so that the entire EAF1 overexpressed is not degraded. This EAF1 was then immunoprecipitated and tested for conjugation with ubiquitin. As shown in Fig. 5.8C, in the presence of DBC1, characteristic polyubiquitylation trails of EAF1 were observed (Lanes 3-5 in the ubiquitin blot in the IP panel). Taken together, these experiments clearly indicate that EAF1 levels are regulated by the ubiquitin-proteasome system, and this degradation of EAF1 is mediated, at least in part, by DBC1.

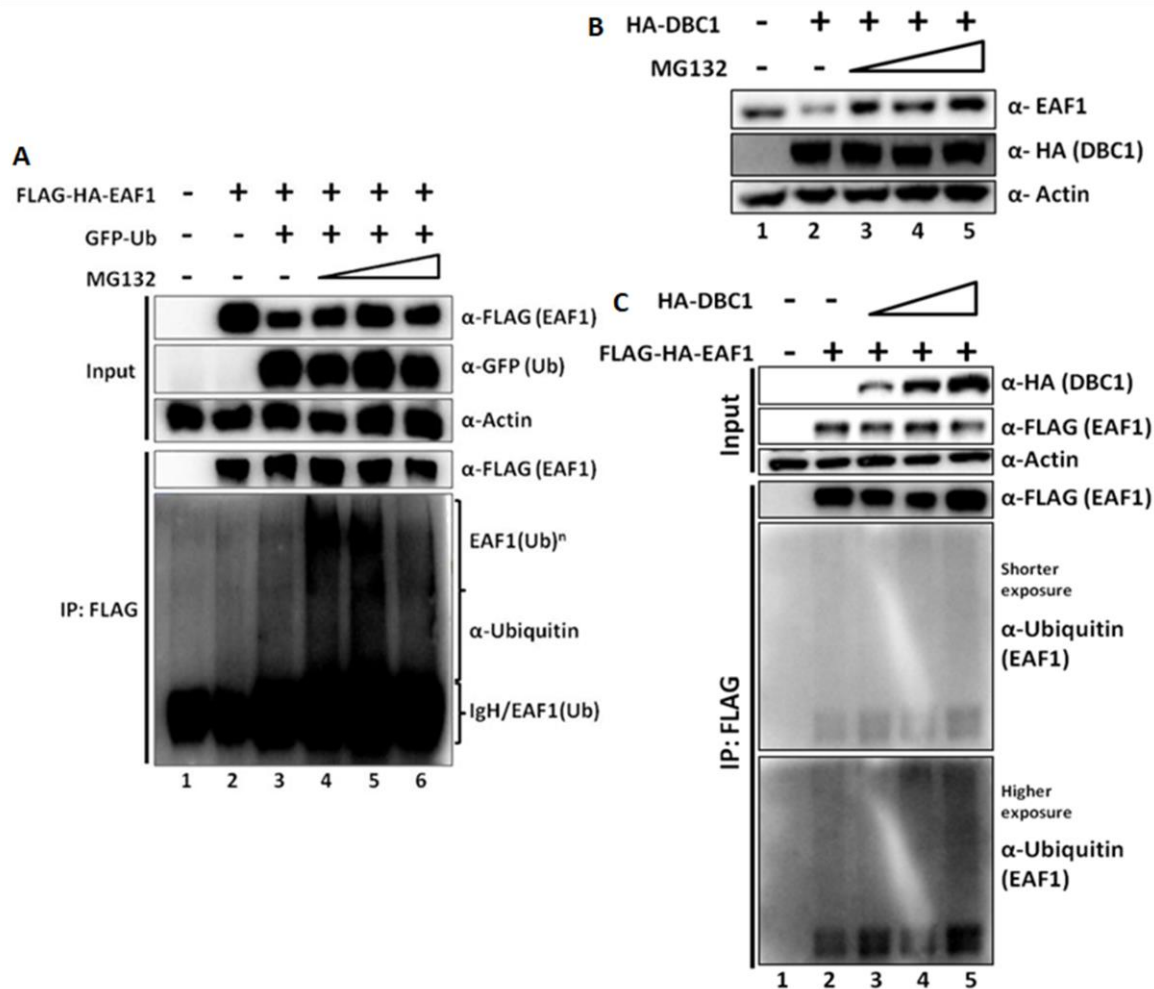


Fig. 5.8 EAF1 undergoes ubiquitin mediated degradation in presence of DBC1

A. The overexpression of ubiquitin causes EAF1 to be polyubiquitylated as shown by these immunoblots
B. DBC1 mediated EAF1 degradation can be rescued by MG132 indicating that EAF1 undergoes proteasome mediated degradation
C. Western blots indicating that in presence of DBC1, EAF1 does indeed get polyubiquitylated within mammalian cells.

5.9 TRIM28 interacts with and ubiquitinates EAF1

Reports of regulation of EAF1 stability and thus its functions by ubiquitination have not been described previously. Our observations that EAF1 is subjected to ubiquitin-mediated degradation, and the presence of DBC1 causes this effect prompted us to search for possible E3 ligases for EAF1. For this, we looked back at the mass spectrometry data identifying EAF1 interactors obtained from a stable cell line expressing FLAG-HA-EAF1. The generation of this cell line and the mass spectrometric analysis of EAF1.com has been described elsewhere[146]. Looking hard at the spectrometric analyses at our disposal, we were intrigued to note the presence of TRIM28 (also known as KAP1) as an EAF1 interactor (Fig. 5.9A). TRIM28 has been well described as a transcriptional repressor by facilitating the

recruitment of CHD3 of the NuRD complex and SETDB1, a lysine methyltransferase. As TRIM28 has also been shown to ubiquitinate p53 leading to its degradation, its presence as an interactor of EAF1 suggested the possibility that it may also act as an E3 ligase of EAF1. But first, we needed to ascertain whether EAF1 and TRIM28 could interact within mammalian cells. For this, we overexpressed TRIM28 within HEK-293T cells for the purpose of immunoprecipitating TRIM28 and its interacting proteins by affinity purification. We then proceeded to check the immunoprecipitated samples for the presence of our target proteins. As shown in Fig. 5.9B, in this reaction performed at physiological salt concentrations, we were content to see that endogenous EAF1 could indeed interact with TRIM28, suggesting a role of this protein in regulating the cellular levels of EAF1. As we had previously observed the potential of DBC1 to cause EAF1 depletion by polyubiquitination, we wondered whether by any chance TRIM28 would also interact with DBC1. Transiently expressing FLAG-DBC1 in HEK-293T cells and purifying its associated proteins by affinity purification, we observed the presence of ELL in the DBC1.com immunoprecipitate, indicating that our reaction conditions were appropriate (Fig. 5.9C). EAF1 was also seen to interact moderately with DBC1. The most interesting observation, however, was that, in our experiments, TRIM28 also appeared to interact stably with DBC1 within mammalian cells (Fig. 5.9C). Like before, a reverse immunoprecipitation of ectopically expressed TRIM28 also confirmed the presence of DBC1 as a novel interactor of this protein (Fig. 5.9D). Thus collectively, we could identify the transcriptional repressor, TRIM28, as a novel interactor of both EAF1 as well as DBC1 within the mammalian cell system.

Next, to test our hypothesis that TRIM28 could act as an E3 ligase for EAF1, we co-expressed EAF1 and TRIM28 within HEK-293T cells and tested for changes in their protein levels. As shown in Fig. 5.9E, in the presence of ectopically expressed TRIM28, the protein levels of EAF1 were seen to decrease substantially (Compare lane 2 vs lane 3 in the EAF1 blot). This decrease in EAF1 expression could be rescued partially by treatment of cells with MG132, indicating clearly that the degradation of EAF1 by TRIM28 must be by the ubiquitin-proteasome pathway (Fig. 5.9E, lane 3 vs lanes 4-6). Subsequently, we also wanted to confirm that the ubiquitylation of EAF1 does indeed occur in the presence of TRIM28. For this, we used an experimental setup similar to as shown in Fig. 5.9D, and immunoprecipitated the EAF1 using anti-FLAG epitope tag agarose beads. Subsequent immunoblotting using an anti-ubiquitin antibody showed that in the presence of TRIM28, EAF1 does indeed get polyubiquitylated, thus confirming the role of TRIM28 as the E3 ligase for EAF1 (Fig. 5.9F, Ubiquitin blot, compare lane 2 vs lanes 4-6). Our results thus identify TRIM28 as a novel interactor of EAF1, and prove that being an E3 ligase, it possesses the potential to cause polyubiquitination and degradation of EAF1 within mammalian cells.

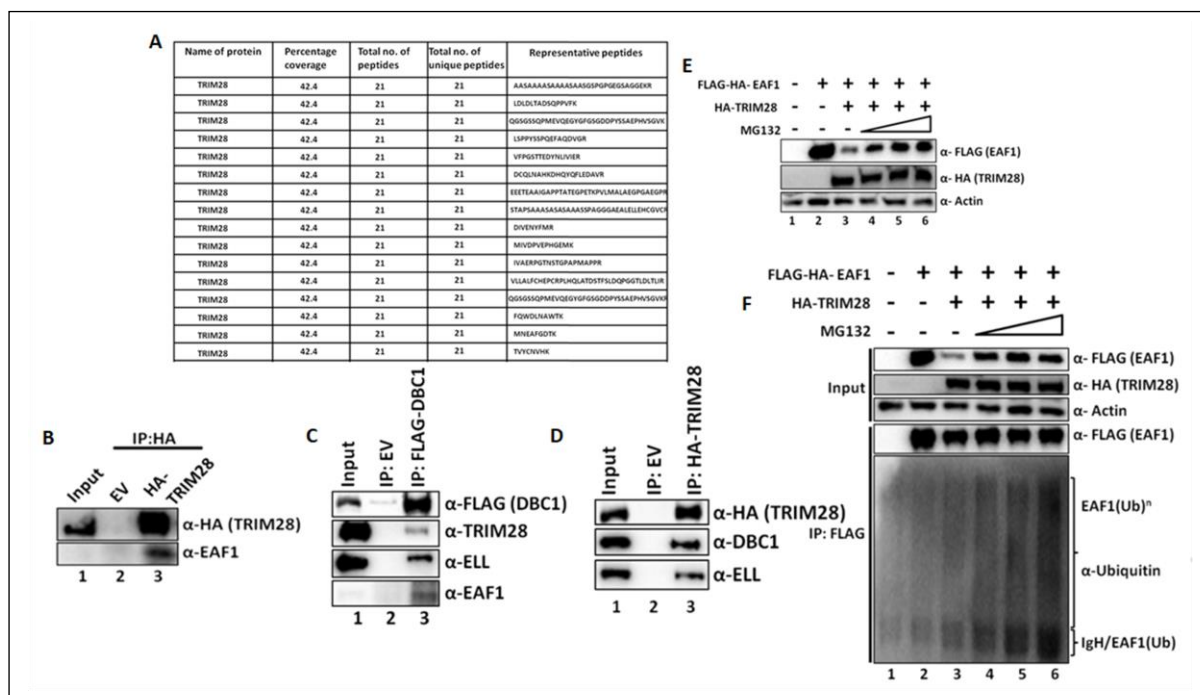


Fig. 5.9 TRIM28 can interact with EAF1 and ubiquitinate it

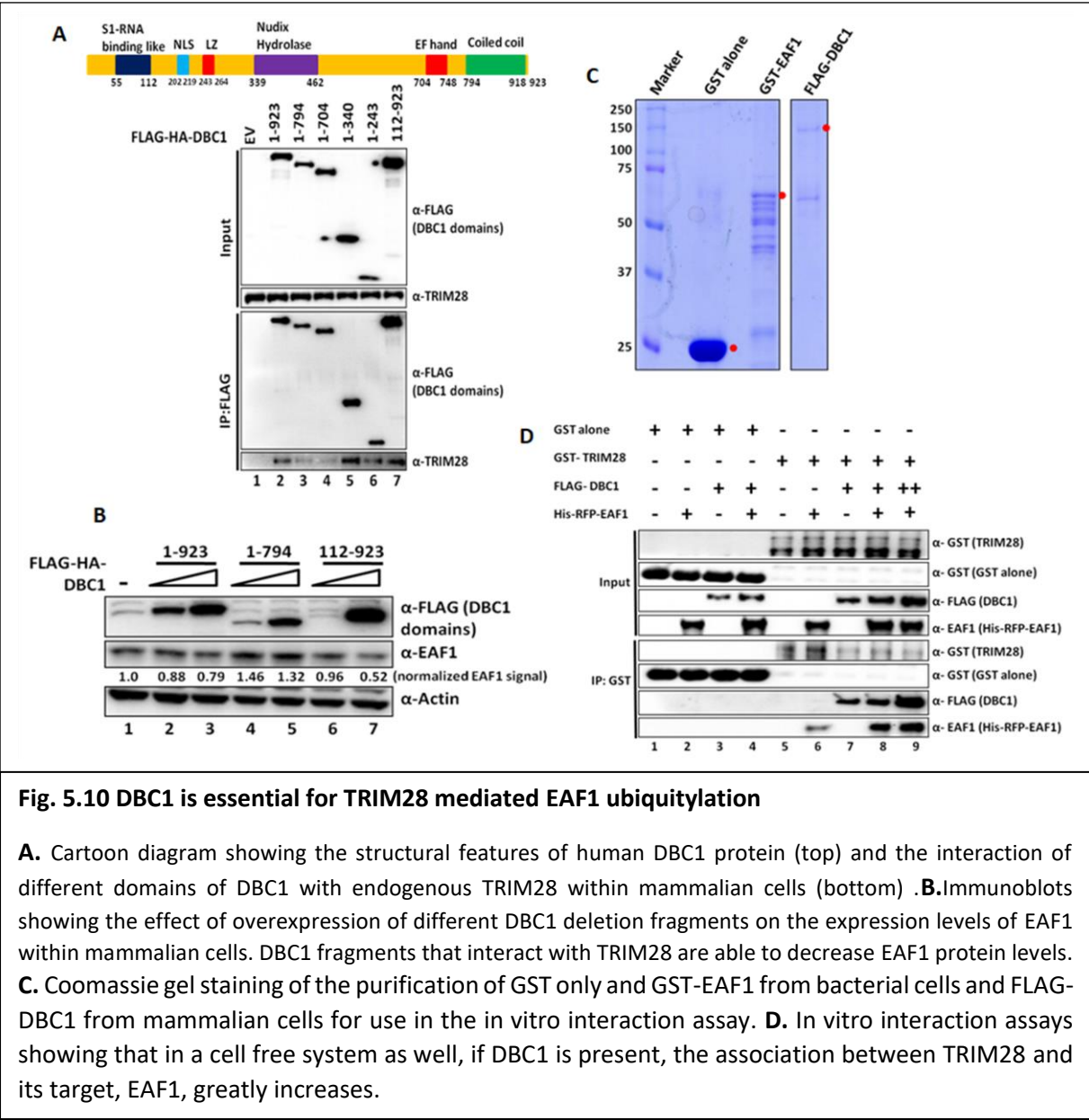
A. Mass spectrometric analysis indicating the presence of TRIM28 peptides in our EAF1.com prep **B.** Immunoprecipitation and subsequent western blot analysis showing a novel interaction between EAF1 and TRIM28 within mammalian cells **C.** TRIM28 also interacts with DBC1, as tested by overexpressing DBC1 and checking for its interacting proteins. **D.** A reciprocal reaction confirms that TRIM28 does indeed associate with DBC1 within mammalian cells. **E.** Western blot analysis indicates that TRIM28 is able to degrade EAF1 within mammalian cells, which can be rescued by the treatment of cells with MG132. **F.** Immunoblots showing that in the presence TRIM28, EAF1 is polyubiquitinated within mammalian cells.

5.10 DBC1 forms a bridge between TRIM28 and EAF1

The presence of DBC1 has been seen to cause ubiquitin-mediated degradation of EAF1. Looking into the mass spectrometric data of EAF1 associated proteins, we noted the presence of TRIM28, known to possess E3 ligase properties. Thus, we hypothesized and confirmed with our experiments that TRIM28 was indeed the E3 ligase for EAF1. However, in our experiments, we had two very interesting observations: (a) TRIM28 interacts with DBC1 within mammalian cells and (b) The effect of both DBC1 and TRIM28 on EAF1 were similar in nature. As DBC1 is not known to possess the capability to cause polyubiquitination of any protein, we postulated that the DBC1's role in this association was that of an adapter protein, or a bridging protein, which would enable the enzyme, TRIM28, to come into close proximity to its target, i.e. EAF1. Hence, the presence of both DBC1, as well as TRIM28, would ideally be necessary for ubiquitin-mediated degradation of EAF1. In order to examine this possibility, we first sought to study in detail, the region of DBC1 which would be necessary for its interaction with TRIM28. For this, first, we generated several deletion mutants of DBC1 based on its several structural regions as shown in Fig. 5.10A (Top schematic). Next, we expressed these deletion fragments of DBC1 along with the full-length protein and tested for the association of these domains with endogenous TRIM28.

Our results indicate that the amino acids between the 244-340 region of DBC1 are essential for its interaction with TRIM28 (Fig. 5.10A, lower panel, compare lane 1 vs lane 5 and lane 6 of the TRIM28 blot in the IP panel). Additionally, we also find that the presence of the N-terminal 111 amino acids of DBC1 is detrimental for its association with TRIM28. In this regard, we observed that removal of these amino acids from the N terminus of DBC1 causes it to interact strongly with TRIM28 (Fig. 5.10A, Compare lane 2 and 3 vs lane 7 of the TRIM28 blot). Now, if our hypothesis is correct, and DBC1 does indeed form a bridgeway between TRIM28 and EAF1, the expression of EAF1 should be reduced in the presence of DBC1 deletion mutants that interact strongly with TRIM28, and remain unaffected when TRIM28-interaction-deficient mutants are transfected. To test this, we introduced increasing concentrations of three plasmids which expressed the full-length DBC1, DBC1 (1-794 amino acids), and the DBC1 (112-923 amino acids) and measured the EAF1 levels within HEK-293T cells. As expected, the transfection of full-length DBC1 brought down EAF1 protein moderately (Fig. 5.10B). However, we obtained some interesting results upon the introduction of the DBC1 (1-794 amino acids) and the DBC1 (112-923 amino acids) fragment. The DBC1 (112-923 amino acids) deletion mutant caused a significant loss of cellular EAF1, decreasing its levels by about 50%. The other deletion fragment transfected, i.e., DBC1 (1-794 amino acids) did not have any effect on the expression levels of endogenous EAF1 (Fig. 5.10B). Incidentally, when we looked back at our earlier data, we found that the DBC1 (112-923 amino acids) mutant interacts quite strongly with TRIM28, whereas the 1-794 amino acid containing fragment fails to do so, indicating that the recruitment of TRIM28 is dependent on DBC1, which finally enables EAF1 degradation by polyubiquitylation. Before drawing any definitive conclusions about this bridging mechanism, however, we wanted to test this hypothesis in a cell-free system, unaffected by the presence of any other contaminating factors. For this, we first purified GST-TRIM28 and His-RFP-EAF1 using a bacterial expression system and FLAG-DBC1 (Fig. 5.10C) after transfection and affinity-based purification from mammalian cells and used them in an *in vitro* direct interaction assay. In this experiment, GST-TRIM28 was kept immobilized on anti-GST agarose beads and incubated with similar concentrations of EAF1 in an appropriate buffer. Also present in this reaction mix were increasing amounts of DBC1. After the reaction was completed, the proteins bound to GST-TRIM28 were analyzed by western blotting. As the blots in Fig. 5.10D suggest, we have observed a stable direct interaction between TRIM28 and EAF1 (See lane 6 of the EAF1 blot). We have also validated our earlier observation of interaction between TRIM28 and DBC1 (See lane 7 of the DBC1 blot). However, as more and more DBC1 was introduced into the reaction mix, greater was the association seen between EAF1 and TRIM28 (Compare lanes 6 vs lanes 8 and 9 of the EAF1 blot in the IP panel). Such clear results irrefutably point to the fact that indeed the direct interaction between EAF1 and TRIM28 is markedly enhanced in the presence of DBC1. Thus, DBC1 most certainly acts as a bridge between these two proteins. and EAF1 interaction with TRIM28, its E3 ligase, requires the

presence of DBC1. Hence, both these proteins, i.e DBC1 and TRIM28 are required for the effective degradation of EAF1 within cells.



5.11 DBC1 and EAF1 function to maintain ELL protein levels within cells

We have performed several experiments addressing the regulation of ELL stability and functions by DBC1, wherein HDAC3 mediated ELL deacetylation and its subsequent Siah1 E3 ligase mediated degradation is prevented by the presence of DBC1. Interestingly, in the presence of EAF1 as well, HDAC3 mediated deacetylation of ELL is prevented, which renders it stable. However, in DBC1

knockdown cells we have observed that the ELL protein levels are significantly reduced, an observation we see when we deplete EAF1 in different cell lines as well. Why then, are these two factors not being able to function in tandem to maintain ELL levels, which is necessary for the maintenance of expression of a wide variety of target genes? This answer came to us with a serendipitous observation. While maintaining DBC1 stable knockdown cells, as a measure of experimental setup integrity, we would check upon every passage, the level of ELL protein. Interestingly, we found that in the initial passages of the stable cells, ELL protein was severely diminished, an observation that we have seen previously as well (Fig. 5.11A, lane 1 vs lanes 3 and 4 of the ELL blot and [245]). However, as we continued maintaining the cells, we saw that the ELL protein levels reverted to almost similar to that of the control scramble cells. This result was unexpected and a hindrance to our experimental needs, but we thought to explore the reason behind this phenomenon. We reasoned that it could be possible that since ELL is essential, and its “protector” DBC1 has been genetically depleted, the other factor responsible for the maintenance of ELL, i.e., EAF1 could be playing a role under these circumstances to ensure normal cellular functions are not hampered. Indeed, testing the EAF1 levels across different passages clearly indicates that as we keep sub-culturing the cells, the EAF1 protein increases by up to 4 folds, which would then act on the cellular ELL and provide it stability (Fig. 5.11A, Compare EAF1 protein across different passages, and ELL variation between lanes 1 and 3-4 vs lanes 5-6). We have also explored whether these changes observed by immunoblotting are a result of perturbations at the transcript level, or are limited to regulation at the level of mature proteins. For this, we performed qRT-PCR analyses for testing ELL, EAF1, and DBC1 mRNA across the different passages. As seen previously, the knockdown of DBC1 does cause a decrease in the protein level of ELL, but its transcripts remain unaffected (Fig. 5.11B, left panel). As DBC1 is stably knocked down by shRNA, the mRNA transcription of DBC1 remains significantly decreased across passages when compared to the control scramble cells (Right panel of Fig. 5.11B). We have seen previously that the overexpression of DBC1 does not have any changes on the mRNA of EAF1, but instead causes its degradation by the ubiquitin-proteasome system. On similar lines, the knockdown of DBC1 does not affect EAF1 transcription (Fig. 5.11B, middle panel), but the increase in EAF1 can be attributed to its decreased TRIM28 mediated degradation because of the absence of the bridging factor, DBC1. Next, we wondered whether this same phenomenon would also occur upon EAF1 knockdown, as ELL levels are also seen to decrease in these cells when compared to the control scramble cells. To examine this, we cultured EAF1 knockdown cells for multiple passages and tested for the protein levels of ELL and DBC1. As shown in Fig. 5.11C, we have achieved almost a knockout of EAF1 with our shRNA constructs. Testing the expression of ELL protein across passages, we found that in the initial passages, ELL protein was quite low, and phenotypically, cells appeared stressed with low proliferation rates. However, when we kept maintaining the cells, with subsequent sub-cultures, the stress granules disappeared and now cells

started looking healthier with increased proliferation capacity. Interestingly, when we tested the ELL protein in the later passages, we found that its levels had increased manifold and almost reached about 80% of its concentration in the scramble control cells (Fig. 5.11C, ELL blot). The reason for the increase in ELL could also be ascertained, and it occurred due to a significant enhancement of the DBC1 protein (See Fig. 5.11C, DBC1 blot, comparing lanes 1 vs 2 vs 3), which causes the stabilization of the ELL by mechanisms illustrated previously. As DBC1 transcription is adversely affected by EAF1, the depletion of endogenous EAF1 by shRNA causes a significant increase in DBC1 mRNA as is reflected in Fig. 5.11D. This translates to enhanced DBC1 protein levels being present, which in turn causes the ELL protein to be stabilized even when EAF1 is not present in the milieu. We were curious also to explore how EAF1 possibly achieves this negative regulation of DBC1, and in this context noted a recent study showing that the presence of EAF1 and EAF2 negatively regulates super elongation complex formation[199]. Thus, it could be possible that when EAF1 or EAF2 associate with ELL, the inefficient SEC formation hinders efficient transcription of DBC1. To test this, we transfected ELL alone and ELL along with EAF1 in HEK-293T cells and tested for the association of known SEC components with ELL under these two conditions. Our immunoprecipitation analyses reveal that when ELL alone is present, it retains its interaction with SEC components such as AFF1, AF9, and CDK9. However, when EAF1 is co-expressed along with ELL, the association of these components critical for RNA Pol II-mediated transcription is lost even though the same amount of ELL is present (Fig. 5.11E, compare lanes 5 vs 6 for the AFF1, AF9, and CDK9 in the IP panel). Thus, the reduced expression of DBC1 may occur due to its inefficient transcription because of the impaired formation of the super elongation complex. Our experiments thus reveal that these two factors, namely DBC1 and EAF1 work alternatively, with the final objective being to maintain steady-state levels of ELL within mammalian cells.

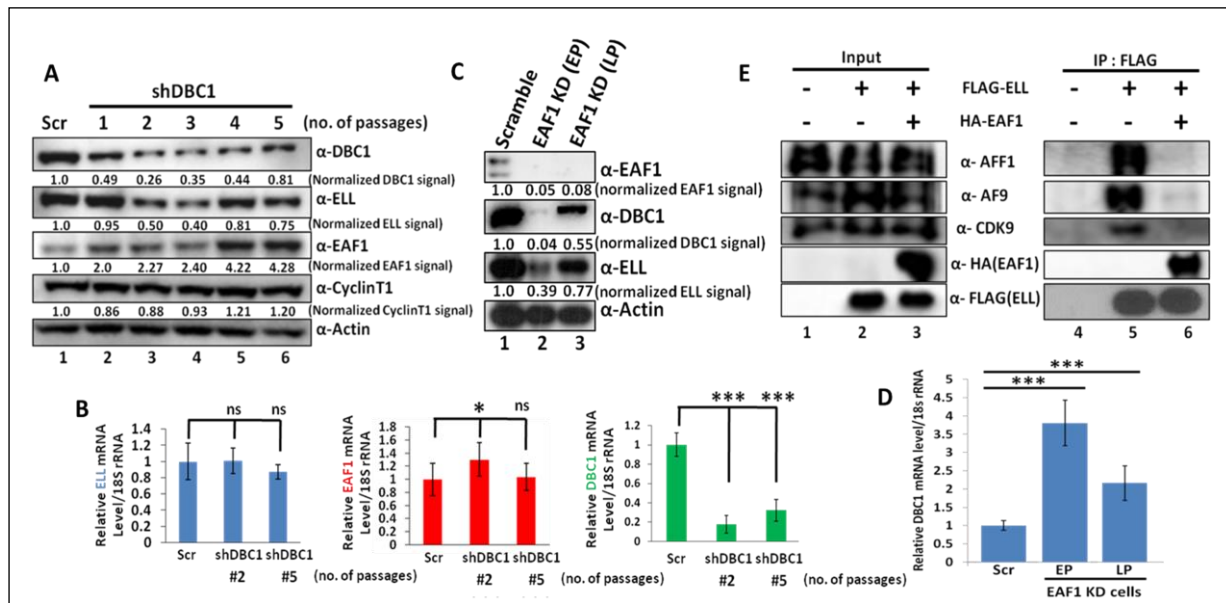


Fig. 5.11 DBC1 and EAF1 work in tandem to maintain ELL stability and thus its function within mammalian cells

A. Immunoblots of the levels of different proteins across different passages upon DBC1 knockdown in a HEK-293T cell line background. **B.** qRT-PCR analyses show that the mRNA levels of both ELL and EAF1 remain unaffected upon DBC1 knockdown in HEK-293T cells. Thus, the DBC1 mediated regulation of these two proteins is restricted to the protein level and there is no effect of DBC1 depletion on the transcription of these proteins. **C.** Western blot analyses show that upon EAF1 knockdown, the ELL decreases initially and then increases with subsequent passages due to enhanced DBC1 expression by virtue of its increased transcription. **D.** qRT-PCR analyses showing that upon knockdown of EAF1, the mRNA levels of DBC1 are enhanced significantly leading to rise in its protein levels. **E.** Immunoprecipitation analyses show that when EAF1 is co-expressed along with ELL, the association of ELL with SEC components (such as AFF1, AF9 and CDK9) decreases significantly.

The data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, ** denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

5.12 DBC1 and EAF1 modulate ELL levels and function in a context-dependent manner

Our extensive analyses in terms of ELL stability and its functions have yielded some interesting results. We have seen that ELL becomes stable and functionally active when acetylated by p300, whereas HDAC3 mediated ELL deacetylation causes its ubiquitylation and subsequent degradation. Now, both DBC1 and EAF1 function to maintain ELL stability and thus protect it from deacetylation and ubiquitination. Thus, the function of both these proteins is to maintain the ELL pool within cells. Interestingly, we have also noted that EAF1 and DBC1 negatively regulate each other, though their function effectively is the same with respect to ELL. These factors taken together makes one wonder

why two different proteins perform similar functions in the same manner. Hence, it is not very unlikely to hypothesize that the differential regulation of ELL by these two proteins occurs in a context-dependent manner. In this regard, if we look back at published literature, a study published in 2014 elegantly describes the role of the EAF1 and ELL (among others) containing Little Elongation Complex (LEC), which plays a very important part in transcription resumption after repair of DNA damage[190]. Thus, EAF1 mediated ELL stability and regulation of its functions could be involved in the DNA damage response. In order to test this, we subjected HEK-293T cells to genotoxic stress by treating the cells with doxorubicin. Doxorubicin is widely used as a chemotherapeutic drug, which intercalates between the DNA strands, thus preventing the progression and activity of topoisomerase II, causing DNA strand breaks. Cells were treated with doxorubicin for different time periods and then the expression of our target proteins was tested by western blotting. Our experimental observations indicate that within 1 hour of treatment, DBC1 protein decreased significantly, whereas both EAF1 as well as ELL levels were seen to increase (Fig. 5.12A, Compare lane 1 vs lane 2 and quantitation below each blot). As the time of treatment progressed, the DBC1 protein was seen to decrease continually and started reverting back to its original levels only about 8 hours after treatment (Compare different lanes in the DBC1 blot). However, as seen in the initial time points, as a result of increased EAF1, the ELL protein was also seen to be significantly enhanced. CyclinT1, taken as a negative control, remained unaffected. Now, as EAF1 was seen to be involved in ELL regulation upon genotoxic stress, for which the initial response is the rapid shutdown of transcription, we wondered that, since DBC1 negatively regulates EAF1, would the DBC1 mediated ELL regulation be necessary for processes in which rapid transcription is required? In this case, we thought to identify stimuli that would result in the rapid induction of transcription, and then test how our target proteins behave when such signals are encountered by the cell. In this aspect, we thought to treat the cells with epidermal growth factor (EGF), a compound known to elicit bursts of transcription within minutes of its introduction within cells. Likewise, as with the doxorubicin treatment, the protein levels of the target factors were measured at different time intervals by western blotting. In this case, we were surprised to note a significant increase in the DBC1 protein, which, as a downstream effect caused the ELL protein also to be moderately stabilized. However, the EAF1, which we have seen earlier to behave opposite to DBC1, was seen to decrease rapidly, indicating that for stimuli requiring rapid transcriptional induction, the DBC1 functions in stabilizing ELL and enables it to cause transcription of its target genes (Fig. 5.12 B, compare the DBC1 blots with EAF1 expression over different time points). Thus, our initial experiments indicate that, when an initial transcriptional pause followed by a delayed restart is required, EAF1 plays a predominant role in ELL stability and functions. In this case, the DBC1 expression is seen to be decreased significantly. However, the DBC1 dependent pathway of ELL-mediated transcriptional regulation is adopted when a rapid induction of transcription is necessary, in which case, to prevent the initial inhibitory effect of

EAF1, we see that its cellular levels get reduced. Thus, a genotoxic stress-dependent transcriptional response involving ELL seems to be regulated by ELL, whilst in the case of growth factor-dependent stimuli, DBC1 functions to modulate ELL actions. To confirm these hypotheses, we decided to perform two simple experiments. First, we treated control Scramble shRNA, shDBC1, and shEAF1 expressing stable knockdown cell lines to doxorubicin for different time points and like done previously, tested for the expression of ELL. As shown in Fig. 5.12C, upon genotoxic stress, ELL protein increased rapidly in the control scramble shRNA cells, similar to the response seen in the HEK-293T cell line (Compare ELL expression between lanes 1-5). A similar fate could be seen for the ELL protein in the shDBC1 cell line as well (See lanes 11-15 in the ELL blot). However, no changes in ELL levels were observed in the shEAF1 cell line. These results clearly indicate that in the DNA damage response, it is the EAF1 that modulates the expression levels of ELL. Thus, in a DBC1 depleted cell line, as EAF1 levels are maintained, the increase in ELL levels occur due to the augmented EAF1 expression. However, as this response is EAF1 dependent, when endogenous EAF1 is knocked down, even though DBC1 is present within the cell, the ELL levels remain unperturbed. Along similar lines, next, we subjected the cells to EGF mediated stimulation as well, and here again, tested for the expression of ELL protein. In this case, however, we observed an opposite response to the one seen for doxorubicin. Upon EGF treatment, ELL was, like before, seen to increase in the control scramble shRNA cell line (Fig. 5.12D, Compare lanes 1-3 in the ELL blot). Similar results were also observed for ELL protein upon exposure of the shEAF1 cells to EGF (See lanes 4-6). However, similar to the effect of doxorubicin seen on ELL protein in the shEAF1 cell line, in this case, ELL levels remained unchanged upon stable knockdown of DBC1 (Fig. 5.12D, lanes 7-9). These results, taken together, clearly indicate that in the case of EGF treatment, DBC1 mediated enhancement of ELL stability and its functions occurs, which is independent of EAF1 dependent ELL regulation. Thus, an intricately interconnected network exists within mammalian cells for the modulation of ELL protein and ensures its proper functioning depending on different contexts and external stimuli. To confirm this mode of DBC1 mediated ELL regulation upon EGF exposure, we carried out another experiment as well. The cellular response to EGF is rapid and occurs within a matter of minutes, inducing the expression of genes known as immediate early genes. These genes include growth factors as well as transcription stimulators such as *FOS* and *JUN*. Immediate early genes are known to affect the induction of late response genes, which see enhanced expression after some hours of treatment and maintain this increased expression for an extended period of time. We thus decided to test the effect of EGF treatment on the mRNA levels of some late response genes (also ELL targets) such as *BCL6*, *CDKN2C*, *CDKN1A*, *CDKN1B* and *CCND1*. qRT-PCR analyses performed on these genes indicate that almost all of them are induced within 4 hours of EGF treatment and stay in a state of heightened transcription for atleast 8 hours (Fig. 5.12E, see the induction in the Scramble panel for the different time points). Our hypothesis indicated that this induction is dependent on DBC1, and our

observations also fall along similar lines, wherein in DBC1 knockdown cells, there is little to no change in mRNA levels of the target genes upon EGF treatment (See the induction kinetics for the DBC1 knockdown panel). In case of EGF-mediated target gene induction in EAF1 knockdown cells, our experiments showed some interesting results. Upon EGF treatment, almost all of the target genes showed heightened transcription, sometimes surpassing levels seen even in the control scramble cells. This response could be attributed to the control of mRNA levels of DBC1 by EAF1. In the shEAF1 cells, due to the presence of increased DBC1, the response to treatment with EGF is also possibly increased to a greater extent (Fig. 5.12E, EAF1 KD panel). This effect is specific for ELL target late response genes only, as *CCNE2*, whose expression is largely unaffected by ELL, remains unperturbed by the knockdown of either DBC1 or EAF1.

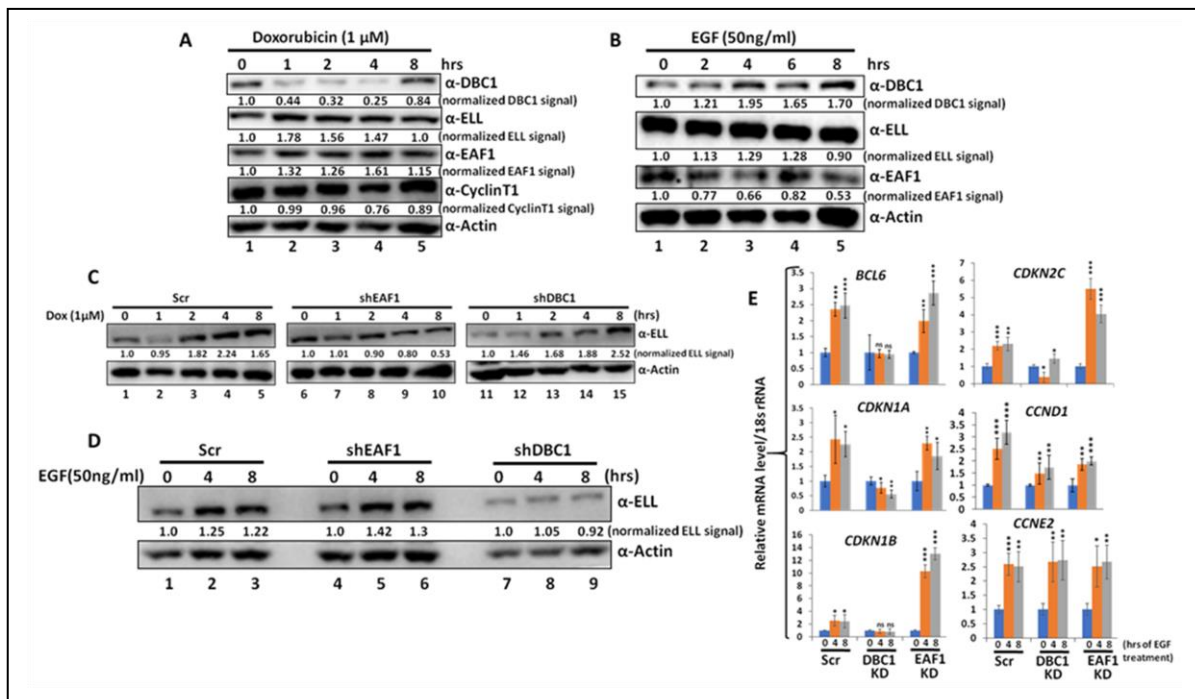


Fig. 5.12 ELL protein levels and its cellular functions are regulated by DBC1 and ELL in a context dependent manner

A. Immunoblots showing the effect of doxorubicin treatment on the protein levels of ELL, EAF1 and DBC1 within HEK-293T cells **B.** Treatment of cells with EGF causes an increase in ELL levels, effected by an increase in the DBC1 protein levels. The EAF1 expression upon EGF treatment is concomitantly decreased. **C.** Western blot analyses indicate towards the EAF1 mediated ELL regulation in case of genotoxic stress, wherein upon doxorubicin treatment, ELL levels are enhanced in shScramble and shDBC1 cells, but unaffected when EAF1 is seen to be knocked down. **D.** A similar assay performed after EGF treatment in these three cell lines show that ELL increases in the shScramble and shEAF1 cell lines. However, as this pathway is regulated by DBC1, upon its knockdown, ELL protein remains stagnant in the shDBC1 cell line. **E.** qRT-PCR analyses show that the effect of EGF on the enhancement of ELL stability is dependent on DBC1 as when DBC1 is stably knocked down, EGF responsive late genes show impaired induction upon EGF treatment. The data is represented as mean $\pm$ S.D. Student's t test was used to calculate the statistical significance of the data in this figure. * denotes $p \leq 0.05$, ** denotes $p \leq 0.01$, *** denotes $p \leq 0.001$, and ns denotes not significant. At least two biological replicates were performed for each experiment.

Our results thus indicate that a complex mechanism exists involving several proteins which function in a contextual manner to regulate the stability and proper functioning of the key elongation factor, ELL within cells. Subsequent studies are still ongoing to delve deeper into this intricate network. A cartoon model of this mode of regulation that has been unraveled till now is shown in Fig. 5.13.

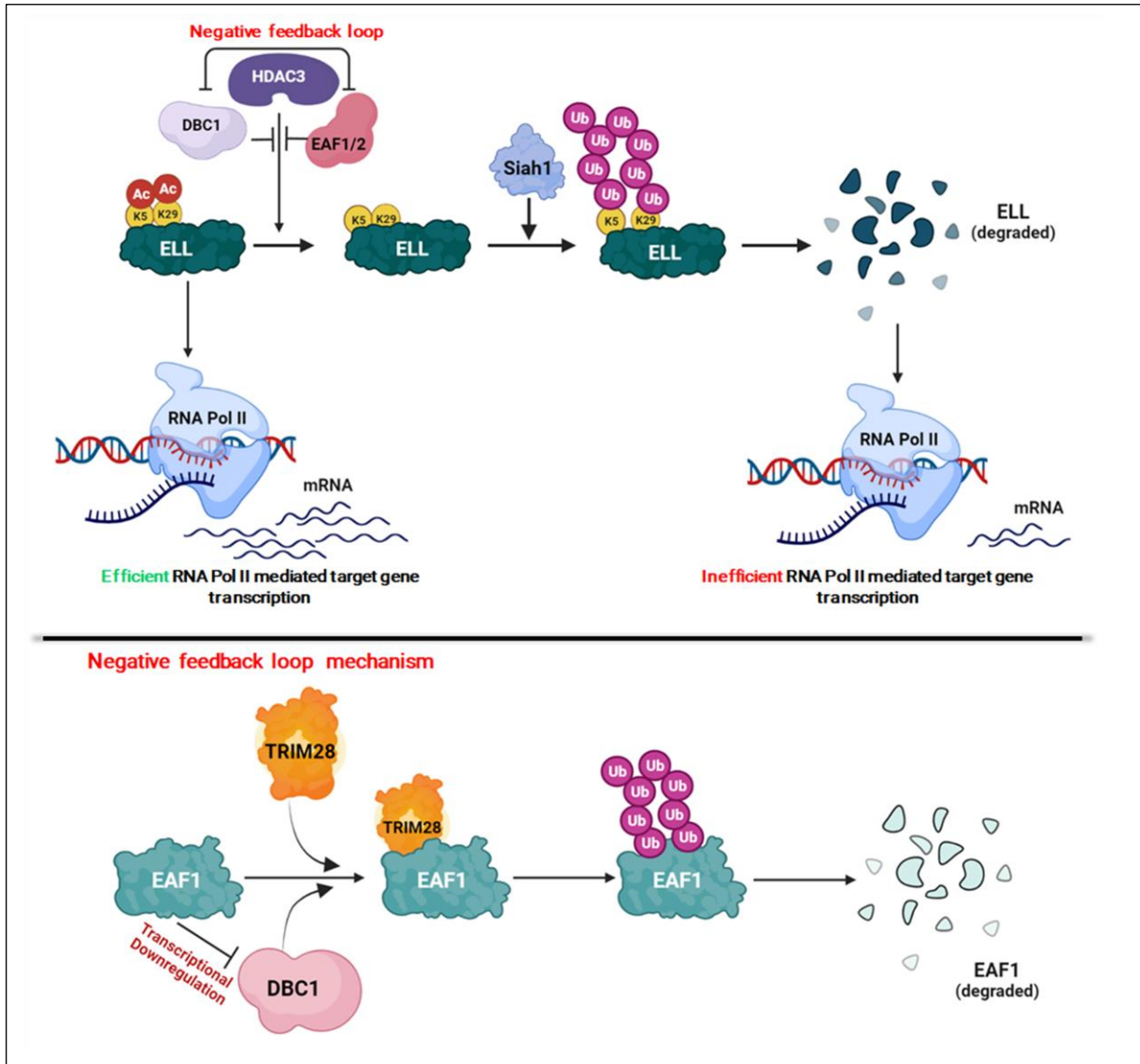


Fig. 5.13 Cartoon diagram depicting the regulation of ELL by the action of EAF1/EAF2 and DBC1 within cells for efficient transcription of target genes

ELL is acetylated at its N-terminus which provides it stability and causes efficient transcription of its target genes within cells, enabling the maintenance of a state of homeostasis. ELL degradation requires a prior deacetylation of the target lysine residues by HDAC3, which facilitates its polyubiquitylation by the E3 ligase Siah1. Both EAF1/ and DBC1 maintain the acetylated state of ELL and thus its levels by competitively inhibiting HDAC3 binding to the N-terminus of ELL. Additionally, a cellular cross-talk exists between DBC1 and EAF1 by which these two proteins regulate each others levels by two distinct mechanisms. The transcription of DBC1 is prevented by EAF1, while DBC1 mediates the interaction between EAF1 and its E3 ligase, TRIM28, thereby causing EAF1 degradation by the proteasome.

Chapter 6

Summary

Summary

Our studies began with the observation of a putative association between ELL, a key RNA Pol II elongation factor, and DBC1, a protein implicated in the regulation of various cellular processes. We first performed biochemical association experiments to verify the presence of direct interaction between ELL and DBC1. Reciprocal interaction analyses for the identification of DBC1 associated proteins yielded surprising results, wherein, including ELL, several other SEC components such as AFF1, EAF1, EAF2, CDK9 and AF9 were found to associate with DBC1, an observation whose basis we would understand later. By means of a reporter gene assay, ELL and DBC1 were found to be capable of acting in an additive manner on RNA Pol II-mediated transcription.

During our luciferase reporter gene assays, a fortuitous observation gave a new direction to our study, when we observed that DBC1 was able to increase ELL protein levels without having an effect on its transcription. As DBC1 was already known to have a similar effect on p53, we hypothesized that ELL may be acetylated within cells and DBC1 may act as its protector by preventing the effect of cellular deacetylases. Indeed, our studies showed that ELL could be acetylated robustly by the histone acetyltransferase, p300, and several deacetylases could interact strongly with both DBC1 and ELL, the most significant interaction being with HDAC3. Additionally, the presence of HDAC3 was seen to result in the degradation of ELL by the ubiquitin-proteasome mediated system.

Interestingly, the overexpression of p300, resulting in the acetylation of ELL, could rescue HDAC3 mediated degradation. This effect was also mirrored by DBC1 by its competitive binding to ELL, which prevented its association with HDAC3, thus preventing its deacetylation and degradation by ubiquitination. Detailed analyses revealed that two lysine residues at the N-terminus of ELL, located at K5, and K29 were sites for p300 mediated acetylation. In the presence of HDAC3, the acetylation marks were removed, paving the way for the conjugation of ubiquitin molecules at these specific residues by the E3 ligase Siah1, finally resulting in the degradation of ELL. The major function of DBC1 thus appears to be the maintenance of ELL protein for the proper expression of ELL target genes, which were found to regulate a wide diversity of cellular processes, such as cell proliferation, growth, metabolism, and RNA Pol II-mediated transcription.

Delving deep into the mechanisms of how the expression of such a wide variety of genes could be regulated by DBC1, we performed chromatin immunoprecipitation assays for recruitment of different factors at the transcription start site of different target genes. We found that upon depletion of DBC1, the recruitment of the TFIID complex for transcription initiation, the functional super elongation complex for elongation, as well as the RNA Pol II for catalyzing the entire process was severely

hindered. Immunoblotting analyses revealed that the reason behind such drastic effects is highly likely due to the role of DBC1 in the regulation of these key processes. For example, upon DBC1 knockdown, the TAF6 protein of the TFIID complex was significantly reduced. As TAF6 acts as a scaffold for the TFIID complex, it is not difficult to think that its absence would prevent the formation of the TFIID, which in turn would result in an incomplete pre-initiation complex (PIC). Likewise, as ELL has been seen to be responsible for the recruitment of RNA Pol II, its decrease due to the absence of DBC1 may result in the reduced Pol II we see at the TSS of the target genes. Additionally, we observed that the knockdown of DBC1 causes the decrease in protein levels of several SEC components, including the AFF1, which is vital for maintaining the integrity of the complex and thus results in a diminished transcriptional output of the respective genes.

Gene ontology analyses revealed that the misregulation of genes regulated by the ELL-DBC1 axis could lead to the development of several diseases, the predominant of which was cancer. As our initial studies were directed towards understanding the role of ELL regulation with particular emphasis on mixed lineage leukemia, we initially thought to focus on this aspect. However, our analyses showed that the misregulation of the DBC1-ELL axis could also potentially lead to the development of Type 2 Diabetes. This would be an interesting and unique aspect for us to focus on, thinking which we sought to identify if this mode of regulation of ELL would be prevalent in this condition. Indeed, protein expression studies from treatment naïve type 2 diabetes patient cells showed that in this condition, levels of cellular DBC1 were reduced, when compared to healthy controls, which also resulted in the decrease of ELL expression. This caused decreased expression of its target genes involved in the metabolism and transport of glucose, thus leading to its accumulation in the peripheral system, a hallmark of diabetes.

We also explored the role of EAF1 and EAF2, two ELL-associated factors, whose role in conjunction with ELL had not been studied in detail. We found by transfection-based assays that EAF1 and EAF2 majorly bind within the N terminal 60 amino acids of ELL, overlapping with the interaction sites of HDAC3. Thus, these two proteins also prevent HDAC3-ELL interaction, thereby protecting ELL and maintaining its steady-state levels within the cell. The mode of regulation observed in these cases appeared to be quite similar to that of DBC1 with respect to ELL, which begged us to ask the question if these two proteins somehow could also regulate each other. Our experiments showed that both DBC1 and EAF1 regulate each other's expression in two different ways. When EAF1 is present in excess, it impedes DBC1 mRNA transcription, while in an opposite manner, DBC1 enables the recruitment of the E3 ubiquitin ligase, TRIM28, and facilitates its association for polyubiquitination and proteasome-mediated degradation of EAF1.

Seeing that both DBC1 and EAF1/EAF2 functioned by two different pathways with the same aim: to maintain ELL steady-state levels within the cell, we wondered if this dual mode of regulation was random or occurred in a context-dependent manner. To study this, we treated cells to conditions conducive as well as unfavorable to growth, i.e: by supplementing growth factors or by the introduction of genotoxic stress. We found that upon treatment of cells with growth factors, EAF1 is reduced, but DBC1 increases, thus resulting in the enhancement of ELL and resulting in proper expression of growth-related genes. Similarly, treatment of cells with genotoxic agents caused an increase in ELL stability mediated by enhanced EAF1, while DBC1 protein expression was severely hindered. Thus, our experiments elucidate two distinct pathways that occur within the cell, under two highly different scenarios, but with the same outcome. The major objective is for the maintenance of the levels of the key elongation factor, ELL, which is achieved by two different networks of proteins, in a context-dependent manner, facilitating the expression of appropriate sets of genes under the effect of distinct stimuli.

Chapter 7

Discussion

Discussion

The Super Elongation Complex (SEC) plays an essential role in mammalian transcription elongation mediated by RNA Pol II, and thus serves as an attractive target for detailed studies on the mechanisms and interplay between factors involved in eukaryotic transcription. The SEC has been shown to contain several component proteins, which dynamically associate with each other in a context-dependent manner[145, 146, 251]. Of all the members of the SEC, Eleven Nineteen Lysine Rich Leukemia (ELL) is the only bona fide RNA Pol II elongation factor, whose presence can stimulate Pol II-mediated transcription elongation[168]. This makes ELL an important factor for studies on transcription regulatory mechanisms.

Our studies on transcription regulation in eukaryotes are conducted in a manner distinct from most studies. While the majority of reports focus on the “effect” of various transcription-associated factors, our work shows the “cause” behind the regulation of functions of the key factor responsible for RNA Pol II-mediated transcriptional elongation, ELL. Our studies show how ELL stability and thus its functions are regulated in a context-dependent manner by the concerted action of several proteins such as DBC1, the lysine acetyltransferase p300, the deacetylase HDAC3, the E3 ubiquitin ligase Siah1, and the ELL-associated factors EAF1 and EAF2.

In the first part of our studies, we focus on understanding the functional implication of stable, direct interaction between ELL, and the versatile protein seen to be involved in various cellular processes, DBC1. Our studies have shown the functional regulation of ELL occurs through two different post-translational modifications. Through detailed experiments conducted within different cells as well as *in vitro* using purified proteins, we show that ELL is post-translationally acetylated by the lysine acetyltransferase p300, which causes its levels to increase within the cell and maintain its functionally active state. Acetylation is important for ELL stability and its functions as depletion of cellular p300 dramatically enhances ELL degradation kinetics as observed by cycloheximide (CHX) chase assays. However, when ELL needs to be reduced, it undergoes polyubiquitination and gets degraded by the effect of the ubiquitin-proteasome system. A recent report has shown that the E3 ligase, Siah1, is often found associated with the SEC[252]. The study showed that two components of the SEC, AFF1, and the homolog of ELL, ELL2 was affected by the presence of this E3 ligase. This study, however, failed to observe any effect of Siah1 on ELL protein levels. Our studies have provided a basis for this observation. We find that since ELL mostly resides in an acetylated state within the cell, its ubiquitylation mediated degradation requires prior deacetylation carried out by the lysine deacetylase, HDAC3. Our detailed mechanistic studies reveal that the lysine residues subjected to acetylation and ubiquitylation are the same, and are located at the N-terminus of ELL at positions K5

and K29. Thus, for efficient ubiquitin conjugation to occur, the acetyl groups from these residues must be removed first, a process that is catalyzed by HDAC3. Because of this two-step process, we observe that while Siah1 alone cannot adversely affect ELL protein levels, in the presence of HDAC3, ELL is rapidly reduced when Siah1 is co-transfected in the assay. The importance of the presence of HDAC3 in this process becomes even more clear with the observations that overexpression of HDAC3 alone rapidly induces ELL degradation, which erstwhile has been observed to be a highly stable protein. In this entire process, the role of DBC1 is that of a “protector” of ELL acetylation, a result it achieves by simply preventing HDAC3-ELL association by competitively binding to the HDAC3 interaction sites on ELL. Thus, in the absence of DBC1, HDAC3 mediated ELL degradation is markedly enhanced, a phenomenon which is easily demonstrated with the observation that ELL protein levels are significantly decreased in DBC1 knockdown cells.

Studies on the physiological implication of this newly described ELL-DBC1 axis show that these two proteins exert significant regulatory effects on the transcription of genes involved in a wide variety of cellular processes. Functional characterization of these genes reveals them to be involved in pathways connected to cell proliferation, cell cycle control, transcriptional regulation, as well as metabolic processes. Our extended studies show, interestingly, that the perturbation of this axis can prove to be the cause of the development of type 2 diabetes, an observation we had while conducting experiments on PBMCs obtained from treatment naïve patients. These PBMCs were found to have reduced DBC1 and ELL levels, which in turn was seen to adversely affect the expression of genes involved in the maintenance of glucose homeostasis. Our study, thus, in part, elucidates a possible mechanism for the development of type 2 diabetes, which now affects about 77 million people in India and 543 million people worldwide, enabling the development of intervention therapies for this highly prevalent debilitating disease.

Our studies also show that other than ELL, the presence of DBC1 is important for the maintenance of protein levels of the TFIID components TAF1, TAF5, TAF6, and TAF7. Additionally, with the knockdown of DBC1, several integral components of the SEC, such as AFF1, AFF4, and CDK9 are seen to be decreased as well. Since TFIID is essential for transcription initiation and bridging initiation with elongation, while the super elongation complex finds importance in Pol II-mediated transcription elongation, the possible regulation of these proteins by DBC1 may prove to be instrumental in understanding the regulation of eukaryotic transcription. Indeed, further studies may demonstrate this novel ELL interacting factor, DBC1 as a master regulator of transcriptional processes.

The second part of our studies focuses on the contextual regulation of ELL protein levels. ELL-associated factors 1 and 2 (EAF1 and EAF2) have been identified as stable interactors of ELL since about two decades ago. However, detailed studies are yet to be conducted to understand how these

two factors work alongside ELL to affect eukaryotic transcription. Our detailed biochemical experiments show that EAF1 and EAF2 also work to stabilize ELL within cells and that the interaction between ELL and these two proteins is essential for this mode of regulation. Like DBC1, EAF1 and EAF2 possess similar binding sites on ELL as HDAC3, and thus compete with it for binding to the N-terminus of ELL, thereby preventing its polyubiquitination mediated degradation. Hence, like DBC1, the depletion of cellular EAF1 and EAF2 causes a significant reduction in ELL protein levels and adversely affects the expression of its target genes, leading to their decreased proliferation and survival potential.

However, we have seen that two different proteins regulate ELL stability and thus its functions by the same mechanism, leading us to ask the question if this regulation occurs in a contextual manner, or is a generalized phenomenon that occurs randomly. In trying to answer this question, we identified the existence of a negative feedback loop that regulates the cellular levels of DBC1 and EAF1. We found, through biochemical assays, that DBC1 and EAF1 can negatively regulate each other's expression by two distinct mechanisms. When EAF1 expression is enhanced, it results in a significantly reduced DBC1 mRNA expression, thus decreasing its protein concentrations within the cell. Likewise, DBC1 acts as a bridge between the E3 ligase TRIM28 (KAP1) and EAF1, mediating their interaction and causing the degradation of EAF1 by polyubiquitination. However, this regulation of DBC1 and EAF1 does not occur randomly but appears to be a carefully crafted mechanism that occurs in the cell upon the perception of different stimuli. When cells are exposed to growth factors, we observe an increase in DBC1 protein with a concomitant enhancement in ELL levels as well. Concurrently, a sharp drop in EAF1 is also observed, possibly effected by the enhanced TRIM28-EAF1 interaction mediated by DBC1. Now, when cells encounter stimuli that are not conducive to growth, such as exposure to genotoxic stress, we observe that there is a linear increase in EAF1, which results in the reduction of DBC1 due to its decreased transcript levels. However, since EAF1 also stabilizes ELL by binding to its N terminal end, the exposure of cells to genotoxic agents such as doxorubicin also leads to an increase in ELL protein. Why these two different pathways are adopted can be plausibly explained by our observations that upon EAF1 overexpression, ELL association with SEC components such as AFF1, AF9 and CDK9 are dramatically reduced. Thus, in the presence of growth-inducing factors, rapid transcription of cell growth and proliferation-related genes, which are mediated by the SEC, are required. As EAF1 would possibly inhibit the SEC assembly, it is plausible that the TRIM28 mediated reduction of EAF1 protein serves as a favorable environment for SEC formation. Similarly, the EAF1 and ELL containing Little Elongation Complex (LEC) has been seen to play an essential role in transcription-coupled DNA repair. The DNA damage response also needs rapid shutdown of transcription, which is probably mediated by the increased EAF1. Enhanced EAF1 would thus sequester ELL away from the SEC for LEC mediated efficient DNA damage response, as well as cause reduction in active transcription as the low DBC1

would result in inefficient TFIID as well as SEC assembly. Our work thus demonstrates an elegant mechanism that occurs within the cell for a highly specific response to an extracellular stimulus, involving the differential regulation of the same factor, ELL, under two varied conditions. The overall aim appears to be the same, i.e: maintenance of ELL stability, but is undertaken by two different factors working along two different pathways.

Negative feedback loop mechanisms have been extensively described as a cross-talk between two different cellular proteins, particularly in the context of regulating the expression of circadian clock-related genes. It is interesting to note that a negative feedback loop mechanism also acts between p53 and poly(A)-specific ribonuclease (PARN), wherein PARN deadenylase keeps p53 mRNA level low under normal cell growth conditions by degradation of its mRNA[253]. Likewise, the Amit group from Weizmann Institute has described negative feedback loop mechanisms between transcription factors involved in the development of different dendritic cell lineages governing chromatin modifications[254]. However, to the best of our knowledge, ours is the first study that mechanistically how two different proteins with widely varying functions within the cell regulate each other's levels for maintaining the expression of a key RNA Polymerase II elongation factor through similar mechanisms in a context-dependent manner. Given the fact that our initial observations have shown that DBC1 can affect the assembly of the TFIID and also affects protein levels of several key SEC components, it is not difficult to speculate that such regulatory mechanisms may exist for the maintenance of other key factors within cells, whose expression may be modulated as per the cellular requirement.

The different members of the SEC have been seen to form stable fusions with the Mixed Lineage Leukemia (MLL) protein, leading to the development of aggressive leukemias. Hence, our initial aim in these studies was to understand the role of these proteins in transcriptional regulation and translate our findings to explain the mechanisms of development of MLL. Our studies indicate that the N-terminus of ELL governs its stability, and we have observed that deletion fragments of ELL devoid of the N terminal 45 amino acids express abundantly and are highly stable. Interestingly, these amino acids are also lost in the MLL-ELL fusion protein, which could be a possible mechanism for its enhanced stability and persistence[255]. Additionally, it is imperative to mention that a previous study has shown that the MLL-EAF1 artificially generated fusion protein has the same capability to immortalize myeloid progenitor cells as MLL-ELL, and MLL-EAF1 induced AML, surprisingly, has a much lesser latency than MLL-ELL[256]. Hence it is highly likely that the ELL interaction and stability maintained by EAF1 may play an essential role in this process. Finally, as DBC1 has been seen to regulate the expression of several proteins that contribute to the formation of highly aggressive leukemias such as

AFF1, it could be interesting to undertake further studies to understand if DBC1 also has any role to play in MLL mediated leukemic transformation.

References

1. Portin, P. and A. Wilkins, *The Evolving Definition of the Term "Gene"*. Genetics, 2017. **205**(4): p. 1353-1364.
2. Koonin, E.V., *Does the central dogma still stand?* Biol Direct, 2012. **7**: p. 27.
3. Kelly, T. and A.J. Callegari, *Dynamics of DNA replication in a eukaryotic cell*. Proc Natl Acad Sci U S A, 2019. **116**(11): p. 4973-4982.
4. Cramer, P., *Organization and regulation of gene transcription*. Nature, 2019. **573**(7772): p. 45-54.
5. Kapp, L.D. and J.R. Lorsch, *The molecular mechanics of eukaryotic translation*. Annu Rev Biochem, 2004. **73**: p. 657-704.
6. Sokabe, M. and C.S. Fraser, *Toward a Kinetic Understanding of Eukaryotic Translation*. Cold Spring Harb Perspect Biol, 2019. **11**(2).
7. Hillier, L.W., et al., *Genomics in C. elegans: so many genes, such a little worm*. Genome Res, 2005. **15**(12): p. 1651-60.
8. Gibcus, J.H. and J. Dekker, *The context of gene expression regulation*. F1000 Biol Rep, 2012. **4**: p. 8.
9. Lee, T.I. and R.A. Young, *Transcriptional regulation and its misregulation in disease*. Cell, 2013. **152**(6): p. 1237-51.
10. Wong, F. and J. Gunawardena, *Gene Regulation in and out of Equilibrium*. Annu Rev Biophys, 2020. **49**: p. 199-226.
11. Sutherland, C. and K.S. Murakami, *An Introduction to the Structure and Function of the Catalytic Core Enzyme of Escherichia coli RNA Polymerase*. EcoSal Plus, 2018. **8**(1).
12. Browning, D.F. and S.J. Busby, *Local and global regulation of transcription initiation in bacteria*. Nat Rev Microbiol, 2016. **14**(10): p. 638-50.
13. Roberts, J.W., *Mechanisms of Bacterial Transcription Termination*. J Mol Biol, 2019. **431**(20): p. 4030-4039.
14. Barba-Aliaga, M., P. Alepuz, and J.E. Perez-Ortin, *Eukaryotic RNA Polymerases: The Many Ways to Transcribe a Gene*. Front Mol Biosci, 2021. **8**: p. 663209.
15. Turowski, T.W. and M. Boguta, *Specific Features of RNA Polymerases I and III: Structure and Assembly*. Front Mol Biosci, 2021. **8**: p. 680090.
16. Turowski, T.W. and D. Tollervey, *Transcription by RNA polymerase III: insights into mechanism and regulation*. Biochem Soc Trans, 2016. **44**(5): p. 1367-1375.

17. Basu, S., A. Nandy, and D. Biswas, *Keeping RNA polymerase II on the run: Functions of MLL fusion partners in transcriptional regulation*. Biochim Biophys Acta Gene Regul Mech, 2020. **1863**(8): p. 194563.
18. Proudfoot, N.J., *Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut*. Science, 2016. **352**(6291): p. aad9926.
19. Le, N.Q.K., et al., *Classifying Promoters by Interpreting the Hidden Information of DNA Sequences via Deep Learning and Combination of Continuous FastText N-Grams*. Front Bioeng Biotechnol, 2019. **7**: p. 305.
20. Haberle, V. and A. Stark, *Eukaryotic core promoters and the functional basis of transcription initiation*. Nat Rev Mol Cell Biol, 2018. **19**(10): p. 621-637.
21. Carninci, P., et al., *Genome-wide analysis of mammalian promoter architecture and evolution*. Nat Genet, 2006. **38**(6): p. 626-35.
22. Vo Ngoc, L., et al., *The human initiator is a distinct and abundant element that is precisely positioned in focused core promoters*. Genes Dev, 2017. **31**(1): p. 6-11.
23. Louder, R.K., et al., *Structure of promoter-bound TFIID and model of human pre-initiation complex assembly*. Nature, 2016. **531**(7596): p. 604-9.
24. Butler, J.E. and J.T. Kadonaga, *The RNA polymerase II core promoter: a key component in the regulation of gene expression*. Genes Dev, 2002. **16**(20): p. 2583-92.
25. Hoskins, R.A., et al., *Genome-wide analysis of promoter architecture in Drosophila melanogaster*. Genome Res, 2011. **21**(2): p. 182-92.
26. Lis, J., *Promoter-associated pausing in promoter architecture and postinitiation transcriptional regulation*. Cold Spring Harb Symp Quant Biol, 1998. **63**: p. 347-56.
27. Rougvie, A.E. and J.T. Lis, *The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged*. Cell, 1988. **54**(6): p. 795-804.
28. Rasmussen, E.B. and J.T. Lis, *In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes*. Proc Natl Acad Sci U S A, 1993. **90**(17): p. 7923-7.
29. Adelman, K. and J.T. Lis, *Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans*. Nat Rev Genet, 2012. **13**(10): p. 720-31.
30. Core, L. and K. Adelman, *Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation*. Genes Dev, 2019. **33**(15-16): p. 960-982.
31. Li, J. and D.S. Gilmour, *Promoter proximal pausing and the control of gene expression*. Curr Opin Genet Dev, 2011. **21**(2): p. 231-5.
32. Maston, G.A., S.K. Evans, and M.R. Green, *Transcriptional regulatory elements in the human genome*. Annu Rev Genomics Hum Genet, 2006. **7**: p. 29-59.

33. Banerji, J., S. Rusconi, and W. Schaffner, *Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences*. Cell, 1981. **27**(2 Pt 1): p. 299-308.
34. Schoenfelder, S. and P. Fraser, *Long-range enhancer-promoter contacts in gene expression control*. Nat Rev Genet, 2019. **20**(8): p. 437-455.
35. Arnold, P.R., A.D. Wells, and X.C. Li, *Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate*. Front Cell Dev Biol, 2019. **7**: p. 377.
36. Field, A. and K. Adelman, *Evaluating Enhancer Function and Transcription*. Annu Rev Biochem, 2020. **89**: p. 213-234.
37. Calo, E. and J. Wysocka, *Modification of enhancer chromatin: what, how, and why?* Mol Cell, 2013. **49**(5): p. 825-37.
38. Wang, X., M.J. Cairns, and J. Yan, *Super-enhancers in transcriptional regulation and genome organization*. Nucleic Acids Res, 2019. **47**(22): p. 11481-11496.
39. He, Y., W. Long, and Q. Liu, *Targeting Super-Enhancers as a Therapeutic Strategy for Cancer Treatment*. Front Pharmacol, 2019. **10**: p. 361.
40. Jia, Y., W.J. Chng, and J. Zhou, *Super-enhancers: critical roles and therapeutic targets in hematologic malignancies*. J Hematol Oncol, 2019. **12**(1): p. 77.
41. Ellmeier, W., S. Sawada, and D.R. Littman, *The regulation of CD4 and CD8 coreceptor gene expression during T cell development*. Annu Rev Immunol, 1999. **17**: p. 523-54.
42. Brand, A.H., G. Micklem, and K. Nasmyth, *A yeast silencer contains sequences that can promote autonomous plasmid replication and transcriptional activation*. Cell, 1987. **51**(5): p. 709-19.
43. Sandmann, T., et al., *A core transcriptional network for early mesoderm development in Drosophila melanogaster*. Genes Dev, 2007. **21**(4): p. 436-49.
44. Riethoven, J.J., *Regulatory regions in DNA: promoters, enhancers, silencers, and insulators*. Methods Mol Biol, 2010. **674**: p. 33-42.
45. Gisselbrecht, S.S., et al., *Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts*. Mol Cell, 2020. **77**(2): p. 324-337 e8.
46. Pang, B. and M.P. Snyder, *Systematic identification of silencers in human cells*. Nat Genet, 2020. **52**(3): p. 254-263.
47. Bire, S., et al., *Mariner Transposons Contain a Silencer: Possible Role of the Polycomb Repressive Complex 2*. PLoS Genet, 2016. **12**(3): p. e1005902.
48. Chopra, V.S., N. Kong, and M. Levine, *Transcriptional repression via antilooping in the Drosophila embryo*. Proc Natl Acad Sci U S A, 2012. **109**(24): p. 9460-4.

49. Gaszner, M. and G. Felsenfeld, *Insulators: exploiting transcriptional and epigenetic mechanisms*. Nat Rev Genet, 2006. **7**(9): p. 703-13.
50. West, A.G., M. Gaszner, and G. Felsenfeld, *Insulators: many functions, many mechanisms*. Genes Dev, 2002. **16**(3): p. 271-88.
51. Parkhurst, S.M., et al., *The Drosophila su(Hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein*. Genes Dev, 1988. **2**(10): p. 1205-15.
52. Zhao, K., C.M. Hart, and U.K. Laemmli, *Visualization of chromosomal domains with boundary element-associated factor BEAF-32*. Cell, 1995. **81**(6): p. 879-89.
53. Gaszner, M., J. Vazquez, and P. Schedl, *The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction*. Genes Dev, 1999. **13**(16): p. 2098-107.
54. Yusufzai, T.M., et al., *CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species*. Mol Cell, 2004. **13**(2): p. 291-8.
55. Lieb, J.D., et al., *Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association*. Nat Genet, 2001. **28**(4): p. 327-34.
56. Narlikar, L. and I. Ovcharenko, *Identifying regulatory elements in eukaryotic genomes*. Brief Funct Genomic Proteomic, 2009. **8**(4): p. 215-30.
57. Hampsey, M., *Molecular genetics of the RNA polymerase II general transcriptional machinery*. Microbiol Mol Biol Rev, 1998. **62**(2): p. 465-503.
58. Yang, V.W., *Eukaryotic transcription factors: identification, characterization and functions*. J Nutr, 1998. **128**(11): p. 2045-51.
59. Singh, H., A.A. Khan, and A.R. Dinner, *Gene regulatory networks in the immune system*. Trends Immunol, 2014. **35**(5): p. 211-8.
60. Fong, A.P. and S.J. Tapscott, *Skeletal muscle programming and re-programming*. Curr Opin Genet Dev, 2013. **23**(5): p. 568-73.
61. Takahashi, K. and S. Yamanaka, *A decade of transcription factor-mediated reprogramming to pluripotency*. Nat Rev Mol Cell Biol, 2016. **17**(3): p. 183-93.
62. Carroll, S.B., *Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution*. Cell, 2008. **134**(1): p. 25-36.
63. Lambert, S.A., et al., *The Human Transcription Factors*. Cell, 2018. **175**(2): p. 598-599.
64. Wunderlich, Z. and L.A. Mirny, *Different gene regulation strategies revealed by analysis of binding motifs*. Trends Genet, 2009. **25**(10): p. 434-40.
65. Khan, A., et al., *JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework*. Nucleic Acids Res, 2018. **46**(D1): p. D260-D266.

66. Matys, V., et al., *TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes*. Nucleic Acids Res, 2006. **34**(Database issue): p. D108-10.
67. Orenstein, Y. and R. Shamir, *A comparative analysis of transcription factor binding models learned from PBM, HT-SELEX and ChIP data*. Nucleic Acids Res, 2014. **42**(8): p. e63.
68. Hume, M.A., et al., *UniPROBE, update 2015: new tools and content for the online database of protein-binding microarray data on protein-DNA interactions*. Nucleic Acids Res, 2015. **43**(Database issue): p. D117-22.
69. Panne, D., T. Maniatis, and S.C. Harrison, *An atomic model of the interferon-beta enhanceosome*. Cell, 2007. **129**(6): p. 1111-23.
70. Mottis, A., L. Mouchiroud, and J. Auwerx, *Emerging roles of the corepressors NCoR1 and SMRT in homeostasis*. Genes Dev, 2013. **27**(8): p. 819-35.
71. Sonenberg, N. and A.C. Gingras, *The mRNA 5' cap-binding protein eIF4E and control of cell growth*. Curr Opin Cell Biol, 1998. **10**(2): p. 268-75.
72. Burkard, K.T. and J.S. Butler, *A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p*. Mol Cell Biol, 2000. **20**(2): p. 604-16.
73. Konarska, M.M., R.A. Padgett, and P.A. Sharp, *Recognition of cap structure in splicing in vitro of mRNA precursors*. Cell, 1984. **38**(3): p. 731-6.
74. Cho, E.J., et al., *mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain*. Genes Dev, 1997. **11**(24): p. 3319-26.
75. Porrua, O. and D. Libri, *Transcription termination and the control of the transcriptome: why, where and how to stop*. Nat Rev Mol Cell Biol, 2015. **16**(3): p. 190-202.
76. Zorio, D.A. and D.L. Bentley, *The link between mRNA processing and transcription: communication works both ways*. Exp Cell Res, 2004. **296**(1): p. 91-7.
77. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
78. Sakharkar, M.K., V.T. Chow, and P. Kanguane, *Distributions of exons and introns in the human genome*. In Silico Biol, 2004. **4**(4): p. 387-93.
79. Kornblihtt, A.R., et al., *Alternative splicing: a pivotal step between eukaryotic transcription and translation*. Nat Rev Mol Cell Biol, 2013. **14**(3): p. 153-65.
80. Baralle, F.E. and J. Giudice, *Alternative splicing as a regulator of development and tissue identity*. Nat Rev Mol Cell Biol, 2017. **18**(7): p. 437-451.
81. Manning, K.S. and T.A. Cooper, *The roles of RNA processing in translating genotype to phenotype*. Nat Rev Mol Cell Biol, 2017. **18**(2): p. 102-114.

82. Shi, Y., *Mechanistic insights into precursor messenger RNA splicing by the spliceosome*. Nat Rev Mol Cell Biol, 2017. **18**(11): p. 655-670.
83. Wilkinson, M.E., C. Charenton, and K. Nagai, *RNA Splicing by the Spliceosome*. Annu Rev Biochem, 2020. **89**: p. 359-388.
84. Dana, H., et al., *Molecular Mechanisms and Biological Functions of siRNA*. Int J Biomed Sci, 2017. **13**(2): p. 48-57.
85. O'Brien, J., et al., *Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation*. Front Endocrinol (Lausanne), 2018. **9**: p. 402.
86. Dexheimer, P.J. and L. Cochella, *MicroRNAs: From Mechanism to Organism*. Front Cell Dev Biol, 2020. **8**: p. 409.
87. Vasudevan, S. and J.A. Steitz, *AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2*. Cell, 2007. **128**(6): p. 1105-18.
88. Bukhari, S.I.A., et al., *A Specialized Mechanism of Translation Mediated by FXR1a-Associated MicroRNP in Cellular Quiescence*. Mol Cell, 2016. **61**(5): p. 760-773.
89. Orom, U.A., F.C. Nielsen, and A.H. Lund, *MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation*. Mol Cell, 2008. **30**(4): p. 460-71.
90. Darricarrere, N., et al., *Function of Piwi, a nuclear Piwi/Argonaute protein, is independent of its slicer activity*. Proc Natl Acad Sci U S A, 2013. **110**(4): p. 1297-302.
91. Wu, X., et al., *The Biogenesis and Functions of piRNAs in Human Diseases*. Mol Ther Nucleic Acids, 2020. **21**: p. 108-120.
92. Weng, W., H. Li, and A. Goel, *Piwi-interacting RNAs (piRNAs) and cancer: Emerging biological concepts and potential clinical implications*. Biochim Biophys Acta Rev Cancer, 2019. **1871**(1): p. 160-169.
93. Yan, H., et al., *piRNA-823 contributes to tumorigenesis by regulating de novo DNA methylation and angiogenesis in multiple myeloma*. Leukemia, 2015. **29**(1): p. 196-206.
94. Law, P.T., et al., *Deep sequencing of small RNA transcriptome reveals novel non-coding RNAs in hepatocellular carcinoma*. J Hepatol, 2013. **58**(6): p. 1165-73.
95. Tang, X., et al., *The Combination of piR-823 and Eukaryotic Initiation Factor 3 B (EIF3B) Activates Hepatic Stellate Cells via Upregulating TGF-beta1 in Liver Fibrogenesis*. Med Sci Monit, 2018. **24**: p. 9151-9165.
96. Weng, W., et al., *Novel evidence for a PIWI-interacting RNA (piRNA) as an oncogenic mediator of disease progression, and a potential prognostic biomarker in colorectal cancer*. Mol Cancer, 2018. **17**(1): p. 16.
97. Yin, J., et al., *piR-823 contributes to colorectal tumorigenesis by enhancing the transcriptional activity of HSF1*. Cancer Sci, 2017. **108**(9): p. 1746-1756.

98. Cheng, J., et al., *piR-823, a novel non-coding small RNA, demonstrates in vitro and in vivo tumor suppressive activity in human gastric cancer cells*. *Cancer Lett*, 2012. **315**(1): p. 12-7.
99. Gu, X., et al., *Exosomal piRNA profiling revealed unique circulating piRNA signatures of cholangiocarcinoma and gallbladder carcinoma*. *Acta Biochim Biophys Sin (Shanghai)*, 2020. **52**(5): p. 475-484.
100. Su, J.F., et al., *piR-823 demonstrates tumor oncogenic activity in esophageal squamous cell carcinoma through DNA methylation induction via DNA methyltransferase 3B*. *Pathol Res Pract*, 2020. **216**(4): p. 152848.
101. Xie, J., et al., *PIWIL1 interacting RNA piR-017061 inhibits pancreatic cancer growth via regulating EFNA5*. *Hum Cell*, 2021. **34**(2): p. 550-563.
102. Fang, S., et al., *NONCODEV5: a comprehensive annotation database for long non-coding RNAs*. *Nucleic Acids Res*, 2018. **46**(D1): p. D308-D314.
103. Statello, L., et al., *Gene regulation by long non-coding RNAs and its biological functions*. *Nat Rev Mol Cell Biol*, 2021. **22**(2): p. 96-118.
104. Lagarde, J., et al., *High-throughput annotation of full-length long noncoding RNAs with capture long-read sequencing*. *Nat Genet*, 2017. **49**(12): p. 1731-1740.
105. Schlackow, M., et al., *Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs*. *Mol Cell*, 2017. **65**(1): p. 25-38.
106. Yin, Y., et al., *U1 snRNP regulates chromatin retention of noncoding RNAs*. *Nature*, 2020. **580**(7801): p. 147-150.
107. Xiang, J.F., et al., *Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus*. *Cell Res*, 2014. **24**(5): p. 513-31.
108. Saldana-Meyer, R., et al., *RNA Interactions Are Essential for CTCF-Mediated Genome Organization*. *Mol Cell*, 2019. **76**(3): p. 412-422 e5.
109. Boque-Sastre, R., et al., *Head-to-head antisense transcription and R-loop formation promotes transcriptional activation*. *Proc Natl Acad Sci U S A*, 2015. **112**(18): p. 5785-90.
110. Arab, K., et al., *GADD45A binds R-loops and recruits TET1 to CpG island promoters*. *Nat Genet*, 2019. **51**(2): p. 217-223.
111. Ariel, F., et al., *R-Loop Mediated trans Action of the APOLO Long Noncoding RNA*. *Mol Cell*, 2020. **77**(5): p. 1055-1065 e4.
112. Jain, A.K., et al., *LncPRESS1 Is a p53-Regulated LncRNA that Safeguards Pluripotency by Disrupting SIRT6-Mediated De-acetylation of Histone H3K56*. *Mol Cell*, 2016. **64**(5): p. 967-981.
113. Wutz, A., *Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation*. *Nat Rev Genet*, 2011. **12**(8): p. 542-53.

114. Csorba, T., et al., *Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization*. Proc Natl Acad Sci U S A, 2014. **111**(45): p. 16160-5.
115. DiStefano, J.K., *The Emerging Role of Long Noncoding RNAs in Human Disease*. Methods Mol Biol, 2018. **1706**: p. 91-110.
116. McGinty, R.K. and S. Tan, *Nucleosome structure and function*. Chem Rev, 2015. **115**(6): p. 2255-73.
117. Muller, M.M., et al., *A two-state activation mechanism controls the histone methyltransferase Suv39h1*. Nat Chem Biol, 2016. **12**(3): p. 188-93.
118. Poepsel, S., V. Kasinath, and E. Nogales, *Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes*. Nat Struct Mol Biol, 2018. **25**(2): p. 154-162.
119. Glancy, E., C. Ciferri, and A.P. Bracken, *Structural basis for PRC2 engagement with chromatin*. Curr Opin Struct Biol, 2021. **67**: p. 135-144.
120. Centore, R.C., et al., *Mammalian SWI/SNF Chromatin Remodeling Complexes: Emerging Mechanisms and Therapeutic Strategies*. Trends Genet, 2020. **36**(12): p. 936-950.
121. Wagner, F.R., et al., *Structure of SWI/SNF chromatin remodeller RSC bound to a nucleosome*. Nature, 2020. **579**(7799): p. 448-451.
122. Spain, M.M., et al., *The RSC complex localizes to coding sequences to regulate Pol II and histone occupancy*. Mol Cell, 2014. **56**(5): p. 653-66.
123. Monahan, B.J., et al., *Fission yeast SWI/SNF and RSC complexes show compositional and functional differences from budding yeast*. Nat Struct Mol Biol, 2008. **15**(8): p. 873-80.
124. Ng, H.H., et al., *Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex*. Genes Dev, 2002. **16**(7): p. 806-19.
125. Willhoft, O. and D.B. Wigley, *INO80 and SWR1 complexes: the non-identical twins of chromatin remodelling*. Curr Opin Struct Biol, 2020. **61**: p. 50-58.
126. Gerhold, C.B. and S.M. Gasser, *INO80 and SWR complexes: relating structure to function in chromatin remodeling*. Trends Cell Biol, 2014. **24**(11): p. 619-31.
127. Gerhold, C.B., M.H. Hauer, and S.M. Gasser, *INO80-C and SWR-C: guardians of the genome*. J Mol Biol, 2015. **427**(3): p. 637-51.
128. Wang, J., et al., *Roles of the INO80 and SWR1 Chromatin Remodeling Complexes in Plants*. Int J Mol Sci, 2019. **20**(18).
129. Bartholomew, B., *ISWI chromatin remodeling: one primary actor or a coordinated effort?* Curr Opin Struct Biol, 2014. **24**: p. 150-5.
130. Li, Y., et al., *The emerging role of ISWI chromatin remodeling complexes in cancer*. J Exp Clin Cancer Res, 2021. **40**(1): p. 346.

131. Biswas, D., R. Dutta-Biswas, and D.J. Stillman, *Chd1 and yFACT act in opposition in regulating transcription*. Mol Cell Biol, 2007. **27**(18): p. 6279-87.
132. Cramer, P., *Eukaryotic Transcription Turns 50*. Cell, 2019. **179**(4): p. 808-812.
133. Patel, A.B., B.J. Greber, and E. Nogales, *Recent insights into the structure of TFIID, its assembly, and its binding to core promoter*. Curr Opin Struct Biol, 2020. **61**: p. 17-24.
134. Compe, E., et al., *TFIIE orchestrates the recruitment of the TFIIH kinase module at promoter before release during transcription*. Nat Commun, 2019. **10**(1): p. 2084.
135. Rimel, J.K. and D.J. Taatjes, *The essential and multifunctional TFIIH complex*. Protein Sci, 2018. **27**(6): p. 1018-1037.
136. Akhtar, M.S., et al., *TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II*. Mol Cell, 2009. **34**(3): p. 387-93.
137. Peng, J., N.F. Marshall, and D.H. Price, *Identification of a cyclin subunit required for the function of Drosophila P-TEFb*. J Biol Chem, 1998. **273**(22): p. 13855-60.
138. Gariglio, P., M. Bellard, and P. Chambon, *Clustering of RNA polymerase B molecules in the 5' moiety of the adult beta-globin gene of hen erythrocytes*. Nucleic Acids Res, 1981. **9**(11): p. 2589-98.
139. Eick, D. and G.W. Bornkamm, *Transcriptional arrest within the first exon is a fast control mechanism in c-myc gene expression*. Nucleic Acids Res, 1986. **14**(21): p. 8331-46.
140. Kao, S.Y., et al., *Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product*. Nature, 1987. **330**(6147): p. 489-93.
141. Paparidis, N.F., M.C. Durvale, and F. Canduri, *The emerging picture of CDK9/P-TEFb: more than 20 years of advances since PITALRE*. Mol Biosyst, 2017. **13**(2): p. 246-276.
142. Anshabo, A.T., et al., *CDK9: A Comprehensive Review of Its Biology, and Its Role as a Potential Target for Anti-Cancer Agents*. Front Oncol, 2021. **11**: p. 678559.
143. Li, Y., et al., *P-TEFb: Finding its ways to release promoter-proximally paused RNA polymerase II*. Transcription, 2018. **9**(2): p. 88-94.
144. Ghosh, K., et al., *Positive Regulation of Transcription by Human ZMYND8 through Its Association with P-TEFb Complex*. Cell Rep, 2018. **24**(8): p. 2141-2154 e6.
145. Luo, Z., C. Lin, and A. Shilatifard, *The super elongation complex (SEC) family in transcriptional control*. Nat Rev Mol Cell Biol, 2012. **13**(9): p. 543-7.
146. Biswas, D., et al., *Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes*. Proc Natl Acad Sci U S A, 2011. **108**(38): p. 15751-6.

147. Yadav, D., et al., *Multivalent Role of Human TFIID in Recruiting Elongation Components at the Promoter-Proximal Region for Transcriptional Control*. Cell Rep, 2019. **26**(5): p. 1303-1317 e7.
148. Tomson, B.N., et al., *Effects of the Paf1 complex and histone modifications on snoRNA 3'-end formation reveal broad and locus-specific regulation*. Mol Cell Biol, 2013. **33**(1): p. 170-82.
149. Richard, P. and J.L. Manley, *Transcription termination by nuclear RNA polymerases*. Genes Dev, 2009. **23**(11): p. 1247-69.
150. Porrua, O., M. Boudvillain, and D. Libri, *Transcription Termination: Variations on Common Themes*. Trends Genet, 2016. **32**(8): p. 508-522.
151. Chen, F.X., E.R. Smith, and A. Shilatifard, *Born to run: control of transcription elongation by RNA polymerase II*. Nat Rev Mol Cell Biol, 2018. **19**(7): p. 464-478.
152. Jonkers, I. and J.T. Lis, *Getting up to speed with transcription elongation by RNA polymerase II*. Nat Rev Mol Cell Biol, 2015. **16**(3): p. 167-77.
153. Kuzmina, A., S. Krasnopolsky, and R. Taube, *Super elongation complex promotes early HIV transcription and its function is modulated by P-TEFb*. Transcription, 2017. **8**(3): p. 133-149.
154. Chou, S., et al., *HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold*. Proc Natl Acad Sci U S A, 2013. **110**(2): p. E123-31.
155. Lin, C., et al., *Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC)*. Genes Dev, 2011. **25**(14): p. 1486-98.
156. Luo, Z., et al., *The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output*. Mol Cell Biol, 2012. **32**(13): p. 2608-17.
157. Takahashi, H., et al., *Human mediator subunit MED26 functions as a docking site for transcription elongation factors*. Cell, 2011. **146**(1): p. 92-104.
158. Van Oss, S.B., C.E. Cucinotta, and K.M. Arndt, *Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation*. Trends Biochem Sci, 2017. **42**(10): p. 788-798.
159. Marschalek, R., *Systematic Classification of Mixed-Lineage Leukemia Fusion Partners Predicts Additional Cancer Pathways*. Ann Lab Med, 2016. **36**(2): p. 85-100.
160. Meyer, C., et al., *The MLL recombinome of acute leukemias in 2017*. Leukemia, 2018. **32**(2): p. 273-284.
161. Chan, A.K.N. and C.W. Chen, *Rewiring the Epigenetic Networks in MLL-Rearranged Leukemias: Epigenetic Dysregulation and Pharmacological Interventions*. Front Cell Dev Biol, 2019. **7**: p. 81.
162. Antunes, E.T.B. and K. Ottersbach, *The MLL/SET family and haematopoiesis*. Biochim Biophys Acta Gene Regul Mech, 2020. **1863**(8): p. 194579.

163. Slany, R.K., *The molecular biology of mixed lineage leukemia*. Haematologica, 2009. **94**(7): p. 984-93.
164. Muntean, A.G. and J.L. Hess, *The pathogenesis of mixed-lineage leukemia*. Annu Rev Pathol, 2012. **7**: p. 283-301.
165. Thirman, M.J., et al., *Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23;p13.1) in acute myeloid leukemia*. Proc Natl Acad Sci U S A, 1994. **91**(25): p. 12110-4.
166. Mitani, K., et al., *Cloning of several species of MLL/MEN chimeric cDNAs in myeloid leukemia with t(11;19)(q23;p13.1) translocation*. Blood, 1995. **85**(8): p. 2017-24.
167. Kanda, Y., et al., *Subcellular localization of the MEN, MLL/MEN and truncated MLL proteins expressed in leukemic cells carrying the t(11;19)(q23;p13.1) translocation*. Int J Hematol, 1997. **66**(2): p. 189-95.
168. Shilatifard, A., et al., *An RNA polymerase II elongation factor encoded by the human ELL gene*. Science, 1996. **271**(5257): p. 1873-6.
169. Shilatifard, A., et al., *Structure and function of RNA polymerase II elongation factor ELL. Identification of two overlapping ELL functional domains that govern its interaction with polymerase and the ternary elongation complex*. J Biol Chem, 1997. **272**(35): p. 22355-63.
170. Shilatifard, A., *Identification and purification of the Holo-ELL complex. Evidence for the presence of ELL-associated proteins that suppress the transcriptional inhibitory activity of ELL*. J Biol Chem, 1998. **273**(18): p. 11212-7.
171. Shilatifard, A., et al., *ELL2, a new member of an ELL family of RNA polymerase II elongation factors*. Proc Natl Acad Sci U S A, 1997. **94**(8): p. 3639-43.
172. Shinobu, N., et al., *Physical interaction and functional antagonism between the RNA polymerase II elongation factor ELL and p53*. J Biol Chem, 1999. **274**(24): p. 17003-10.
173. Khattak, S., et al., *dELL, a drosophila homologue of transcription elongation factor ELL (Eleven-nineteen Lysine rich Leukemia), is required for early development*. Cell Biochem Funct, 2002. **20**(2): p. 119-27.
174. Gerber, M., et al., *Drosophila ELL is associated with actively elongating RNA polymerase II on transcriptionally active sites in vivo*. EMBO J, 2001. **20**(21): p. 6104-14.
175. Eissenberg, J.C., et al., *dELL is an essential RNA polymerase II elongation factor with a general role in development*. Proc Natl Acad Sci U S A, 2002. **99**(15): p. 9894-9.
176. Gerber, M.A., A. Shilatifard, and J.C. Eissenberg, *Mutational analysis of an RNA polymerase II elongation factor in Drosophila melanogaster*. Mol Cell Biol, 2005. **25**(17): p. 7803-11.
177. Sakurai, K., et al., *Inhibition of the canonical Wnt signaling pathway in cytoplasm: a novel property of the carboxyl terminal domains of two Xenopus ELL genes*. Zoolog Sci, 2004. **21**(4): p. 407-16.

178. Pascual-Le Tallec, L., et al., *The elongation factor ELL (eleven-nineteen lysine-rich leukemia) is a selective coregulator for steroid receptor functions*. Mol Endocrinol, 2005. **19**(5): p. 1158-69.
179. Zhou, J., et al., *Elongation factor ELL (Eleven-Nineteen Lysine-rich Leukemia) acts as a transcription factor for direct thrombospondin-1 regulation*. J Biol Chem, 2009. **284**(28): p. 19142-52.
180. Shell, S.A., et al., *Increased phosphorylation of the carboxyl-terminal domain of RNA polymerase II and loading of polyadenylation and cotranscriptional factors contribute to regulation of the ig heavy chain mRNA in plasma cells*. J Immunol, 2007. **179**(11): p. 7663-73.
181. Martincic, K., et al., *Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing*. Nat Immunol, 2009. **10**(10): p. 1102-9.
182. Santos, P., et al., *Transcriptional and epigenetic regulation of B cell development*. Immunol Res, 2011. **50**(2-3): p. 105-12.
183. Liu, L., et al., *ELL is an HIF-1alpha partner that regulates and responds to hypoxia response in PC3 cells*. Prostate, 2010. **70**(7): p. 797-805.
184. Chen, Y., et al., *ELL targets c-Myc for proteasomal degradation and suppresses tumour growth*. Nat Commun, 2016. **7**: p. 11057.
185. Byun, J.S., et al., *Dynamic bookmarking of primary response genes by p300 and RNA polymerase II complexes*. Proc Natl Acad Sci U S A, 2009. **106**(46): p. 19286-91.
186. Byun, J.S., et al., *ELL facilitates RNA polymerase II pause site entry and release*. Nat Commun, 2012. **3**: p. 633.
187. Smith, E.R., et al., *Regulation of the transcriptional activity of poised RNA polymerase II by the elongation factor ELL*. Proc Natl Acad Sci U S A, 2008. **105**(25): p. 8575-9.
188. Fufa, T.D., et al., *The Tax oncogene enhances ELL incorporation into p300 and P-TEFb containing protein complexes to activate transcription*. Biochem Biophys Res Commun, 2015. **465**(1): p. 5-11.
189. Smith, E.R., et al., *The little elongation complex regulates small nuclear RNA transcription*. Mol Cell, 2011. **44**(6): p. 954-65.
190. Mourgues, S., et al., *ELL, a novel TFIIH partner, is involved in transcription restart after DNA repair*. Proc Natl Acad Sci U S A, 2013. **110**(44): p. 17927-32.
191. Simone, F., et al., *EAF1, a novel ELL-associated factor that is delocalized by expression of the MLL-ELL fusion protein*. Blood, 2001. **98**(1): p. 201-9.
192. Li, M., et al., *Expression of murine ELL-associated factor 2 (EAF2) is developmentally regulated*. Dev Dyn, 2003. **228**(2): p. 273-80.

193. Simone, F., et al., *ELL-associated factor 2 (EAF2), a functional homolog of EAF1 with alternative ELL binding properties*. Blood, 2003. **101**(6): p. 2355-62.
194. Kong, S.E., et al., *ELL-associated factors 1 and 2 are positive regulators of RNA polymerase II elongation factor ELL*. Proc Natl Acad Sci U S A, 2005. **102**(29): p. 10094-8.
195. Xiao, W., F. Jiang, and Z. Wang, *ELL binding regulates U19/EAF2 intracellular localization, stability, and transactivation*. Prostate, 2006. **66**(1): p. 1-12.
196. Hahn, J., et al., *Apoptosis induction and growth suppression by U19/EAF2 is mediated through its ELL-binding domain*. Prostate, 2007. **67**(2): p. 146-53.
197. Xiao, W., et al., *U19/EAF2 knockout causes lung adenocarcinoma, B-cell lymphoma, hepatocellular carcinoma and prostatic intraepithelial neoplasia*. Oncogene, 2008. **27**(11): p. 1536-44.
198. Yu, X., et al., *Regulation of tumor suppressor EAF2 polyubiquitination by ELL1 and SIAH2 in prostate cancer cells*. Oncotarget, 2016. **7**(20): p. 29245-54.
199. Liu, R., et al., *ELL-associated factors EAF1/2 negatively regulate HIV-1 transcription through inhibition of Super Elongation Complex formation*. Biochim Biophys Acta Gene Regul Mech, 2020. **1863**(5): p. 194508.
200. Hamaguchi, M., et al., *DBC2, a candidate for a tumor suppressor gene involved in breast cancer*. Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13647-52.
201. Johnson, G.S., P. Rajendran, and R.H. Dashwood, *CCAR1 and CCAR2 as gene chameleons with antagonistic duality: Preclinical, human translational, and mechanistic basis*. Cancer Sci, 2020. **111**(10): p. 3416-3425.
202. Rishi, A.K., et al., *Identification and characterization of a cell cycle and apoptosis regulatory protein-1 as a novel mediator of apoptosis signaling by retinoid CD437*. J Biol Chem, 2003. **278**(35): p. 33422-35.
203. Zhao, W., et al., *Negative regulation of the deacetylase SIRT1 by DBC1*. Nature, 2008. **451**(7178): p. 587-90.
204. Kim, J.E., J. Chen, and Z. Lou, *DBC1 is a negative regulator of SIRT1*. Nature, 2008. **451**(7178): p. 583-6.
205. Kang, H., et al., *Peptide switch is essential for Sirt1 deacetylase activity*. Mol Cell, 2011. **44**(2): p. 203-13.
206. Boily, G., et al., *Sirt1 regulates energy metabolism and response to caloric restriction in mice*. PLoS One, 2008. **3**(3): p. e1759.
207. Boutant, M. and C. Canto, *SIRT1 metabolic actions: Integrating recent advances from mouse models*. Mol Metab, 2014. **3**(1): p. 5-18.

208. Chang, H.C. and L. Guarente, *SIRT1 and other sirtuins in metabolism*. Trends Endocrinol Metab, 2014. **25**(3): p. 138-45.
209. Li, Z., et al., *Inhibition of SUV39H1 methyltransferase activity by DBC1*. J Biol Chem, 2009. **284**(16): p. 10361-6.
210. Vaquero, A., et al., *SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation*. Nature, 2007. **450**(7168): p. 440-4.
211. Vandel, L., et al., *Transcriptional repression by the retinoblastoma protein through the recruitment of a histone methyltransferase*. Mol Cell Biol, 2001. **21**(19): p. 6484-94.
212. Chini, C.C., et al., *HDAC3 is negatively regulated by the nuclear protein DBC1*. J Biol Chem, 2010. **285**(52): p. 40830-7.
213. Magni, M., G. Buscemi, and L. Zannini, *Cell cycle and apoptosis regulator 2 at the interface between DNA damage response and cell physiology*. Mutat Res Rev Mutat Res, 2018. **776**: p. 1-9.
214. Fu, J., et al., *Deleted in breast cancer 1, a novel androgen receptor (AR) coactivator that promotes AR DNA-binding activity*. J Biol Chem, 2009. **284**(11): p. 6832-40.
215. Trauernicht, A.M., et al., *Modulation of estrogen receptor alpha protein level and survival function by DBC-1*. Mol Endocrinol, 2007. **21**(7): p. 1526-36.
216. Koyama, S., et al., *Repression of estrogen receptor beta function by putative tumor suppressor DBC1*. Biochem Biophys Res Commun, 2010. **392**(3): p. 357-62.
217. Yu, E.J., et al., *Reciprocal roles of DBC1 and SIRT1 in regulating estrogen receptor alpha activity and co-activator synergy*. Nucleic Acids Res, 2011. **39**(16): p. 6932-43.
218. Garapaty, S., et al., *Identification and characterization of a novel nuclear protein complex involved in nuclear hormone receptor-mediated gene regulation*. J Biol Chem, 2009. **284**(12): p. 7542-52.
219. Chini, C.C., et al., *DBC1 (Deleted in Breast Cancer 1) modulates the stability and function of the nuclear receptor Rev-erbalpha*. Biochem J, 2013. **451**(3): p. 453-61.
220. Close, P., et al., *DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation*. Nature, 2012. **484**(7394): p. 386-9.
221. Escande, C., et al., *Deleted in breast cancer-1 regulates SIRT1 activity and contributes to high-fat diet-induced liver steatosis in mice*. J Clin Invest, 2010. **120**(2): p. 545-58.
222. Escande, C., et al., *Deleted in breast cancer 1 limits adipose tissue fat accumulation and plays a key role in the development of metabolic syndrome phenotype*. Diabetes, 2015. **64**(1): p. 12-22.
223. Qiang, L., et al., *Brown remodeling of white adipose tissue by SirT1-dependent deacetylation of Ppargamma*. Cell, 2012. **150**(3): p. 620-32.

224. Nin, V., et al., *Deleted in breast cancer 1 (DBC1) protein regulates hepatic gluconeogenesis*. J Biol Chem, 2014. **289**(9): p. 5518-27.
225. Kong, S., et al., *DBC1 is a suppressor of B cell activation by negatively regulating alternative NF-kappaB transcriptional activity*. J Immunol, 2014. **193**(11): p. 5515-24.
226. Kong, S., et al., *Deleted in Breast Cancer 1 Suppresses B Cell Activation through RelB and Is Regulated by IKKalpha Phosphorylation*. J Immunol, 2015. **195**(8): p. 3685-93.
227. Rudra, D., et al., *Transcription factor Foxp3 and its protein partners form a complex regulatory network*. Nat Immunol, 2012. **13**(10): p. 1010-9.
228. Beier, U.H., et al., *Sirtuin-1 targeting promotes Foxp3+ T-regulatory cell function and prolongs allograft survival*. Mol Cell Biol, 2011. **31**(5): p. 1022-9.
229. Gao, Y., et al., *Inflammation negatively regulates FOXP3 and regulatory T-cell function via DBC1*. Proc Natl Acad Sci U S A, 2015. **112**(25): p. E3246-54.
230. Bolderson, E., et al., *Kruppel-associated Box (KRAB)-associated co-repressor (KAP-1) Ser-473 phosphorylation regulates heterochromatin protein 1beta (HP1-beta) mobilization and DNA repair in heterochromatin*. J Biol Chem, 2012. **287**(33): p. 28122-31.
231. Hu, C., et al., *Roles of Kruppel-associated Box (KRAB)-associated Co-repressor KAP1 Ser-473 Phosphorylation in DNA Damage Response*. J Biol Chem, 2012. **287**(23): p. 18937-52.
232. Magni, M., et al., *CCAR2/DBC1 is required for Chk2-dependent KAP1 phosphorylation and repair of DNA damage*. Oncotarget, 2015. **6**(19): p. 17817-31.
233. Li, J., et al., *A conserved NAD(+) binding pocket that regulates protein-protein interactions during aging*. Science, 2017. **355**(6331): p. 1312-1317.
234. Ray Chaudhuri, A. and A. Nussenzweig, *The multifaceted roles of PARP1 in DNA repair and chromatin remodelling*. Nat Rev Mol Cell Biol, 2017. **18**(10): p. 610-621.
235. Cramer, P., et al., *Structure of eukaryotic RNA polymerases*. Annu Rev Biophys, 2008. **37**: p. 337-52.
236. Ferreira, R., et al., *Targeting the RNA Polymerase I Transcription for Cancer Therapy Comes of Age*. Cells, 2020. **9**(2).
237. Russell, J. and J.C. Zomerdijs, *The RNA polymerase I transcription machinery*. Biochem Soc Symp, 2006(73): p. 203-16.
238. Schier, A.C. and D.J. Taatjes, *Structure and mechanism of the RNA polymerase II transcription machinery*. Genes Dev, 2020. **34**(7-8): p. 465-488.
239. Salzberg, S.L., *Open questions: How many genes do we have?* BMC Biol, 2018. **16**(1): p. 94.
240. Chang, C.W., et al., *Identification of human housekeeping genes and tissue-selective genes by microarray meta-analysis*. PLoS One, 2011. **6**(7): p. e22859.

241. Reilly, S.K. and J.P. Noonan, *Evolution of Gene Regulation in Humans*. Annu Rev Genomics Hum Genet, 2016. **17**: p. 45-67.
242. Giguere, S.S., et al., *The Proteomic Profile of Deleted in Breast Cancer 1 (DBC1) Interactions Points to a Multifaceted Regulation of Gene Expression*. Mol Cell Proteomics, 2016. **15**(3): p. 791-809.
243. Hiraike, H., et al., *Identification of DBC1 as a transcriptional repressor for BRCA1*. Br J Cancer, 2010. **102**(6): p. 1061-7.
244. Able, A.A., A.J. Richard, and J.M. Stephens, *Loss of DBC1 (CCAR2) affects TNFalpha-induced lipolysis and Glut4 gene expression in murine adipocytes*. J Mol Endocrinol, 2018. **61**(4): p. 195-205.
245. Basu, S., et al., *DBC1, p300, HDAC3, and Siah1 coordinately regulate ELL stability and function for expression of its target genes*. Proc Natl Acad Sci U S A, 2020. **117**(12): p. 6509-6520.
246. Kumari, N., et al., *AFF1 acetylation by p300 temporally inhibits transcription during genotoxic stress response*. Proc Natl Acad Sci U S A, 2019. **116**(44): p. 22140-22151.
247. Wang, L., et al., *The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation*. Science, 2011. **333**(6043): p. 765-9.
248. Mi, H., et al., *Large-scale gene function analysis with the PANTHER classification system*. Nat Protoc, 2013. **8**(8): p. 1551-66.
249. Huang, D.W., et al., *DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists*. Nucleic Acids Res, 2007. **35**(Web Server issue): p. W169-75.
250. Sathishkumar, C., et al., *Augmentation of histone deacetylase 3 (HDAC3) epigenetic signature at the interface of proinflammation and insulin resistance in patients with type 2 diabetes*. Clin Epigenetics, 2016. **8**: p. 125.
251. Smith, E., C. Lin, and A. Shilatifard, *The super elongation complex (SEC) and MLL in development and disease*. Genes Dev, 2011. **25**(7): p. 661-72.
252. Liu, M., et al., *The ubiquitin ligase Siah1 controls ELL2 stability and formation of super elongation complexes to modulate gene transcription*. Mol Cell, 2012. **46**(3): p. 325-34.
253. Devany, E., et al., *Positive and negative feedback loops in the p53 and mRNA 3' processing pathways*. Proc Natl Acad Sci U S A, 2013. **110**(9): p. 3351-6.
254. Bornstein, C., et al., *A negative feedback loop of transcription factors specifies alternative dendritic cell chromatin States*. Mol Cell, 2014. **56**(6): p. 749-62.

- 255. Luo, R.T., et al., *The elongation domain of ELL is dispensable but its ELL-associated factor 1 interaction domain is essential for MLL-ELL-induced leukemogenesis*. Mol Cell Biol, 2001. **21**(16): p. 5678-87.
- 256. Polak, P.E., et al., *ELL and EAF1 are Cajal body components that are disrupted in MLL-ELL leukemia*. Mol Biol Cell, 2003. **14**(4): p. 1517-28.

DBC1, p300, HDAC3, and Siah1 coordinately regulate ELL stability and function for expression of its target genes

Subham Basu^a, Mahesh Barad^a, Dipika Yadav^a, Arijit Nandy^a, Bidisha Mukherjee^b, Jit Sarkar^c, Partha Chakrabarti^c, Satinath Mukhopadhyay^b, and Debabrata Biswas^{a,1}

^aLaboratory of Transcription Biology, Molecular Genetics Division, Council of Scientific and Industrial Research (CSIR)–Indian Institute of Chemical Biology (IICB), Kolkata 700032, India; ^bDepartment of Endocrinology and Metabolism, Institute of Postgraduate Medical Education & Research (IPGM&R), Kolkata 700020, India; and ^cMetabolic Diseases Laboratory, Cell Biology and Physiology Division, CSIR–IICB, Kolkata 700032, India

Edited by Joan W. Conaway, Stowers Institute for Medical Research, Kansas City, MO, and accepted by Editorial Board Member Brenda A. Schulman February 7, 2020 (received for review July 20, 2019)

Among all of the Super Elongation Complex (SEC) components, ELL1 (also known as ELL) is the only bona fide elongation factor that directly stimulates transcription elongation by RNA polymerase II. However, the mechanism(s) of functional regulation of ELL1 (referred to as ELL hereafter), through its stabilization, is completely unknown. Here, we report a function of human DBC1 in regulating ELL stability involving HDAC3, p300, and Siah1. Mechanistically, we show that p300-mediated site-specific acetylation increases, whereas HDAC3-mediated deacetylation decreases, ELL stability through polyubiquitylation by the E3 ubiquitin ligase Siah1. DBC1 competes with HDAC3 for the same binding sites on ELL and thus increases its acetylation and stability. Knockdown of DBC1 reduces ELL levels and expression of a significant number of genes, including those involved in glucose metabolism. Consistently, Type 2 diabetes patient-derived peripheral blood mononuclear cells show reduced expression of DBC1 and ELL and associated key target genes required for glucose homeostasis. Thus, we describe a pathway of regulating stability and functions of key elongation factor ELL for expression of diverse sets of genes, including ones that are linked to Type 2 diabetes pathogenesis.

transcription | ELL | acetylation | ubiquitylation | DBC1

During the last few decades, studies in the field of transcriptional regulation were focused on detailed understanding of factors involved, and their role, in regulation of the initiation step. Although transcriptional initiation plays a rate-limiting step for controlling expression of a significant number of genes, studies in the last several years have also shown a role for postinitiation steps, especially at the step of promoter proximal pausing, in overall regulation of transcription of genes that are important for development (1).

Among the elongation factors/complexes, human Super Elongation Complex (SEC) has been proposed to have multiple elongation factors in a megadalton complex (2, 3). Predominant players of SEC are AF9, ENL, ELL/ELL2, AFF1, AFF4, and P-TEFb complex (a heterodimer of CyclinT1 and CDK9). Excepting P-TEFb complex, other SEC components are frequently fused with the N terminus of Mixed Lineage Leukemia 1 (MLL1, hereafter MLL) to give rise to MLL fusion proteins and cause leukemogenesis (4).

Among all of the SEC components, ELL is the only bona fide elongation factor that directly stimulates transcription elongation by RNA polymerase II (hereafter Pol II) in vitro by reducing the rate of Pol II stalling during elongation (5–7). ELL interacts with ELL-associated factors (EAF1/2) that directly stimulate ELL-mediated transcription elongation by Pol II (2, 7–10). Besides SEC, ELL also associates with EAF1, KIAA0947 (aka ICE1), and NARG2 (aka ICE2) to form Little Elongation Complex (LEC) that functions in regulating small nuclear RNA gene transcription by Pol II as well as in the recruitment of TFIIF during transcription-coupled DNA repair (11–13). Although mechanisms of ELL2 degradation by the E3 ubiquitin ligase, Siah1, have been described by two earlier studies (14, 15), regulation of ELL function through its stabilization is completely

unknown. In this regard, we were intrigued by the strong association of Deleted in Breast Cancer 1 (DBC1) in our ELL.com mass spectrometric analysis (2) and *SI Appendix, Fig. S1A*.

Human DBC1 was first described as a negative regulator of deacetylase activity of the histone deacetylase (HDAC), SIRT1, thus protecting the acetylation of tumor suppressor p53, a key modification required for its function within mammalian cells (16, 17). DBC1 has also been shown to regulate p53 stability through its competitive binding with E3 ubiquitin ligase MDM2 (18). Physiologically, DBC1 knockout mice show high fat diet-induced liver steatosis through regulation of SIRT1 activity as well as an increase in blood glucose level through increased gluconeogenesis (19, 20). A recent study has also shown a role for the Nudix homology domain of DBC1 in regulating PARP binding in a NAD⁺ concentration-dependent manner (15, 21). However, beyond its role in regulating p53 activity by controlling SIRT1 deacetylase functions, the role of DBC1 in regulating other key transcription factors have not been documented in detail.

In this study, we describe a role of DBC1 in regulation of ELL stability, and thus its function(s) within mammalian cells that also

Significance

Human RNA polymerase II (Pol II) transcribes messenger RNA from underlying DNA sequences for the purpose of generating proteins for proper biological functions. One of the key steps of Pol II-driven transcription is elongation, which is directly stimulated by transcription factors such as ELL and affects expression of vast number of genes. However, mechanisms by which the overall functions of ELL are regulated through its stabilization are completely unknown. In this study, we decipher a mechanism of regulation of ELL stability, and thus its functions through multiple factors involving DBC1, HDAC3, p300, and Siah1 for regulation of expression of diverse sets of genes that are required for several physiological processes, including maintenance of glucose homeostasis in human.

Author contributions: S.B., M.B., D.Y., and D.B. designed research; S.B., M.B., D.Y., and A.N. performed research; S.B., M.B., A.N., B.M., J.S., P.C., and S.M. contributed new reagents/analytic tools; S.B., M.B., D.Y., and D.B. analyzed data; and S.B. and D.B. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. J.W.C. is a guest editor invited by the Editorial Board.

This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

Data deposition: Raw uncropped images related to this paper have been deposited in the Mendeley Data repository (<https://doi.org/10.17632/hy84kd5hwj.1>).

¹To whom correspondence may be addressed. Email: dbiswas@iicb.res.in.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1912375117/-DCSupplemental>.

First published March 9, 2020.

involves the activity of critical regulators such as p300, HDAC3, and Siah1. Our detailed studies have uncovered a pathway of transcriptional regulation involving the p300/DBC1/HDAC3/Siah1/ELL axis within mammalian cells that may have an implication for proper expression of key genes required for several physiological processes, including glucose homeostasis, in human.

Results

Interaction between ELL and DBC1. We reported association of SEC components with ELL in our earlier study (2). Along with SEC components, we also observed association of DBC1 with ELL in our mass spectrometric analysis (SI Appendix, Fig. S1A). Subsequent immunoprecipitation (IP) and blotting analysis using nuclear extract from a stable cell line expressing FLAG-HA-ELL protein confirmed ELL interaction with DBC1 as well as other known ELL interactors (Fig. 1A). IP of endogenous ELL using ELL-specific antibody and subsequent blotting analyses further confirmed ELL interaction with DBC1 in an endogenous context (Fig.

1B). Reciprocal IP using nuclear extract from a FLAG-HA-DBC1 expressing stable cell line (SI Appendix, Fig. S1B) and subsequent mass spectrometric and immunoblotting analyses further confirmed DBC1 interaction with ELL (Fig. 1C and D). Interestingly, besides ELL, both of the analyses also showed DBC1 interaction with other SEC components (Fig. 1C and D). Subsequently, using purified recombinant DBC1 and ELL proteins through their expression in baculoviral and bacterial expression system, respectively (Fig. 1E), in vitro interaction analysis further confirmed direct interaction of DBC1 with GST-ELL, but not GST alone (Fig. 1F, lane 3 vs. lane 4). Therefore, based on these analyses, we conclude that human DBC1 and ELL interact directly with each other both in vitro as well as in vivo within mammalian cells.

Next, we were interested in identifying ELL domain(s) that would interact with DBC1. Coimmunoprecipitation (co-IP) analyses using ELL deletion constructs cloned in a mammalian expression vector (Fig. 1G, Middle) showed that fragments containing the elongation domain (45 to 373) interacted with

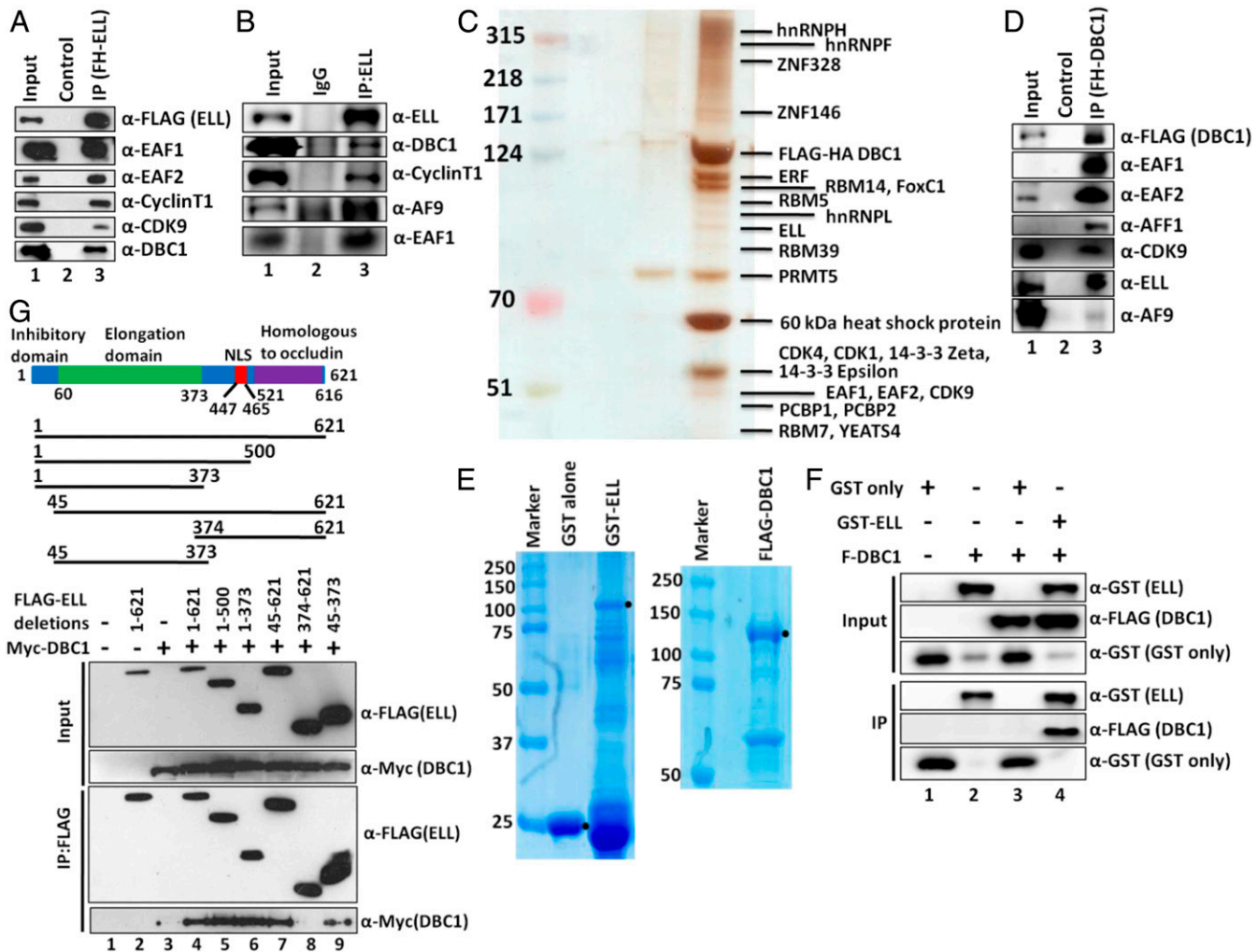


Fig. 1. DBC1 and ELL interact both in vitro and in vivo within mammalian cells. (A and B) IP and Western blotting analysis showing identification of DBC1 as an interactor of (A) ectopic and (B) endogenous ELL along with other known interacting components (IgG: immunoglobulin G). (C) Purification of DBC1-associated protein complex from a stable cell line that ectopically expresses DBC1 as FLAG-HA-tagged. Eluted proteins were run on 4 to 12% gradient gel and silver-stained for their visualization. Individual bands were excised, and proteins were identified by mass spectrometric analysis. (D) IP and Western blotting analysis for identification of SEC component association with DBC1, and FLAG-DBC1 (using baculoviral expression system). (E) Coomassie staining of purified recombinant GST, GST-ELL (using bacterial expression system), and FLAG-DBC1 (using baculoviral expression system). GST-ELL shows significant degradation in our purification. (F) In vitro direct interaction assay with purified proteins showing direct interaction of DBC1 with GST-ELL but not GST alone. (G) Cartoon diagram showing domains of ELL that are important for regulating its functions (Upper) (NLS: nuclear localization signal). Depiction of ELL domains that have been used for its interaction assay with DBC1 by co-IP analysis (Middle). Western blot analysis showing interaction of indicated domains of ELL with DBC1 within mammalian cells by co-IP analysis (Lower).

DBC1 (Fig. 1G, IP panel, lanes 4 to 7 and lane 9), whereas elongation-deficient domain (374 to 621) failed to do so (Fig. 1G, lane 8). Therefore, we conclude that the elongation domain of ELL is involved in its interaction with DBC1 within mammalian cells.

DBC1 and ELL Positively Regulate Transcription of Chromosomally Integrated Reporter Gene Expression. Toward understanding the functional significance of the role of DBC1 and ELL in regulation of Pol II-mediated transcription, we initially sought to use a chromosomally integrated reporter gene expression system. This system contains the adenovirus major late promoter followed by five copies of GAL4 DNA binding sites upstream of a reporter luciferase gene (22, 23). Since the reporter gene is expressed in context of chromatin, as opposed to a naked DNA template, the expression analysis reflects function of any given protein/complex similar to their native target genes *in vivo*. Our initial analysis showed that, with increasing expression, DBC1 strongly stimulated target reporter gene activation in the presence of activator GAL4-VP16 (*SI Appendix, Fig. S2A*). This effect is direct, since DBC1 failed to show any effect on the expression of GAL4-VP16 (*SI Appendix, Fig. S2B*). Further, consistent with its role in positive regulation of transcription, we observed a strong effect of ELL in stimulating reporter gene expression in the presence of GAL4-VP16 (*SI Appendix, Fig. S2C*). Therefore, we conclude that both the human DBC1 and ELL have positive roles in regulation of target chromosomally integrated reporter gene expression.

Our subsequent analysis through coexpression of increasing DBC1 and constant ELL further showed increased reporter gene activity compared with ELL alone (*SI Appendix, Fig. S2D*). The observed effect is additive rather than synergistic. Similar results with increasing ELL expression in the presence of constant DBC1 (*SI Appendix, Fig. S2E*) led us to conclude that both DBC1 and ELL positively regulate transcription of chromosomally integrated reporter gene in an additive manner.

DBC1 Stabilizes ELL Protein within Mammalian Cells. The interesting and consistent observation of increased expression of ELL protein upon increasing coexpression of DBC1 (*SI Appendix, Fig. S2D*, compare lane 4 with lanes 5 to 8), despite using a constant amount of ELL expression construct in our reporter assay, further prompted us to investigate whether DBC1 could have any role(s) in stabilizing ELL protein. As shown in Fig. 2A, increased expression of DBC1 resulted in a marked increase in expression of cotransfected ELL protein (Fig. 2A and *Lower* quantitation), but not control GFP (*SI Appendix, Fig. S3A*). Although, through qRT-PCR analysis, increased DBC1 messenger RNA (mRNA) was observed upon increased DBC1 cotransfection (*SI Appendix, Fig. S3B, Upper*), our failure to observe any significant effect on ELL mRNA expression (*SI Appendix, Fig. S3B, Lower*) indicated an overall DBC1 effect on ELL protein level only and ruled out any effect through increased ELL transcription. Our subsequent analysis further showed an effect of increased DBC1 expression on increasing endogenous ELL protein levels in both 293T and HCT116 cells (Fig. 2B and *SI Appendix, Fig. S3C*, respectively), thus ruling out a cell type-specific effect of DBC1 on ELL stability. Therefore, we conclude that human DBC1 stabilizes ELL protein within mammalian cells in a cell type-independent manner.

HDAC3 Interacts with ELL and Causes Its Destabilization. For a deeper mechanistic understanding of ELL stabilization, and since DBC1 regulates p53 functions through its association with SIRT1 (16–18), we initially tested whether ELL would also show an interaction with SIRT1. As shown in *SI Appendix, Fig. S3D*, in our co-IP analysis, we have failed to observe any interaction between ELL and SIRT1. However, subsequent co-IP analyses of ELL with other HDACs showed specific and strong interactions of ELL with HDAC1, HDAC3, and HDAC5 (Fig. 2C, lanes 4, 6, and 8), in which HDAC3 showed stronger interaction than the others. Interestingly,

and consistent with an earlier study (24), similar analyses with DBC1 also showed specific and strong interaction with HDAC1 and HDAC3 (Fig. 2D, lanes 4 and 6). Therefore, we conclude that both ELL and DBC1 interact with HDAC3 more strongly than with other HDACs within mammalian cells, and subsequent studies were designed toward deeper understanding of the role of HDAC3 in regulation of DBC1-mediated ELL functions within mammalian cells.

Contrary to the effect of DBC1 on ELL, increased expression of HDAC3 resulted in decreased expression of ectopic ELL (Fig. 2E) without affecting its mRNA expression from the transfected plasmid as measured by qRT-PCR (*SI Appendix, Fig. S3E*). In fact, increased expression of HDAC3 caused an increase in ELL mRNA expression in our assay (*SI Appendix, Fig. S3E*; compare lane 2 with lanes 3 to 6). Similar effect of HDAC3 on ELL protein level in HCT116 cells (*SI Appendix, Fig. S3F*) also ruled out a cell type-specific effect. The effect of HDAC3 on ELL destabilization is very specific, since we failed to observe any significant effect of other HDACs when coexpressed with ELL (Fig. 2F). Further analysis through coexpression of SIRT1-7 proteins along with ELL also failed to show any specific and significant effect on ELL stability (*SI Appendix, Fig. S3G*). Thus, we conclude that human HDAC3 specifically destabilizes ELL protein without affecting transcription of ELL mRNA within mammalian cells. Consistent with a role of HDAC3 in regulation of ELL functions, we also observed endogenous HDAC3 interaction with ectopic ELL (*SI Appendix, Fig. S3H*) as well as endogenous ELL interaction with ectopic HDAC3 (*SI Appendix, Fig. S3I*).

Next, we addressed whether overexpression of HDAC3 would enhance ELL degradation kinetics within mammalian cells. Consistent with an earlier study showing ELL (ELL1) as a stable protein (14), our initial cycloheximide (CHX) chase assay showed unchanged levels of ELL and DBC1 even after 24 h of chase (Fig. 2G). However, overexpression of HDAC3 caused rapid degradation of ELL (Fig. 2H, quantitation in *Lower*) and thus led us to conclude that the stable human ELL protein is degraded significantly faster in the presence of overexpressed HDAC3.

Stable knockdown of HDAC3 by two different specific short hairpin RNAs (shRNAs) (Fig. 2I) (25) caused an increase in endogenous ELL protein level (Fig. 2I; compare lane 1 with lanes 2 to 3) without altering its mRNA level (*SI Appendix, Fig. S3J*), and thus further ruled out any artifacts in our observation using overexpressed HDAC3. Thus, we conclude that HDAC3 regulates degradation of the ectopic as well as endogenous ELL within mammalian cells.

DBC1 and HDAC3 Act in Opposition to Stabilize ELL Protein. So far, our analyses have shown opposing roles of DBC1 and HDAC3 toward ELL stabilization. Next, we addressed whether overexpression of DBC1 would rescue HDAC3-mediated ELL destabilization. As shown in Fig. 3A, although overexpression of HDAC3 destabilizes ELL (lane 3), concomitant expression of DBC1 rescued ELL protein from HDAC3-mediated degradation (compare lane 3 with lanes 4 to 6 and quantitation in *Lower*). Consistent with this opposing role, simultaneous coexpression of DBC1 significantly reduced ELL degradation kinetics in a CHX chase assay in the presence of ectopically expressed HDAC3 (Fig. 3B; compare lanes 4 to 6 with lanes 7 to 9 and quantitation in *Lower*). Therefore, based on all of these results, we conclude that human DBC1 and HDAC3 act in opposition, wherein DBC1 stabilizes and HDAC3 destabilizes ELL protein.

DBC1 and HDAC3 Compete with Each Other for ELL Binding. For further understanding the opposing roles of DBC1 and HDAC3 in regulating ELL stability, and primarily based on earlier observations showing DBC1 and MDM2 regulating p53 stability through their competitive binding (18), we wondered whether similar mechanisms would also exist in our study. To address this hypothesis, we initially identified ELL domain(s) that would

interact with HDAC3. Co-IP analysis showed that, whereas an N-terminal fragment containing elongation domain fully retained HDAC3 interaction similar to that of full-length ELL (Fig. 3C, lane 6), a C-terminal fragment (374 to 621) failed to do so (Fig. 3C, lane 8). Thus, we conclude that human ELL interacts with HDAC3 through its N-terminal elongation domain. A similar interaction pattern of ELL and DBC1 (Fig. 1G) further indicated the existence of competition between these two factors for their binding to ELL. Consistently, whereas, in the absence of ectopically expressed DBC1, ELL interacted strongly with HDAC3 (Fig.

3D, lane 1), increasing expression and binding of ectopic DBC1 with ELL resulted in a concomitant reduction of HDAC3 binding (Fig. 3D; compare lane 1 with lanes 2 to 4). Toward obtaining direct evidence of similar competitive binding, we purified HDAC3 through its overexpression in mammalian cells and purification using a high salt-based purification method (Fig. 3E). Subsequent *in vitro* interaction showed that only purified recombinant GST-ELL interacted with HDAC3, not GST alone (Fig. 3F; compare lane 3 with lanes 4 to 5 in IP panel). Further assays using purified DBC1 (Fig. 1E) clearly showed that, whereas GST-ELL retained its strong

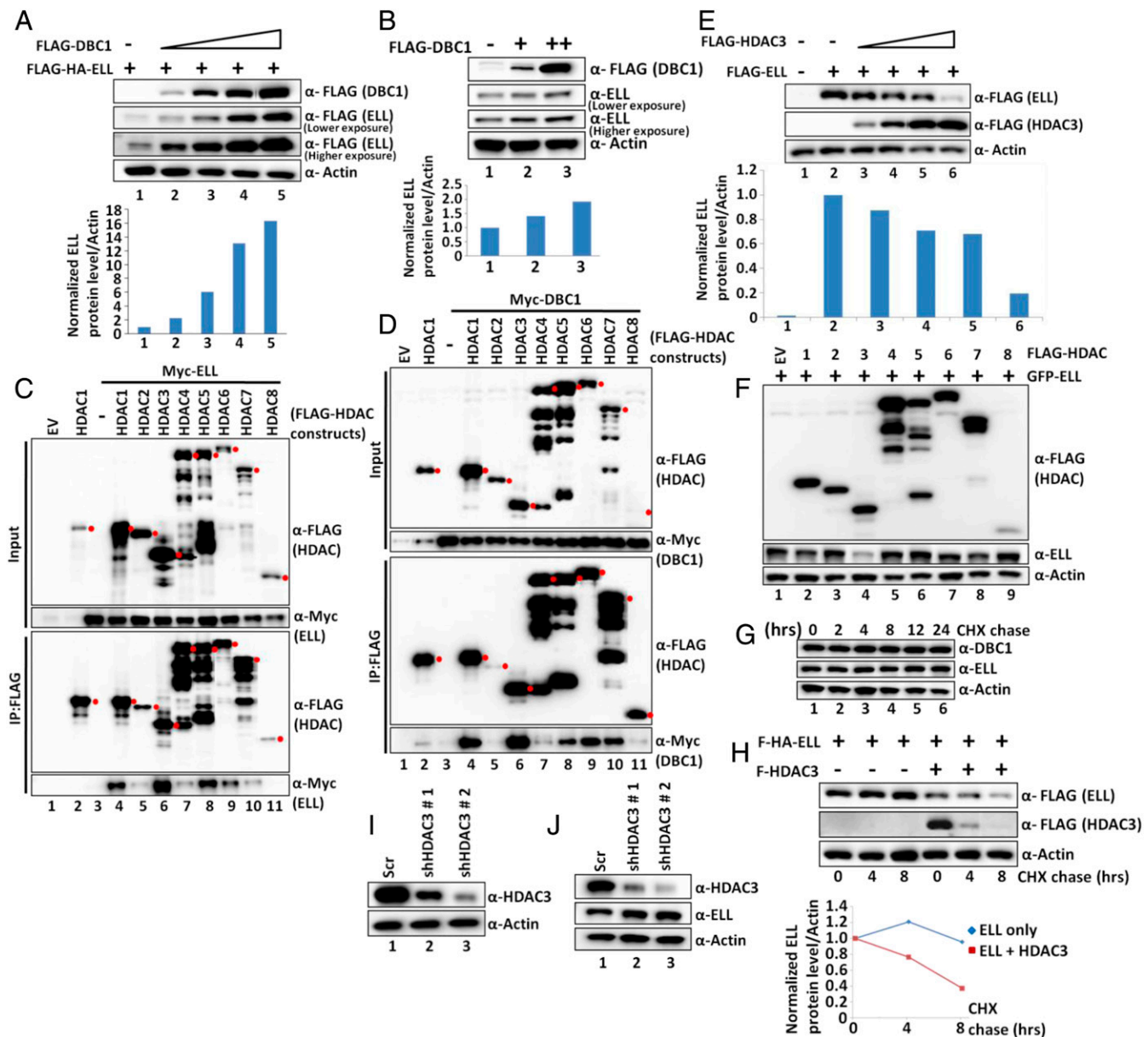


Fig. 2. DBC1 and HDAC3 act in opposition to regulate ELL stability within mammalian cells. (A and B) Western blot analysis showing the effect of overexpression of DBC1 on (A) ectopic and (B) endogenous expression of ELL in 293T cells (Upper) and quantitation of ELL protein relative to Actin (Lower). (C and D) Co-IP and Western blot analysis showing interactions of HDAC1-8 with (C) ELL and (D) DBC1 in 293T cells. Because of the unavailability of human HDAC2, murine HDAC2 was used in this assay. Red filled circles indicate target protein bands, and others are either degradation or nonspecific proteins. (E) Western blot analysis showing the effect of overexpression of HDAC3 on ectopic expression of ELL in 293T cells (Upper) and quantitation of ELL relative to Actin (Lower). (F) Western blot analysis showing effect of overexpression of HDAC1-8 on ectopic expression of ELL in 293T cells. (G) CHX chase assay showing degradation kinetics of endogenous DBC1 and ELL at different time points after addition of CHX in 293T cells. (H) CHX chase assay showing degradation kinetics of ectopically expressed ELL in presence or absence of ectopically expressed HDAC3 at different time points after addition of CHX in 293T cells (Upper) and quantitation of ELL degradation over time (Lower). (I and J) Western blotting analysis showing (I) the stable knockdown of HDAC3 and (J) its effect on expression of endogenous ELL by using two different shRNAs in 293T cells.

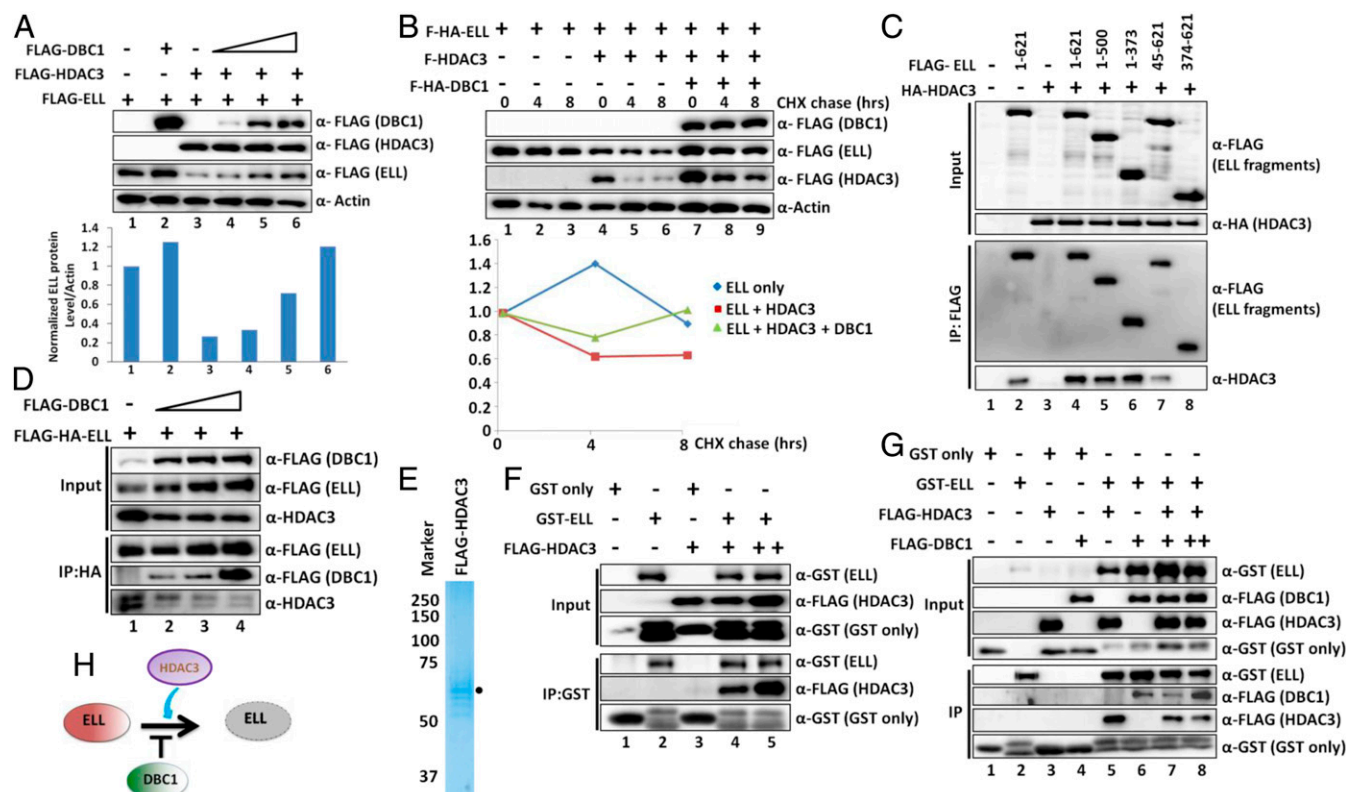


Fig. 3. Competitive binding between DBC1 and HDAC3 to ELL. (A) Western blot analysis showing the rescue of HDAC3-mediated ELL degradation through concomitant overexpression of DBC1 (Upper) and quantitation of ELL protein relative to Actin (Lower). (B) CHX chase assay showing reduced degradation kinetics of ectopically expressed ELL mediated by HDAC3 in presence of DBC1 at different time points after addition of CHX in 293T cells (Upper) and quantitation of relative ELL degradation over time (Lower). (C) Western blot analysis showing the interaction of indicated ELL domains with HDAC3 within mammalian cells by co-IP analysis. (D) IP and Western blot analysis showing reduced interaction of HDAC3 with ELL upon increasing expression and binding of DBC1 to ELL within 293T cells. (E) Purification of HDAC3 through its overexpression in mammalian 293T cells. (F) In vitro direct interaction assay with purified proteins showing direct interaction of HDAC3 with GST-ELL but not GST alone. (G) In vitro direct binding assay, using purified proteins, showing competitive binding between HDAC3 and DBC1 to ELL. (H) Cartoon diagram depicting opposing roles of DBC1 and HDAC3 in regulation of ELL stability, wherein DBC1 stabilizes and HDAC3 destabilizes ELL protein.

interaction with HDAC3 (Fig. 3G, lane 5 in IP panel), addition of DBC1 and its binding strongly inhibited HDAC3 binding with GST-ELL (Fig. 3G; compare lane 5 with lanes 7 to 8 in IP panel). Therefore, all of these results led us to conclude that DBC1 and HDAC3 compete for the same binding site for their interaction with ELL and thus perform opposing functions in regulating ELL stability within mammalian cells (Fig. 3H).

p300-Mediated Acetylation Increases ELL Stability. Since HDAC3 destabilizes ELL at the protein level, we hypothesized that ELL protein may be destabilized by the deacetylation activity of HDAC3, whereas acetylation would increase its stability. Toward addressing this hypothesis, we initially tested whether ELL is acetylated within mammalian cells. Coexpression of ELL with various histone acetyl transferases and subsequent IP and blotting analysis using pan-acetyl lysine-specific antibody showed ELL acetylation by p300 and MOF (Fig. 4A, lanes 4 and 6) and not by PCAF or GCN5. Because of its maximum effect, we further focused on the role of p300-mediated ELL acetylation in regulating its functions. In our subsequent experiments, we observed increased ELL acetylation upon increased expression of p300 within mammalian cells, by cotransfection analysis (SI Appendix, Fig. S4A). For in vitro analysis of direct acetylation by p300, we purified recombinant p300 HAT domain through its expression in bacterial system (SI Appendix, Fig. S4B) (26). As shown in Fig. 4B, whereas purified recombinant ELL failed to show any acetylation signal by itself (lane 3), addition of p300 in the reaction showed strong ELL acetylation (compare lane 3 with lanes 4 to 6). Thus, we conclude

that human p300 directly acetylates ELL in vitro as well as in vivo within mammalian cells.

Interestingly, contrary to the destabilization effect of HDAC3, concomitant expression of p300 increases ectopic ELL expression from a plasmid (SI Appendix, Fig. S4C, quantitation in Lower). A parallel analysis showed a robust effect of p300 overexpression on increasing endogenous ELL protein levels (Fig. 4C) without affecting its mRNA expression (SI Appendix, Fig. S4D). Based on these evidences, we conclude that p300-mediated acetylation increases ELL stability within mammalian cells.

Next, we knocked down endogenous p300 using two specific shRNAs (SI Appendix, Fig. S4E) (27) and observed a concomitant reduction in expression of endogenous ELL protein (Fig. 4D) without affecting its mRNA level (SI Appendix, Fig. S4F). A similar result is also observed for endogenous ELL protein levels upon p300 knockdown in HCT116 cells (SI Appendix, Fig. S4G), thus ruling out a cell type-specific effect. IP and subsequent blotting analysis also showed a reduced acetylation level of endogenous ELL and thus further confirmed a role for p300 in acetylating ELL (SI Appendix, Fig. S4H) in an endogenous context. Subsequent CHX chase analysis using these cells clearly showed faster degradation kinetics of endogenous ELL upon p300 knockdown than of control scramble cells (Fig. 4E). All of these analyses clearly confirm the role of p300 in stabilizing ELL within mammalian cells.

HDAC3 Deacetylates ELL In Vitro and In Vivo. To address opposing functions of p300-mediated acetylation and HDAC3-mediated

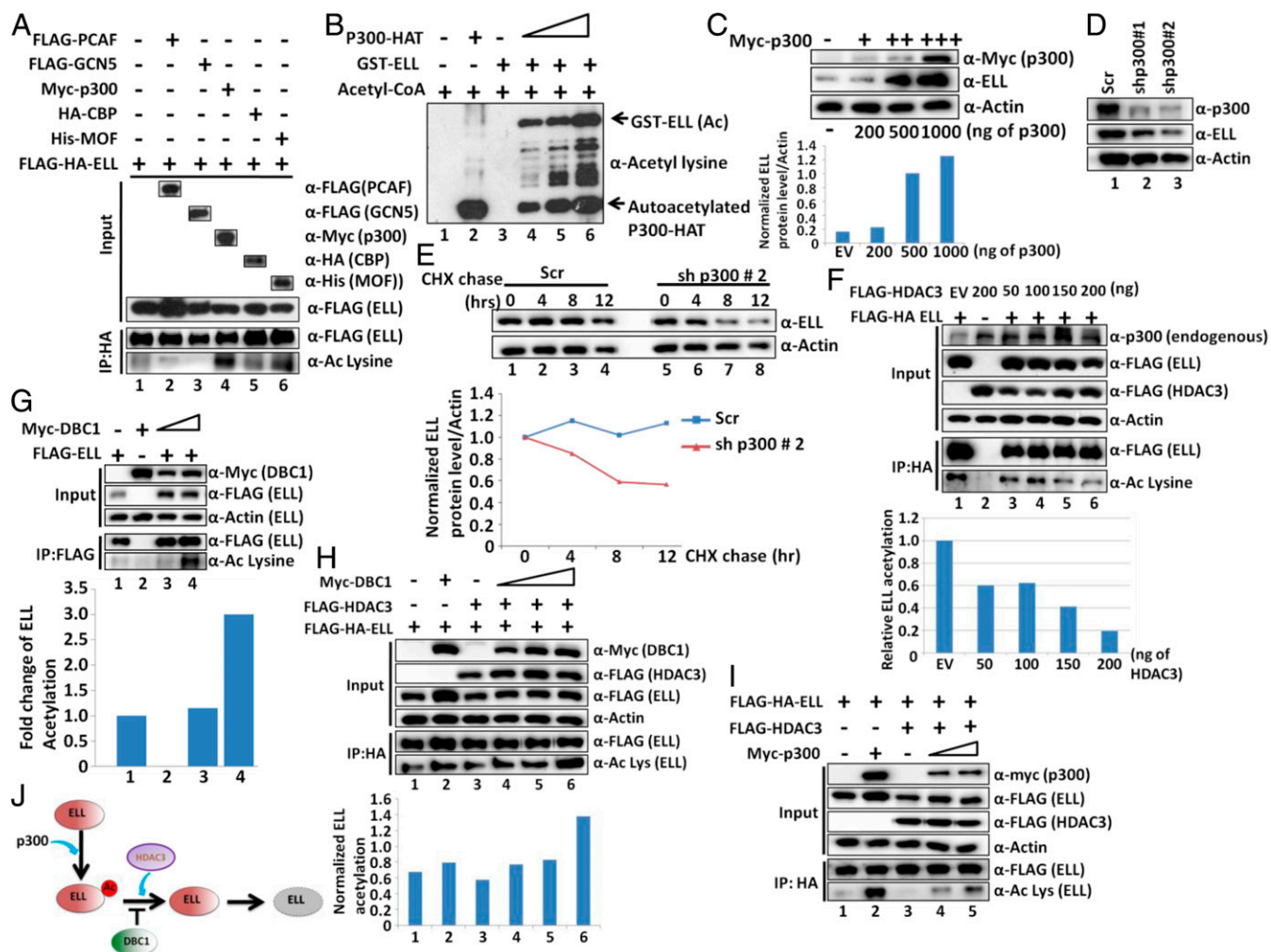


Fig. 4. The p300-mediated acetylation stabilizes ELL, whereas DBC1 and HDAC3 act in an opposite manner to regulate ELL acetylation and stability. (A) Immunoblotting analyses showing robust ELL acetylation specifically by p300 within 293T cells through immunoblotting using pan-acetyl lysine-specific antibody. (B) In vitro HAT assay and subsequent Western blotting showing ELL acetylation by p300 using purified recombinant p300 HAT domain and GST-ELL (SI Appendix, Fig. S4B). (C) Western blot analysis showing the effect of overexpression of p300 on expression of endogenous ELL in 293T cells (Upper) and quantitation of ELL protein level relative to Actin (Lower). (D) Western blot analysis showing the effect of stable p300 knockdown, by two different shRNA constructs, on endogenous ELL protein level within 293T cells. (E) CHX chase assay showing degradation kinetics of ELL in a p300 knockdown cell line when compared to control scramble cells (Upper). Quantitation of ELL protein degradation relative to Actin over time (Lower). (F) Western blot analysis showing deacetylation of ELL through overexpression of HDAC3 within 293T cells (Upper). Quantitation of relative ELL acetylation (Lower). (G) Western blot analysis showing increased ELL acetylation upon overexpression of DBC1 in 293T cells. (H) Western blot analysis showing rescue of acetylation of ELL, which is being deacetylated by HDAC3, through concomitant overexpression of DBC1 (Upper). Quantitation of relative ELL acetylation (Lower). (I) Western blot analysis showing rescue of acetylation of ELL, which is being deacetylated by HDAC3, through concomitant overexpression of p300. (J) Cartoon diagram depicting role of p300-mediated acetylation in regulation of ELL stability involving DBC1 and HDAC3.

deacetylation in regulation of ELL stability, we initially wondered whether ELL would be deacetylated by HDAC3. Coexpression and IP analysis showed that, with increased expression of HDAC3, the ELL acetylation level is concomitantly reduced (Fig. 4F, compare lane 1 with lanes 3 to 6). Since endogenous p300 levels remain unchanged upon increased expression of HDAC3 in our assay (Fig. 4F, Upper, top row, input panel), an indirect effect of reduced p300 expression on overall decrease in ELL acetylation can be ruled out. Our subsequent in vitro deacetylation assay using purified and acetylated ELL showed modest decrease in ELL acetylation in the presence of HDAC3 (SI Appendix, Fig. S4I), and provided direct evidence of HDAC3-mediated ELL deacetylation. Further, pulling down endogenous ELL using ELL-specific antibody from control scramble and HDAC3 knockdown cells (Fig. 2I) and subsequent blotting analysis clearly showed an increase in endogenous ELL acetylation upon HDAC3 knockdown within

mammalian cells (SI Appendix, Fig. S4J). Based on these results, we conclude that, whereas p300 acetylates ELL, acetylated ELL is subjected to deacetylation by HDAC3 in vitro as well as in vivo within mammalian cells.

DBC1 and p300 Act in Opposition to HDAC3 to Regulate ELL Acetylation. Since our results indicate an opposing function of DBC1 and HDAC3 in binding to ELL and regulating its stability, we wondered whether DBC1 would act against HDAC3 to increase ELL acetylation. Our initial experiments showed increased acetylation of ectopically expressed ELL when coexpressed with increasing DBC1 (Fig. 4G; compare lane 1 with lanes 3 to 4). Further, we addressed whether DBC1 could rescue HDAC3-mediated ELL deacetylation. As shown in Fig. 4H, whereas cotransfection of ELL with HDAC3 reduced its acetylation level (compare lane 1 with lane 3), concomitant expression of DBC1 rescued it (Fig. 4H; compare

lane 3 with lanes 4 to 6). Similar results are also obtained upon overexpression of p300 within mammalian cells (Fig. 4I; compare lane 3 with lanes 4 to 5). Further, DBC1 overexpression markedly decreased endogenous ELL interaction with HDAC3 and increased its acetylation level (SI Appendix, Fig. S4K). Therefore, based on all of these results, we conclude that both DBC1 and p300 act against HDAC3 to regulate ELL acetylation, leading to its stability (Fig. 4J).

HDAC3-Mediated ELL Destabilization Involves Ubiquitin Proteasome-Mediated Degradation System. For deeper understanding of the mechanisms involved in HDAC3-mediated ELL degradation, we initially used MG132 that inhibits ubiquitin proteasome-mediated degradation of target proteins and observed significant ELL stabilization from HDAC3-mediated degradation (Fig. 5A; compare lane 3 with lanes 4 to 6). This observation indicated a role for proteasome-mediated ubiquitylation in this degradation process. Consistent with this, whereas ectopically expressed ELL did not

show much ubiquitylation in our assay (Fig. 5B, lane 1), coexpression of HDAC3 markedly increased ELL polyubiquitylation, as observed by immunoblotting using ubiquitin-specific antibody (compare lane 1 with lane 3). As expected, addition of MG132 further increased ELL polyubiquitylation (Fig. 5B, lane 4) and thus confirmed a role of HDAC3 in increasing ELL ubiquitylation. Subsequent analysis with increasing HDAC3 expression showed dose-dependent ELL destabilization (Fig. 5C, input lanes) and concomitant increased polyubiquitylation (Fig. 5C, lane 1 vs. lanes 3 to 6). This also correlates well with ELL acetylation wherein reduced acetylation shows concomitant increased polyubiquitylation (lane 1 vs. lanes 3 to 6). Therefore, we conclude that HDAC3-mediated ELL deacetylation promotes its polyubiquitylation and thus causes its degradation.

DBC1 and p300 Increase ELL Acetylation and Decrease Its Poly-Ubiquitylation. Based on our earlier observations (Figs. 3 and 4), we wondered whether both DBC1 and p300 would reduce

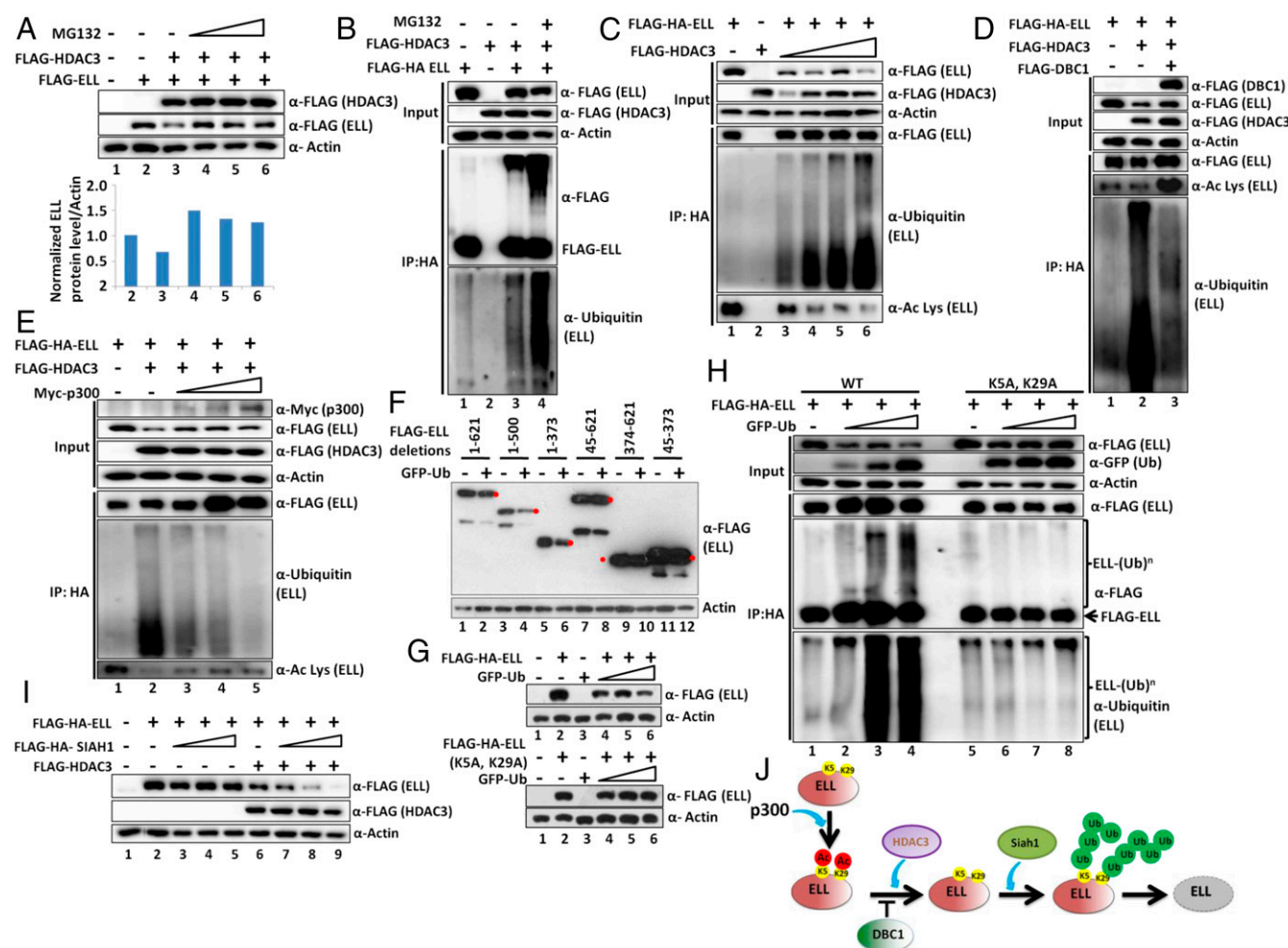


Fig. 5. Poly-ubiquitylation of ELL at N-terminal K5 and K29 residues in presence of HDAC3 dictates ELL degradation by Siah1 within mammalian cells. (A) Western blot analysis showing rescue of HDAC3-mediated ELL degradation through addition of ubiquitin proteasome inhibitor MG132 (Upper) and quantitation of ELL level (Lower). (B) Western blot analysis showing increased polyubiquitylation of ELL by coexpression of HDAC3. The polyubiquitylation signal is further enhanced upon addition of MG132. (C) Western blot analysis showing a correlation of increased polyubiquitylation and decreased acetylation of ELL mediated by increased expression of HDAC3. (D and E) Western blot analysis showing strong correlation of increased ELL acetylation and reduced polyubiquitylation in presence of HDAC3 through concomitant overexpression of (D) DBC1 and (E) p300. (F) Western blot analysis showing the specific domain of ELL that shows sensitivity to ubiquitin-mediated degradation. Red filled circles indicate target protein bands, and others are either degradation or nonspecific proteins. (G) Western blot analysis showing sensitivity of ELL (WT) and ELL (K5A, K29A) mutant toward ubiquitin-mediated degradation. (H) Western blot analysis showing reduced polyubiquitylation of ELL (K5A, K29A) mutant in presence of overexpressed ubiquitin. (I) Western blot analysis showing increased ELL degradation by the ubiquitin E3 ligase Siah1 only in presence of HDAC3. (J) Cartoon diagram depicting the role of p300-mediated acetylation, DBC1-mediated acetylation protection, HDAC3-mediated deacetylation, and Siah1-mediated polyubiquitylation in overall regulation of ELL stability within mammalian cells.

HDAC3-mediated ELL polyubiquitylation. As shown in Fig. 5D, and consistent with our earlier observations, coexpression of DBC1 results in rescue of ELL destabilization by HDAC3 (input blot). Further IP analysis showed an increase in ELL acetylation and simultaneous reduced polyubiquitylation (Fig. 5D; compare lane 2 with lane 3). A similar experiment with p300 also showed rescue of HDAC3-mediated ELL degradation (Fig. 5E, input blots) that correlates well with its increased acetylation and decreased polyubiquitylation (Fig. 5E; compare lane 2 with lanes 3 to 5). Therefore, these results led us to conclude that p300-mediated ELL acetylation is protected from HDAC3-mediated deacetylation by DBC1 through its competitive binding to ELL. Increased ELL acetylation further causes decreased polyubiquitylation resulting in ELL stabilization.

DBC1 and p300 Stabilize ELL against Ubiquitin-Mediated Degradation.

For addressing the correlation between acetylation and ubiquitylation of ELL, we initially wondered whether ELL protein would be degraded by increased expression of ubiquitin. Cotransfection analysis showed that, with overexpression of ubiquitin, the ectopically expressed ELL protein is also degraded (*SI Appendix, Fig. S5 A and B*, lane 5 vs. lane 3). Interestingly, overexpression of both DBC1 and p300 significantly rescued ubiquitin-mediated ELL degradation (*SI Appendix, Fig. S5 A and B*; compare lane 5 with lanes 6 to 8). Further IP analysis showed that both DBC1 and p300-mediated ELL stabilization (input lanes) strongly correlated with a concomitant increase in ELL acetylation as observed in *SI Appendix, Fig. S5 C and D* (compare lane 2 with lane 3 in both figures) that also strongly correlated with its decreased polyubiquitylation (*SI Appendix, Fig. S5C*, lane 2 vs. lane 3, and *SI Appendix, Fig. S5D*, lane 2 vs. lane 3). Therefore, based on these results, we conclude that both p300 and DBC1 increase ELL acetylation, which, in turn, stabilizes ELL protein through its reduced polyubiquitylation.

N-Terminal K5 and K29 of ELL Are Key Targets for p300-Mediated Acetylation and Ubiquitin-Mediated Degradation. Next, we addressed specific residue(s) of ELL that would be targeted for p300-mediated acetylation. Our initial in vitro acetylation of different purified ELL fragments (*SI Appendix, Fig. S5 E, Top*) by p300 showed that, upon deletion of the N-terminal 60 amino acids, p300-mediated ELL acetylation was significantly reduced (*SI Appendix, Fig. S5 E, Bottom*; compare lane 3 with lane 9 in both input GST blot and IP acetyl lysine blot). Further analysis showed the presence of only two key conserved lysine residues (K5 and K29) across multiple species (*SI Appendix, Fig. S5F*). Subsequent in vitro acetylation studies showed marked reduction in acetylation of the K5A, K29A mutant as compared to the wild type (WT) by p300 (*SI Appendix, Fig. S5G*; compare lanes 2 to 4 with lanes 6 to 8). Consistently, the ELL (K5A, K29A) mutant also showed reduced acetylation when coexpressed with p300 within mammalian cells (*SI Appendix, Fig. S5H*; compare lanes 2 to 3 with lanes 5 to 6). All of these analyses clearly show a role of ELL K5 and K29 residues as being key targets for p300-mediated acetylation both in vitro and in vivo within mammalian cells.

Next, we addressed whether these two key acetylation sites within ELL would also be targeted for ubiquitin-mediated degradation within mammalian cells. As shown in Fig. 5F, ELL fragments containing the N-terminal 45 amino acids are sensitive, whereas fragments losing those amino acids are completely resistant, to ubiquitin-mediated degradation (lanes 1 to 6 vs. lanes 7 to 12). Subsequent analysis using ELL (K5A, K29A) mutant construct showed complete resistance, whereas a parallel analysis using WT showed sensitivity to degradation by ubiquitin (Fig. 5G) and thus indicated that the N-terminal K5 and K29 of ELL are key targets for ubiquitin-mediated degradation. Our subsequent analysis also showed significant reduction in polyubiquitylation of ELL (K5A, K29A) mutant when compared to the WT (Fig. 5H, lanes 2

to 4 vs. lanes 6 to 8) in the presence of overexpressed ubiquitin. Consistent with a role for HDAC3-mediated deacetylation in increasing ELL ubiquitylation, coexpression of HDAC3 with ELL (K5A, K29A) mutant also showed significantly reduced ubiquitylation when compared to the WT (*SI Appendix, Fig. S5I*, compare lanes 2 to 4 with lanes 6 to 8). Further, overexpression of DBC1 also failed to show any increase in acetylation and resulting stabilization of ELL (K5A, K29A) mutant when compared to the WT (*SI Appendix, Fig. S5J*; compare lane 2 with lanes 4 to 6 and lane 8 with lanes 10 to 12). Thus, based on all of these results, we conclude that N-terminal K5 and K29 of ELL are key targets for p300-mediated acetylation and ubiquitin-mediated degradation of ELL. DBC1 and HDAC3 competitively bind to ELL and thus regulate ELL acetylation and its stability through modulating its polyubiquitylation. Although ELL (K5A, K29A) mutant showed enhanced stability, it showed decreased luciferase reporter gene activity compared with the WT (*SI Appendix, Fig. S6A*). This result further indicates additional role(s) of these two key residues in overall transcriptional regulation by ELL.

E3 Ubiquitin Ligase Siah1-Mediated ELL Degradation Requires HDAC3.

Next, we were interested in identifying the E3 ubiquitin ligase that would be involved in ELL degradation within mammalian cells. An extensive literature survey showed that some of the SEC components, including ELL2 but not ELL, are targeted by Siah1 E3 ubiquitin ligase for their stability and regulation of target gene expression (14, 28). Consistently, we also failed to observe ELL degradation by Siah1 alone (Fig. 5I, lane 2 vs. lanes 3 to 5). We reasoned that, since degradation of ELL would also require its deacetylation by HDAC3, presence of HDAC3 is a prerequisite for Siah1-mediated degradation of ELL. As shown in Fig. 5I, in the presence of HDAC3, increased expression of Siah1 (*SI Appendix, Fig. S6 B, Lower*) resulted in increased degradation of ELL (lanes 3 to 5 vs. lanes 7 to 9) without affecting ELL mRNA expression (*SI Appendix, Fig. S6 B, Upper*). Thus, we conclude that human ELL protein is subjected to Siah1-mediated degradation only in the presence of HDAC3.

Based on all of these results, we conclude that, by targeting the same residues, p300-mediated acetylation causes stability of human ELL protein through reduced polyubiquitylation, whereas HDAC3-mediated deacetylation destabilizes ELL through increased polyubiquitylation. DBC1, by virtue of possessing the same binding site for ELL interaction, competes with HDAC3 and prevents ELL deacetylation and subsequent Siah1-mediated degradation (Fig. 5J).

DBC1-Mediated ELL Stabilization Is Required for Expression of Diverse Sets of Genes.

Toward identifying the native target genes whose expression would be regulated by the newly identified DBC1–p300–HDAC3–Siah1–ELL axis, we initially intended to identify DBC1 and ELL coregulated genes. We noted two earlier studies that described an effect of ELL and DBC1 knockdown on global down-regulation of target genes in 293T cells (GSE34104 for ELL knockdown and GSE35480 for DBC1 knockdown) (29, 30). Considering the top 10,000 down-regulated genes, our analysis showed an overlap of a significant number of genes (~37%) being co-down-regulated by both DBC1 and ELL (*SI Appendix, Fig. S6C*). Gene ontology (GO) analyses further showed a considerable fraction (24%) of these genes predicted to be involved in regulation of metabolic processes (*SI Appendix, Fig. S6D*). This is also consistent with the functional role of DBC1 that has been described by several studies until now (31).

For validating our high-throughput analysis using specific target genes, we generated two different stable DBC1 knockdown cell lines using shRNA constructs specific for DBC1 (*SI Appendix, Fig. S6E*). Interestingly, knockdown of DBC1 also resulted in decreased ELL protein level (*SI Appendix, Fig. S6E*) without affecting its RNA expression (Fig. 6A, ELL mRNA

expression). Similar results were also observed on ELL protein level upon DBC1 knockdown in HCT116 cells (*SI Appendix, Fig. S6F*). Further RNA expression analysis showed that, upon DBC1 knockdown, the majority of selected target genes were down-regulated as shown in Fig. 6A. These target genes include key factors regulating cell proliferation (e.g., *CCND1*, *CCNE2*, *CDK6*, *PCNA*) as well as the basal transcription machinery (*MED12*, *MED15*). The effect of DBC1 on target gene expression is specific since the same knockdown cells did not show any effect on expression of nontarget genes such as *NEDD9*, *PTPRF*, *MYO10*, *NFκB2*, and *NAV2* (Fig. 6A). Consistent with the effect of DBC1 knockdown on reduced expression of cell proliferation-related genes, our subsequent assay showed reduced proliferation of 293T cells upon knockdown of DBC1 by two different shRNAs (*SI Appendix, Fig. S7A*).

Next, we addressed whether reduced ELL expression in DBC1 knockdown cells would result in its reduced recruitment, thus affecting downstream recruitment of SEC, on identified target genes. Our chromatin immunoprecipitation (ChIP) analyses

showed reduced ELL recruitment at the transcription start site (TSS) on selected target genes upon DBC1 knockdown (Fig. 6B). Interestingly, and consistent with an earlier study showing a role of ELL in recruiting SEC components at the TSS region (29), we also observed significant down-regulation of recruitment of other SEC components, for example, CDK9. Further, ChIP analysis showed a strong reduction in Pol II recruitment as well as the presence of its Ser2 and Ser5 phosphorylated form, a hallmark of transcriptionally engaged elongating Pol II, on all target genes that we have tested upon DBC1 knockdown. The overall effect is target gene-specific since, in a parallel analysis, DBC1 knockdown failed to cause any impaired recruitment of all of the target factors on two nontarget *NEDD9* and *PTPRF* genes (*SI Appendix, Fig. S6G*) that we have tested. These results led us to conclude that, upon DBC1 knockdown, recruitment of SEC components along with Pol II at the TSS region of target genes are impaired (Fig. 6B).

Next, we addressed whether preinitiation complex (PIC) formation on target genes would also be impaired upon DBC1

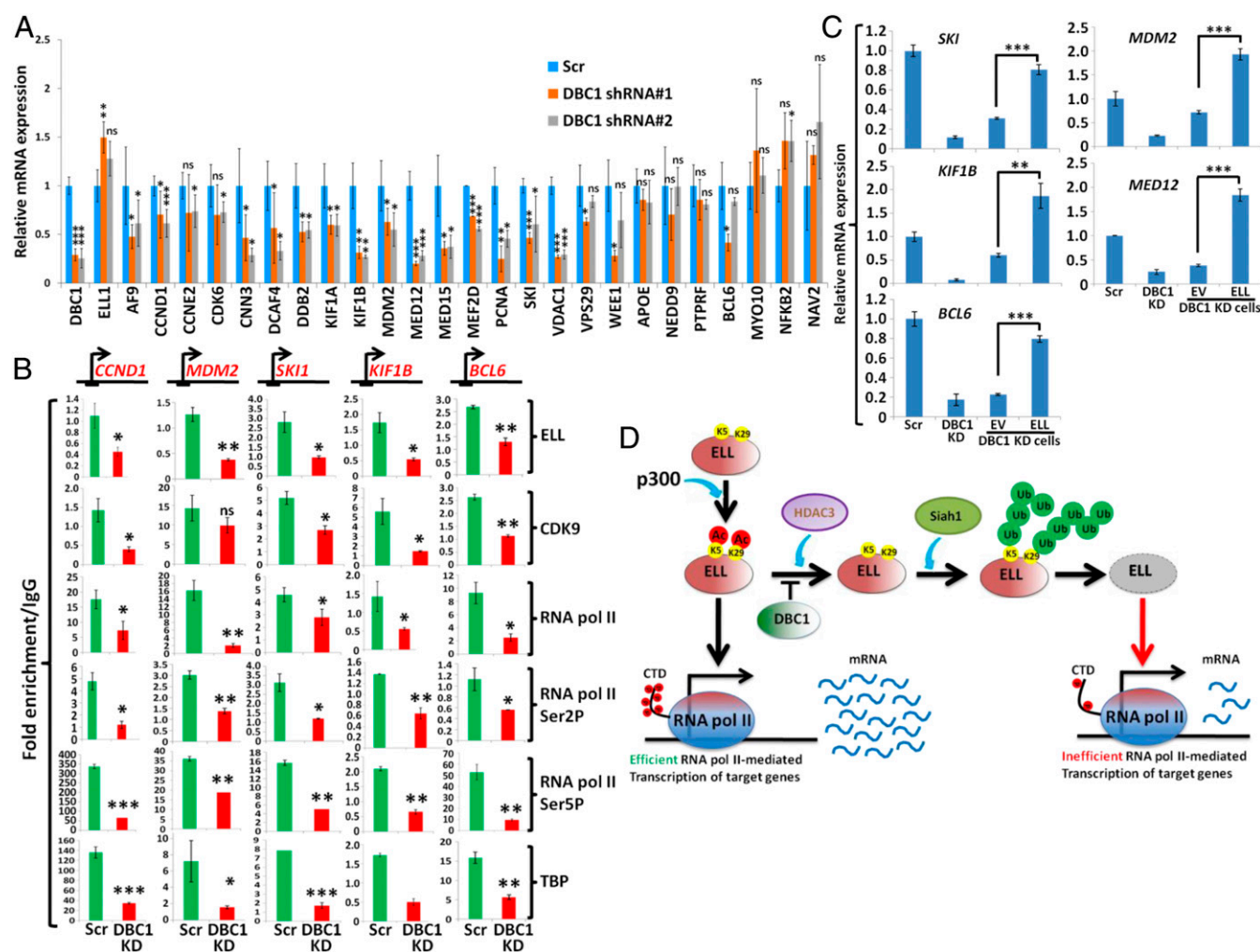


Fig. 6. Down-regulation of target gene expression upon DBC1 knockdown is critically dependent on ELL. (A) The qRT-PCR analysis showing the effect of DBC1 knockdown, by using two different shRNAs, on basal level expression of target genes when compared to control scramble knockdown cells. (B) ChIP analysis showing the recruitment of target factors at the TSS region of indicated target genes upon DBC1 knockdown. (C) The qRT-PCR analysis showing the effect of reexpression of ELL on expression of selected target genes in DBC1 knockdown cells. (D) Cartoon diagram depicting mechanistic understanding of the role of DBC1, p300, HDAC3, and Siah1 in regulating ELL expression and downstream effect on expression of target genes. DBC1 knockdown results in simultaneous down-regulation of ELL by the action of HDAC3 deacetylase and Siah1-mediated polyubiquitylation. Reduced ELL level thus causes impaired expression of its target genes. In all of our statistical analyses, one-tailed Student's *t* test was used to calculate the statistical significance of the data, wherein * denotes $P \leq 0.05$, ** denotes $P \leq 0.01$, *** denotes $P \leq 0.001$, and ns denotes not significant.

knockdown. We performed TBP ChIP (as part of TFIID), as a candidate factor for PIC assembly, and also observed its impaired recruitment only on the target genes (Fig. 6B), but not on the two nontarget genes (SI Appendix, Fig. S6G). Consistent with their reduced recruitment, immunoblotting analysis confirmed reduced expression of the majority of TAF subunits of TFIID upon DBC1 knockdown (SI Appendix, Fig. S7B). Importantly, further analysis showed that RNA expression of a few TAF subunits, including TAF3 and TAF6, were impaired (SI Appendix, Fig. S7C). Since expression of TAF6 plays a key role in overall expression and assembly of other TAF subunits within TFIID (32), a down-regulation of TAF6 by DBC1 knockdown possibly results in overall impaired TFIID assembly and its recruitment and subsequent Pol II recruitment for target gene activation.

We next addressed whether the overall effect of DBC1 knockdown on transcriptional down-regulation of selected target genes is indeed a result of down-regulation of ELL expression, by reexpressing ELL in DBC1 knockdown cells. As shown in Fig. 6C and SI Appendix, Fig. S7D, while the empty vector failed to restore target gene expression, reexpression of ELL restored transcription in the majority of the genes that we have tested. Thus, we conclude that, upon DBC1 knockdown, simultaneous down-regulation of ELL results in transcriptional down-regulation of the target genes (Fig. 6D). Restoration of transcription of target genes by overexpression of ELL strongly suggests that ELL level is rate limiting in transcriptional down-regulation of target genes in DBC1 knockdown cells.

Down-Regulation of Expression of DBC1, ELL, and Concomitant Glucose Homeostasis-Related Target Genes in Peripheral Blood Mononuclear Cells Isolated from Type 2 Diabetes Patients. Toward understanding a role for DBC1-mediated ELL stabilization in regulation of physiological processes, we further analyzed the set of metabolic genes that show down-regulation of expression upon knockdown of both DBC1 and ELL (SI Appendix, Fig. S6D). GO analysis predicted that a significant number of these genes are involved in metabolic processes, including nucleic acid, lipid, and glucose metabolism (Fig. 7A). Interestingly, the same analysis also predicted that misregulation of expression of a significant number of these coregulated genes (~12%) is involved in pathogenesis of Type 2 diabetes (Fig. 7B). Consistent with this prediction, we observed significant down-regulation of a number of genes that are required for insulin secretion as well as glucose uptake pathways, including *GLUT1*, *ATF2*, and cAMP responsive proteins (*CREB1* and *CREB5*) upon DBC1 knockdown in 293T cells (Fig. 7C). Similar to our earlier results, we also observed impaired SEC as well as Pol II and TFIID recruitment at the TSS region of the *GLUT1* gene (Fig. 7D). Further, restoration of *GLUT1* gene expression through overexpression of ELL strongly indicated a role of ELL in the overall regulation of expression of *GLUT1* (Fig. 7E) in 293T cells. Thus, we conclude that DBC1-mediated ELL stabilization is required for proper expression of target genes that are strongly implicated in maintaining glucose metabolism in humans.

For further understanding a role for this pathway in glucose metabolism-related diseases, we noted a recent study that showed a strong association of transcriptional down-regulation of DBC1 and up-regulation of HDAC3 mRNA in peripheral blood mononuclear cells (PBMCs) of Type 2 diabetes patients (33) without providing any mechanistic insight. For understanding the underlying mechanisms, we collected PBMCs from a cohort of 30 control (healthy) and 30 treatment-naïve patients and analyzed the level of target proteins being expressed in these samples. As shown in Fig. 7F, and consistent with the earlier study showing down-regulation at the RNA level (33), we also observed significant down-regulation of both DBC1 and ELL protein expression in the patient samples as compared to healthy controls (Fig. 7F). Interestingly, and consistent with our data in 293T cells, we failed to observe any significant

down-regulation in ELL RNA level in the patient samples when compared to healthy controls (Fig. 7G, ELL panel). These data further indicated a key role of DBC1 in maintaining ELL protein level within patient samples. Contrary to the reported study showing up-regulation of HDAC3 RNA level in Type 2 diabetes patients (33), we did not observe significant changes in HDAC3 at the protein level. Thus, based on our mechanistic understanding from 293T cells, we conclude that the observed down-regulation of ELL protein level could as well be a result of down-regulation of DBC1 protein leading to increased ELL degradation by HDAC3 that remains unchanged between healthy and patient cohort samples.

Next, we asked whether DBC1 down-regulation and simultaneous reduced ELL level in patient PBMCs would also correlate with decreased expression of glucose metabolism-related genes as observed in 293T cells. As shown in Fig. 7G, RNAs extracted from patient samples showed significant down-regulation of expression of some of the key glucose uptake and metabolism-related genes such as *GLUT1*, *CREB5*, and *GLP1R*. Since these factors are well known for their role in maintaining glucose homeostasis in humans, a reduction of their expression upon reduced DBC1 expression could potentially lead to imbalance in cellular glucose level in the peripheral system as observed in our study.

Overall, our results in this study uncover important mechanistic insights into ELL stabilization within mammalian cells involving DBC1, p300, HDAC3, and E3 ubiquitin ligase, Siah1. Among these players, DBC1 provides an additional layer of stabilization to ELL by protecting its acetylation and thus increasing its stability. A down-regulation of DBC1 further leads to increased ELL deacetylation by HDAC3 and its subsequent degradation by Siah1. Reduced ELL levels contribute to decreased expression of diverse sets of genes, including the ones that are required for glucose homeostasis, in 293T cells. Strong correlation of these results with PBMCs isolated from Type 2 diabetes patients further indicates possible involvement of these mechanisms in overall disease pathogenesis (Fig. 7H). Further studies, using a large cohort of patients as well as animal models, would be required for in-depth understanding of the overall effect of reduced DBC1 and ELL expression and its implications in Type 2 diabetes pathogenesis.

Discussion

In this study, we describe an important pathway in regulating the stability, and thus function, of a key SEC component, that is, ELL. Our results show a multilayered regulation involving DBC1, acetyltransferase p300, deacetylase HDAC3, and an E3 ubiquitin ligase, Siah1. The overall mechanism of regulation is presented in Fig. 6D.

Two earlier studies described a role for human E3 ubiquitin ligase Siah1, but not Siah2, in regulating cellular levels of other SEC components such as ELL2 and AFF1 (14, 28). The study by Liu et al. (14) failed to observe ELL degradation by Siah1, whereas, in the same assay, ELL2 and AFF1 (to a less extent) showed extreme sensitivity. Our results provide a mechanistic explanation for their observation. We have clearly shown that ELL degradation by Siah1 requires concomitant HDAC3-mediated deacetylation. Thus, our study provides a paradigm for detailed understanding of the maintenance of cellular levels of all of these factors for proper ELL-mediated transcriptional response within mammalian cells during exposure to various environmental stimuli.

With our studies, we also show a role for DBC1, acting as an “acetylation protector” for ELL through competitive binding with HDAC3. Since the target for acetylation and Siah1-mediated polyubiquitination happen to be on the same residues, interaction with DBC1 further protects ELL from HDAC3-mediated deacetylation and subsequent degradation by Siah1. This multiple layer of regulation of ELL, making it an extremely stable protein, is reflected in our CHX chase assay, wherein we have failed to observe any significant ELL degradation even after 24 h of chase (Fig. 2G). Besides ELL, DBC1 also showed

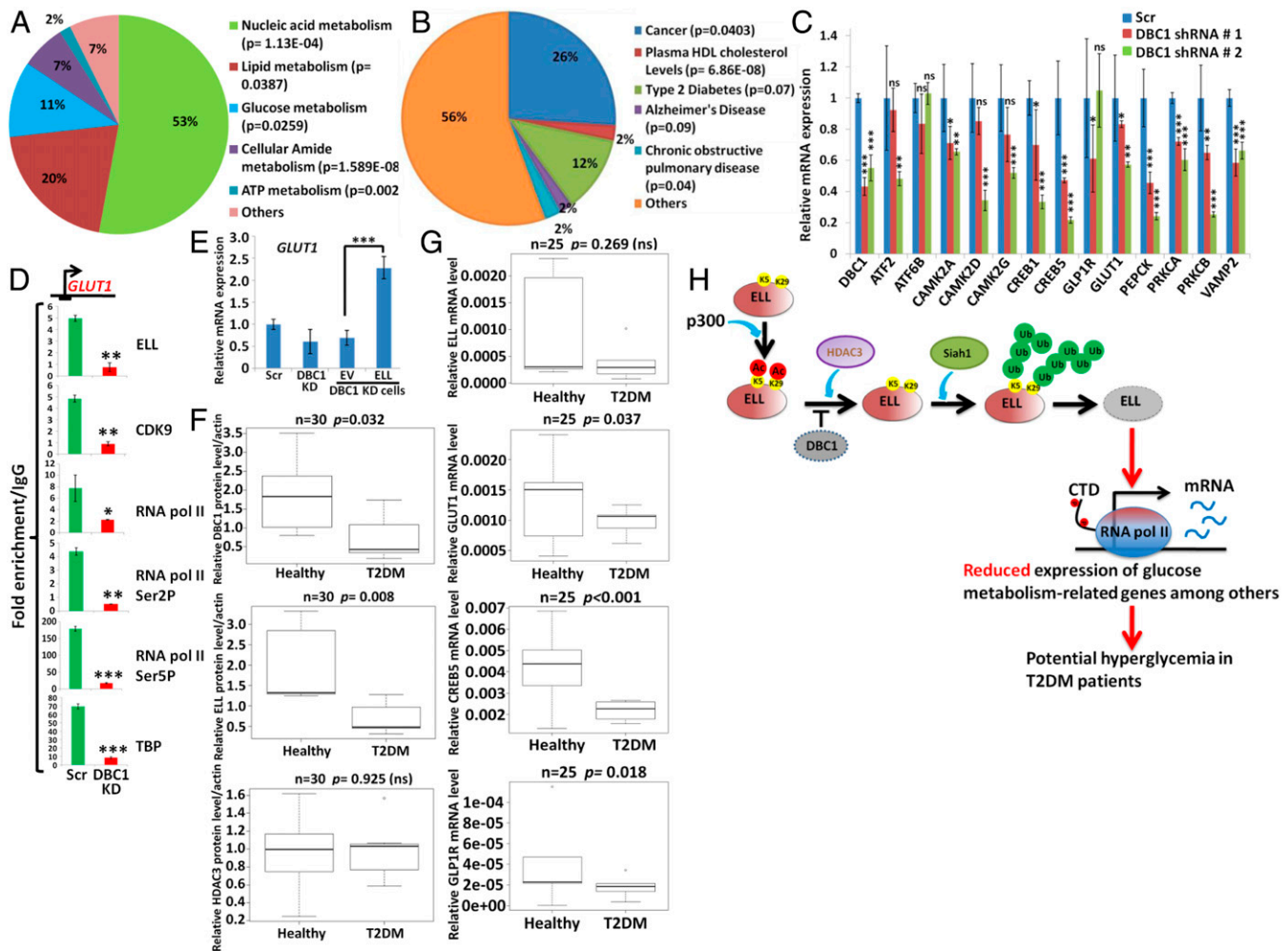


Fig. 7. DBC1-mediated ELL stabilization is required for expression of genes that are associated with glucose homeostasis in humans. (A) GO analysis showing the different types of metabolic processes controlled by DBC1-ELL coregulated genes that are predicted to be involved in controlling metabolic processes (24% as shown in *SI Appendix*, Fig. S6D). (B) GO analysis predicting the involvement of DBC1-ELL coregulated genes (*SI Appendix*, Fig. S6C) in various human diseases. (C) The qRT-PCR analysis showing the effect of DBC1 knockdown on the basal level expression of various glucose metabolism-related genes compared to control scramble knockdown cells. (D) ChIP analysis showing the recruitment of several target factors at the TSS region of *GLUT1* gene upon knockdown of DBC1. (E) The qRT-PCR analysis showing the effect of reexpression of ELL in DBC1 knockdown cells on basal level expression of *GLUT1* gene in 293T cells. (F) Quantitation of the level of indicated target proteins (DBC1, ELL, and HDAC3) in healthy and T2DM patient cohort samples ($n = 30$ for each) as measured by Western blot analysis. Actin was used as loading control for each sample. (G) The qRT-PCR analysis showing the expression of different target genes (at RNA level) in healthy and patient cohort samples ($n = 25$). For the ease of dealing with total number of samples for subsequent qRT-PCR analysis, equal concentrations of RNA sample from five individuals were pooled (in an unbiased way) and mixed together to form a group. A total of five groups each of healthy and patient RNA samples were tested for presence of target RNAs. (H) Cartoon diagram showing the overall model of the role of p300-mediated acetylation, DBC1-mediated acetylation protection, HDAC3-mediated deacetylation, and subsequent Siah1-mediated degradation in maintaining ELL level for proper expression of target genes that are strongly associated with glucose homeostasis in healthy individuals. A down-regulation of upstream DBC1 in these processes leads to decreased ELL level through HDAC3-mediated degradation and thus possibly leads to decreased expression of key genes required for glucose homeostasis in healthy individuals, thus indicating a possible involvement of this mechanism of hyperglycemia in Type 2 diabetes patients. In all of our statistical analyses, one-tailed Student's *t* test was used to calculate the statistical significance of the data, wherein * denotes $P \leq 0.05$, **denotes $P \leq 0.01$, *** denotes $P \leq 0.001$, and ns denotes not significant.

interaction with other SEC components such as EAF1/2, CDK9, and AFF1 as evidenced by our IP analyses (Fig. 1D). Interestingly, DBC1 knockdown also reduced cellular levels of all of these proteins without affecting their mRNA expression (*SI Appendix*, Fig. S7E and F). It would be interesting to further study whether DBC1 would also regulate abundance of other SEC components by the same mechanism as discussed in this study.

From our study, we also uncover an unexplored—yet interesting—role of ELL in regulation of expression of key genes that are required for maintaining several physiological processes, including glucose homeostasis. Notably, deregulation of a substantial number of genes that are coregulated by DBC1 and ELL has been predicted to be involved in giving rise to several diseases,

including Type 2 diabetes (Fig. 7B). The strong correlation of effect of reduced level of DBC1 on expression of glucose homeostasis-related genes, both in 293T cells and PBMCs isolated from Type 2 diabetes patients, further indicates the possible importance of our study toward a mechanistic understanding of Type 2 diabetes pathogenesis. Nevertheless, these interesting observations have to be studied in detail for a deeper mechanistic understanding of the role of these mechanisms in overall Type 2 diabetes pathogenesis.

Materials and Methods

Details of all materials regarding list of plasmids, primers for RNA and ChIP analyses, and antibodies used for our studies can be found in *SI Appendix*.

Further, methods detailing cell culture techniques, creation of different plasmid constructs, stable cell line generation, mass spectrometry analysis, nuclear extract preparation, IP analysis, protein complex purification from nuclear extract, luciferase assay, CHX chase assay, recombinant protein purification, in vitro acetylation assay, in vitro deacetylation assay, GO analysis, ChIP analysis, RNA analysis using qRT-PCR, and PBMC isolation can also be found in [SI Appendix](#).

Clinical Studies and Collection of Blood Samples. A total of 78 individuals were recruited from the IPGME&R, Kolkata, following the American Diabetes Association criteria (34) for the clinical studies described. Of these, 34 were treatment-naïve Type 2 diabetes mellitus patients, and 32 were healthy, age-matched controls. The research study was approved by the Institutional Ethics and Research Committee of IPGME&R and CSIR-ICB, and all test subjects

provided informed consent. Blood samples were collected from patients with Type2 diabetes and age- and gender-matched healthy controls after obtaining written informed consent from each study participant in his or her own vernacular.

Data Availability. All of the data for this study are included in the manuscript and [SI Appendix](#).

ACKNOWLEDGMENTS. This work was supported by Wellcome-Trust Department of Biotechnology India Alliance Intermediate Fellowship (IA/I/14/1/501287 awarded to D.B.). S.B., M.B., and D.Y. are recipients of University Grants Commission and CSIR Senior Research Fellowships. The HDAC and SIRT plasmids were obtained from Addgene, and we deeply acknowledge the sharing by the concerned authors.

1. K. Adelman, J. T. Lis, Promoter-proximal pausing of RNA polymerase II: Emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731 (2012).
2. D. Biswas *et al.*, Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15751–15756 (2011).
3. Z. Luo, C. Lin, A. Shilatifard, The super elongation complex (SEC) family in transcriptional control. *Nat. Rev. Mol. Cell Biol.* **13**, 543–547 (2012).
4. A. V. Krivosov, S. A. Armstrong, MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer* **7**, 823–833 (2007).
5. A. Shilatifard, W. S. Lane, K. W. Jackson, R. C. Conaway, J. W. Conaway, An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* **271**, 1873–1876 (1996).
6. A. Shilatifard, Identification and purification of the Holo-ELL complex. Evidence for the presence of ELL-associated proteins that suppress the transcriptional inhibitory activity of ELL. *J. Biol. Chem.* **273**, 11212–11217 (1998).
7. S. E. Kong, C. A. Banks, A. Shilatifard, J. W. Conaway, R. C. Conaway, ELL-associated factors 1 and 2 are positive regulators of RNA polymerase II elongation factor ELL. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10094–10098 (2005).
8. F. Simone *et al.*, EAF1, a novel ELL-associated factor that is delocalized by expression of the MLL-ELL fusion protein. *Blood* **98**, 201–209 (2001).
9. P. E. Polak, F. Simone, J. J. Kaberlein, R. T. Luo, M. J. Thirman, ELL and EAF1 are Cajal body components that are disrupted in MLL-ELL leukemia. *Mol. Biol. Cell* **14**, 1517–1528 (2003).
10. F. Simone, R. T. Luo, P. E. Polak, J. J. Kaberlein, M. J. Thirman, ELL-associated factor 2 (EAF2), a functional homolog of EAF1 with alternative ELL binding properties. *Blood* **101**, 2355–2362 (2003).
11. E. R. Smith *et al.*, The little elongation complex regulates small nuclear RNA transcription. *Mol. Cell* **44**, 954–965 (2011).
12. D. Hu *et al.*, The little elongation complex functions at initiation and elongation phases of snRNA gene transcription. *Mol. Cell* **51**, 493–505 (2013).
13. S. Mourgues *et al.*, ELL, a novel TFIIF partner, is involved in transcription restart after DNA repair. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17927–17932 (2013).
14. M. Liu, J. Hsu, C. Chan, Z. Li, Q. Zhou, The ubiquitin ligase Siah1 controls ELL2 stability and formation of super elongation complexes to modulate gene transcription. *Mol. Cell* **46**, 325–334 (2012).
15. D. Yu, R. Liu, G. Yang, Q. Zhou, The PARP1-Siah1 axis controls HIV-1 transcription and expression of Siah1 substrates. *Cell Rep.* **23**, 3741–3749 (2018).
16. J. E. Kim, J. Chen, Z. Lou, DBC1 is a negative regulator of SIRT1. *Nature* **451**, 583–586 (2008).
17. W. Zhao *et al.*, Negative regulation of the deacetylase SIRT1 by DBC1. *Nature* **451**, 587–590 (2008).
18. B. Qin *et al.*, DBC1 functions as a tumor suppressor by regulating p53 stability. *Cell Rep.* **10**, 1324–1334 (2015).
19. C. Escande *et al.*, Deleted in breast cancer-1 regulates SIRT1 activity and contributes to high-fat diet-induced liver steatosis in mice. *J. Clin. Invest.* **120**, 545–558 (2010).
20. V. Nin, C. C. Chini, C. Escande, V. Capellini, E. N. Chini, Deleted in breast cancer 1 (DBC1) protein regulates hepatic gluconeogenesis. *J. Biol. Chem.* **289**, 5518–5527 (2014).
21. J. Li *et al.*, A conserved NAD⁺ binding pocket that regulates protein-protein interactions during aging. *Science* **355**, 1312–1317 (2017).
22. K. Ghosh *et al.*, Positive regulation of transcription by human ZMYND8 through its association with P-TEFb complex. *Cell Rep.* **24**, 2141–2154.e6 (2018).
23. H. Jiang *et al.*, Regulation of transcription by the MLL2 complex and MLL complex-associated AKAP95. *Nat. Struct. Mol. Biol.* **20**, 1156–1163 (2013).
24. C. C. Chini, C. Escande, V. Nin, E. N. Chini, HDAC3 is negatively regulated by the nuclear protein DBC1. *J. Biol. Chem.* **285**, 40830–40837 (2010).
25. T. Ishizuka, M. A. Lazar, The N-CoR/histone deacetylase 3 complex is required for repression by thyroid hormone receptor. *Mol. Cell. Biol.* **23**, 5122–5131 (2003).
26. N. Kumari, M. A. Hassan, X. Lu, R. G. Roeder, D. Biswas, AFF1 acetylation by p300 temporally inhibits transcription during genotoxic stress response. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22140–22151 (2019).
27. L. Wang *et al.*, The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science* **333**, 765–769 (2011).
28. P. L. Oliver, E. Bitoun, J. Clark, E. L. Jones, K. E. Davies, Mediation of Af4 protein function in the cerebellum by Siah proteins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14901–14906 (2004).
29. J. S. Byun *et al.*, ELL facilitates RNA polymerase II pause site entry and release. *Nat. Commun.* **3**, 633 (2012).
30. P. Close *et al.*, DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation. *Nature* **484**, 386–389 (2012).
31. E. N. Chini, C. C. Chini, V. Nin, C. Escande, Deleted in breast cancer-1 (DBC-1) in the interface between metabolism, aging and cancer. *Biosci. Rep.* **33**, e00058 (2013).
32. K. J. Wright, M. T. Marr 2nd, R. Tjian, TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12347–12352 (2006).
33. C. Sathishkumar *et al.*, Augmentation of histone deacetylase 3 (HDAC3) epigenetic signature at the interface of proinflammation and insulin resistance in patients with type 2 diabetes. *Clin. Epigenetics* **8**, 125 (2016).
34. American Diabetes Association, 2. Classification and diagnosis of diabetes: Standards of Medical Care in Diabetes—2018. *Diabetes Care* **41** (suppl. 1), S13–S27 (2018).



Keeping RNA polymerase II on the run: Functions of MLL fusion partners in transcriptional regulation

Subham Basu^{a,1}, Arijit Nandy^{b,1}, Debabrata Biswas^{a,*}

^aLaboratory of Transcription Biology, Molecular Genetics Division, CSIR-Indian Institute of Chemical Biology, 4, Raja S. C. Mullick Road, Kolkata 32, India

^bAcademy of Scientific and Innovative Research (AcSIR), Ghaziabad, 201002, India

ABSTRACT

Since the identification of key MLL fusion partners as transcription elongation factors regulating expression of *HOX* cluster genes during hematopoiesis, extensive work from the last decade has resulted in significant progress in our overall mechanistic understanding of role of MLL fusion partner proteins in transcriptional regulation of diverse set of genes beyond just the *HOX* cluster. In this review, we are going to detail overall understanding of role of MLL fusion partner proteins in transcriptional regulation and thus provide mechanistic insights into possible MLL fusion protein-mediated transcriptional misregulation leading to aberrant hematopoiesis and leukemogenesis.

1. Introduction

The curious case of “lineage switching” in several cases of leukemia has been reported as early as in 1984. Five patients with acute lymphoblastic leukemia (ALL) at the St. Jude Children's Research Hospital in Memphis, Tennessee had suddenly converted from ALL to acute non-lymphoblastic leukemia (ANLL). Likewise, a single patient with ANLL had converted to ALL [1]. Mirro *et al.*, from the same hospital, concurred with these reported findings in light of their observations that in some patients with typical morphological features of ALL, certain myeloid lineage associated surface markers were also being expressed in these transformed cells, and thus they were not responding favorably to therapies typically used in ALL [2]. The presence of both these surface markers in a single occurrence of the disease led to the coining of the term “mixed lineage” leukemia. It had already been speculated in 1981 that the causative reason for leukemic transformations in patients may be chromosomal defects [3]. In 1992, researchers identified that some cases of acute lymphoblastic leukemia, especially in children, were occurring possibly due to chromosomal translocations involving a breakpoint at the 11q23 site that encodes a protein similar in sequence to the *Drosophila trithorax* gene [4]. This gene, named MLL1 or Mixed Lineage Leukemia 1 encodes a protein of 3969 amino acids and contains several characteristic domains. At the N-terminus, it possesses three AT hooks that enable it to bind to DNA, a CXXC motif for binding to nonmethylated CpG nucleotides as well as the PAF1 complex, a PHD domain which enables Histone H3 lysine methylation recognition, followed by a bromodomain that can recognize acetylated lysines and two

FYR domains for mediating protein-protein interaction [5–12]. At the C-terminus, it possesses a transactivation domain, mostly serving as a scaffold for interaction of transcription factors, and a highly conserved SET domain which serves as a histone H3 lysine 4 methyltransferase [13,14]. Besides, its interacting partner WDR5 also interacts with histone H3 and may serve to recruit MLL for further histone methylation [15]. It is perhaps the SET domain which is most essential for MLL function, i.e. regulating expression of the *HOX* gene cluster during mammalian embryonic development [13].

The gigantic MLL protein is cleaved by Taspase1, generating an N-terminal 320 kDa fragment and a C-terminal 180 kDa fragment [16]. It has been suggested that this cleavage provides stability to the N-terminal fragment and results in nuclear localization for enhancing the availability of this fragment for its target genes [17]. Several studies on patients affected with leukemia led to the discovery that the MLL N-terminal fragment is able to form in-frame fusion with several different genes across chromosomes [18–33]. These lead to the development of approximately 70% of all infant ALL and 35–50% of infant AML, with a survival period of less than 5 years [34,35]. As of 2017, more than 150 genes have been reported to fuse to MLL (atlasgeneticsoncology.org). Of these, the most frequent MLL fusion partners observed in clinical cases are AFF1, ENL, AF9, AF10, ELL, AF6, AF17 and SEPT6 [36]. In a comprehensive study carried out in 2017 among 1420 ALL patients, 96% of reported MLL fusions comprised of MLL-AFF1 (57%), MLL-ENL (18%), MLL-AF9 (13%), MLL-AF10 (4%), and MLL-AF6 (2%). Additionally, 83% of the reported MLL fusions in AML patients were found to harbor MLL-AF9 (30%), MLL-AF10 (16%), MLL-ELL (11%), MLL-AF6

* Corresponding author.

E-mail address: dbiswas@iicb.res.in (D. Biswas).

¹ These authors contributed equally to this work

(8%), MLL-ENL (4%), and MLL-SEPT6 (2%) [37]. In order to elucidate the molecular mechanisms behind MLL-fusion gene-mediated transcriptional deregulation affecting expression of key genes such as *HOXA5-11*, *MEIS1*, *p27*, *MEF2C*, *EPHA7* and the *mir-17-92* cluster, several groups sought to identify proteins associated with commonly occurring MLL-chimera [38–47]. These studies showed that the MLL-AFF9 and MLL-ENL-interacting partners included AFF1, AFF4 and DOT1L, an enzyme catalyzing histone H3 Lysine79 methylation, found to be critical for leukemogenesis by MLL fusions [48–52]. What was mostly surprising is that even though MLL-AFF4 fusions rarely occur, AFF4 was found to be a part of all the chimeric complexes. Additionally, and interestingly, it appeared as if the most common MLL fusion partners, though encoded from various different chromosomal regions, often exist as part of a large macromolecular complex. All these combined observations led to the postulation of the presence of a single large super elongation complex (SEC) consisting of AFF4, AFF9, ENL, ELL, EAF1, and the RNA polymerase II (Pol II, hereafter) C-terminal domain (CTD) kinase P-TEFb.

Several further studies helped in elucidating the roles played by this elongation complex in transcriptional regulation. The SEC has mostly been implicated in mediating elongation of Pol II in response to several different stimuli. These may include genes involved in stress responses, rapid response genes, developmentally controlled genes or in genes implicated in several diseases [53,54]. Perhaps, the most well studied role of the SEC is in the release of promoter proximally paused Pol II to facilitate its entry into productive elongation [55]. It has also been well established that in genes where Pol II is not paused, SEC still plays a role in Pol II elongation. In HIV1 pathogenesis as well, SEC recruitment has been found to be essential for Tat transactivation property [56,57]. The presence of the most common MLL fusion partners in a single complex, which play important roles in transcriptional regulation mediated by Pol II, make them attractive targets for detailed studies. Indeed, several studies over the last few years have focused on understanding the role of these proteins in various cellular processes that are either dependent on or independent of MLL fusions. This has also led to the development of several small molecule inhibitors for therapeutic interventions for the treatment of MLL fusion-mediated leukemia itself and also for other diseases caused by mis-regulation of these factors [35,58–60]. The current review focuses on the normal transcriptional activity of MLL fusion partner protein containing complexes and provides future perspectives for their functional regulation and their relevance in disease pathogenesis.

2. Overview of transcription by RNA polymerase II

As this review is primarily based on transcriptional regulation by proteins which fuse to the MLL1, we must always keep in mind that these proteins are involved in the core transcriptional processes. Hence, initially, it is necessary to have a brief outline of the steps in eukaryotic transcription and a subsequent discussion on the role(s) of the MLL fusion partners in these steps.

Eukaryotic transcription initiation starts with the modification of chromatin for the release of DNA from its histone bound form to a free structure. Several families of ATP-dependent chromatin remodeling complexes have been described, which function in tandem for removal of histones and decompaction of DNA: SWI/SNF, ISWI, CHD, RSC, and INO80/SWR1 [61]. Three co-activator complexes with the ability to acetylate chromatin also exist in eukaryotes: SAGA (SPT-ADA-GCN5 acetylase), TFTC (TATA-binding-protein-free TAFII-containing complex), and STAGA (SPT-3-TAFII31-GCN5L acetylase) [62]. The Mediator co-activator complex functions to “mediate” signals between transcriptional activators located at enhancer regions and make them accessible to the promoter region for pre-initiation complex (PIC) formation [63]. A simplified view of these regulations and factors involved is presented in Fig. 1.

In the first step of transcription initiation, TATA binding protein

(TBP) of the TFIID complex binds to the promoter DNA. This is followed by TFIIB and TFIIA, which facilitate DNA binding to the TBP [64,65]. Subsequent recruitment of Pol II followed by TFIIF and Mediator complexes enables Pol II stability and helps in assuming conformations conducive to elongation [66,67]. Further recruitment of TFIIE followed by TFIIH [68,69], results in formation of complete pre-initiation complex (PIC) assembly. The XPB subunit of TFIIH possesses ATPase and translocase activity. It interacts with the DNA, and in presence of ATP, enables its “opening” using its 5′-3′ translocase activity, leading the DNA to enter into the Pol II active site. Once XPB dissociates, the TFIIB, TFIIE and TFIIF keep the open DNA complex stabilized [70]. Unphosphorylated Pol II has an affinity for the Mediator complex, and thus, for promoter escape, it must be phosphorylated [71,72]. This is accomplished by the TFIIH subunit CDK7, which phosphorylates Pol II at the Ser5 and Ser7 in the heptad repeats of its CTD [73,74]. Once transcription initiation occurs, Pol II synthesizes nascent RNA approximately 20–60 nucleotides in length and then pauses [75–79]. Several factors contribute to Pol II pausing, the most prevalent of them being the DRB sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF) [80]. Other factors involved may include Gdown1, KAP1/TRIM28 and, as most recently described, FACT [81–83]. Pol II release from pausing for rapid expression of target genes may be affected by developmental signals, environmental stimuli, inflammatory signals, stress, and epigenetic modifications and serves as an important checkpoint for transcriptional regulation [84].

The release of paused Pol II into productive elongation is mediated by positive transcription elongation factor b (P-TEFb), a heterodimer of the kinase module CDK9 and CyclinT1/2. P-TEFb is responsible for phosphorylating the Pol II CTD at the Ser2 position of its heptad repeats. P-TEFb also phosphorylates the Spt5 subunit of DSIF and NELF-E subunit of the NELF complex, causing NELF-E dissociation and progression of the Pol II into productive elongation [85]. Several factors are involved in P-TEFb recruitment from its inactive form present as the 7SK snRNA complex (7SK snRNP). These include BRD4, MYC, NF- κ B, G-Actin, p53, ER- α , MED23 and ZMYND8 [86–92]. The KAP1/TRIM28 protein has also been reported to affect transcription elongation via recruitment of the 7SK snRNA complex to promoter proximal sites [93]. Interestingly, BRD4 and the SEC are critical for the formation of the active P-TEFb complex. MED1, MED23 and the Tat-SF1 play a role in BRD4-mediated P-TEFb recruitment, whereas for the SEC, this phenomenon involves PAF1c, MED26, and TFIID [94,95]. The composition of SEC is not rigid since results from our group and others have shown dynamic association between SEC components depending on stages of transcription as well as gene targets [85,96]. Because the coordinated function of the SEC is absolutely essential for proper transcriptional output, we will thus, be focusing on the functional regulation of different SEC components individually and implication of their mis-regulation in various human diseases.

3. Interaction of MLL fusion partners with various other factors

Detailed list of various interacting proteins with MLL fusion partners is mentioned below in Table 1:

4. MLL fusion partners in transcriptional regulation

4.1. AFF1 and AFF4

The AFF4/EMR2 (AFF) family consists of four members (AFF4/FMR2 Family Member 1–4; AFF1–4) [229–233]. All four members have been shown to possess a highly conserved serine-rich transactivation domain which has been shown to be involved in acute leukemia [234–237] (Fig. 2A). Among these 4 members, only AFF1 has been reported as a MLL-fusion protein in majority of the leukemia cases. Others have been reported to occur as MLL fusions, but on very rare occasions [35,37,238]. AFF1 was initially found as a MLL fusion partner but AFF4

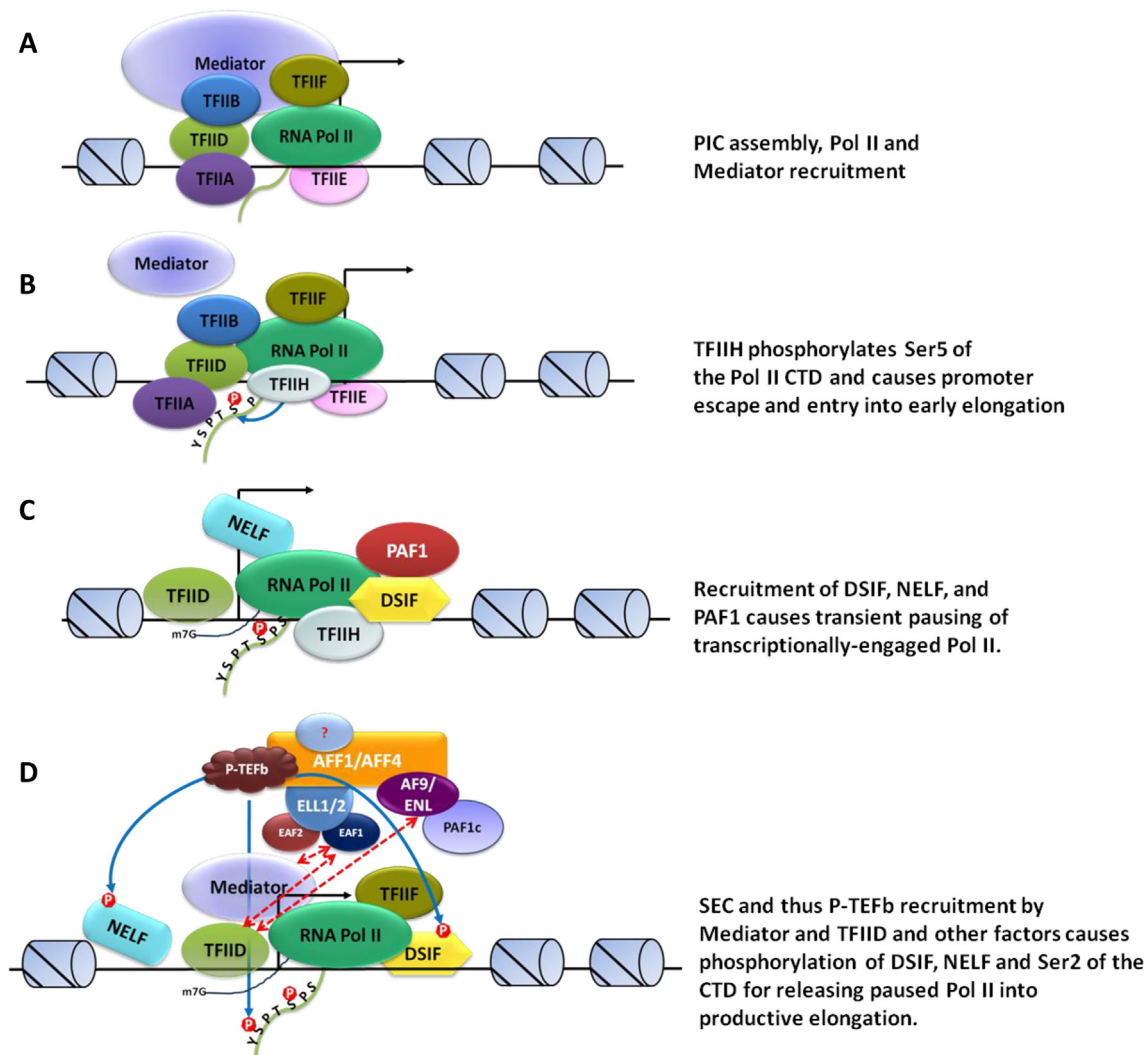


Fig. 1. An outline of the basic steps involved in eukaryotic transcription initiation and its transition into productive elongation.

A) Assembly of the pre-initiation complex along with Pol II and the Mediator complex at the promoter region for transcription initiation.

B) Phosphorylation of the Ser5 of the Pol II CTD by CDK7 of the TFIIH complex enables RNA Pol II promoter escape.

C) Concerted action of DSIF, NELF and PAF1 cause Pol II to transiently “pause” after travelling between 20 and 60 nucleotides of the transcription start site (TSS).

D) P-TEFb complex recruitment through recruiting SEC by Mediator and the TFIID complex as well as others (designated as “?”) lead to phosphorylation of DSIF, NELF and Ser2 of the CTD and thus causes release of paused Pol II and marks its entry into productive elongation.

was isolated as a P-TEFb-interacting partner in acute lymphoblastic leukemia [229,239].

AFF1 and AFF4 (AFF1/4 hereafter) are structurally very similar (Fig. 2A). Both are elongated, flexible scaffold proteins. More than 90% of these proteins is predicted to be intrinsically disordered [98,240]. They are rich in hydrophilic amino acids which favor such structure and have patches of conserved hydrophobic amino acids positioned along their length for binding to partner proteins [98]. Both AFF1 and AFF4 act as scaffolds to assemble P-TEFb, AF9 or ENL, ELL or ELL2/3 to form the SEC [47,96]. HIV Tat protein has been shown to work in close association with AFF1/4 to recruit both P-TEFb and ELL/ELL2 for transcriptional activation of the HIV provirus. The ability of Tat to recruit two such strong potential transcriptional activators (mostly in the context of SEC) makes it a very powerful activator for HIV proviral transcription [241].

Interacting regions on AFF4 have been closely mapped wherein, P-TEFb binds to the extreme N-terminal region (1–73 amino acids), ELL2 binds between the 301–350 amino acid region and ENL/AF9 binds between the 761–774 amino acid region [241] (Fig. 2A). Although the ENL and ELL2 binding regions of AFF4 consist of short hydrophobic

stretches, for binding to CyclinT1, AFF4 has multiple hydrophobic patches arranged over a larger region. AFF4 has been shown to snake around CyclinT1 and bind it through multiple amino acid pair interactions along an extensive surface. However, multiple studies, including ours, have shown that AFF4 does not make direct contact with CDK9 to bind P-TEFb [96,99,241]. Instead, Tat directly binds to a less open pocket created by AFF4 and P-TEFb and enhances the affinity of AFF4 for P-TEFb [98,99]. AFF1/4 partner proteins have also been shown to have disordered regions through which they bind to AFF1 or AFF4. Mutual coupling of these disordered regions among the two binding partners provide enough hydrophobic interaction interface to stabilize the interaction and also causes local folding of the binding regions of both partners [98,99]. This coupled folding has been well studied for AF9 which is discussed later [98,242]. Different SEC components bind to AFF1 or AFF4 in a dynamic and stoichiometric fashion [53,96,112]. The binding of one component can be independent of the other which provides a basis for somewhat independent and dynamic interaction of AFF1/4 with other SEC components. These interactions may get modulated by post-translational modifications at different interaction regions of AFF1/4 at different cellular contexts or for different

Table 1
MLL fusion partners and their key interacting proteins.

MLL fusion partners	Interacting proteins
AFF1/AFF4 AF9/ENL	Tat [56,57,97–100], SIAH1 [101,102], Fused in sarcoma(FUS) [103], DDX6 [104], Selectivity Factor1 (SL1) [105], Vitamin D receptor (VDR) [106], DOT1L [107–109], PAF1 [108], CBX8 [108–110], BCL6 co-repressor [111], PAF1 [112], Hsp90 [113], Dot1a [114–118], TAF5 and TAF6 (as a part of TFIID complex) [95], Mineralocorticoid receptor(MR) [116], serum- and glucocorticoid-induced kinase-1 (Sgk1) [117], RING finger protein 2 (Rnf2) [119], Diversin [120], SWI/SNF complex [121], Bmi1 [122], Ku70 [122]
ELL	PAF1 [98], p300 [123], KIAA0947 [124], NARG2 [124], Dishevelled (Dvl) [125], Mineralocorticoid receptor(MR) [126], HIF1- α [127], c-Myc [128], E2F1 [129], HDAC1 [129], CDK7 (as a part of TFIIF) [130]
EAF1/2	MED26 (as a part of mediator complex) [131,132], TAF7, TAF8 and TAF9 [95], β -catenin transcription complex components (β -catenin, Axin, Tcf, c-Jun) [133], GATA1 [134], Spi1 [134], p53 [135], TGF- β transcriptional factors(Smad2, Smad3, Smad4, Foxh1) [135], pVHL [136], HIF-1 α [137], FB1 [138], Ku70/Ku80 [139]
AF10 AF17 AF6	DOT1L [107,140], HP1 [141], SYT(synovial sarcoma associated protein) [142], Follistatin-related gene(FLRG) [143], GAS41 [144] Dot1a [118,145] ZO-1[146–148], Fam Deubiquitinating Enzyme [149,150], Ponsin [151], Nectin/PRR [152], Afadin DIL domain-interacting protein (ADIP) [153,154], Junctional adhesion molecule (JAM) [148], SPA-1 (Rap1 GTPase-activating protein) [155,156], Kalirin-7 [157], Connexin36 [158], AMPA receptor [159], Usp9x [160], SHP2 [161], Profilin [162], F-Actin [163], α -Catenin [164]
P-TEFb	MBP [165], Retinoblastoma protein(Rb) [165], HIV1 Tat protein [56,166–173], DRB Sensitivity Inducing Factor (DSIF) [174–176], Negative transcription elongation factor (NELF) [177], CTD of RPB1 [178,179], TFIIF [180], HEXIM1 (as a part of 7SKsnRNP) [181–186], BRD4 [92,187–192], Myc [193,194], ZMYND8 [87], PPAR- γ [195], Class II transactivator (CIITA) [196], NF- κ B [86], Granulin [197], Tax [198], Aryl hydrocarbon receptor [199], Androgen receptor (AR) [200], VP16 [201], PIE-1 [202], Runx1 [203], Pcm1 [204], Skp2 [205–207], UBR5 [208], p300 [209,210], GCN5 [211], P/CAF [211], HDAC3 (as a part of N-CoR) [209], SIRT7 [212], SIRT2 [213,214], STAT1 [214], STAT3 [215], PIAS1 [216], TRIM28 [93,217], Xrn2 [218], TIP60 [219], REV-ERB α [220], Protein phosphatase 1 (PP1) [221], G-Actin [90], Autoimmune response element (AIRE) [222], Estrogen receptor- α [89], MED23 [91], MMP9 [223], LDB1 [224], TAF7 [225], CDK2 [226], Hsp70 and Hsp90 [227], p53 [228]

sets of genes [53,243]. The C-terminal occludin homology domain of ELL2 has been shown to bind AFF4 [241]. Interestingly, the N-terminus of ELL2 directly interacts with PAF1 which raises the possibility that ELL2 may connect the SEC with the PAF complex [98]. Similar interactions of AFF1 with other SEC components have also been shown [52,54]. The C-terminal end of AFF1 or AFF4 does not interact with any of the SEC components and thus, may associate with other uncharacterized partners which could be important in leukemogenesis as this region remains intact in MLL fusions [52,96]. EAF1 and EAF2 may also interact dynamically with AFF1/4 or sometimes may remain outside the AFF1 or AFF4 containing SEC [53,96].

AFF1 is a highly unstable protein inside the cell at least in part due to its proteasomal degradation by the E3 ubiquitin ligase SIAH1 [102]. The importance of this rapid degradation can be appreciated from the fact that AFF1 fused to MLL gets protected from this degradation and the resultant fusion protein is very stable and highly oncogenic. The overexpression of AFF1 itself can transform cells [101,102]. Bitoun *et al.*, have shown that the AFF1 not only acts merely as a scaffold but also stimulates the activity of P-TEFb [244]. In a negative feedback loop mechanism, the same study has also shown that P-TEFb-mediated AFF1 phosphorylation causes a reduction in its transactivation potential. Binding of AF9/ ENL protects AFF1 from SIAH1-mediated degradation through the displacement of SIAH1 [244]. SIAH1 and AF9/ENL bind far apart on AFF1, suggesting an allosteric change in AFF1 upon binding to AF9/ENL resulting in the inefficient binding of SIAH1 and thus leading to stabilization. Phosphorylation of AF9/ENL by P-TEFb promotes their proteasomal degradation, possibly because of their dissociation from AFF1 [244].

Elegant biochemical analyses from many groups have shown HIV Tat to be present in a single complex with SEC components for HIV proviral activation [56,57,100]. He *et al.*, have shown that majority of the AFF4 and ELL2 remains in a P-TEFb associated complex inside the cell. This is a separate complex from two previously described P-TEFb containing complexes (Brd4-P-TEFb and 7SK snRNP) [181,182,187,189]. Although ELL2 interacts with P-TEFb in the absence of Tat, the presence of Tat increases their interaction. ELL2 (but not ELL) and Tat synergistically increase HIV transcription in a TAR-dependent manner in which Tat increases the association between ELL2 and P-TEFb, most likely by increasing the level of ELL2 through its protection from proteasomal degradation [56]. The study has also shown that the kinase activity as well as its association with ELL2 (through Tat or AFF4), P-TEFb helps in ELL2 stabilization.

Huasong Lu *et al.*, have shown AFF1 as a *bona fide* component of the 7SK snRNP [100]. They demonstrated that through its direct binding to P-TEFb, AFF1 interacts with the rest of the 7SK snRNP components and gets sequestered into the inactive 7SK snRNP complex [100]. Upon appropriate stimulation, AFF1 cooperates with Tat to enhance its affinity towards CyclinT1 and excludes HEXIM1. This enhances the ability of Tat to extract P-TEFb from the 7SK snRNP by direct binding to the CyclinT1 and also orchestrates the entire SEC recruitment for activating HIV transcription. Upon cellular stress, e.g. flavopiridol treatment, AFF1-P-TEFb complex gets released from the 7SK snRNP and associates with Brd4 (as a separate complex from SEC) to induce the expression of primary response genes [100].

AFF1 and AFF4 have both overlapping as well as non-overlapping functions inside the cell. It has been shown that in the absence of AFF4, AFF1 can only bind to CDK9 but not to the other SEC components like ELL1, ELL2, AF9, ENL, whereas, knockdown of AFF1 does not affect the binding of AFF4 with other SEC components [245]. These observations hint towards the fact that AFF1 could be a part of only a minor subset of SEC-related complexes and AFF4 is the major determinant of SEC formation. This work also indicates that whatever AFF1 interaction is observed with other SEC components could partially be due to AFF1 and AFF4 co-residence in one complex, as both these proteins have been shown to form heterodimers [246]. Whereas AFF1-SEC is more potent in extraction of P-TEFb from the inactive complex and HIV transactivation in conjunction with Tat, AFF4-SEC is more involved in stress mediated responses e.g. HSP70 induction upon heat shock [97]. Also, in the context of cellular genes, AFF1 and AFF4 control a diverse set of non-overlapping genes with distinct sets of functions along with regulating some common genes [97]. In this regard, FUS protein has been shown to inhibit AFF4 and ELL2-mediated activation of HIV transcription [103]. DDX6, a DEAD-box RNA helicase has also been shown to bind to both 7SKsnRNA and AFF1 [104]. DDX6 helps in the release of P-TEFb from the 7SK snRNA complex and hands it over to AFF1-SEC for efficient gene activation. Interestingly DDX6 has been shown to be upregulated in many different cancer types where aberrant gene expression is a common feature [104]. Recently, RBM7, another RNA binding protein has been found to have the ability to release P-TEFb by interacting with 7SK snRNP upon genotoxic stress for transcriptional upregulation during early DNA damage response (DDR) genes [247]. It is to be seen whether in this case as well, P-TEFb is handed over to the AFF1-SEC or not.

AFF1, through its serine rich domain recruits Selectivity Factor1



Fig. 2. Cartoon diagrams depicting the domain architecture of the most common MLL fusion partner proteins and their interactors for regulation of transcription. The abbreviation used for marking the important domains and their full nomenclature are mentioned in the inset panel.

(SL1 consists of TBP and four other TAF proteins namely TAF1A-D) on to the promoter of ribosomal genes and mRNA encoding genes for PIC formation to regulate Pol I- and Pol II-driven transcription respectively [105]. The working model proposed in this study suggests that ENL (and also possibly AF9) directly binds to chromatin via YEATS domain and recruits AFF1 (or AFF4) and P-TEFb. AFF1/AFF4 then recruits the SL1 complex through interaction with its TAF1C component and ultimately recruits Pol II [105]. Further, our recent study has also shown that p300-mediated dynamic AFF1 acetylation is a key mechanism for transcriptional inhibition in response to genotoxic stress [243]. AFF1 complex has been shown to directly interact with the vitamin D receptor (VDR) and promotes transcription of the *ALOX5* gene in a calcitriol-dependent manner [106]. Thus, besides its role as a transcriptional regulator through its presence within the SEC, AFF1 and AFF4 could also regulate transcription within mammalian cells through their interactions with multiple other factors (Table 1).

4.2. AF9 and ENL

AF9 and its homolog ENL were first identified as MLL fusion partners [248,249]. AF9 and ENL form similar complexes with other proteins but do not reside in the same complex [112]. AF9/ENL (AF9 or ENL) majorly form three distinct complexes by binding to key components in each complex via their C-terminal ANC1 homology domain (AHD) (Fig. 2B,C). As part of SEC, AF9/ENL recruits P-TEFb along with ELL for the release of paused Pol II for transcriptional activation of target genes [47,52,57,96]. Further, as a separate complex, AF9/ENL recruits DOT1L to transcriptionally poised regions on chromatin. DOT1L is a known H3K79 methyltransferase whose catalytic activity helps in exclusion of SIRT1 mediated histone deacetylation and prevents epigenetic silencing for maintaining transcriptional activation [52,107,108]. AF9/ENL also forms a transcription repression module via direct interaction with the polycomb repressive complex 1 (PRC1) component CBX8. Upon recruitment, the PRC1-mediated ubiquitylation of histone H2A K119 marks the chromatin region for repression [110,250,251]. Interestingly, in case of MLL fusions, SEC and DOT1L-mediated gene activation is predominantly observed, whereas, repressor complex-mediated transcriptional inhibition is completely absent, thus resulting in hyperactivation of MLL target genes [252]. ENL/AF9 complex with DOT1L has been shown to inhibit SEC-dependent transcription and thus suggests a competition between DOT1L and SEC components for AF9/ENL-mediated target gene activation [112]. ENL/AF9 also associates with the human PAF1c complex via their N-terminal YEATS domain (Fig. 2B,C), and through this interaction, ENL/AF9 associates with chromatin-bound Pol II and eventually recruits the whole SEC on chromatin [112].

AF9/ENL has been shown to get recruited to a subset of MLL target gene promoters. This could be through AF9/ENL interaction with PAF1 via its YEATS domain and PAF1 interaction with the CXXC domain of MLL [11,12,112]. This suggests a mechanism of AF9/ENL function in transcriptional upregulation of target genes by both wild type MLL and MLL-fusion proteins [11,12,112]. ENL YEATS domain-mediated PAF1 recruitment has also been shown to oppose CBX8-mediated transcriptional repression [108].

The AHD region has been shown to be important for MLL-AF9-mediated leukemia [253,254]. Binding of AFF1 to AF9 induces folding of AHD to attain a suitable structure to accommodate AFF1. AF9 AHD normally consists mostly of random coils, but, upon binding to AFF1, this region folds into mixed α -helix and β -sheet structures [242]. This 'mutual synergistic folding' provides a large interface to generate extensive intermolecular interaction for high affinity binding between two partner proteins while maintaining a high amount of entropy in the whole complex [242]. AF9-interacting partners also have been shown to possess disordered regions and interactions between two disordered regions from two partners causes mutual stabilization [242]. Interestingly AF9 AHD attains similar fold after binding to any of its partners

and this may be the reason for AF9 binding to all of them in a mutually exclusive manner [52,111,242]. This mutual synergistic folding provides a unique opportunity for two proteins to form a higher order complex [255]. The competition for the same binding site allows rapid exchange among binding partners which is crucial for dynamic regulation of transcription [256]. AF9 AHD, thus, may enable very dynamic interaction with its binding partners depending on their local concentration, and possibly post translational modifications of both the AF9 AHD and the interacting domains of the binding partners.

Among all the SEC components, AF9 and ENL have the capacity to directly bind to chromatin via their YEATS domain [109,257,258]. AF9 binds strongly to acetylated H3K9 and with a lesser affinity to acetylated H3K27 and H3K18, but not to unmodified or methylated H3 [109]. ENL, on the other hand, binds more strongly to acetylated H3K27 and with a lesser affinity to acetylated H3K9 [257,258]. This may be one of the basis of AF9/ENL mediated recruitment of other factors for transcriptional upregulation only at the chromatin region that is primed for high amount of transcription. Active promoter regions are marked by H3K9 acetylation and active enhancers by H3K27 acetylation. As the YEATS domain of ENL shows higher preference towards H3K27Ac over H3K9Ac, it could be possible that AF9 recruits SEC and DOT1L mostly at the promoter region and ENL could do the same at the enhancer region [259–262]. This provides a basis for non-redundant functions of AF9 and ENL as they reside in separate complexes [96,112]. This non-redundancy can also be appreciated from the fact that ENL, but not its homolog AF9, is required for maintenance of acute leukemia [257,258]. Along similar lines, it is also interesting to note that MLL-ENL predominantly gives rise to lymphoid cell type and MLL-AF9 gives rise to mostly myeloid cell type in the resultant leukemia [112,263–265]. Also, in this context, AF9 has been shown to regulate E/Meg cell fate possibly by regulating the expression of E/Meg affiliated gene expression programs, particularly like GATA1 [266]. At the structural level, YEATS domain takes up a conformation of immunoglobulin fold consisting of eight β -strands which make up an aromatic cage lined with serine/threonine residues to bind acetyl-lysine on histones which is a different mode of recognition than that is adopted by bromodomains or tandem PHD fingers motifs [109,267–269]. H3K9 acetylation may be carried out by PCAF or GCN5 along with p300 as treating cells with a selective PCAF/GCN5 inhibitor has been shown to reduce AF9 occupancy on the *Myc* gene [109]. Through this mechanism AF9 has been shown to recruit DOT1L on chromatin for H3K79me3 deposition [109].

Although p300 carries out both histone acetylation and crotonylation (depending on the cellular crotonyl-CoA availability), crotonylation provides higher transcriptional output than acetylation [270,271]. Just like acetylation, in case of histone crotonylation, the crotonyl functional group is transferred from crotonyl-CoA on to specific lysine amino acids of N-terminal tails of histones through covalent bonding. These two marks are not mutually exclusive. Rather they can be present in the same actively transcribed genomic regions [272]. AF9 through its YEATS domain, and TAF1 through its bromodomain have been shown to recognize histone crotonylation [271,273,274]. AF9 YEATS domain shows a higher affinity towards crotonylated-lysine over acetylated-lysine, whereas, TAF1 bromodomain shows the opposite preference [271,273,274]. Unlike acetyl-lysine, crotonyl-lysine, because of its double bond side chain, can engage in π - π aromatic stacking with the aromatic amino acids of the YEATS domain core and can engage in more hydrophobic interactions. All these provide the basis for higher affinity of the YEATS domain towards the crotonyl-lysine. This π - π aromatic stacking is absent in case of bromodomain. Also, recognition of acetyl-lysine by a bromodomain is majorly dependent on hydrogen bonding in which water molecules in the bromodomain cleft play a major role. Due to an extended conformation of crotonyl-lysine, water molecules cannot accommodate it in a very energetically favourable way in the bromodomain cleft. All these make crotonyl-lysine a less desired candidate for the bromodomain [271,273,274]. All these

different acyl-modifications at different histones and at different residues combined with varying affinity of different readers towards them, give cells enormous plasticity for gene expression depending on different context and availability of different acyl-CoAs.

ENL has also been shown to be a part of a low abundance SWI/SNF complex and is a homolog of yeast SWI/SNF subunit ANC1/TFG3 [121]. Interestingly MLL-ENL has been shown to associate and co-operate with SWI/SNF complexes to upregulate transcription of the *HOXA7* gene which is important for MLL-ENL-mediated oncogenic transformation [121]. ENL, possibly through its YEATS domain, helps in the recruitment of the SWI/SNF complex to the promoter region of the target genes to activate their transcription [121].

4.3. ELL

MLL-ELL leukemia accounts for approximately 11% of childhood and 9% of adult MLL fusion-mediated AML cases [275], however, is not widely prevalent in ALL. ELL was discovered almost simultaneously by two independent groups as the gene that fused to the N-terminus of MLL in the t(11;19)(q23;p13.1) translocation [26,27]. In 1996, the Conaway group reported the purification of an ~80 kDa “Pol II transcription elongation factor” from rat liver, which was able to stimulate transcription elongation from a DNA template *in vitro* [276]. Sequencing of the protein peptides confirmed its identity as ELL, and established its role as a member of the Pol II elongation complex. Microscopic studies on ELL localization in human, mouse, and rat cell lines identified ELL being restricted to the nucleus, albeit in a diffused pattern. This observation is important since Pol II has a similar expression pattern suggesting a strong correlation between ELL and Pol II functions [277]. Further in-depth studies demonstrated that ELL had the potential to stimulate as well as inhibit Pol II-mediated transcription [278]. Using an *in vitro* assay system, Shilatfard *et al.*, reported that ELL, when added after PIC assembly on a DNA template, stimulated the transcription rate [276]. However, and surprisingly, addition of ELL prior to assembly of the PIC on the DNA template led to a significant decrease in the number of transcripts produced, without affecting the rate of transcription. Subsequent experiments showed that ELL was capable of binding both Pol II and TFIIB at the promoter, and possessed the ability to bind Pol II strongly, thus preventing its entry into the functional PIC.

Contrary to earlier speculations, ELL alone may not be able to stimulate transcription and may possess several strongly interacting partners affecting its overall elongation efficiency. Towards identifying factors that would enhance ELL elongation activity, Shilatfard *et al.*, used a combination of several chromatographic and fractionation steps to isolate ELL from rat livers and was found to exist in a large complex, approximately 210 kDa in size [279]. This ELL complex was shown to contain at least three other proteins, namely ELL Associated Proteins 20, 30 and 45 (EAP20, EAP30 and EAP45) based on the molecular weight at which they migrated on SDS-PAGE. Interestingly, this ELL complex was unable to inhibit transcription *in vitro* irrespective of the order in which it was added to the reaction [280]. This observation paved the way for understanding the mechanism of possible deregulation of transcription occurring in MLL-ELL leukemia as these EAPs either functionally rendered the ELL N-terminus mediated transcriptional inhibition defunct, or physically interacted with, and prevented the association of ELL with Pol II at the promoter. Consistent with these observations, within an interval of two years, two ELL-associated factors were identified and reported [281,282]. Designated as ELL Associated Factors 1 and 2 (EAF1 and EAF2), both were found to interact with the N-terminus of ELL (Fig. 2D). EAF1, additionally, was also seen to interact with the extreme C-terminus of ELL (Fig. 2D). Microscopic studies revealed that ELL-EAF1 existed as speckles within the nucleus, separate from MLL and other elongation factors such as CBP, AFF1, and phosphorylated Pol II [281]. This observation suggested a different role of this complex separate from transcription elongation.

Since ELL was first identified as an MLL fusion partner, it is also

interesting to note that the minimal C-terminal region of ELL interacting with EAF1 was shown to be capable of transforming and immortalizing murine hematopoietic progenitor cells [283]. Strikingly, even though ELL and EAF1 were found to interact strongly (upto ~1.2 M NaCl), the presence of MLL-ELL could delocalize EAF1 from its characteristic speckled foci that has been observed within mammalian cells [281]. Further studies by the same group were able to determine that both ELL and EAF1 are mostly localised to Cajal bodies inside the nucleus [284]. The role of Cajal bodies in the biogenesis of snRNPs has been extensively studied in the last two decades (summarized in [285]). Cajal bodies have been found to actively participate in snRNA transcription, processing and quality control. Given the exclusive cellular localization of ELL and EAF1 to these nuclear organelles, it suggested a possible role of these two proteins in snRNA biogenesis. Indeed, Smith *et al.*, used affinity chromatography followed by multidimensional protein identification technology (MudPIT) to identify proteins interacting with the C-terminus of ELL in *Drosophila* S2 cells. Along with other known interactors of ELL, the group identified two unique proteins that shared homology with human proteins: ICE1 (interacts with C-terminus of ELL 1, similar to human KIAA0947) and ICE2 (similar to human NARG2) [124]. EAF1 was also shown to be a part of this purified protein complex that excluded out other ELL interactors. Knockdown of ELL and ICE1 decreased snRNA levels. Further ChIP analyses with ICE1 and ICE2 showed the specific recruitment of this minimal complex containing ELL, EAF1, ICE1 and ICE2 to snRNA genes. This complex, named the Little Elongation Complex (LEC), subsequently was reported to contain ICE1 as the scaffold and ELL as the Pol II elongation factor. The functions of the other components, ICE2 and EAF1/2 remained largely unknown, suggesting a role of these proteins in other cellular processes [286]. Whilst studying the role of EAF1 and EAF2 by transcription assays, it has been discovered that both these proteins are able to stimulate the rate of transcription elongation *in vitro* by Pol II only when ELL is present [287]. However, both EAF1 and EAF2 were reported to be distinct from EAP20, EAP30 and EAP45. The identities and functions of these proteins, which are able to overcome ELL-mediated transcriptional inhibition by limiting Pol II recruitment, are a subject of speculation till date.

Though P-TEFb has been proposed to be the complex required for Pol II to transition into productive elongation from its stalled state [288,289], Smith *et al.*, proposed that this phenomenon required the presence of ELL as well [290]. Using RNAi, this group showed that knockdown of both ELL and EAF1 in *Drosophila* led to a loss of viability in the progeny. Additionally, ELL depletion was found to be essential for target gene expression, as well as Ser-2-phosphorylation on the CTD of Pol II. It was becoming increasingly clear that ELL required the presence of other partner proteins for performing its functions as a Pol II elongation factor. With the objective of addressing this question, Lin *et al.*, purified ELL containing complexes from mammalian cells [47]. Identification of the ELL associated proteins revealed not only the presence of known interactors like EAF1 and EAF2, but a host of other proteins involved in the transcription process. These include AFF1, AFF4, AF9, ENL and also the P-TEFb complex (Fig. 2D). However, a detailed mechanistic study carried out by our group and others showed that the MLL fusion partner proteins do not necessarily exist in a single large macromolecular complex [96,97]. In contrast, these studies demonstrated the presence of several distinct complexes with common subunits. For example, the ELL containing complex contained EAF1, EAF2, P-TEFb and AFF1, but lacked AF9.

ELL was also reported to interact with the N-terminus of mineralocorticoid receptor in human cells and increase its transcriptional activity [126]. Interestingly, glucocorticoid receptor activity was found to be repressed strongly by the presence of ELL. Whilst performing mechanistic studies on the role of the different domains of ELL in transcription, Zhou and colleagues found that a C-terminal fragment of ELL had the potential to stimulate transcription of a synthetic reporter gene to levels greater than the full-length protein itself [291].

Furthermore, they identified that ELL acted similar to a transcription factor for the *TSP1* gene by binding to a specific region at its promoter and regulating its mRNA expression [291]. An interesting study published in 2010 showed ELL binding to HIF1- α at both hypoxic and normoxic conditions [127]. Contrary to its role as a transcription activator, ELL in this case, was found to decrease HIF1- α protein levels. The authors, however, could not provide a plausible mechanism behind this behaviour of ELL [127]. Along similar lines, a study showed that ELL may act as a tumor suppressor by acting as an E3 ubiquitin ligase for the proto-oncogene c-Myc [128]. By virtue of this newly discovered property, ELL was reported to inhibit c-Myc dependent cell proliferation and also suppress HCT116 cell-mediated xenograft tumor growth in nude mice.

The possibility of ELL having a role in DNA damage was first reported by Zhang et al. who showed that ELL and E2F1 work in a tightly regulated “loop” upon genotoxic stress [129]. ELL transcription, and thus its protein levels were found to increase upon exposure to genotoxic reagents in an E2F1 dependent manner. E2F1 activity, in turn, was inhibited by ELL via the recruitment of HDAC1 leading to an anti-apoptotic response (Fig. 2D). In transcription coupled DNA repair as well, ELL plays an important role as a member of the LEC [130]. In UV-induced DNA damage, ELL interacts directly with CDK7 of the TFIIH complex and results in the recruitment of the LEC to DNA damage sites. It does not directly take part in the repair process *per se*, but regulates the resumption of transcription after completion of repair of damaged DNA. Therefore, human ELL participates in multiple physiological processes through regulation of cellular events including transcription and DNA repair.

4.4. AF10

The AF10 gene encodes a 1085 amino acid protein, part of which was found fused to MLL in a t(10;11)(p12;q23) translocation first observed in three patients at St. Bartholomew's Hospital Medical College, London. The breakpoints in AF10 involve its zinc finger domain which is replaced by a leucine zipper region and thus lead to the formation of several heterodimers that deregulate target gene expression [292,293] (Fig. 2E). AF10 was found to contain an N-terminal LAP/PHD finger motif adjacent to a Cysteine rich region (CRR) which had the potential to assemble AF10 into stable tetrameric forms *in vitro* (Fig. 2E). The extended LAP/PHD motif mediates protein-protein interactions and presence of an AT-hook enables AF10 binding to cruciform DNA, thus possibly playing a role in chromatin structure and transcription regulation [294]. The first report describing a function of this gene was in 2001, wherein, it was reported that a mutation in the *Drosophila* homolog of this gene, named *Dalf*, affected expression pattern of the *EVE* gene, leading to impaired growth during embryogenesis [295]. Further report showed that *Drosophila* AF10 (dAF10) was able to interact directly with heterochromatin protein 1 (HP1) which led to speculation of role of dAF10 in stabilizing heterochromatin complex formation or acting as a “scaffold” for recruitment of other silencing proteins [141].

AF10 may also act as a transcription factor by virtue of its interaction with the synovial sarcoma associated protein SYT [142]. SYT does not possess DNA binding motifs, but may however act as a transcription co-activator. It probably does so by virtue of its interaction with AF10, which possesses the ability to bind to DNA. This may lead to an AF10-mediated SYT effect on several target genes leading to deregulation of transcription. The FLRG protein, which acts as a transcriptional co-activator, has also been reported to interact with AF10 for transcriptional activation as observed from a reporter gene [143]. Another possible mechanism for AF10-mediated transcription regulation emerged from a study from Brian Young's group who demonstrated an interaction of AF10 with GAS41 (Fig. 2E). GAS41 interacts with INI1, the human homolog of the yeast SNF5 protein in the SWI/SNF remodeling complex [144]. Interestingly, GAS41 is homologous to

yeast ANC1 sharing similarity with human AF9 and ENL. Additionally, the presence of INI1 in AF10 immunoprecipitates suggested that AF10, GAS41, and INI1 may exist in a larger multiprotein complex that act together for transcriptional regulation.

The role of the DOT1L protein has been extensively studied in the context of leukemia [296,297]. In fact, DOT1L inhibitors are currently in clinical trials as potential drug candidates for leukemia treatment [298]. Along with MLL, AF10 is also known to fuse to the CALM protein, and this fusion leads to leukemogenic transformation by DOT1L mediated H3K79 methylation and upregulation of *HOXA5* [299,300]. The MLL-AF10 fusion protein also can recruit DOT1L to the *HOX* cluster of genes via its LZ motif (Fig. 2E). Mohan *et al.*, identified a DOTL-containing complex consisting of AF10, AF17, ENL, AF9, Skp1, TRRAP and β -catenin [107]. Because of the presence of β -catenin, the group hypothesized that the Wnt pathway could be regulated by this complex. As expected, H3K79 methylation was affected with DOT1L inhibition leading to the misregulation in expression of Wnt-target genes. A subsequent study showed an essential role of DOT1L-AF10 in activating Wnt target genes in colorectal cancer [301]. Consequently, depleting DOT1L-AF10 led to significant reduction in Wnt target genes and also checked the proliferation of intestinal epithelial cells [301]. Avgousti *et al.*, have demonstrated that the PH domain of AF10 (Fig. 2E), enables it to bind to chromatin at promoters overlapping with histone H3 Lysine 4 methylation signatures. They, thus, speculated that via this binding, several effector proteins may be recruited for regulating gene expression [302]. AF10, as a DOT1L cofactor, regulates DOT1L mediated H3K79 methylation, and Pol II-mediated transcription [303].

AF10 also has a PHD finger-Zn knuckle-PHD finger region (PZP), which acts as a reader for unmodified H3K27 [304] (Fig. 2E). This recognition serves as a pre-requisite for DOT1L mediated H3K79 methylation. Indeed, modifications at the H3K27 site prevent H3K79 methylation events [304]. H3K79 methylation also requires a pre-ubiquitination event at Histone H2B [305,306]. DOT1L-AF10 leads to an increase in H3K79 methylation and a decrease in H2BUB and thus slows down Pol II, causing it to pause [140]. The authors thus proposed that since H2BUB is associated with facilitation of transcription, DOT1L-AF10 functioned as a regulatory mechanism to control the expression of highly expressing genes.

4.5. AF17

AF17 was named so as it was first observed in t(11;17) (q23;21) translocation in two acute myeloid leukemia patients. Containing several zinc fingers and a leucine zipper dimerization motif (Fig. 2F), which is retained in the MLL-AF17 protein chimera, these structural features have been presumed to cause the protein to behave as a transcription factor and also take part in protein-protein interactions [307]. These sequences have also been observed in AF10 and BR140, and a *C. elegans* gene named *CEZF*. What is even more interesting to note is that AF10 and AF17 bear 93% similarity as per their amino acid sequence, whereas BR140 bears resemblance to TAF1 of the TFIID complex. Thus, it is not surprising to note that like AF10, a point mutation in the *Drosophila* homolog of AF17 causes impaired growth, as has been mentioned previously. Additionally, like AF10, AF17 has also been reported to be a part of the Wnt signalling pathway as a target of the β -catenin-Tcf/LEF complex [308]. Over-expression of AF17 enhanced colony forming potential in NIH-3 T3 cells, whereas, suppression of AF17 caused SW480 cells to be arrested at the G₂-M transition phase. In contrast, over-expression of AF17 in DLD1 colorectal adenocarcinoma cells reduced the G₂-M transition time and accelerated the cell cycle. The authors correlated these findings and implicated them in colorectal cancer progression because of the high expression of both AF17 and β -catenin.

The VHL protein, which acts as a major tumor suppressor, has been reported to bind to Jade1, thus stabilizing it and enabling its action as a potential tumor suppressor in renal cancer [309]. Incidentally Jade1

possesses an extended PHD domain, similar to AF10 and AF17, leading to the speculation that VHL may also bind to and affect the expression, and thus functions, of these two transcription factors. Perhaps the most detailed studies on AF17 functioning deal with the regulation of expression of the epithelial sodium (Na^+) channel- α (ENaC- α). It has been shown that the Dot1a-AF9 complex represses the ENaC gene in mice [114]. Detailed studies into how this is achieved were performed by the Zhang group at the University of Texas Health Centre [118]. In their studies, they found that AF17 competes with AF9 in binding to Dot1a, leading to a decrease in H3K79 methylation due to Dot1a export from the nucleus to the cytoplasm, resulting in accentuated expression of ENaC- α [118,145,310]. Further studies from the same group demonstrated that AF17 was crucial for the maintenance of blood pressure and sodium levels, as AF17 deficient mice were found to have increased urine volumes, higher sodium excretion and thus lower blood pressure levels [310]. The authors attributed this largely to decreased ENaC- α expression because of Dot1a-mediated H3K79 dimethylation, suggesting a role of AF17 in maintaining normal blood pressure levels [310]. As mentioned previously, Mohan and colleagues, whilst trying to identify the members of a complex with DOT1 as its key component, observed the complex to contain AF17, in addition to other proteins [107]. AF17 was also present in eluates of AF10, AF9 and ENL. Unlike AF10, AF17 depletion did not affect DOT1-mediated H3K79 methylation [304]. Silencing of the *Drosophila* analogue of AF10 and AF17, named *Alhambra*, led to a reduction in the *Senseless* gene expression in the wing imaginal discs [107]. However, further studies would be required for detailed understanding the role of this protein in relation to the Dot1 complex, or independent of it in transcriptional regulation.

4.6. AF6

Like most leukemic fusion partners, AF6, also known as Afadin or MLLT4, was first identified as an effector of Ras [311]. It contains kinase and myosin-like-cargo binding domains and is able to bind H-Ras, K-Ras, R-Ras2 or Rap in a GTP dependent manner [311] (Fig. 2G). AF6 exists in two forms, the longer isoform is expressed in epithelial cells and the shorter isoform is expressed exclusively in the brain [312]. In one of the first studies on AF6 function, it was observed that Ras activation inhibits AF6 interaction with the ZO-1 protein and is responsible for maintaining tight junctions in cells [146,147,313]. Thus, AF6 could possibly be a part of the complex regulating cell to cell contacts and Ras signalling that may serve to enable cell migration. AF6 stability at these junctions is, in part, effected by the Fam deubiquitinating enzyme, which also accumulates at cell-cell adhesion sites [149]. Fam has also been reported to prevent the ubiquitination of β -catenin [314]. The combined effect of Fam on both these two proteins maintains cleavage rate and cell adhesion in embryos [150]. Additionally, the AF6-interacting protein Ponsin was also found localised to cell-cell junctions, wherein, it interacted with vinculin to directly maintain cell-matrix adhesion [151].

It was becoming increasingly clear that AF6 may serve as a “recruiter” for proteins responsible for maintaining cell-cell adherent junctions through effecting the recruitment of the Nectin protein, which displayed the propensity to effect cadherin based adhesion [152,315]. On similar lines, the junction adhesion molecule (JAM), an integral membrane protein, also required the interaction between AF6 as well as ZO-1 to be localised to intracellular junctions [148]. Indeed, null mutations of AF6 in mouse embryos caused death within 10 days of coitus, possibly due to improper placenta formation [316]. In the context of neural cell development, the localization of atypical PKC isotype-specific-interacting protein (ASIP) at the luminal surface of neuroepithelium is dependent on β -catenin and AF6 [317]. AF6 not only has been reported to be a “recruiter” of proteins at cell-cell adhesions, but also has been shown to have roles in controlling integrin-mediated cell-cell adhesion via interaction with SPA-1, a Rap1 GTPase-activating protein [155]. AF6 accumulation in the nucleus coincided with histone H3

phosphorylation at serine 10 (H3S10p), along with dendritic remodeling of the spine in rat embryonic cells [318]. Within mouse testis, AF6 is observed to express in supporting as well as in steroidogenic cells, where its function is regulated by deubiquitylation mediated stabilization by the Usp9x enzyme, in a mechanism similar to Fam as described above [160]. Whilst studying mesenchymal-epithelial transition in mouse kidneys, Kanazawa *et al.*, reported that AF6 is required for hepatocyte nuclear factor 4 α (Hnf4a)-induced epithelial and reduced mesenchymal gene expression. Within the set of target genes, the expression of AF6 was found to be significantly increased, establishing it as an epithelial marker [319]. AF6 is also required for actin stress fibres in rat podocyte cells through its activation of the Rho family small GTPase RhoA via the RhoA/ROCK signalling pathway, resulting in myosin phosphorylation [320]. Latest reports on AF6 show that other than maintaining cell polarity and adhesion, AF6 is also implicated in insulin sensitivity by regulating the IRS1/AKT kinase-mediated insulin signalling pathway via its modulation of SHP2 tyrosine phosphatase activity [161]. AF6 was found to be higher in diabetic mice, and its knockdown ablated insulin resistance, showing a role of AF6 in metabolic regulation. Thus, until now, unlike the role of other MLL fusion partners primarily in transcriptional regulation, the role of AF6 in cellular processes has primarily been described independent of transcriptional regulation.

5. Lesser known MLL fusion proteins in leukemogenic transformation

Although this review is primarily focused on those MLL fusion proteins which by themselves act as transcription factors, we must acknowledge some of the other lesser known MLL fusion proteins for which no direct role in transcription has been described, but can still give rise to leukemic phenotype upon being fused with the MLL protein. These are mostly cytoplasmic proteins which can dimerize or oligomerize through their self-association motifs (e.g. coiled coil region, leucine zipper motif) and thus enable the resultant MLL-fusion protein to self-associate, a property that has been shown by many groups to be an alternative mechanism to generate leukemia [293,307,321]. AF10, AF17, AF3p21, LCX, SEPTING, EEN, and among many others retain their self-association motifs in the resultant fusion proteins [307,321–325]. Chi Wai So *et al.*, have elegantly shown that two cytoplasmic proteins, AF1p/Eps15 and GAS7, when fused with MLL, give rise to acute leukemia [326,327]. They reported that the dimerization modules of those two proteins, when fused to MLL, are sufficient for leukemogenic transformation. Mary Ellen Martin *et al.*, have also shown that induced dimerization of truncated MLL by fusing it with FKBP12 gives rise to leukemic phenotype [328]. Although exact mechanism of this dimerization induced leukemia is still not clear, the authors suggest that dimerization of MLL gives a platform for other transcriptional activators to assemble and up-regulate leukemia associated genes.

6. P-TEFb, an integral component of the SEC

It is widely believed that misregulation of transcription caused by the aberrant activity by MLL fusion proteins, is a leading cause of leukemia development. Notably, the “elongation” of Pol II is mediated by two post translational phosphorylation events, initially at the fifth, and subsequently the second Serine residue of its highly disordered CTD repeat region. Whilst the first phosphorylation at the fifth Serine (Ser–5P) is mediated by the CDK7 of the TFIIH complex, the succeeding Serine2 phosphorylation is responsible for driving Pol II into productive elongation as well as termination. This critical modification is catalyzed by the P-TEFb. Though either subunit of P-TEFb complex does not fuse to MLL, its activity is believed to be critical for development of MLL and other diseases. This has also led to the development and usage of several P-TEFb inhibitors for leukemia treatment [329,330]. Keeping all these critical points in mind, a brief note on transcription regulation by

P-TEFb is also being included in this review.

The saga of P-TEFb began in 1992 with the conjecture by Marshall and Price about the existence of a positive transcription elongation factor (P-TEF), which would not be associated with the transcription initiation complex, but be responsible for the transition of Pol II from early elongation to the productive elongation phase [331]. Furthermore, the same group reported in subsequent studies that P-TEFb comprises of a heterodimer of two unknown proteins, which phosphorylates the CTD of Pol II in a manner distinct from TFIID and is sensitive to DRB treatment [288]. Meanwhile, P-TEFb was reported as an activator of the HIV Tat protein [332]. David Price's group reported that the kinase, erstwhile known as PITALRE, was actually the catalytically active smaller subunit of P-TEFb and that it was capable of phosphorylating the Pol II CTD [172]. Further, and almost simultaneously, two groups were able to identify the larger subunit of PTEF-b as CyclinT and the smaller kinase as CDK9 [333–335]. In the first of many studies aimed at understanding the mechanism of action of P-TEFb in Pol II elongation, it was observed that this heterodimer counteracts the negative effect of DSIF, a protein reported earlier to repress transcription [176]. P-TEFb could phosphorylate DSIF and some other unknown factors, leading to de-repression of the early elongation phase of transcription. In the case of HIV infection, as shown by Zhou *et al.*, the Tat protein enables the CDK9 to phosphorylate the Ser-5 as well as the Ser-2 of the CTD, thus nullifying the requirement of two separate kinase complexes for TAR synthesis [178]. Because most of the initial reports on P-TEFb-mediated transcription elongation was restricted to the mechanisms of HIV infection, Lis *et al.*, took a different approach for studying P-TEFb effect on other native genes. This group intended to study the mechanism of P-TEFb action on heat shock loci in *Drosophila* cells [336]. This group reported that P-TEFb was recruited rapidly to heat shock responsive genes upon stimulation and was not initially associated with the proximally paused Pol II. Only upon heat shock, P-TEFb swiftly associated with Pol II, and this association enabled its escape into productive elongation. CDK9 autophosphorylation was found to be critical for its nuclear localization and formation of the P-TEFb complex with CyclinT1 [337]. Subsequent studies, also in *Drosophila* cells, showed that CDK9 depletion resulted in decreased level of both H3K4 mono and trimethylation that are associated with active enhancers and gene promoters respectively [179]. Additionally, H3K36 dimethylation as well as recruitment of the Pol II elongation factor ELL was reduced, suggesting that P-TEFb-mediated CTD phosphorylation may be necessary for these events to occur [179]. Majority of cellular P-TEFb complex remains as an inactive 7SK snRNP complex comprised of HEXIM1, LARP7, MePCE and 7SK snRNA [182,183]. In another study, the Zhou group also showed that HIV-1 Tat prevented the formation of the 7SK snRNP as well as was able to directly prevent HEXIM1-CyclinT1 interaction, thus enabling activation of the HIV-1 promoter [338]. Several studies have reported that upon stress or HIV infection, the AFF1-P-TEFb was released from the 7SK snRNP and transferred to Tat and BRD4 for transcriptional activation of target genes [100].

Initial studies suggested that the bromodomain protein BRD4 could recruit P-TEFb to the promoter, and enable it to contact the Mediator in a manner similar to HIV1-mediated P-TEFb recruitment [189]. In the NF- κ B pathway, RelA phosphorylated at Ser276 also recruits P-TEFb for the activation of NF- κ B dependent genes such as *IL-8* and *Gro- β* [339]. The MEK-ERK pathway increases the nuclear pool of CDK9, enabling the formation of more P-TEFb, and facilitates its recruitment to immediate early genes in a thyroid-releasing hormone-dependent manner [340]. A recent study has also demonstrated that the Mediator subunit MED23 could interact with CDK9 and recruit P-TEFb for maintaining a basal level of transcription [91]. Our recent study also showed a role of the erstwhile known transcription repressor ZMYND8 in recruiting P-TEFb via CyclinT1, through its ability to dimerize and thus switch between transcription repression and activation potential [87]. The recruitment of P-TEFb appears to be a complex and tightly regulated phenomenon as described by Lu *et al.*, [94]. Their study proposed that

multiple P-TEFbs exist for the release of paused Pol II. The first P-TEFb is recruited via BRD4-MED1-MED23-TatSF1 for phosphorylating the Spt5 subunit of DSIF. The second P-TEFb is recruited by SEC-containing AFF1 or AFF4 for the phosphorylation of NELF-E via MED26. The third P-TEFb recruitment is done by AFF1-AF9-SEC via PAF1c for the phosphorylation of NELF-A and NELF-E. The fourth and last P-TEFb is recruited to PAF1c via AFF4-AF9-SEC to phosphorylate the Pol II CTD. The recruitment of P-TEFb is thus a concerted mechanism, involving various different factors in different cellular complexes, which possess the ability to function independently, and yet play a critical role in overall functional activation of target genes.

7. Recruitment of SEC at the promoter proximal region for transcriptional regulation

The first evidence of activator-dependent P-TEFb complex recruitment was discussed in the seminal work from the Young lab showing Myc-dependent P-TEFb complex recruitment at the promoter proximal region for regulation of release of paused Pol II into productive elongation [194]. Subsequent to the discovery of SEC, MED26-dependent recruitment of the SEC was discussed by the Conaway lab. In their study, the authors showed that the N-terminus of MED26 interacts with components of both TFIID and the EAF1 component of SEC [131]. Based on their experimental data, the authors postulated a dual role of MED26 in connecting events at promoter proximal region. The initial event relates to the interaction of MED26 with TFIID leading to the recruitment of the Mediator at the promoter region. The authors further hypothesized that upon binding at the promoter region, a conformational switch of the N-terminus leads to its dissociation from TFIID and thus becomes free to engage with EAF1 for SEC recruitment at the promoter proximal region. Following this study, the YEATS domain of AF9 was shown to interact with PAF1 complex, thus suggesting PAF1 complex-dependent SEC recruitment for transcriptional regulation of target genes [112]. However, direct evidence for the role of key initiation factors in regulation of the recruitment of SEC components was not provided until our study showed a comprehensive role of the human TFIID in regulation of SEC recruitment for helping in the transition from initiation to elongation of Pol II during transcriptional activation [95]. As discussed below, this study changes our overall understanding of the role of TFIID beyond just controlling events at promoter region only.

7.1. TFIID as novel regulator of SEC recruitment at promoter proximal region

TFIID complex is a part of general transcription factors (GTFs) and is the first complex to get recruited at the promoter region for PIC assembly and to recruit Pol II. TBP component of TFIID binds to the TATA box and recruits the whole complex in an activator dependent manner. In majority of the cases, TFIID functions have been studied with respect to PIC formation at the promoter proximal region. However, several recent cryo-EM structures of transcriptionally-engaged TFIID showed the presence of several TAF subunits well within the coding region from the transcription start site (TSS) [341,342]. These observations indicate an additional role of TFIID in regulating events close to the promoter proximal region. Further, consistent with this observation, results from the Zeitlinger lab also showed a change in position of TBP subunit of TFIID within the coding region upon transcriptional activation [343].

Perhaps, the first indication of a role of TFIID in regulation of functions of SEC components was through our initial studies, wherein, we showed a SEC subunit-specific interaction of TFIID within mammalian cells [96]. Specific observation that only the AF9 component interacted with TFIID components, further suggested specificity of this interaction and potential role(s) of these interactions in transcriptional regulation at the promoter proximal region. Further following up these observations, we have recently shown a novel role of human TFIID in

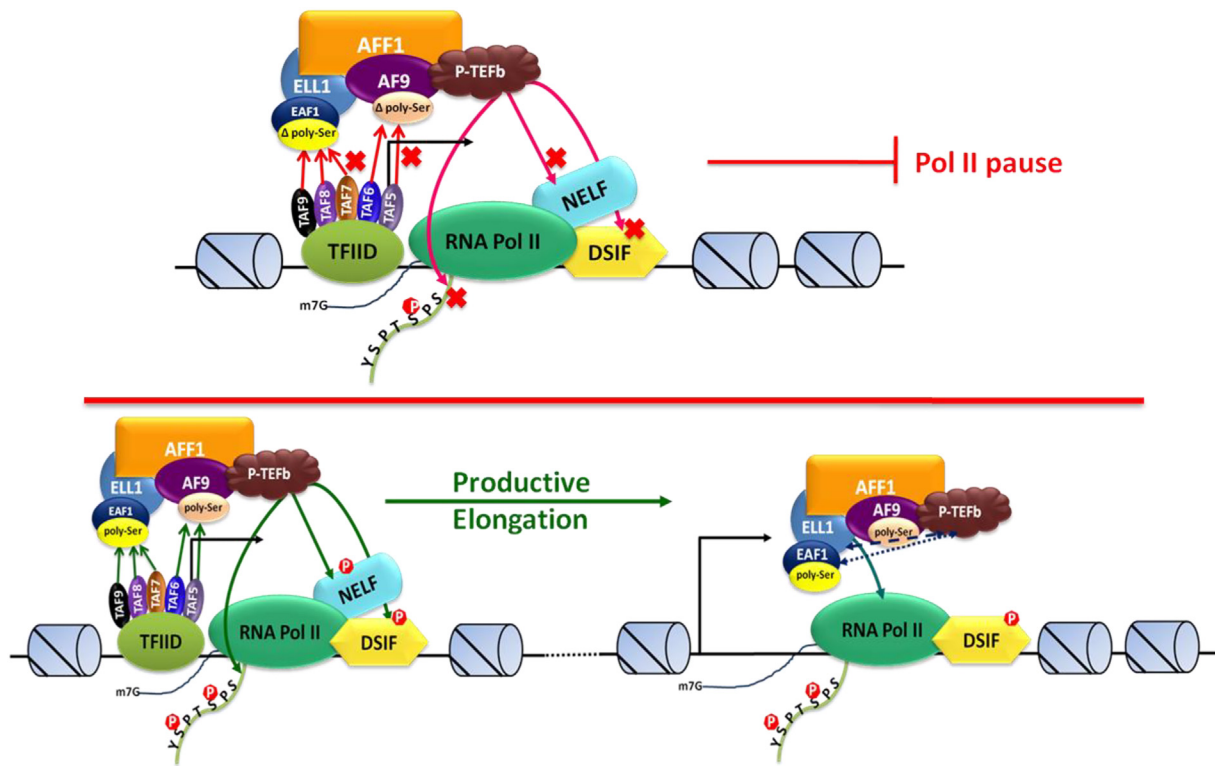


Fig. 3. Cartoon diagram showing the role of TFIID in bridging the gap between transcription initiation and elongation.

A) Absence of poly-Serine regions in AF9 and EAF1 components of the SEC prevent its interaction with TFIID and impair its recruitment to the paused Pol II (top panel). Absence of P-TEFb further leads to impaired phosphorylation of DSIF, NELF as well as CTD Ser2 and thus increases Pol II pausing at promoter proximal region.

B) Interaction of AF9 and EAF1 (via their poly-Serine rich domain) with specific TFIID subunits enables the recruitment of P-TEFb (through SEC) to the paused Pol II. The recruited P-TEFb is now able to phosphorylate DSIF, NELF and Serine2 of the Pol II CTD (bottom panel). P-TEFb mediated phosphorylation, combined with the elongation activity of different SEC components causes Pol II to progress into the productive elongation phase.

recruiting SEC components at the promoter proximal region for transcriptional activation (Fig. 3) [95]. Our initial report of AF9 interaction with TFIID subunits were further supported by detailed immunoprecipitation analyses that clearly showed varying degree of interaction of TFIID with different SEC components. Interestingly, among all the SEC subunits that were tested, only AF9 and EAF1 showed interaction with all TAF subunits that were tested in our assay, thus indicating the existence of direct interaction between these two SEC subunits and TFIID. Indeed, our *in vitro* interaction assays with purified proteins further supported their direct interactions.

Next, the most interesting observation was that both the SEC subunits i.e. AF9 and EAF1, showed TFIID interactions in a specific domain-dependent manner. Earlier reports suggested the presence of stretch(es) of poly-Serine (poly-Ser, hereafter) residues within several SEC components that act as transcriptional activation domains [282,283]. However, the precise role of these domains was completely unknown. In our quest to obtain detailed mechanistic insights into the overall regulation of TFIID-mediated SEC components' recruitment, we were surprised to identify poly-Ser regions of AF9 and EAF1 that directly interact with TFIID both *in vitro* as well as *in vivo* within mammalian cells (Fig. 3). It was even more fascinating to observe that both these proteins use non-overlapping domains to interact with the SEC and TFIID. These observations led us to hypothesize that the human TFIID can act as a recruiter of the SEC at the promoter proximal region for transcriptional activation.

Using an *in vitro* recruitment assay of SEC components, we have clearly shown TFIID-dependent SEC component recruitment on a DNA template. Using these analyses, we have also clearly shown a role of poly-Ser domain in TFIID-dependent SEC recruitment on template DNA. Identification of specific TAF subunits that interact with specific SEC

components, further reinforces the idea that the overall regulation involves multiple interactions between components of TFIID and SEC that manifests in transcriptional activation (Fig. 3). All these interactions within mammalian cells are indeed MED26-independent since a knockdown of MED26 had little effect on overall TFIID-SEC interaction within mammalian cells.

The effect of overall TFIID-dependent SEC recruitment on expression of native target genes was further substantiated through the use of AF9 knockdown cells. Our results clearly showed that poly-Ser domain-dependent AF9-TFIID interaction is critically important for target gene expression and cell survival. Overall, our study has opened up a completely new paradigm of eukaryotic transcriptional regulation involving TFIID, wherein, this multi-subunit complex also plays additional roles in critical transcriptional regulation beyond just the promoter region (Fig. 3). This interesting study is bound to open up question about the dynamic regulation of these interactions and associated events that lead to the release of TFIID-bound SEC from the promoter proximal region for moving along with elongating Pol II in the coding region.

8. MLL fusion partners in other diseases

Since SEC components are likely to play a universal role in regulating expression of significant number of genes, it is quite conceivable that misregulation of functions of these factors would also be involved in several other diseases. As the role of SEC components in HIV pathogenesis has been well discussed in the earlier section of this review, we shall be focusing on the role of SEC components in cancer and other diseases. Also role of MLL fusions in leukemia will be discussed in detail by other groups in this special edition and is beyond the scope of this review.

The role of AFF1 in Purkinje cell survival and function, possibly through maintenance of insulin-like growth factor 1 has been described by multiple studies [344,345]. Mutations in AFF1 cause region-specific loss of Purkinje cells due to cell death in cerebellum of mouse and results in jerky movements and ataxic gait (robotic mouse). It also causes complete and bilateral cataract [344,346,347]. AFF1 gene polymorphism has been associated with Systemic Lupus Erythematosus (SLE) in many different ethnic groups [348–350]. Further, AFF1 gene polymorphism has also been observed in disorders related to circulating lipid metabolism including coronary artery disease [351].

AF9 is associated with schizophrenia and bipolar disorder [352]. AF9 acts as a proto-oncogene in prostate cancer through the action of miR-564 that acts as a tumor suppressor by targeting AF9 to down-regulate its expression [353]. Missense mutations in the YEATS domain of ENL have been shown to be highly associated with several cases of Wilms tumor [354,355]. Recent reported study has shown a role for ENL self-association-dependent increased SEC recruitment is a key mechanism for mutant ENL-mediated Wilms tumor pathogenesis [355].

Unlike AF9 and ENL, the EAF1 and EAF2 have primarily been described as tumor suppressors. Androgen stimulates the expression of both inducers and nullifiers of cell proliferation in the prostate to maintain a transcriptional homeostasis. EAF2 falls under the category of androgen-induced tumor suppressor. EAF2 overexpression has been shown to induce apoptosis in various cancerous cell lines as well as in xenografts [356]. This apoptosis induction requires new protein synthesis which may be a result of their increased transcriptional output caused by EAF2. Interestingly, in prostate tumor patient samples, EAF2 has been shown to be downregulated either genetically or epigenetically resulting in progression of prostate cancer [356]. EAF2 knockout mice show high occurrence of prostate intraepithelial neoplasia, lung adenocarcinoma, hepatocellular carcinoma, B-cell lymphoma as well as cardiac cell hypertrophy [357]. All these suggest a role of EAF2 in tumor suppression and maintenance of homeostasis for optimal cell growth [357]. Like EAF2, conditional knock out of EAF1 also gives rise to prostate intraepithelial neoplasia in mice [358]. Loss of EAF2 and PTEN (another tumor suppressor) has been shown to synergistically promote prostate cancer in mouse models, suggesting a functional connection of EAF2 with other tumor suppressor proteins [359]. Although the mechanistic understanding is not clear, EAF2 has also been shown to take part in apoptosis after UV mediated damage by upregulating the expression of PUMA and NOXA to reduce the chances of UV damage-induced carcinogenesis [360].

As AF6 is mostly associated with cell-cell adhesion junctions, its loss is associated with cell migration. Indeed, Fournier and colleagues have shown that in three different breast cancer cell lines, knockdown of AF6 decreased cell polarity while increasing cell proliferation and migration, possibly due to the negative regulation of the Src kinase and the RAS/MAPK pathways [361]. Low level of expression of AF6 in human pancreatic cancer cells is also a major cause for their proliferation and metastasis potential. In this cancer, decrease of AF6 contributed to an increased expression of *Snail* and *N-cadherin* genes. AF6 was also shown to cause decreased metastasis by inhibiting Dvl2/FOXO1-dependent *Snail* transcription [362]. A screen to identify the most common genes associated with glioblastoma multiforme (GBM) implicated AF6 as one of thirteen critical genes for GBM prognosis. However, the study did not provide any mechanism behind this observation [363].

Lin and colleagues initially described a role of AF17 in colorectal carcinogenesis [364]. Their study found that AF17 is a downstream target of the β -catenin/Tcf pathway. Like β -catenin, AF17 expression was also found to be increased in this cancer. Between the HR+ /HER2+ and the HR+ /HER2- groups of patients in breast cancer, the expression of AF17 was found to vary markedly. AF17 expressed 28% more in HR+ /HER2+ phenotype, relating to a higher hazard ratio and possibly, lower survival [365].

CDK9 involvement in cancer prognosis relates to its affinity to two key proteins, BRD4 and Myc. Myc addiction in different cancers has

been very well characterized, and thus, it is not surprising that several CDK9 antagonists have been designed for the therapy of Myc driven cancers [366]. Higher CDK9 expression coincides with greater differentiation of neuroblastomas, and CDK9 itself has been shown to influence neural differentiation [367]. In TNF- α mediated MMP9 activation, which leads to increased metastasis, P-TEFb recruitment to MMP9 is stimulated by TNF- α for its increased expression [223]. Overexpression of CyclinT1, leading to the formation of P-TEFb is able to induce transformation of NIH-3 T3 cells *in vitro* and also enhance oncogenic potential in mice possibly by acting through the Rb/E2F1 pathway *in vivo* [368]. As P-TEFb is also involved in regulation of the Wnt pathway, it is not surprising that it plays in integral part in Wnt-mediated proliferation of head and neck cancer cells [369]. In androgen-dependent castration-resistant prostate cancer, a set of enhancer associated RNAs bind to and activate P-TEFb and thus causing enhanced target gene transcription [370]. Other than its role in the development of leukemias and lymphomas, CDK9 has also been implicated in the development of several cancers of the cervix, lung, ovaries, colon and melanoma [371]. Thus, it is not surprising that several CDK9 inhibitors are now being designed for therapeutic intervention in cancers and other diseases.

9. Concluding remarks

Work from last several years have made significant progress in the overall understanding of key MLL fusion partners and thus, SEC-mediated transcriptional regulation and its potential involvement in several human diseases. However, the overall concept of functional regulation of these key proteins through an unified mechanism by residing in a giant megadalton SEC does not well correlate with results using mouse models and patient samples. Gradually, it is becoming clear that the overall impact of MLL fusion partner-mediated transcriptional regulation can as well be significantly regulated through an SEC-independent manner. Therefore, multiple mechanisms could play a role in overall transcriptional (mis)regulation by MLL fusion proteins and thus contribute to generation of specific types of leukemia as observed for different fusion partners. Therefore, significant future studies would be directed towards the detailed mechanistic understanding of the role of each of the fusion partners in diverse transcriptional regulation in SEC-dependent or -independent manner and their implication in leukemic pathogenesis by corresponding MLL fusion proteins. Further, the novel concept of the role of liquid-liquid demixing-mediated phase separation of transcription factors and its implication in transcriptional regulation is of immense interest in SEC-mediated transcriptional regulation as well. Of note, majority of SEC components show presence of intrinsically disordered regions that could play important roles in phase separation and transcriptional regulation. Significant future studies would be directed towards addressing these key questions and understanding the role of MLL fusion proteins in overall (mis)regulation of these processes and their implication in leukemic pathogenesis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by Department of Science and Technology (DST), Govt. of India extramural grant EMR/2016/001593 awarded to DB. SB is a recipient of UGC Senior Research Fellowship and AN is a recipient of CSIR Junior Research Fellowship. We also deeply acknowledge assistance from other members of Biswas lab for their critical comments on this manuscript.

References

- [1] S. Stass, J. Mirro, S. Melvin, C.H. Pui, S.B. Murphy, D. Williams, Lineage switch in acute leukemia, *Blood* 64 (1984) 701–706.
- [2] J. Mirro, G.R. Kitchingman, D.L. Williams, S.B. Murphy, T.F. Zipf, S.A. Stass, Mixed lineage leukemia: the implications for hematopoietic differentiation, *Blood* 68 (1986) 597–599.
- [3] J.J. Yunis, C.D. Bloomfield, K. Ensrud, All patients with acute nonlymphocytic leukemia may have a chromosomal defect, *N. Engl. J. Med.* 305 (1981) 135–139.
- [4] M. Djabali, L. Selleri, P. Parry, M. Bower, B.D. Young, G.A. Evans, A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias, *Nat. Genet.* (2) (1992) 113–118.
- [5] C.M. Macrini, M.S. Pombo-de-Oliveira, A.M. Ford, G. Alves, MLL AT-hook sequence is strongly conserved in infant acute leukemia with or without MLL gene rearrangement, *Leukemia* 17 (2003) 1432–1433.
- [6] T. Cierpicki, L.E. Risner, J. Grembecka, S.M. Lukasik, R. Popovic, M. Omonkowska, D.D. Shultz, N.J. Zeleznik-Le, J.H. Bushweller, Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia, *Nat. Struct. Mol. Biol.* 17 (2010) 62–68.
- [7] M. Ali, R.A. Hom, W. Blakeslee, L. Ikenouye, T.G. Kutateladze, Diverse functions of PHD fingers of the MLL/KMT2 subfamily, *Biochim. Biophys. Acta* 1843 (2014) 366–371.
- [8] M.S. Cosgrove, A. Patel, Mixed lineage leukemia: a structure-function perspective of the MLL1 protein, *FEBS J.* 277 (2010) 1832–1842.
- [9] B. Pless, C. Oehm, S. Knauer, R.H. Stauber, T. Dinger, R. Marschalek, The heterodimerization domains of MLL-FYRN and FYRC are potential target structures in t(4;11) leukemia, *Leukemia* 25 (2011) 663–670.
- [10] J.L. Hess, MLL: a histone methyltransferase disrupted in leukemia, *Trends Mol. Med.* 10 (2004) 500–507.
- [11] T.A. Milne, J. Kim, G.G. Wang, S.C. Stadler, V. Basur, S.J. Whitcomb, Z. Wang, A.J. Ruthenburg, K.S. Elenitoba-Johnson, R.G. Roeder, C.D. Allis, Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis, *Mol. Cell* 38 (2010) 853–863.
- [12] A.G. Muntean, J. Tan, K. Sitwala, Y. Huang, J. Bronstein, J.A. Connelly, V. Basur, K.S. Elenitoba-Johnson, J.L. Hess, The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis, *Cancer Cell* 17 (2010) 609–621.
- [13] T.A. Milne, S.D. Briggs, H.W. Brock, M.E. Martin, D. Gibbs, C.D. Allis, J.L. Hess, MLL targets SET domain methyltransferase activity to Hox gene promoters, *Mol. Cell* 10 (2002) 1107–1117.
- [14] Y. Zhang, A. Mittal, J. Reid, S. Reich, S.J. Gamblin, J.R. Wilson, Evolving catalytic properties of the MLL family SET domain, *Structure* 23 (2015) 1921–1933.
- [15] J.J. Song, R.E. Kingston, WDR5 interacts with mixed lineage leukemia (MLL) protein via the histone H3-binding pocket, *J. Biol. Chem.* 283 (2008) 35258–35264.
- [16] J.J. Hsieh, E.H. Cheng, S.J. Korsmeyer, Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression, *Cell* 115 (2003) 293–303.
- [17] J.J. Hsieh, P. Ernst, H. Erdjument-Bromage, P. Tempst, S.J. Korsmeyer, Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization, *Mol. Cell. Biol.* 23 (2003) 186–194.
- [18] M. Reichel, E. Gillert, I. Breitenlohner, R. Repp, J. Greil, J.D. Beck, G.H. Fey, R. Marschalek, Rapid isolation of chromosomal breakpoints from patients with t(4;11) acute lymphoblastic leukemia: implications for basic and clinical research, *Cancer Res.* 59 (1999) 3357–3362.
- [19] O.A. Bernard, R. Berger, Molecular basis of 11q23 rearrangements in hematopoietic malignant proliferations, *Genes Chromosom. Cancer* 13 (1995) 75–85.
- [20] A.V. Krivtsov, D. Twomey, Z. Feng, M.C. Stubbs, Y. Wang, J. Faber, J.E. Levine, J. Wang, W.C. Hahn, D.G. Gilliland, T.R. Golub, S.A. Armstrong, Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9, *Nature* 442 (2006) 818–822.
- [21] J. Corral, I. Lavenir, H. Impey, A.J. Warren, A. Forster, T.A. Larson, S. Bell, A.N. McKenzie, G. King, T.H. Rabbitts, An MLL-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes, *Cell* 85 (1996) 853–861.
- [22] A. Hagemeijer, K. Hahlen, W. Sizoo, J. Abels, Translocation (9;11)(p21;q23) in three cases of acute monoblastic leukemia, *Cancer Genet. Cytogenet.* 5 (1982) 95–105.
- [23] R. Prasad, Y. Gu, H. Alder, T. Nakamura, O. Canaani, H. Saito, K. Huebner, R.P. Gale, P.C. Nowell, K. Kuriyama, et al., Cloning of the ALL-1 fusion partner, the AF-6 gene, involved in acute myeloid leukemias with the t(6;11) chromosome translocation, *Cancer Res.* 53 (1993) 5624–5628.
- [24] V.A. Alexandrov, V.G. Bespalov, A.S. Petrov, D.N. Troyan, M. Lidaks, Study of post-natal effect of chemopreventive agents on ethylnitrosourea-induced transplacental carcinogenesis in rats. III. Inhibitory action of indomethacin, voltaren, theophylline and epsilon-aminocaproic acid, *Carcinogenesis* 17 (1996) 1935–1939.
- [25] Y. Tsujimoto, J. Yunis, L. Onorato-Showe, J. Erikson, P.C. Nowell, C.M. Croce, Molecular cloning of the chromosomal breakpoint of B-cell lymphomas and leukemias with the t(11;14) chromosome translocation, *Science* 224 (1984) 1403–1406.
- [26] K. Mitani, Y. Kanda, S. Ogawa, T. Tanaka, J. Inazawa, Y. Yazaki, H. Hirai, Cloning of several species of MLL/MEN chimeric cDNAs in myeloid leukemia with t(11;19)(q23;p13.1) translocation, *Blood* 85 (1995) 2017–2024.
- [27] M.J. Thirman, D.A. Levitan, H. Kobayashi, M.C. Simon, J.D. Rowley, Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23;p13.1) in acute myeloid leukemia, *Proc. Natl. Acad. Sci. U. S. A.* 91 (1994) 12110–12114.
- [28] J.E. Rubnitz, B.M. Camitta, H. Mahmoud, S.C. Raimondi, A.J. Carroll, M.J. Borowitz, J.J. Shuster, M.P. Link, D.J. Pullen, J.R. Downing, F.G. Behm, C.H. Pui, Childhood acute lymphoblastic leukemia with the MLL-ENL fusion and t(11;19)(q23;p13.3) translocation, *J. Clin. Oncol.* 17 (1999) 191–196.
- [29] B.B. Zeisig, T. Milne, M.P. Garcia-Cuellar, S. Schreiner, M.E. Martin, U. Fuchs, A. Borkhardt, S.K. Chanda, J. Walker, R. Soden, J.L. Hess, R.K. Slany, Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization, *Mol. Cell. Biol.* 24 (2004) 617–628.
- [30] P.M. Ayton, M.L. Cleary, Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins, *Oncogene* 20 (2001) 5695–5707.
- [31] T. Taki, H. Ohnishi, K. Shinohara, M. Sako, F. Bessho, M. Yanagisawa, Y. Hayashi, AF17q25, a putative septin family gene, fuses the MLL gene in acute myeloid leukemia with t(11;17)(q23;q25), *Cancer Res.* 59 (1999) 4261–4265.
- [32] K. Suzuki, S. Shimizu, N. Nemoto, N. Takei, T. Taki, T. Nagasawa, Identification of a chromosomal breakpoint and detection of a novel form of an MLL-AF17 fusion transcript in acute monocytic leukemia with t(11;17)(q23;q21), *Int. J. Hematol.* 82 (2005) 38–41.
- [33] B.B. Zeisig, S. Schreiner, M.P. Garcia-Cuellar, R.K. Slany, Transcriptional activation is a key function encoded by MLL fusion partners, *Leukemia* 17 (2003) 359–365.
- [34] G. Mann, A. Attarbaschi, M. Schrappe, P. De Lorenzo, C. Peters, I. Hann, G. De Rossi, M. Felice, B. Lausen, T. Leblanc, T. Szczepanski, A. Ferster, G. Janka-Schaub, J. Rubnitz, L.B. Silverman, J. Stary, M. Campbell, C.K. Li, R. Suppliah, A. Biondi, A. Vora, M.G. Valsecchi, R. Pieters, G. Interfant-99 Study, Improved outcome with hematopoietic stem cell transplantation in a poor prognostic subgroup of infants with mixed-lineage-leukemia (MLL)-rearranged acute lymphoblastic leukemia: results from the Interfant-99 study, *Blood* 116 (2010) 2644–2650.
- [35] A.C. Winters, K.M. Bernt, MLL-rearranged Leukemias—an update on science and clinical approaches, *Front. Pediatr.* 5 (2017) 4.
- [36] R. Marschalek, Systematic classification of mixed-lineage leukemia fusion partners predicts additional cancer pathways, *Ann. Lab. Med.* 36 (2016) 85–100.
- [37] C. Meyer, T. Burmeister, D. Groger, G. Tsaur, L. Fechina, A. Renneville, R. Sutton, N.C. Venn, M. Emerenciano, M.S. Pombo-de-Oliveira, C. Barbieri Blunck, B. Almeida Lopes, J. Zuna, J. Trka, P. Ballerini, H. Lapillonne, M. De Braekeleer, G. Cazzaniga, L. Corral Abascal, V.H.J. van der Velden, E. Delabesse, T.S. Park, S.H. Oh, M.L.M. Silva, T. Lund-Aho, V. Juvonen, A.S. Moore, O. Heidenreich, J. Vormoor, E. Zerkalenkova, Y. Olshanskaya, C. Bueno, P. Menendez, A. Teigler-Schlegel, U. Zur Stadt, J. Lentes, G. Gohring, A. Kustanovich, O. Aleinikova, B.W. Schafer, S. Kubetzko, H.O. Madsen, B. Gruhn, X. Duarte, P. Gameiro, E. Lippert, A. Bidet, J.M. Cayuela, E. Clappier, C.N. Alonso, C.M. Zwaan, M.M. van den Heuvel-Eibrink, S. Izraeli, L. Trakhtenbrot, P. Archer, J. Hancock, A. Moricke, J. Alten, M. Schrappe, M. Stanulla, S. Strehl, A. Attarbaschi, M. Dworzak, O.A. Haas, R. Panzer-Grumayer, L. Sedek, T. Szczepanski, A. Caye, L. Suarez, H. Cave, R. Marschalek, The MLL recombinome of acute leukemias in 2017, *Leukemia* 32 (2018) 273–284.
- [38] A.R. Kumar, Q. Li, W.A. Hudson, W. Chen, T. Sam, Q. Yao, E.A. Lund, B. Wu, B.J. Kowal, J.H. Kersey, A role for MEIS1 in MLL-fusion gene leukemia, *Blood* 113 (2009) 1756–1758.
- [39] H. Liu, E.H. Cheng, J.J. Hsieh, MLL fusions: pathways to leukemia, *Cancer Biol. Ther.* 8 (2009) 1204–1211.
- [40] T.A. Milne, Y. Dou, M.E. Martin, H.W. Brock, R.G. Roeder, J.L. Hess, MLL associates specifically with a subset of transcriptionally active target genes, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 14765–14770.
- [41] K. Orlovsky, A. Kalinkovich, T. Rozovskaia, E. Shezen, T. Itkin, H. Alder, H.G. Ozer, L. Carramusa, A. Avigdor, S. Volinia, A. Buchberg, A. Mazo, O. Kollet, C. Largman, C.M. Croce, T. Nakamura, T. Lapidot, E. Canaani, Down-regulation of homeobox genes MEIS1 and HOXA in MLL-rearranged acute leukemia impairs engraftment and reduces proliferation, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 7956–7961.
- [42] Q.F. Wang, G. Wu, S. Mi, F. He, J. Wu, J. Dong, R.T. Luo, R. Mattison, J.J. Kaberlein, S. Prabhakar, H. Ji, M.J. Thirman, MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome, *Blood* 117 (2011) 6895–6905.
- [43] J. Zhang, C.S. Seet, C. Sun, J. Li, D. You, A. Volk, P. Breslin, X. Li, W. Wei, Z. Qian, N.J. Zeleznik-Le, Z. Zhang, J. Zhang, p27kip1 maintains a subset of leukemia stem cells in the quiescent state in murine MLL-leukemia, *Mol. Oncol.* 7 (2013) 1069–1082.
- [44] M. Schwieger, A. Schuler, M. Forster, A. Engelmann, M.A. Arnold, R. Delwel, P.J. Valk, J. Lohler, R.K. Slany, E.N. Olson, C. Stocking, Homing and invasiveness of MLL/ENL leukemic cells is regulated by MEF2C, *Blood* 114 (2009) 2476–2488.
- [45] H. Nakanishi, T. Nakamura, E. Canaani, C.M. Croce, ALL1 fusion proteins induce deregulation of EphA7 and ERK phosphorylation in human acute leukemias, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 14442–14447.
- [46] S. Mi, Z. Li, P. Chen, C. He, D. Cao, A. Elkahoul, J. Lu, L.A. Pelloso, M. Wunderlich, H. Huang, R.T. Luo, M. Sun, M. He, M.B. Neilly, N.J. Zeleznik-Le, M.J. Thirman, J.C. Mulloy, P.P. Liu, J.D. Rowley, J. Chen, Aberrant overexpression and function of the miR-17-92 cluster in MLL-rearranged acute leukemia, *Proc. Natl. Acad. Sci. U. S. A.* 107 (2010) 3710–3715.
- [47] C. Lin, E.R. Smith, H. Takahashi, K.C. Lai, S. Martin-Brown, L. Florens, M.P. Washburn, J.W. Conaway, R.C. Conaway, A. Shilatifard, 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 (2010) 429–437.
- [48] K.M. Bernt, S.A. Armstrong, A role for DOT1L in MLL-rearranged leukemias, *Epigenomics* 3 (2011) 667–670.
- [49] K.M. Bernt, N. Zhu, A.U. Sinha, S. Vempati, J. Faber, A.V. Krivtsov, Z. Feng,

- N. Punt, A. Daigle, L. Bullinger, R.M. Pollock, V.M. Richon, A.L. Kung, S.A. Armstrong, MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L, *Cancer Cell* 20 (2011) 66–78.
- [50] A.T. Nguyen, J. He, O. Taranova, Y. Zhang, Essential role of DOT1L in maintaining normal adult hematopoiesis, *Cell Res.* 21 (2011) 1370–1373.
- [51] A.T. Nguyen, O. Taranova, J. He, Y. Zhang, DOT1L, the H3K79 methyltransferase, is required for MLL-AF9-mediated leukemogenesis, *Blood* 117 (2011) 6912–6922.
- [52] A. Yokoyama, M. Lin, A. Nares, I. Kitabayashi, M.L. Cleary, A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription, *Cancer Cell* 17 (2010) 198–212.
- [53] Z. Luo, C. Lin, A. Shilatifard, The super elongation complex (SEC) family in transcriptional control, *Nat. Rev. Mol. Cell Biol.* 13 (2012) 543–547.
- [54] E. Smith, C. Lin, A. Shilatifard, The super elongation complex (SEC) and MLL in development and disease, *Genes Dev.* 25 (2011) 661–672.
- [55] C. Lin, A.S. Garrett, B. De Kumar, E.R. Smith, M. Gogol, C. Seidel, R. Krumlauf, A. Shilatifard, Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC), *Genes Dev.* 25 (2011) 1486–1498.
- [56] N. He, M. Liu, J. Hsu, Y. Xue, S. Chou, A. Burlingame, N.J. Krogan, T. Alber, Q. Zhou, 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 (2010) 428–438.
- [57] B. Sobhian, N. Laguet, A. Yatim, M. Nakamura, Y. Levy, R. Kiernan, M. Benkirane, HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 75K snRNP, *Mol. Cell* 38 (2010) 439–451.
- [58] A.K.N. Chan, C.W. Chen, Rewiring the epigenetic networks in MLL-rearranged leukemias: epigenetic dysregulation and pharmacological interventions, *Front Cell Dev Biol* 7 (2019) 81.
- [59] S. He, T.J. Senter, J. Pollock, C. Han, S.K. Upadhyay, T. Purohit, R.D. Gogliotti, C.W. Lindsley, T. Cierpicki, S.R. Stauffer, J. Grembecka, High-affinity small-molecule inhibitors of the menin-mixed lineage leukemia (MLL) interaction closely mimic a natural protein-protein interaction, *J. Med. Chem.* 57 (2014) 1543–1556.
- [60] R. Marschalek, MLL leukemia and future treatment strategies, *Arch Pharm (Weinheim)* 348 (2015) 221–228.
- [61] M.Y. Mazina, N.E. Vorobyeva, The role of ATP-dependent chromatin remodeling complexes in regulation of genetic processes, *Genetika* 52 (2016) 529–540.
- [62] E. Martinez, V.B. Palhan, A. Tjernberg, E.S. Lyman, A.M. Gamper, T.K. Kundu, B.T. Chait, R.G. Roeder, Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo, *Mol. Cell Biol.* 21 (2001) 6782–6795.
- [63] J. Soutourina, Transcription regulation by the mediator complex, *Nat. Rev. Mol. Cell Biol.* 19 (2018) 262–274.
- [64] Q. Liu, S.E. Gabriel, K.L. Roinick, R.D. Ward, K.M. Arndt, Analysis of TFIIF function in vivo: evidence for a role in TATA-binding protein recruitment and gene-specific activation, *Mol. Cell Biol.* 19 (1999) 8673–8685.
- [65] R.O. Weinzierl, S.C. Wiesler, Revealing the functions of TFIIF, *Transcription* 2 (2011) 254–257.
- [66] C. Bernecky, P. Grob, C.C. Ebmeier, E. Nogales, D.J. Taatjes, Molecular architecture of the human mediator-RNA polymerase II-TFIIF assembly, *PLoS Biol.* 9 (2011) e1000603.
- [67] D.S. Luse, Rethinking the role of TFIIF in transcript initiation by RNA polymerase II, *Transcription* 3 (2012) 156–159.
- [68] Y. Ohkuma, S. Hashimoto, C.K. Wang, M. Horikoshi, R.G. Roeder, Analysis of the role of TFIIE in basal transcription and TFIIF-mediated carboxy-terminal domain phosphorylation through structure-function studies of TFIIE- α , *Mol. Cell Biol.* 15 (1995) 4856–4866.
- [69] T. Okamoto, S. Yamamoto, Y. Watanabe, T. Ohta, F. Hanaoka, R.G. Roeder, Y. Ohkuma, Analysis of the role of TFIIE in transcriptional regulation through structure-function studies of the TFIIE β subunit, *J. Biol. Chem.* 273 (1998) 19866–19876.
- [70] J.K. Rimel, D.J. Taatjes, The essential and multifunctional TFIIF complex, *Protein Sci.* 27 (2018) 1018–1037.
- [71] B.W. Guidi, G. Bjornsdottir, D.C. Hopkins, L. Lacomis, H. Erdjument-Bromage, P. Tempst, L.C. Myers, Mutual targeting of mediator and the TFIIF kinase Kin28, *J. Biol. Chem.* 279 (2004) 29114–29120.
- [72] T.M. Sogaard, J.Q. Svejstrup, Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator, *J. Biol. Chem.* 282 (2007) 14113–14120.
- [73] M.S. Akhtar, M. Heidemann, J.R. Tietjen, D.W. Zhang, R.D. Chapman, D. Eick, A.Z. Ansari, TFIIF kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II, *Mol. Cell* 34 (2009) 387–393.
- [74] K. Glover-Cutter, S. Laroche, B. Erickson, C. Zhang, K. Shokat, R.P. Fisher, D.L. Bentley, TFIIF-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II, *Mol. Cell Biol.* 29 (2009) 5455–5464.
- [75] P. Gariglio, M. Bellard, P. Chambon, Clustering of RNA polymerase B molecules in the 5' moiety of the adult beta-globin gene of hen erythrocytes, *Nucleic Acids Res.* 9 (1981) 2589–2598.
- [76] A.E. Rougvie, J.T. Lis, The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged, *Cell* 54 (1988) 795–804.
- [77] E.B. Rasmussen, J.T. Lis, In vivo transcriptional pausing and cap formation on three *Drosophila* heat shock genes, *Proc. Natl. Acad. Sci. U. S. A.* 90 (1993) 7923–7927.
- [78] L.J. Strobl, D. Eick, Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-myc in vivo, *EMBO J.* 11 (1992) 3307–3314.
- [79] M. Palangat, T.I. Meier, R.G. Keene, R. Landick, Transcriptional pausing at +62 of the HIV-1 nascent RNA modulates formation of the TAR RNA structure, *Mol. Cell* 1 (1998) 1033–1042.
- [80] Y.H. Ping, T.M. Rana, DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation, *J. Biol. Chem.* 276 (2001) 12951–12958.
- [81] H. Bunch, X. Zheng, A. Burkholder, S.T. Dillon, S. Motola, G. Birrane, C.C. Ebmeier, S. Levine, D. Fargo, G. Hu, D.J. Taatjes, S.K. Calderwood, TRIM28 regulates RNA polymerase II promoter-proximal pausing and pause release, *Nat. Struct. Mol. Biol.* 21 (2014) 876–883.
- [82] B. Cheng, T. Li, P.B. Rahl, T.E. Adamson, N.B. Loudas, J. Guo, K. Varzavand, J.J. Cooper, X. Hu, A. Gnatt, R.A. Young, D.H. Price, Functional association of Gdown1 with RNA polymerase II poised on human genes, *Mol. Cell* 45 (2012) 38–50.
- [83] T.T. Tettey, X. Gao, W. Shao, H. Li, B.A. Story, A.D. Chitsazan, R.L. Glaser, Z.H. Goode, C.W. Seidel, R.C. Conaway, J. Zeitlinger, M. Blanchette, J.W. Conaway, A role for FACT in RNA polymerase II promoter-proximal pausing, *Cell Rep.* 27 (e3777) (2019) 3770–3779.
- [84] X. Liu, W.L. Kraus, X. Bai, Ready, pause, go: regulation of RNA polymerase II pausing and release by cellular signaling pathways, *Trends Biochem. Sci.* 40 (2015) 516–525.
- [85] I. Jonkers, J.T. Lis, Getting up to speed with transcription elongation by RNA polymerase II, *Nat. Rev. Mol. Cell Biol.* 16 (2015) 167–177.
- [86] M. Barboric, R.M. Nissen, S. Kanazawa, N. Jabrane-Ferrat, B.M. Peterlin, NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II, *Mol. Cell* 8 (2001) 327–337.
- [87] K. Ghosh, M. Tang, N. Kumari, A. Nandy, S. Basu, D.P. Mall, K. Rai, D. Biswas, Positive regulation of transcription by human ZMYND8 through its association with P-TEFb complex, *Cell Rep.* 24 (2018) 2141–2154 e2146.
- [88] S. Kanazawa, L. Soucek, G. Evan, T. Okamoto, B.M. Peterlin, c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis, *Oncogene* 22 (2003) 5707–5711.
- [89] P. Mitra, L.A. Pereira, Y. Drabsch, R.G. Ramsay, T.J. Gonda, Estrogen receptor- α recruits P-TEFb to overcome transcriptional pausing in intron 1 of the MYB gene, *Nucleic Acids Res.* 40 (2012) 5988–6000.
- [90] T. Qi, W. Tang, L. Wang, L. Zhai, L. Guo, X. Zeng, G-actin participates in RNA polymerase II-dependent transcription elongation by recruiting positive transcription elongation factor b (P-TEFb), *J. Biol. Chem.* 286 (2011) 15171–15181.
- [91] W. Wang, X. Yao, Y. Huang, X. Hu, R. Liu, D. Hou, R. Chen, G. Wang, Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb, *Transcription* 4 (2013) 39–51.
- [92] Z. Yang, N. He, Q. Zhou, Brd4 recruits P-TEFb to chromosomes at late mitosis to promote G1 gene expression and cell cycle progression, *Mol. Cell Biol.* 28 (2008) 967–976.
- [93] R.P. McNamara, J.E. Reeder, E.A. McMillan, C.W. Bacon, J.L. McCann, I. D'Orso, KAP1 recruitment of the 75K snRNP complex to promoters enables transcription elongation by RNA polymerase II, *Mol. Cell* 61 (2016) 39–53.
- [94] X. Lu, X. Zhu, Y. Li, M. Liu, B. Yu, Y. Wang, M. Rao, H. Yang, K. Zhou, Y. Wang, Y. Chen, M. Chen, S. Zhuang, L.F. Chen, R. Liu, R. Chen, Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II, *Nucleic Acids Res.* 44 (2016) 6853–6867.
- [95] D. Yadav, K. Ghosh, S. Basu, R.G. Roeder, D. Biswas, Multivalent role of human TFIID in recruiting elongation components at the promoter-proximal region for transcriptional control, *Cell Rep.* 26 (2019) 1303–1317 e1307.
- [96] D. Biswas, T.A. Milne, V. Basur, J. Kim, K.S. Elenitoba-Johnson, C.D. Allis, R.G. Roeder, Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) 15751–15756.
- [97] H. Lu, Z. Li, W. Zhang, U. Schulze-Gahmen, Y. Xue, Q. Zhou, Gene target specificity of the Super Elongation Complex (SEC) family: how HIV-1 tat employs selected SEC members to activate viral transcription, *Nucleic Acids Res.* 43 (2015) 5868–5879.
- [98] S. Chou, H. Upton, K. Bao, U. Schulze-Gahmen, A.J. Samelson, N. He, A. Nowak, H. Lu, N.J. Krogan, Q. Zhou, T. Alber, HIV-1 Tat recruits transcription elongation factors dispersed along a flexible AFF4 scaffold, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) E123–E131.
- [99] U. Schulze-Gahmen, H. Upton, A. Birnberg, K. Bao, S. Chou, N.J. Krogan, Q. Zhou, T. Alber, The AFF4 scaffold binds human P-TEFb adjacent to HIV tat, *Elife* 2 (2013) e00327.
- [100] H. Lu, Z. Li, Y. Xue, U. Schulze-Gahmen, J.R. Johnson, N.J. Krogan, T. Alber, Q. Zhou, AFF1 is a ubiquitous P-TEFb partner to enable Tat extraction of P-TEFb from 75K snRNP and formation of SECs for HIV transactivation, *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) E15–E24.
- [101] A. Bursen, S. Moritz, A. Gaussmann, S. Moritz, T. Dingermann, R. Marschalek, Interaction of AF4 wild-type and AF4.MLL fusion protein with SIAH proteins: indication for t(4;11) pathobiology? *Oncogene* 23 (2004) 6237–6249.
- [102] P.L. Oliver, E. Bitoun, J. Clark, E.L. Jones, K.E. Davies, Mediation of Af4 protein function in the cerebellum by Siah proteins, *Proc. Natl. Acad. Sci. U. S. A.* 101 (2004) 14901–14906.
- [103] S. Krasnopolsky, L. Marom, R.A. Victor, A. Kuzmina, J.C. Schwartz, K. Fujinaga, R. Taube, Fused in sarcoma silences HIV gene transcription and maintains viral latency through suppressing AFF4 gene activation, *Retrovirology* 16 (2019) 16.
- [104] F. Muck, S. Bracharz, R. Marschalek, DDX6 transfers P-TEFb kinase to the AF4/AF4N (AFF1) super elongation complex, *Am J Blood Res* 6 (2016) 28–45.
- [105] H. Okuda, A. Kanai, S. Ito, H. Matsui, A. Yokoyama, AF4 uses the SL1 components

- of RNAP1 machinery to initiate MLL fusion- and AEP-dependent transcription, *Nat. Commun.* 6 (2015) 8869.
- [106] K. Ahmad, B. Scholz, R. Capelo, I. Schweighofer, A.S. Kahnt, R. Marschalek, D. Steinhilber, AF4 and AF4-MLL mediate transcriptional elongation of 5-lipoxygenase mRNA by 1, 25-dihydroxyvitamin D₃, *Oncotarget* 6 (2015) 25784–25800.
- [107] M. Mohan, H.M. Herz, Y.H. Takahashi, C. Lin, K.C. Lai, Y. Zhang, M.P. Washburn, L. Florens, A. Shilatifard, Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom), *Genes Dev.* 24 (2010) 574–589.
- [108] K. Hetzner, M.P. Garcia-Cuellar, C. Buttner, R.K. Slany, The interaction of ENL with PAF1 mitigates polycomb silencing and facilitates murine leukemogenesis, *Blood* 131 (2018) 662–673.
- [109] Y. Li, H. Wen, Y. Xi, K. Tanaka, H. Wang, D. Peng, Y. Ren, Q. Jin, S.Y. Dent, W. Li, H. Li, X. Shi, AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation, *Cell* 159 (2014) 558–571.
- [110] C.S. Hemenway, A.C. de Erkenez, G.C. Gould, The polycomb protein MPC3 interacts with AF9, an MLL fusion partner in t(9;11)(p22;q23) acute leukemias, *Oncogene* 20 (2001) 3798–3805.
- [111] R.S. Srinivasan, A.C. de Erkenez, C.S. Hemenway, The mixed lineage leukemia fusion partner AF9 binds specific isoforms of the BCL-6 corepressor, *Oncogene* 22 (2003) 3395–3406.
- [112] N. He, C.K. Chan, B. Sobhian, S. Chou, Y. Xue, M. Liu, T. Alber, M. Benkirane, Q. Zhou, Human polymerase-associated factor complex (PAF_c) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin, *Proc. Natl. Acad. Sci. U. S. A.* 108 (2011) E636–E645.
- [113] J.J. Lin, C.S. Hemenway, Hsp90 directly modulates the spatial distribution of AF9/MLLT3 and affects target gene expression, *J. Biol. Chem.* 285 (2010) 11966–11973.
- [114] W. Zhang, X. Xia, M.R. Reisenauer, C.S. Hemenway, B.C. Kone, Dot1a-AF9 complex mediates histone H3 Lys-79 hypermethylation and repression of ENaC α in an aldosterone-sensitive manner, *J. Biol. Chem.* 281 (2006) 18059–18068.
- [115] W. Zhang, Z. Yu, H. Wu, L. Chen, Q. Kong, B.C. Kone, An AF9 cis-element directly targets Dot1a to mediate transcriptional repression of the α ENaC gene, *Am. J. Physiol. Ren. Physiol.* 304 (2013) F367–F375.
- [116] X. Zhang, Q. Zhou, L. Chen, S. Berger, H. Wu, Z. Xiao, D. Pearce, X. Zhou, W. Zhang, Mineralocorticoid receptor antagonizes Dot1a-AF9 complex to increase α ENaC transcription, *Am. J. Physiol. Ren. Physiol.* 305 (2013) F1436–F1444.
- [117] W. Zhang, X. Xia, M.R. Reisenauer, T. Rieg, F. Lang, D. Kuhl, V. Vallon, B.C. Kone, Aldosterone-induced Sgk1 relieves Dot1a-AF9-mediated transcriptional repression of epithelial Na⁺ channel α , *J. Clin. Invest.* 117 (2007) 773–783.
- [118] M.R. Reisenauer, M. Anderson, L. Huang, Z. Zhang, Q. Zhou, B.C. Kone, A.P. Morris, G.D. Lesage, S.E. Dryer, W. Zhang, AF17 competes with AF9 for binding to Dot1a to up-regulate transcription of epithelial Na⁺ channel α , *J. Biol. Chem.* 284 (2009) 35659–35669.
- [119] Z.Y. Yu, Q. Kong, B.C. Kone, Physical and functional interaction of Rnf2 with AF9 regulates basal and aldosterone-stimulated transcription of the α -ENaC gene in a renal collecting duct cell line, *Biosci. Rep.* 33 (2013).
- [120] R. Hariabaskar, M. Putz, B. Schupp, K. Skouloudaki, A. Bietenbeck, G. Walz, T. Schafer, The planar cell polarity (PCP) protein Diversin translocates to the nucleus to interact with the transcription factor AF9, *Biochem. Biophys. Res. Commun.* 387 (2009) 212–217.
- [121] Z. Nie, Z. Yan, E.H. Chen, S. Sechi, C. Ling, S. Zhou, Y. Xue, D. Yang, D. Murray, E. Kanakubo, M.L. Cleary, W. Wang, Novel SWI/SNF chromatin-remodeling complexes contain a mixed-lineage leukemia chromosomal translocation partner, *Mol. Cell. Biol.* 23 (2003) 2942–2952.
- [122] A. Ui, Y. Nagaura, A. Yasui, Transcriptional elongation factor ENL phosphorylated by ATM recruits polycomb and switches off transcription for DSB repair, *Mol. Cell* 58 (2015) 468–482.
- [123] J.S. Byun, T.D. Fufa, C. Wakano, A. Fernandez, C.M. Haggerty, M.H. Sung, K. Gardner, ELL facilitates RNA polymerase II pause site entry and release, *Nat. Commun.* 3 (2012) 633.
- [124] E.R. Smith, C. Lin, A.S. Garrett, J. Thornton, N. Mohaghegh, D. Hu, J. Jackson, A. Saraf, S.K. Swanson, C. Seidel, L. Florens, M.P. Washburn, J.C. Eissenberg, A. Shilatifard, The little elongation complex regulates small nuclear RNA transcription, *Mol. Cell* 44 (2011) 954–965.
- [125] K. Sakurai, T. Michiue, A. Kikuchi, M. Asashima, Inhibition of the canonical Wnt signaling pathway in cytoplasm: a novel property of the carboxyl terminal domains of two Xenopus ELL genes, *Zool. Sci.* 21 (2004) 407–416.
- [126] L. Pascual-Le Tallec, F. Simone, S. Viengchareun, G. Meduri, M.J. Thirman, M. Lombes, The elongation factor ELL (eleven-nineteen lysine-rich leukemia) is a selective coregulator for steroid receptor functions, *Mol. Endocrinol.* 19 (2005) 1158–1169.
- [127] L. Liu, J. Ai, W. Xiao, J. Liu, Y. Wang, D. Xin, Z. He, Y. Guo, Z. Wang, ELL is an HIF-1 α partner that regulates and responds to hypoxia response in PC3 cells, *Prostate* 70 (2010) 797–805.
- [128] Y. Chen, C. Zhou, W. Ji, Z. Mei, B. Hu, W. Zhang, D. Zhang, J. Wang, X. Liu, G. Ouyang, J. Zhou, W. Xiao, ELL targets c-Myc for proteasomal degradation and suppresses tumour growth, *Nat. Commun.* 7 (2016) 11057.
- [129] W. Zhang, W. Ji, X. Liu, G. Ouyang, W. Xiao, ELL inhibits E2F1 transcriptional activity by enhancing E2F1 deacetylation via recruitment of histone deacetylase 1, *Mol. Cell. Biol.* 34 (2014) 765–775.
- [130] S. Mourgues, V. Gautier, A. Lagarou, C. Bordier, A. Mourcet, J. Slingerland, L. Kaddoum, F. Coin, W. Vermeulen, A. Gonzales de Peredo, M. Monsarrat, P.O. Mari, G. Giglia-Mari, ELL, a novel TFIIF partner, is involved in transcription restart after DNA repair, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 17927–17932.
- [131] H. Takahashi, T.J. Parmely, S. Sato, C. Tomomori-Sato, C.A. Banks, S.E. Kong, H. Szutorisz, S.K. Swanson, S. Martin-Brown, M.P. Washburn, L. Florens, C.W. Seidel, C. Lin, E.R. Smith, A. Shilatifard, R.C. Conaway, J.W. Conaway, Human mediator subunit MED26 functions as a docking site for transcription elongation factors, *Cell* 146 (2011) 92–104.
- [132] H. Takahashi, I. Takigawa, M. Watanabe, D. Anwar, M. Shibata, C. Tomomori-Sato, S. Sato, A. Ranjan, C.W. Seidel, T. Tsukiyama, W. Mizushima, M. Hayashi, Y. Ohkawa, J.W. Conaway, R.C. Conaway, S. Hatakeyama, MED26 regulates the transcription of snRNA genes through the recruitment of little elongation complex, *Nat. Commun.* 6 (2015) 5941.
- [133] J.X. Liu, D. Zhang, X. Xie, G. Ouyang, X. Liu, Y. Sun, W. Xiao, Eaf1 and Eaf2 negatively regulate canonical Wnt/ β -catenin signaling, *Development* 140 (2013) 1067–1078.
- [134] B. Hu, W. Zhang, X. Feng, W. Ji, X. Xie, W. Xiao, Zebrafish eaf1 suppresses foxo3b expression to modulate transcriptional activity of gata1 and sp1 in primitive hematopoiesis, *Dev. Biol.* 388 (2014) 81–93.
- [135] J.X. Liu, Q.H. Xu, S. Li, X. Yu, W. Liu, G. Ouyang, T. Zhang, L.L. Chen, Transcriptional factors Eaf1/2 inhibit endoderm and mesoderm formation via suppressing TGF- β signaling, *Biochim Biophys Acta Gene Regul Mech* 1860 (2017) 1103–1116.
- [136] W. Xiao, J. Ai, G. Habermacher, O. Volpert, X. Yang, A.Y. Zhang, J. Hahn, X. Cai, Z. Wang, U19/Eaf2 binds to and stabilizes Von Hippel-Lindau protein, *Cancer Res.* 69 (2009) 2599–2606.
- [137] Z. Chen, X. Liu, Z. Mei, Z. Wang, W. Xiao, EAF2 suppresses hypoxia-induced factor 1 α transcriptional activity by disrupting its interaction with coactivator CBP/p300, *Mol. Cell. Biol.* 34 (2014) 1085–1099.
- [138] F. Jiang, J. Ai, W. Xiao, Z. Wang, FB1, an E2A fusion partner in childhood leukemia, interacts with U19/EAF2 and inhibits its transcriptional activity, *Cancer Lett.* 253 (2007) 265–272.
- [139] J. Ai, L.E. Pascal, L. Wei, Y. Zang, Y. Zhou, X. Yu, Y. Gong, S. Nakajima, J.B. Nelson, A.S. Levine, L. Lan, Z. Wang, EAF2 regulates DNA repair through Ku70/Ku80 in the prostate, *Oncogene* 36 (2017) 2054–2065.
- [140] G. Cecere, S. Hoersch, M.B. Jensen, S. Dixit, A. Grishok, The ZFP-1(AF10)/DOT-1 complex opposes H2B ubiquitination to reduce Pol II transcription, *Mol. Cell* 50 (2013) 894–907.
- [141] B. Linder, N. Gerlach, H. Jackle, The Drosophila homolog of the human AF10 is an HP1-interacting suppressor of position effect variegation, *EMBO Rep.* 2 (2001) 211–216.
- [142] D.R. de Bruijn, N.R. dos Santos, J. Thijssen, M. Balemans, S. Debernardi, B. Linder, B.D. Young, A. Geurts van Kessel, The synovial sarcoma associated protein SYT interacts with the acute leukemia associated protein AF10, *Oncogene* 20 (2001) 3281–3289.
- [143] S. Forissier, D. Razanajaona, A.S. Ay, S. Martel, L. Bartholin, R. Rimokh, AF10-dependent transcription is enhanced by its interaction with FLRG, *Biol. Cell.* 99 (2007) 563–571.
- [144] S. Debernardi, A. Bassini, L.K. Jones, T. Chaplin, B. Linder, D.R. de Bruijn, E. Meese, B.D. Young, The MLL fusion partner AF10 binds GAS41, a protein that interacts with the human SWI/SNF complex, *Blood* 99 (2002) 275–281.
- [145] H. Wu, L. Chen, Q. Zhou, W. Zhang, AF17 facilitates Dot1a nuclear export and upregulates ENaC-mediated Na⁺ transport in renal collecting duct cells, *PLoS One* 6 (2011) e27429.
- [146] T. Yamamoto, N. Harada, K. Kano, S. Taya, E. Canaani, Y. Matsuura, A. Mizoguchi, C. Ide, K. Kaibuchi, The Ras target AF-6 interacts with ZO-1 and serves as a peripheral component of tight junctions in epithelial cells, *J. Cell Biol.* 139 (1997) 785–795.
- [147] T. Sakisaka, H. Nakanishi, K. Takahashi, K. Mandai, M. Miyahara, A. Satoh, K. Takaishi, Y. Takai, Different behavior of l-afadin and neurabin-II during the formation and destruction of cell-cell adherens junction, *Oncogene* 18 (1999) 1609–1617.
- [148] K. Ebnet, C.U. Schulz, M.K. Meyer Zu Brickwedde, G.G. Pendl, D. Vestweber, Junctional adhesion molecule interacts with the PDZ domain-containing proteins AF-6 and ZO-1, *J. Biol. Chem.* 275 (2000) 27979–27988.
- [149] S. Taya, T. Yamamoto, K. Kano, Y. Kawano, A. Iwamatsu, T. Tsuchiya, K. Tanaka, M. Kanai-Azuma, S.A. Wood, J.S. Mattick, K. Kaibuchi, The Ras target AF-6 is a substrate of the fam deubiquitinating enzyme, *J. Cell Biol.* 142 (1998) 1053–1062.
- [150] M. Pantaleon, M. Kanai-Azuma, J.S. Mattick, K. Kaibuchi, P.L. Kaye, S.A. Wood, FAM deubiquitylating enzyme is essential for preimplantation mouse embryonic development, *Mech. Dev.* 109 (2001) 151–160.
- [151] K. Mandai, H. Nakanishi, A. Satoh, K. Takahashi, K. Satoh, H. Nishioka, A. Mizoguchi, Y. Takai, Ponsin/SH3P12: an l-afadin- and vinculin-binding protein localized at cell-cell and cell-matrix adherens junctions, *J. Cell Biol.* 144 (1999) 1001–1017.
- [152] K. Takahashi, H. Nakanishi, M. Miyahara, K. Mandai, K. Satoh, A. Satoh, H. Nishioka, J. Aoki, A. Nomoto, A. Mizoguchi, Y. Takai, Nectin/PRR: an immunoglobulin-like cell adhesion molecule recruited to cadherin-based adherens junctions through interaction with Afadin, a PDZ domain-containing protein, *J. Cell Biol.* 145 (1999) 539–549.
- [153] M. Asada, K. Irie, K. Morimoto, A. Yamada, W. Ikeda, M. Takeuchi, Y. Takai, ADIP, a novel Afadin- and alpha-actinin-binding protein localized at cell-cell adherens junctions, *J. Biol. Chem.* 278 (2003) 4103–4111.
- [154] T. Ooshio, K. Irie, K. Morimoto, A. Fukuhara, T. Imai, Y. Takai, Involvement of LMO7 in the association of two cell-cell adhesion molecules, nectin and E-cadherin, through afadin and alpha-actinin in epithelial cells, *J. Biol. Chem.* 279 (2004) 31365–31373.
- [155] L. Su, M. Hattori, M. Moriyama, N. Murata, M. Harazaki, K. Kaibuchi, N. Minato, AF-6 controls integrin-mediated cell adhesion by regulating Rap1 activation through the specific recruitment of Rap1GTP and SPA-1, *J. Biol. Chem.* 278 (2003) 15232–15238.

- [156] Z. Xie, R.L. Hugarin, P. Penzes, Activity-dependent dendritic spine structural plasticity is regulated by small GTPase Rap1 and its target AF-6, *Neuron* 48 (2005) 605–618.
- [157] Z. Xie, H. Photowala, M.E. Cahill, D.P. Srivastava, K.M. Woolfrey, C.Y. Shum, R.L. Hugarin, P. Penzes, Coordination of synaptic adhesion with dendritic spine remodeling by AF-6 and kalinin-7, *J. Neurosci.* 28 (2008) 6079–6091.
- [158] X. Li, B.D. Lynn, J.I. Nagy, The effector and scaffolding proteins AF6 and MUPP1 interact with connexin36 and localize at gap junctions that form electrical synapses in rodent brain, *Eur. J. Neurosci.* 35 (2012) 166–181.
- [159] D.P. Srivastava, B.A. Copits, Z. Xie, R. Huda, K.A. Jones, S. Mukherji, M.E. Cahill, J.E. VanLeeuwen, K.M. Woolfrey, I. Rafalovich, G.T. Swanson, P. Penzes, Afadin is required for maintenance of dendritic structure and excitatory tone, *J. Biol. Chem.* 287 (2012) 35964–35974.
- [160] T. Sato, Y. Kanai, T. Noma, M. Kanai-Azuma, S. Taya, T. Matsui, M. Ishii, H. Kawakami, M. Kurohmaru, K. Kaibuchi, S.A. Wood, Y. Hayashi, A close correlation in the expression patterns of Af-6 and Usp9x in Sertoli and granulosa cells of mouse testis and ovary, *Reproduction* 128 (2004) 583–594.
- [161] C. Dai, X. Wang, Y. Wu, Y. Xu, S. Zhuo, M. Qi, W. Ji, L. Zhan, Polarity protein AF6 controls hepatic glucose homeostasis and insulin sensitivity by modulating IRS1/AKT insulin pathway in an SHP2-dependent manner, *Diabetes* 68 (2019) 1577–1590.
- [162] B. Boettner, E.E. Govek, J. Cross, L. Van Aelst, The junctional multidomain protein AF-6 is a binding partner of the Rap1A GTPase and associates with the actin cytoskeletal regulator profilin, *Proc. Natl. Acad. Sci. U. S. A.* 97 (2000) 9064–9069.
- [163] M. Lorgier, K. Moelling, Regulation of epithelial wound closure and intercellular adhesion by interaction of AF6 with actin cytoskeleton, *J. Cell Sci.* 119 (2006) 3385–3398.
- [164] S. Pokutta, F. Drees, Y. Takai, W.J. Nelson, W.I. Weiss, Biochemical and structural definition of the I-afadin- and actin-binding sites of alpha-catenin, *J. Biol. Chem.* 277 (2002) 18868–18874.
- [165] X. Grana, A. De Luca, N. Sang, Y. Fu, P.P. Claudio, J. Rosenblatt, D.O. Morgan, A. Giordano, PITARE, a nuclear CDC2-related protein kinase that phosphorylates the retinoblastoma protein in vitro, *Proc. Natl. Acad. Sci. U. S. A.* 91 (1994) 3834–3838.
- [166] J. Gu, N.D. Babayeva, Y. Suwa, A.G. Baranovskiy, D.H. Price, T.H. Tahirou, Crystal structure of HIV-1 Tat complexed with human P-TEFb and AFF4, *Cell Cycle* 13 (2014) 1788–1797.
- [167] T.H. Tahirou, N.D. Babayeva, K. Varzavand, J.J. Cooper, S.C. Sedore, D.H. Price, Crystal structure of HIV-1 Tat complexed with human P-TEFb, *Nature* 465 (2010) 747–751.
- [168] D.H. Price, P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II, *Mol. Cell. Biol.* 20 (2000) 2629–2634.
- [169] S.C. Sedore, S.A. Byers, S. Biglione, J.P. Price, W.J. Maury, D.H. Price, Manipulation of P-TEFb control machinery by HIV: recruitment of P-TEFb from the large form by Tat and binding of HEXIM1 to TAR, *Nucleic Acids Res.* 35 (2007) 4347–4358.
- [170] B.J. Krueger, K. Varzavand, J.J. Cooper, D.H. Price, The mechanism of release of P-TEFb and HEXIM1 from the 7SK snRNP by viral and cellular activators includes a conformational change in 7SK, *PLoS One* 5 (2010) e12335.
- [171] Y.H. Ping, T.M. Rana, Tat-associated kinase (P-TEFb): a component of transcription preinitiation and elongation complexes, *J. Biol. Chem.* 274 (1999) 7399–7404.
- [172] Y. Zhu, T. Pe'ery, J. Peng, Y. Ramanathan, N. Marshall, T. Marshall, B. Amendt, M.B. Mathews, D.H. Price, Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro, *Genes Dev.* 11 (1997) 2622–2632.
- [173] K. Fujinaga, T.P. Cujec, J. Peng, J. Garriga, D.H. Price, X. Grana, B.M. Peterlin, The ability of positive transcription elongation factor B to transactivate human immunodeficiency virus transcription depends on a functional kinase domain, cyclin T1, and tat, *J. Virol.* 72 (1998) 7154–7159.
- [174] J.B. Kim, P.A. Sharp, Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase, *J. Biol. Chem.* 276 (2001) 12317–12323.
- [175] Y. Liu, L. Warfield, C. Zhang, J. Luo, J. Allen, W.H. Lang, J. Ranish, K.M. Shokat, S. Hahn, Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex, *Mol. Cell. Biol.* 29 (2009) 4852–4863.
- [176] T. Wada, T. Takagi, Y. Yamaguchi, D. Watanabe, H. Handa, Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro, *EMBO J.* 17 (1998) 7395–7403.
- [177] K. Fujinaga, D. Irwin, Y. Huang, R. Taube, T. Kurosu, B.M. Peterlin, Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element, *Mol. Cell. Biol.* 24 (2004) 787–795.
- [178] M. Zhou, M.A. Halanski, M.F. Radonovich, F. Kashanchi, J. Peng, D.H. Price, J.N. Brady, Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription, *Mol. Cell. Biol.* 20 (2000) 5077–5086.
- [179] J.C. Eissenberg, A. Shilatifard, N. Dorokhov, D.E. Michener, Cdk9 is an essential kinase in *Drosophila* that is required for heat shock gene expression, histone methylation and elongation factor recruitment, *Mol. Gen. Genomics.* 277 (2007) 101–114.
- [180] M. Zhou, S. Nekhai, D.C. Bharucha, A. Kumar, H. Ge, D.H. Price, J.M. Egly, J.N. Brady, TFIIF inhibits CDK9 phosphorylation during human immunodeficiency virus type 1 transcription, *J. Biol. Chem.* 276 (2001) 44633–44640.
- [181] J.H. Yik, R. Chen, R. Nishimura, J.L. Jennings, A.J. Link, Q. Zhou, Inhibition of P-TEFb (CDK9/cyclin T) kinase and RNA polymerase II transcription by the co-ordinated actions of HEXIM1 and 7SK snRNA, *Mol. Cell* 12 (2003) 971–982.
- [182] Z. Yang, Q. Zhu, K. Luo, Q. Zhou, The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription, *Nature* 414 (2001) 317–322.
- [183] V.T. Nguyen, T. Kiss, A.A. Michels, O. Bensaude, 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes, *Nature* 414 (2001) 322–325.
- [184] A.A. Michels, V.T. Nguyen, A. Faldut, V. Labas, M. Edwards, F. Bonnet, L. Lania, O. Bensaude, MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner, *Mol. Cell. Biol.* 23 (2003) 4859–4869.
- [185] L. Kobbi, E. Demey-Thomas, F. Braye, F. Proux, O. Kolesnikova, J. Vinh, A. Poterszman, O. Bensaude, An evolutionary conserved Hexim1 peptide binds to the Cdk9 catalytic site to inhibit P-TEFb, *Proc. Natl. Acad. Sci. U. S. A.* 113 (2016) 12721–12726.
- [186] A. Schulte, N. Czudnochowski, M. Barboric, A. Schonichen, D. Blazek, B.M. Peterlin, M. Geyer, Identification of a cyclin T-binding domain in Hexim1 and biochemical analysis of its binding competition with HIV-1 Tat, *J. Biol. Chem.* 280 (2005) 24968–24977.
- [187] M.K. Jang, K. Mochizuki, M. Zhou, H.S. Jeong, J.N. Brady, K. Ozato, The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription, *Mol. Cell* 19 (2005) 523–534.
- [188] M. Zhou, K. Huang, K.J. Jung, W.K. Cho, Z. Klase, F. Kashanchi, C.A. Pise-Masison, J.N. Brady, Bromodomain protein Brd4 regulates human immunodeficiency virus transcription through phosphorylation of CDK9 at threonine 29, *J. Virol.* 83 (2009) 1036–1044.
- [189] Z. Yang, J.H. Yik, R. Chen, N. He, M.K. Jang, K. Ozato, Q. Zhou, Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4, *Mol. Cell* 19 (2005) 535–545.
- [190] F. Itzen, A.K. Greifenberg, C.A. Bosken, M. Geyer, Brd4 activates P-TEFb for RNA polymerase II CTD phosphorylation, *Nucleic Acids Res.* 42 (2014) 7577–7590.
- [191] M.C. Patel, M. Debrosse, M. Smith, A. Dey, W. Huynh, N. Sarai, T.D. Heightman, T. Tamura, K. Ozato, BRD4 coordinates recruitment of pause release factor P-TEFb and the pausing complex NELF/DSIF to regulate transcription elongation of interferon-stimulated genes, *Mol. Cell. Biol.* 33 (2013) 2497–2507.
- [192] D.A. Bisgrove, T. Mahmoudi, P. Henklein, E. Verdin, Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 13690–13695.
- [193] S.R. Eberhardy, P.J. Farnham, Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter, *J. Biol. Chem.* 277 (2002) 40156–40162.
- [194] P.B. Rahl, C.Y. Lin, A.C. Seila, R.A. Flynn, S. McQuinn, C.B. Burge, P.A. Sharp, R.A. Young, c-Myc regulates transcriptional pause release, *Cell* 141 (2010) 432–445.
- [195] I. Iankova, R.K. Petersen, J.S. Annicotte, C. Chavey, J.B. Hansen, I. Kratchmarova, D. Sarruf, M. Benkirane, K. Kristiansen, L. Fajas, Peroxisome proliferator-activated receptor gamma recruits the positive transcription elongation factor b complex to activate transcription and promote adipogenesis, *Mol. Endocrinol.* 20 (2006) 1494–1505.
- [196] S. Kanazawa, T. Okamoto, B.M. Peterlin, Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV infection, *Immunity* 12 (2000) 61–70.
- [197] M. Hoque, T.M. Young, C.G. Lee, G. Serrero, M.B. Mathews, T. Pe'ery, The growth factor granulin interacts with cyclin T1 and modulates P-TEFb-dependent transcription, *Mol. Cell. Biol.* 23 (2003) 1688–1702.
- [198] M. Zhou, H. Lu, H. Park, J. Wilson-Chiru, R. Linton, J.N. Brady, Tat interacts with P-TEFb in a novel manner to stimulate human T-lymphotropic virus type 1 transcription, *J. Virol.* 80 (2006) 4781–4791.
- [199] Y. Tian, S. Ke, M. Chen, T. Sheng, Interactions between the aryl hydrocarbon receptor and P-TEFb. Sequential recruitment of transcription factors and differential phosphorylation of C-terminal domain of RNA polymerase II at cyp1a1 promoter, *J. Biol. Chem.* 278 (2003) 44041–44048.
- [200] D.K. Lee, H.O. Duan, C. Chang, Androgen receptor interacts with the positive elongation factor P-TEFb and enhances the efficiency of transcriptional elongation, *J. Biol. Chem.* 276 (2001) 9978–9984.
- [201] T. Kurosu, B.M. Peterlin, VP16 and ubiquitin; binding of P-TEFb via its activation domain and ubiquitin facilitates elongation of transcription of target genes, *Curr. Biol.* 14 (2004) 1112–1116.
- [202] F. Zhang, M. Barboric, T.K. Blackwell, B.M. Peterlin, A model of repression: CTD analogs and PIE-1 inhibit transcriptional elongation by P-TEFb, *Genes Dev.* 17 (2003) 748–758.
- [203] H. Jiang, F. Zhang, T. Kurosu, B.M. Peterlin, Runx1 binds positive transcription elongation factor b and represses transcriptional elongation by RNA polymerase II: possible mechanism of CD4 silencing, *Mol. Cell. Biol.* 25 (2005) 10675–10683.
- [204] A. Guiguen, S. Soutourina, M. Dewez, L. Tafforeau, M. Dieu, M. Raes, J. Vandenhoute, M. Werner, D. Hermand, Recruitment of P-TEFb (Cdk9-Pch1) to chromatin by the cap-methyl transferase Pcm1 in fission yeast, *EMBO J.* 26 (2007) 1552–1559.
- [205] M. Barboric, F. Zhang, M. Besenica, A. Plemenitas, B.M. Peterlin, Ubiquitylation of Cdk9 by Skp2 facilitates optimal Tat transactivation, *J. Virol.* 79 (2005) 11135–11141.
- [206] J. Garriga, S. Bhattacharya, J. Calbo, R.M. Marshall, M. Truongcao, D.S. Haines, X. Grana, CDK9 is constitutively expressed throughout the cell cycle, and its steady-state expression is independent of SKP2, *Mol. Cell. Biol.* 23 (2003) 5165–5173.
- [207] R.E. Kiernan, S. Emiliani, K. Nakayama, A. Castro, J.C. Labbe, T. Lorca, K. Nakayama, K. Benkirane, Interaction between cyclin T1 and SCF(SKP2)

- targets CDK9 for ubiquitination and degradation by the proteasome, *Mol. Cell. Biol.* 21 (2001) 7956–7970.
- [208] M. Cojocaru, A. Bouchard, P. Cloutier, J.J. Cooper, K. Varzavand, D.H. Price, B. Coulombe, Transcription factor IIS cooperates with the E3 ligase UBR5 to ubiquitinate the CDK9 subunit of the positive transcription elongation factor B, *J. Biol. Chem.* 286 (2011) 5012–5022.
- [209] J. Fu, H.G. Yoon, J. Qin, J. Wong, Regulation of P-TEFb elongation complex activity by CDK9 acetylation, *Mol. Cell. Biol.* 27 (2007) 4641–4651.
- [210] S. Cho, S. Schroeder, K. Kaehele, H.S. Kwon, A. Pedal, E. Herker, M. Schnoelzer, M. Ott, Acetylation of cyclin T1 regulates the equilibrium between active and inactive P-TEFb in cells, *EMBO J.* 28 (2009) 1407–1417.
- [211] A. Sabo, M. Lusica, A. Cereseto, M. Giacca, Acetylation of conserved lysines in the catalytic core of cyclin-dependent kinase 9 inhibits kinase activity and regulates transcription, *Mol. Cell. Biol.* 28 (2008) 2201–2212.
- [212] M.F. Blank, S. Chen, F. Poetz, M. Schnolzer, R. Voit, I. Grummt, SIRT7-dependent deacetylation of CDK9 activates RNA polymerase II transcription, *Nucleic Acids Res.* 45 (2017) 2675–2686.
- [213] H. Zhang, S.H. Park, B.G. Pantazides, O. Karpiuk, M.D. Warren, C.W. Hardy, D.M. Duong, S.J. Park, H.S. Kim, A. Vassilopoulos, N.T. Seyfried, S.A. Johnsen, D. Gius, D.S. Yu, SIRT2 directs the replication stress response through CDK9 deacetylation, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 13546–13551.
- [214] E.M. Kosciuzuk, S. Mehrotra, D. Saleiro, B. Kroczyńska, B. Majchrzak-Kita, P. Lisowski, C. Driehaus, A. Rogalska, A. Turner, T. Lienhoop, D. Gius, E.N. Fish, A. Vassilopoulos, L.C. Platanias, Sirtuin 2-mediated deacetylation of cyclin-dependent kinase 9 promotes STAT1 signaling in type I interferon responses, *J. Biol. Chem.* 294 (2019) 827–837.
- [215] S. Giraud, A. Hurlstone, S. Avril, O. Coqueret, Implication of BRG1 and cdk9 in the STAT3-mediated activation of the p21waf1 gene, *Oncogene* 23 (2004) 7391–7398.
- [216] F. Yu, G. Shi, S. Cheng, J. Chen, S.Y. Wu, Z. Wang, N. Xia, Y. Zhai, Z. Wang, Y. Peng, D. Wang, J.X. Du, L. Liao, S.Z. Duan, T. Shi, J. Cheng, C.M. Chiang, J. Li, J. Wong, SUMO suppresses and MYC amplifies transcription globally by regulating CDK9 sumoylation, *Cell Res.* 28 (2018) 670–685.
- [217] X. Ma, T. Yang, Y. Luo, L. Wu, Y. Jiang, Z. Song, T. Pan, B. Liu, G. Liu, J. Liu, F. Yu, Z. He, W. Zhang, J. Yang, L. Liang, Y. Guan, X. Zhang, L. Li, W. Cai, X. Tang, S. Gao, K. Deng, H. Zhang, TRIM28 promotes HIV-1 latency by SUMOylating CDK9 and inhibiting P-TEFb, *Elife* 8 (2019).
- [218] M. Sanso, R.S. Levin, J.J. Lipp, V.Y. Wang, A.K. Greifengberg, E.M. Quezada, A. Ali, A. Ghosh, S. Larochele, T.M. Rana, M. Geyer, L. Tong, K.M. Shokat, R.P. Fisher, P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates, *Genes Dev.* 30 (2016) 117–131.
- [219] P. Brauns-Schubert, F. Schubert, M. Wissler, M. Weiss, L. Schlicher, S. Bessler, M. Safavi, C. Miething, C. Borner, T. Brummer, U. Maurer, CDK9-mediated phosphorylation controls the interaction of TIP60 with the transcriptional machinery, *EMBO Rep.* 19 (2018) 244–256.
- [220] J. Ou, H. Li, P. Qiu, Q. Li, H.C. Chang, Y.C. Tang, CDK9 modulates circadian clock by attenuating REV-ERB α activity, *Biochem. Biophys. Res. Commun.* 513 (2019) 967–973.
- [221] P.K. Parua, G.T. Booth, M. Sanso, B. Benjamin, J.C. Tanny, J.T. Lis, R.P. Fisher, A Cdk9-PP1 switch regulates the elongation-termination transition of RNA polymerase II, *Nature* 558 (2018) 460–464.
- [222] I. Oven, N. Brdiczka, J. Kohoutek, T. Vaupotic, M. Narat, B.M. Peterlin, AIRE recruits P-TEFb for transcriptional elongation of target genes in medullary thymic epithelial cells, *Mol. Cell. Biol.* 27 (2007) 8815–8823.
- [223] B. Shan, Y. Zhuo, D. Chin, C.A. Morris, G.F. Morris, J.A. Lasky, Cyclin-dependent kinase 9 is required for tumor necrosis factor- α -stimulated matrix metalloproteinase-9 expression in human lung adenocarcinoma cells, *J. Biol. Chem.* 280 (2005) 1103–1111.
- [224] N. Meier, S. Krpic, P. Rodriguez, J. Strouboulis, M. Monti, J. Krijgsveld, M. Gering, R. Patient, A. Hostert, F. Grosfeld, Novel binding partners of Ldb1 are required for haematopoietic development, *Development* 133 (2006) 4913–4923.
- [225] A. Gegonne, J.D. Weissman, H. Lu, M. Zhou, A. Dasgupta, R. Ribble, J.N. Brady, D.S. Singer, TFIID component TAF7 functionally interacts with both TFIIF and P-TEFb, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 5367–5372.
- [226] D. Breuer, A. Kotelkin, T. Ammosova, N. Kumari, A. Ivanov, A.V. Ilatovskiy, M. Beullens, P.R. Roane, M. Bollen, M.G. Petukhov, F. Kashanchi, S. Nekhai, CDK2 regulates HIV-1 transcription by phosphorylation of CDK9 on serine 90, *Retrovirology* 9 (2012) 94.
- [227] B. O'Keeffe, Y. Fong, D. Chen, S. Zhou, Q. Zhou, Requirement for a kinase-specific chaperone pathway in the production of a Cdk9/cyclin T1 heterodimer responsible for P-TEFb-mediated tat stimulation of HIV-1 transcription, *J. Biol. Chem.* 275 (2000) 279–287.
- [228] S.K. Radhakrishnan, A.L. Gartel, CDK9 phosphorylates p53 on serine residues 33, 315 and 392, *Cell Cycle* 5 (2006) 519–521.
- [229] T. Taki, H. Kano, M. Taniwaki, M. Sako, M. Yanagisawa, Y. Hayashi, AF5q31, a newly identified AF4-related gene, is fused to MLL in infant acute lymphoblastic leukemia with ins(5;11)(q31;q13q23), *Proc. Natl. Acad. Sci. U. S. A.* 96 (1999) 14535–14540.
- [230] L. Chakrabarti, S.J. Knight, A.V. Flannery, K.E. Davies, A candidate gene for mild mental handicap at the FRAXE fragile site, *Hum. Mol. Genet.* 5 (1996) 275–282.
- [231] A.R. von Bergh, H.B. Beverloo, P. Rombout, E.R. van Wering, M.H. van Weel, G.C. Beverstock, P.M. Kluin, R.M. Slater, E. Schuurings, LAF4, an AF4-related gene, is fused to MLL in infant acute lymphoblastic leukemia, *Genes Chromosom. Cancer* 35 (2002) 92–96.
- [232] I. Nilsson, K. Lochner, G. Siegler, J. Greil, J.D. Beck, G.H. Fey, R. Marschalek, Exon/intron structure of the human ALL-1 (MLL) gene involved in translocations to chromosomal region 11q23 and acute leukaemias, *Br. J. Haematol.* 93 (1996) 966–972.
- [233] I. Nilsson, M. Reichel, M.G. Ennas, R. Greim, C. Knorr, G. Siegler, J. Greil, G.H. Fey, R. Marschalek, Exon/intron structure of the human AF-4 gene, a member of the AF-4/LAF-4/FMR-2 gene family coding for a nuclear protein with structural alterations in acute leukaemia, *Br. J. Haematol.* 98 (1997) 157–169.
- [234] R. Prasad, T. Yano, C. Sorio, T. Nakamura, R. Rallapalli, Y. Gu, D. Leshkowitz, C.M. Croce, E. Canaani, Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia, *Proc. Natl. Acad. Sci. U. S. A.* 92 (1995) 12160–12164.
- [235] C. Ma, L.M. Staudt, LAF-4 encodes a lymphoid nuclear protein with transactivation potential that is homologous to AF-4, the gene fused to MLL in t(4;11) leukemias, *Blood* 87 (1996) 734–745.
- [236] K. Baskaran, F. Erfurth, G. Taborn, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, P.M. Iannaccone, P.H. Dumer, Cloning and developmental expression of the murine homolog of the acute leukemia proto-oncogene AF4, *Oncogene* 15 (1997) 1967–1978.
- [237] M.A. Hillman, J. Gecz, Fragile XE-associated familial mental retardation protein 2 (FMR2) acts as a potent transcription activator, *J. Hum. Genet.* 46 (2001) 251–259.
- [238] A. Daser, T.H. Rabbitts, Extending the repertoire of the mixed-lineage leukemia gene MLL in leukemogenesis, *Genes Dev.* 18 (2004) 965–974.
- [239] M.C. Estable, M.H. Naghavi, H. Kato, H. Xiao, J. Qin, A. Vahline, R.G. Roeder, MCEF, the newest member of the AF4 family of transcription factors involved in leukemia, is a positive transcription elongation factor-b-associated protein, *J. Biomed. Sci.* 9 (2002) 234–245.
- [240] M.C. Good, J.G. Zalatan, W.A. Lim, Scaffold proteins: hubs for controlling the flow of cellular information, *Science* 332 (2011) 680–686.
- [241] S. Qi, Z. Li, U. Schulze-Gahmen, G. Stjepanovic, Q. Zhou, J.H. Hurley, Structural basis for ELL2 and AFF4 activation of HIV-1 proviral transcription, *Nat. Commun.* 8 (2017) 14076.
- [242] B.I. Leach, A. Kuntimaddi, C.R. Schmidt, T. Cierpicki, S.A. Johnson, J.H. Bushweller, Leukemia fusion target AF9 is an intrinsically disordered transcriptional regulator that recruits multiple partners via coupled folding and binding, *Structure* 21 (2013) 176–183.
- [243] N. Kumari, M.A. Hassan, X. Lu, R.G. Roeder, D. Biswas, AFF1 acetylation by p300 temporally inhibits transcription during genotoxic stress response, *Proc. Natl. Acad. Sci. U. S. A.* 116 (2019) 22140–22151.
- [244] E. Bitoun, P.L. Oliver, K.E. Davies, The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling, *Hum. Mol. Genet.* 16 (2007) 92–106.
- [245] Z. Li, H. Lu, Q. Zhou, A minor subset of super elongation complexes plays a predominant role in reversing HIV-1 latency, *Mol. Cell. Biol.* 36 (2016) 1194–1205.
- [246] Y. Chen, P. Cramer, Structure of the super-elongation complex subunit AFF4 C-terminal homology domain reveals requirements for AFF homo- and hetero-dimerization, *J. Biol. Chem.* 294 (2019) 10663–10673.
- [247] A. Bugai, A.J.C. Quaresma, C.C. Friedel, T. Lenasi, R. Duster, C.R. Sibley, K. Fujinaga, P. Kukanja, T. Hennig, M. Blasius, M. Geyer, J. Ule, L. Dolken, M. Barboric, P-TEFb activation by RBM7 shapes a pro-survival transcriptional response to genotoxic stress, *Mol. Cell* 74 (e210) (2019) 254–267.
- [248] H.G. Super, P.L. Strissel, O.M. Sobulo, D. Burian, S.C. Reshmi, B. Roe, N.J. Zeleznik-Le, M.O. Diaz, J.D. Rowley, Identification of complex genomic breakpoint junctions in the t(9;11) MLL-AF9 fusion gene in acute leukemia, *Genes Chromosom. Cancer* 20 (1997) 185–195.
- [249] D.C. Tkachuk, S. Kohler, M.L. Cleary, Involvement of a homolog of *Drosophila* trithorax by 11q23 chromosomal translocations in acute leukemias, *Cell* 71 (1992) 691–700.
- [250] D. Mueller, C. Bach, D. Zeisig, M.P. Garcia-Cuellar, S. Monroe, A. Sreekumar, R. Zhou, A. Nesvizhskii, A. Chinnaiyan, J.L. Hess, R.K. Slany, A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification, *Blood* 110 (2007) 4445–4454.
- [251] M.P. Garcia-Cuellar, S.A. Schreiner, M. Birke, M. Hamacher, G.H. Fey, R.K. Slany, ENL, the MLL fusion partner in t(11;19), binds to the c-Abl interactor protein 1 (ABI1) that is fused to MLL in t(10;11)+, *Oncogene* 19 (2000) 1744–1751.
- [252] H. Okuda, B. Stanojevic, A. Kanai, T. Kawamura, S. Takahashi, H. Matsui, A. Takaori-Kondo, A. Yokoyama, Cooperative gene activation by AF4 and DOT1L drives MLL-rearranged leukemia, *J. Clin. Invest.* 127 (2017) 1918–1931.
- [253] G. Mitterbauer, C. Zimmer, C. Fonatsch, O. Haas, R. Thalhammer-Scherrer, I. Schwarzwinger, P. Kalhs, U. Jaeger, K. Lechner, C. Mannhalter, Monitoring of minimal residual leukemia in patients with MLL-AF9 positive acute myeloid leukemia by RT-PCR, *Leukemia* 13 (1999) 1519–1524.
- [254] C.M. Palermo, C.A. Bennett, A.C. Winters, C.S. Hemenway, The AF4-mimetic peptide, PFWT, induces necrotic cell death in MV4-11 leukemia cells, *Leuk. Res.* 32 (2008) 633–642.
- [255] S.J. Demarest, M. Martinez-Yamout, J. Chung, H. Chen, W. Xu, H.J. Dyson, R.M. Evans, P.E. Wright, Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators, *Nature* 415 (2002) 549–553.
- [256] R.N. De Guzman, M.A. Martinez-Yamout, H.J. Dyson, P.E. Wright, Interaction of the TAZ1 domain of the CREB-binding protein with the activation domain of CITED2: regulation by competition between intrinsically unstructured ligands for non-identical binding sites, *J. Biol. Chem.* 279 (2004) 3042–3049.
- [257] L. Wan, H. Wen, Y. Li, J. Lyu, Y. Xi, T. Hoshii, J.K. Joseph, X. Wang, Y.E. Loh, M.A. Erb, A.L. Souza, J.E. Bradner, L. Shen, W. Li, H. Li, C.D. Allis, S.A. Armstrong, X. Shi, ENL links histone acetylation to oncogenic gene expression in acute myeloid leukaemia, *Nature* 543 (2017) 265–269.
- [258] M.A. Erb, T.G. Scott, B.E. Li, H. Xie, J. Paulk, H.S. Seo, A. Souza, J.M. Roberts, S. Dastjerdi, D.L. Buckley, N.E. Sanjana, O. Shalem, B. Nabet, R. Zeid, N.K. Offei-

- Addo, S. Dhe-Paganon, F. Zhang, S.H. Orkin, G.E. Winter, J.E. Bradner, Transcription control by the ENL YEATS domain in acute leukaemia, *Nature* 543 (2017) 270–274.
- [259] D.T. Zeisig, C.B. Bittner, B.B. Zeisig, M.P. Garcia-Cuellar, J.L. Hess, R.K. Slany, The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin, *Oncogene* 24 (2005) 5525–5532.
- [260] S. Spicuglia, L. Vanhille, Chromatin signatures of active enhancers, *Nucleus* 3 (2012) 126–131.
- [261] S. Bonn, R.P. Zinzen, C. Girardot, E.H. Gustafson, A. Perez-Gonzalez, N. Delhomme, Y. Ghavi-Helm, B. Wilczynski, A. Riddell, E.E. Furlong, Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development, *Nat. Genet.* 44 (2012) 148–156.
- [262] C. Lin, A.S. Garruss, Z. Luo, F. Guo, A. Shilatifard, The RNA Pol II elongation factor ELL3 marks enhancers in ES cells and primes future gene activation, *Cell* 152 (2013) 144–156.
- [263] C.L. Dobson, A.J. Warren, R. Pannell, A. Forster, I. Lavenir, J. Corral, A.J. Smith, T.H. Rabbitts, The mll-AF9 gene fusion in mice controls myeloproliferation and specifies acute myeloid leukaemogenesis, *EMBO J.* 18 (1999) 3564–3574.
- [264] J. Corral, A. Forster, S. Thompson, F. Lampert, Y. Kaneko, R. Slater, W.G. Kroes, C.E. van der Schoot, W.D. Ludwig, A. Karpas, et al., Acute leukemias of different lineages have similar MLL gene fusions encoding related chimeric proteins resulting from chromosomal translocation, *Proc. Natl. Acad. Sci. U. S. A.* 90 (1993) 8538–8542.
- [265] S. Iida, M. Seto, K. Yamamoto, H. Komatsu, A. Tojo, S. Asano, N. Kamada, Y. Ariyoshi, T. Takahashi, R. Ueda, MLLT3 gene on 9p22 involved in t(9;11) leukemia encodes a serine/proline rich protein homologous to MLLT1 on 19p13, *Oncogene* 8 (1993) 3085–3092.
- [266] C. Pina, G. May, S. Soneji, D. Hong, T. Enver, MLLT3 regulates early human erythroid and megakaryocytic cell fate, *Cell Stem Cell* 2 (2008) 264–273.
- [267] E.K. Shanle, F.H. Andrews, H. Meriesh, S.L. McDaniel, R. Dronamraju, J.V. DiFiore, D. Jha, G.G. Wozniak, J.B. Bridges, J.L. Kerschner, K. Krajewski, G.M. Martin, A.J. Morrison, T.G. Kutateladze, B.D. Strahl, Association of Taf14 with acetylated histone H3 directs gene transcription and the DNA damage response, *Genes Dev.* 29 (2015) 1795–1800.
- [268] C. Dhalluin, J.E. Carlson, L. Zeng, C. He, A.K. Aggarwal, M.M. Zhou, Structure and ligand of a histone acetyltransferase bromodomain, *Nature* 399 (1999) 491–496.
- [269] L. Zeng, Q. Zhang, S. Li, A.N. Plotnikov, M.J. Walsh, M.M. Zhou, Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b, *Nature* 466 (2010) 258–262.
- [270] B.R. Sabari, Z. Tang, H. Huang, V. Yong-Gonzalez, H. Molina, H.E. Kong, L. Dai, M. Shimada, J.R. Cross, Y. Zhao, R.G. Roeder, C.D. Allis, Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation, *Mol. Cell* 58 (2015) 203–215.
- [271] Y. Li, B.R. Sabari, T. Panchenko, H. Wen, D. Zhao, H. Guan, L. Wan, H. Huang, Z. Tang, Y. Zhao, R.G. Roeder, X. Shi, C.D. Allis, H. Li, Molecular coupling of histone crotonylation and active transcription by AF9 YEATS domain, *Mol. Cell* 62 (2016) 181–193.
- [272] M. Tan, H. Luo, S. Lee, F. Jin, J.S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, Z. Lu, Z. Ye, Q. Zhu, J. Wysocka, Y. Ye, S. Khochbin, B. Ren, Y. Zhao, Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification, *Cell* 146 (2011) 1016–1028.
- [273] E.M. Flynn, O.W. Huang, F. Poy, M. Oppikofer, S.F. Bellon, Y. Tang, A.G. Cochran, A subset of human bromodomains recognizes butyryllysine and crotonyllysine histone peptide modifications, *Structure* 23 (2015) 1801–1814.
- [274] Q. Zhang, L. Zeng, C. Zhao, Y. Ju, T. Konuma, M.M. Zhou, Structural insights into histone crotonyl-lysine recognition by the AF9 YEATS domain, *Structure* 24 (2016) 1606–1612.
- [275] A.V. Krivtsov, S.A. Armstrong, MLL translocations, histone modifications and leukaemia stem-cell development, *Nat. Rev. Cancer* 7 (2007) 823–833.
- [276] A. Shilatifard, W.S. Lane, K.W. Jackson, R.C. Conaway, J.W. Conaway, An RNA polymerase II elongation factor encoded by the human ELL gene, *Science* 271 (1996) 1873–1876.
- [277] M.J. Thirman, E.B. Diskin, S.S. Bin, H.S. Ip, J.M. Miller, M.C. Simon, Developmental analysis and subcellular localization of the murine homologue of ELL, *Proc. Natl. Acad. Sci. U. S. A.* 94 (1997) 1408–1413.
- [278] A. Shilatifard, D. Haque, R.C. Conaway, J.W. Conaway, Structure and function of RNA polymerase II elongation factor ELL. Identification of two overlapping ELL functional domains that govern its interaction with polymerase and the ternary elongation complex, *J. Biol. Chem.* 272 (1997) 22355–22363.
- [279] A. Shilatifard, Identification and purification of the Holo-ELL complex. Evidence for the presence of ELL-associated proteins that suppress the transcriptional inhibitory activity of ELL, *J. Biol. Chem.* 273 (1998) 11212–11217.
- [280] A. Shilatifard, Identification and purification of the Holo-ELL complex - evidence for the presence of ELL-associated proteins that suppress the transcriptional inhibitory activity of ELL, *J. Biol. Chem.* 273 (1998) 11212–11217.
- [281] F. Simone, P.E. Polak, J.J. Kaberlein, R.T. Luo, D.A. Levitan, M.J. Thirman, EAF1, a novel ELL-associated factor that is delocalized by expression of the MLL-ELL fusion protein, *Blood* 98 (2001) 201–209.
- [282] F. Simone, R.T. Luo, P.E. Polak, J.J. Kaberlein, M.J. Thirman, ELL-associated factor 2 (EAF2), a functional homolog of EAF1 with alternative ELL binding properties, *Blood* 101 (2003) 2355–2362.
- [283] R.T. Luo, C. Lavau, C. Du, F. Simone, P.E. Polak, S. Kawamata, M.J. Thirman, The elongation domain of ELL is dispensable but its ELL-associated factor 1 interaction domain is essential for MLL-ELL-induced leukemogenesis, *Mol. Cell. Biol.* 21 (2001) 5678–5687.
- [284] P.E. Polak, F. Simone, J.J. Kaberlein, R.T. Luo, M.J. Thirman, ELL and EAF1 are Cajal body components that are disrupted in MLL-ELL leukemia, *Mol. Biol. Cell* 14 (2003) 1517–1528.
- [285] D. Stanek, Cajal bodies and snRNPs - friends with benefits, *RNA Biol.* 14 (2017) 671–679.
- [286] D. Hu, E.R. Smith, A.S. Garruss, N. Mohaghegh, J.M. Varberg, C. Lin, J. Jackson, X. Gao, A. Saraf, L. Florens, M.P. Washburn, J.C. Eissenberg, A. Shilatifard, The little elongation complex functions at initiation and elongation phases of snRNA gene transcription, *Mol. Cell* 51 (2013) 493–505.
- [287] S.E. Kong, C.A. Banks, A. Shilatifard, J.W. Conaway, R.C. Conaway, ELL-associated factors 1 and 2 are positive regulators of RNA polymerase II elongation factor ELL, *Proc. Natl. Acad. Sci. U. S. A.* 102 (2005) 10094–10098.
- [288] N.F. Marshall, D.H. Price, Purification of P-TEFb, a transcription factor required for the transition into productive elongation, *J. Biol. Chem.* 270 (1995) 12335–12338.
- [289] A. Saunders, L.J. Core, J.T. Lis, Breaking barriers to transcription elongation, *Nat. Rev. Mol. Cell Biol.* 7 (2006) 557–567.
- [290] E.R. Smith, B. Winter, J.C. Eissenberg, A. Shilatifard, Regulation of the transcriptional activity of poised RNA polymerase II by the elongation factor ELL, *Proc. Natl. Acad. Sci. U. S. A.* 105 (2008) 8575–8579.
- [291] J. Zhou, X. Feng, B. Ban, J. Liu, Z. Wang, W. Xiao, Elongation factor ELL (eleven-nineteen lysine-rich leukemia) acts as a transcription factor for direct thrombospondin-1 regulation, *J. Biol. Chem.* 284 (2009) 19142–19152.
- [292] V. Saha, T. Chaplin, A. Gregorini, P. Ayton, B.D. Young, The leukemia-associated-protein (LAP) domain, a cysteine-rich motif, is present in a wide range of proteins, including MLL, AF10, and MLLT6 proteins, *Proc. Natl. Acad. Sci. U. S. A.* 92 (1995) 9737–9741.
- [293] T. Chaplin, O. Bernard, H.B. Beverloo, V. Saha, A. Hagemeijer, R. Berger, B.D. Young, The t(10;11) translocation in acute myeloid leukemia (M5) consistently fuses the leucine zipper motif of AF10 onto the HRX gene, *Blood* 86 (1995) 2073–2076.
- [294] B. Linder, R. Newman, L.K. Jones, S. Debernardi, B.D. Young, P. Freemont, C.P. Verrijzer, V. Saha, Biochemical analyses of the AF10 protein: the extended LAP/PHD-finger mediates oligomerisation, *J. Mol. Biol.* 299 (2000) 369–378.
- [295] S.M. Bahri, W. Chia, X. Yang, The Drosophila homolog of human AF10/AF17 leukemia fusion genes (Dalf) encodes a zinc finger/leucine zipper nuclear protein required in the nervous system for maintaining EVE expression and normal growth, *Mech. Dev.* 100 (2001) 291–301.
- [296] X. Wang, C.W. Chen, S.A. Armstrong, The role of DOT1L in the maintenance of leukemia gene expression, *Curr. Opin. Genet. Dev.* 36 (2016) 68–72.
- [297] K. Wood, M. Tellier, S. Murphy, DOT1L and H3K79 methylation in transcription and genomic stability, *Biomolecules* 8 (2018).
- [298] E.M. Stein, M.S. Tallman, Mixed lineage rearranged leukaemia: pathogenesis and targeting DOT1L, *Curr. Opin. Hematol.* 22 (2015) 92–96.
- [299] L. Chen, A.J. Deshpande, D. Banka, K.M. Bernt, S. Dias, C. Buske, E.J. Olhava, S.R. Daigle, V.M. Richon, R.M. Pollock, S.A. Armstrong, Abrogation of MLL-AF10 and CALM-AF10-mediated transformation through genetic inactivation or pharmacological inhibition of the H3K79 methyltransferase Dot1l, *Leukemia* 27 (2013) 813–822.
- [300] Y. Okada, Q. Jiang, M. Lemieux, L. Jeannotte, L. Su, Y. Zhang, Leukaemic transformation by CALM-AF10 involves upregulation of Hoxa5 by hDOT1L, *Nat. Cell Biol.* 8 (2006) 1017–1024.
- [301] T. Mahmoudi, S.F. Boj, P. Hatzis, V.S. Li, N. Taouatas, R.G. Vries, H. Teunissen, H. Begthel, J. Korving, S. Mohammed, A.J. Heck, H. Clevers, The leukemia-associated Mllt10/Af10-Dot1l are Tcf4/beta-catenin coactivators essential for intestinal homeostasis, *PLoS Biol.* 8 (2010) e1000539.
- [302] D.C. Avgousti, G. Cecere, A. Grishok, The conserved PHD1-PHD2 domain of ZFP-1/AF10 is a discrete functional module essential for viability in *Caenorhabditis elegans*, *Mol. Cell. Biol.* 33 (2013) 999–1015.
- [303] A.J. Deshpande, A. Deshpande, A.U. Sinha, L. Chen, J. Chang, A. Cihan, M. Fazio, C.W. Chen, N. Zhu, R. Koche, L. Dzhekevich, G. Ibanez, S. Dias, D. Banka, A. Krivtsov, M. Luo, R.G. Roeder, J.E. Bradner, K.M. Bernt, S.A. Armstrong, AF10 regulates progressive H3K79 methylation and Hox gene expression in diverse AML subtypes, *Cancer Cell* 26 (2014) 896–908.
- [304] S. Chen, Z. Yang, A.W. Wilkinson, A.J. Deshpande, S. Sidoli, K. Krajewski, B.D. Strahl, B.A. Garcia, S.A. Armstrong, D.J. Patel, O. Gozani, The PZP domain of AF10 senses unmodified H3K27 to regulate DOT1L-mediated methylation of H3K79, *Mol. Cell* 60 (2015) 319–327.
- [305] J. Kim, M. Guermah, R.K. McGinty, J.S. Lee, Z. Tang, T.A. Milne, A. Shilatifard, T.W. Muir, R.G. Roeder, RAD6-mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells, *Cell* 137 (2009) 459–471.
- [306] R.K. McGinty, J. Kim, C. Chatterjee, R.G. Roeder, T.W. Muir, Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation, *Nature* 453 (2008) 812–816.
- [307] R. Prasad, D. Leshkowitz, Y. Gu, H. Alder, T. Nakamura, H. Saito, K. Huebner, R. Berger, C.M. Croce, E. Canaani, Leucine-zipper dimerization motif encoded by the AF17 gene fused to ALL-1 (MLL) in acute leukemia, *Proc. Natl. Acad. Sci. U. S. A.* 91 (1994) 8107–8111.
- [308] Y.M. Lin, K. Ono, S. Satoh, H. Ishiguro, M. Fujita, N. Miwa, T. Tanaka, T. Tsunoda, K.C. Yang, Y. Nakamura, Y. Furukawa, Identification of AF17 as a downstream gene of the beta-catenin/T-cell factor pathway and its involvement in colorectal carcinogenesis, *Cancer Res.* 61 (2001) 6345–6349.
- [309] M.I. Zhou, H. Wang, R.L. Foy, J.J. Ross, H.T. Cohen, Tumor suppressor von Hippel-Lindau (VHL) stabilization of Jade-1 protein occurs through plant homeodomains and is VHL mutation dependent, *Cancer Res.* 64 (2004) 1278–1286.
- [310] L. Chen, H. Wu, O.M. Pochynyuk, M.R. Reisenauer, Z. Zhang, L. Huang, O.L. Zaika, M. Mamenko, W. Zhang, Q. Zhou, M. Liu, Y. Xia, W. Zhang, Af17 deficiency

- increases sodium excretion and decreases blood pressure, *J. Am. Soc. Nephrol.* 22 (2011) 1076–1086.
- [311] C.P. Ponting, D.R. Benjamin, A novel family of Ras-binding domains, *Trends Biochem. Sci.* 21 (1996) 422–425.
- [312] M. Buchert, C. Poon, J.A. King, T. Baechli, G. D'Abaco, F. Hollande, C.M. Hovens, AF6/s-afadin is a dual residency protein and localizes to a novel subnuclear compartment, *J. Cell. Physiol.* 210 (2007) 212–223.
- [313] T. Yamamoto, N. Harada, Y. Kawano, S. Taya, K. Kaibuchi, In vivo interaction of AF-6 with activated Ras and ZO-1, *Biochem. Biophys. Res. Commun.* 259 (1999) 103–107.
- [314] S. Taya, T. Yamamoto, M. Kanai-Azuma, S.A. Wood, K. Kaibuchi, The deubiquitinating enzyme fam interacts with and stabilizes beta-catenin, *Genes Cells* 4 (1999) 757–767.
- [315] K. Tachibana, H. Nakanishi, K. Mandai, K. Ozaki, W. Ikeda, Y. Yamamoto, A. Nagafuchi, S. Tsukita, Y. Takai, Two cell adhesion molecules, nectin and cadherin, interact through their cytoplasmic domain-associated proteins, *J. Cell Biol.* 150 (2000) 1161–1176.
- [316] A.B. Zhadanov, D.W. Provance Jr., C.A. Speer, J.D. Coffin, D. Goss, J.A. Blixt, C.M. Reichert, J.A. Mercer, Absence of the tight junctional protein AF-6 disrupts epithelial cell-cell junctions and cell polarity during mouse development, *Curr. Biol.* 9 (1999) 880–888.
- [317] N. Manabe, S. Hirai, F. Imai, H. Nakanishi, Y. Takai, S. Ohno, Association of ASIP/mPAR-3 with adherens junctions of mouse neuroepithelial cells, *Dev. Dyn.* 225 (2002) 61–69.
- [318] J.E. VanLeeuwen, I. Rafalovich, K. Sellers, K.A. Jones, T.N. Griffith, R. Huda, R.J. Miller, D.P. Srivastava, P. Penzes, Coordinated nuclear and synaptic shuttling of afadin promotes spine plasticity and histone modifications, *J. Biol. Chem.* 289 (2014) 10831–10842.
- [319] T. Kanazawa, O. Ichii, S. Otsuka, Y. Namiki, Y. Hashimoto, Y. Kon, Hepatocyte nuclear factor 4 alpha is associated with mesenchymal-epithelial transition in developing kidneys of C57BL/6 mice, *J. Vet. Med. Sci.* 73 (2011) 601–607.
- [320] K. Saito, T. Shiino, H. Kurihara, Y. Harita, S. Hattori, Y. Ohta, Afadin regulates RhoA/rho-associated protein kinase signaling to control formation of actin stress fibers in kidney podocytes, *Cytoskeleton (Hoboken)* 72 (2015) 146–156.
- [321] K. Sano, Structure of AF3p21, a new member of mixed lineage leukemia (MLL) fusion partner proteins-implication for MLL-induced leukemogenesis, *Leuk. Lymphoma* 42 (2001) 595–602.
- [322] J.F. DiMartino, P.M. Ayton, E.H. Chen, C.C. Naftzger, B.D. Young, M.L. Cleary, The AF10 leucine zipper is required for leukemic transformation of myeloid progenitors by MLL-AF10, *Blood* 99 (2002) 3780–3785.
- [323] R. Ono, T. Taki, T. Taketani, M. Taniwaki, H. Kobayashi, Y. Hayashi, LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23), *Cancer Res.* 62 (2002) 4075–4080.
- [324] R. Ono, T. Taki, T. Taketani, H. Kawaguchi, M. Taniwaki, T. Okamura, K. Kawa, R. Hanada, M. Kobayashi, Y. Hayashi, SEPTIN6, a human homologue to mouse Septin6, is fused to MLL in infant acute myeloid leukemia with complex chromosomal abnormalities involving 11q23 and Xq24, *Cancer Res.* 62 (2002) 333–337.
- [325] H. Liu, B. Chen, H. Xiong, Q.H. Huang, Q.H. Zhang, Z.G. Wang, B.L. Li, Z. Chen, S.J. Chen, Functional contribution of EEN to leukemogenic transformation by MLL-EEN fusion protein, *Oncogene* 23 (2004) 3385–3394.
- [326] C.W. So, H. Karsunky, E. Passegue, A. Cozzio, L.L. Weissman, M.L. Cleary, MLL-GAS7 transforms multipotent hematopoietic progenitors and induces mixed lineage leukemias in mice, *Cancer Cell* 3 (2003) 161–171.
- [327] C.W. So, M. Lin, P.M. Ayton, E.H. Chen, M.L. Cleary, Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias, *Cancer Cell* 4 (2003) 99–110.
- [328] M.E. Martin, T.A. Milne, S. Bloyer, K. Galoian, W. Shen, D. Gibbs, H.W. Brock, R. Slany, J.L. Hess, Dimerization of MLL fusion proteins immortalizes hematopoietic cells, *Cancer Cell* 4 (2003) 197–207.
- [329] S. Boffo, A. Damato, L. Alfano, A. Giordano, CDK9 inhibitors in acute myeloid leukemia, *J. Exp. Clin. Cancer Res.* 37 (2018) 36.
- [330] L. Lyle, N. Daver, Current and emerging therapies for patients with acute myeloid leukemia: a focus on MCL-1 and the CDK9 pathway, *Am. J. Manag. Care* 24 (2018) S356–S365.
- [331] N.F. Marshall, D.H. Price, Control of formation of two distinct classes of RNA polymerase II elongation complexes, *Mol. Cell. Biol.* 12 (1992) 2078–2090.
- [332] Q. Zhou, D. Chen, E. Pierstorff, K. Luo, Transcription elongation factor P-TEFb mediates tat activation of HIV-1 transcription at multiple stages, *EMBO J.* 17 (1998) 3681–3691.
- [333] J. Peng, N.F. Marshall, D.H. Price, Identification of a cyclin subunit required for the function of *Drosophila* P-TEFb, *J. Biol. Chem.* 273 (1998) 13855–13860.
- [334] J. Peng, Y. Zhu, J.T. Milton, D.H. Price, Identification of multiple cyclin subunits of human P-TEFb, *Genes Dev.* 12 (1998) 755–762.
- [335] P. Wei, M.E. Garber, S.M. Fang, W.H. Fischer, K.A. Jones, A novel CDK9-associated C-type cyclin interacts directly with HIV-1 tat and mediates its high-affinity, loop-specific binding to TAR RNA, *Cell* 92 (1998) 451–462.
- [336] J.T. Lis, P. Mason, J. Peng, D.H. Price, J. Werner, P-TEFb kinase recruitment and function at heat shock loci, *Genes Dev.* 14 (2000) 792–803.
- [337] G. Napolitano, B. Majello, L. Lania, Catalytic activity of Cdk9 is required for nuclear co-localization of the Cdk9/cyclin T1 (P-TEFb) complex, *J. Cell. Physiol.* 197 (2003) 1–7.
- [338] M. Barboric, J.H. Yik, N. Czudnochowski, Z. Yang, R. Chen, X. Contreras, M. Geyer, B. Matija Peterlin, Q. Zhou, Tat competes with HEXIM1 to increase the active pool of P-TEFb for HIV-1 transcription, *Nucleic Acids Res.* 35 (2007) 2003–2012.
- [339] D.E. Nowak, B. Tian, M. Jamaluddin, I. Boldogh, L.A. Vergara, S. Choudhary, A.R. Brasier, RelA Ser276 phosphorylation is required for activation of a subset of NF-kappaB-dependent genes by recruiting cyclin-dependent kinase 9/cyclin T1 complexes, *Mol. Cell. Biol.* 28 (2008) 3623–3638.
- [340] T. Fujita, S. Ryser, I. Piuze, W. Schlegel, Up-regulation of P-TEFb by the MEK1-extracellular signal-regulated kinase signaling pathway contributes to stimulated transcription elongation of immediate early genes in neuroendocrine cells, *Mol. Cell. Biol.* 28 (2008) 1630–1643.
- [341] M.A. Cianfrocco, G.A. Kassavetis, P. Grob, J. Fang, T. Juven-Gershon, J.T. Kadonaga, E. Nogales, Human TFIID binds to core promoter DNA in a re-organized structural state, *Cell* 152 (2013) 120–131.
- [342] C. Plaschka, M. Hantsche, C. Dienemann, C. Burzinski, J. Plitzko, P. Cramer, Transcription initiation complex structures elucidate DNA opening, *Nature* 533 (2016) 353–358.
- [343] W. Shao, J. Zeitlinger, Paused RNA polymerase II inhibits new transcriptional initiation, *Nat. Genet.* 49 (2017) 1045–1051.
- [344] A.M. Isaacs, P.L. Oliver, E.L. Jones, A. Jeans, A. Potter, B.H. Hovik, P.M. Nolan, L. Vizor, P. Glenister, A.K. Simon, I.C. Gray, N.K. Spurr, S.D. Brown, A.J. Hunter, K.E. Davies, A mutation in AF4 is predicted to cause cerebellar ataxia and cataracts in the robotic mouse, *J. Neurosci.* 23 (2003) 1631–1637.
- [345] E. Bitoun, M.J. Finelli, P.L. Oliver, S. Lee, K.E. Davies, AF4 is a critical regulator of the IGF-1 signaling pathway during Purkinje cell development, *J. Neurosci.* 29 (2009) 15366–15374.
- [346] P.L. Oliver, D.A. Keays, K.E. Davies, Behavioural characterisation of the robotic mouse mutant, *Behav. Brain Res.* 181 (2007) 239–247.
- [347] E. Bitoun, K.E. Davies, The robotic mouse: understanding the role of AF4, a co-factor of transcriptional elongation and chromatin remodelling, in purkinje cell function, *Cerebellum* 8 (2009) 175–183.
- [348] E. Arabi, M. Garshabi, A.R. Jamshidi, R. Khalesi, N. Ahmadzadeh, M. Akbarian, M. Mahmoudi, Association study of an AFF1 gene polymorphism (rs340630) with Iranian systemic lupus erythematosus patients, *Acta Reumatol. Port.* 41 (2016) 68–73.
- [349] H. Cen, R.X. Leng, W. Wang, M. Zhou, C.C. Feng, G.M. Chen, R. Li, H.F. Pan, X.P. Li, D.Q. Ye, Association of AFF1 rs340630 and AFF3 rs10865035 polymorphisms with systemic lupus erythematosus in a Chinese population, *Immunogenetics* 64 (2012) 935–938.
- [350] Y. Okada, K. Shimane, Y. Kochi, T. Tahira, A. Suzuki, K. Higasa, A. Takahashi, T. Horita, T. Atsumi, T. Ishii, A. Okamoto, K. Fujio, M. Hirakata, H. Amano, Y. Kondo, S. Ito, K. Takada, A. Mimori, K. Saito, M. Kamachi, Y. Kawaguchi, K. Ikari, O.W. Mohammed, K. Matsuda, C. Terao, K. Ohmura, K. Myouzen, N. Hosono, T. Tsunoda, N. Nishimoto, T. Mimori, F. Matsuda, Y. Tanaka, T. Sumida, H. Yamanaka, Y. Takasaki, T. Koike, T. Horiuchi, K. Hayashi, M. Kubo, N. Kamatani, R. Yamada, Y. Nakamura, K. Yamamoto, A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus erythematosus in Japanese, *PLoS Genet.* 8 (2012) e1002455.
- [351] D.M. Waterworth, S.L. Ricketts, K. Song, L. Chen, J.H. Zhao, S. Ripatti, Y.S. Aulchenko, W. Zhang, X. Yuan, N. Lim, J. Luan, S. Ashford, E. Wheeler, E.H. Young, D. Hadley, J.R. Thompson, P.S. Braund, T. Johnson, M. Struchalin, I. Surakka, R. Luben, K.T. Khaw, S.A. Rodwell, R.J. Loos, S.M. Boekholdt, M. Inouye, P. Deloukas, P. Elliott, D. Schlessinger, S. Sanna, A. Scuteri, A. Jackson, K.L. Mohlke, J. Tuomilehto, R. Roberts, A. Stewart, Y.A. Kesaniemi, R.W. Mahley, S.M. Grundy, C. Wellcome Trust Case Control, W. McArdle, L. Cardon, G. Waeber, P. Vollenweider, J.C. Chambers, M. Boehnke, G.R. Abecasis, V. Salomaa, M.R. Jarvelin, A. Ruokonen, I. Barroso, S.E. Epstein, H.H. Hakonarson, D.J. Rader, M.P. Reilly, J.C. Witteman, A.S. Hall, N.J. Samani, K. O'Malley, P. Barter, C.M. van Duijn, J.S. Kooner, L. Peltonen, N.J. Wareham, R. McPherson, V. Mooser, M.S. Sandhu, Genetic variants influencing circulating lipid levels and risk of coronary artery disease, *Arterioscler. Thromb. Vasc. Biol.* 30 (2010) 2264–2276.
- [352] L. Zhan, J.R. Kerr, M.J. Lafuente, A. Maclean, M.V. Chibalina, B. Liu, B. Burke, S. Bevan, J. Nasir, Altered expression and coregulation of dopamine signalling genes in schizophrenia and bipolar disorder, *Neuropathol. Appl. Neurobiol.* 37 (2011) 206–219.
- [353] F.J. Meng, F.M. Meng, H.X. Wu, X.F. Cao, miR-564 inhibited metastasis and proliferation of prostate cancer by targeting MLLT3, *Eur. Rev. Med. Pharmacol. Sci.* 21 (2017) 4828–4834.
- [354] E.J. Perlman, S. Gadd, S.T. Arold, A. Radhakrishnan, D.S. Gerhard, L. Jennings, V. Huff, J.M. Guidry Auvel, T.M. Davidson, J.S. Dome, D. Meerzaman, C.H. Hsu, C. Nguyen, J. Anderson, Y. Ma, A.J. Mungall, R.A. Moore, M.A. Marra, C.G. Mullighan, J. Ma, D.A. Wheeler, O.A. Hampton, J.M. Gastier-Foster, N. Ross, M.A. Smith, MLLT1 YEATS domain mutations in clinically distinctive Favourable histology Wilms tumours, *Nat. Commun.* 6 (2015) 10013.
- [355] L. Wan, S. Chong, F. Xuan, A. Liang, X. Cui, L. Gates, T.S. Carroll, Y. Li, L. Feng, G. Chen, S.P. Wang, M.V. Ortiz, S.K. Daley, X. Wang, H. Xuan, A. Kentsis, T.W. Muir, R.G. Roeder, H. Li, W. Li, R. Tjian, H. Wen, C.D. Allis, Impaired cell fate through gain-of-function mutations in a chromatin reader, *Nature* 577 (2020) 121–126.
- [356] W. Xiao, Q. Zhang, F. Jiang, M. Pins, J.M. Kozlowski, Z. Wang, Suppression of prostate tumor growth by U19, a novel testosterone-regulated apoptosis inducer, *Cancer Res.* 63 (2003) 4698–4704.
- [357] W. Xiao, Q. Zhang, G. Habermacher, X. Yang, A.Y. Zhang, X. Cai, J. Hahn, J. Liu, M. Pins, L. Doglio, R. Dhir, J. Gingrich, Z. Wang, U19/Eaf2 knockout causes lung adenocarcinoma, B-cell lymphoma, hepatocellular carcinoma and prostatic intraepithelial neoplasia, *Oncogene* 27 (2008) 1536–1544.
- [358] L.E. Pascal, F. Su, D. Wang, J. Ai, Q. Song, Y. Wang, K.J. O'Malley, B. Cross, L.H. Rigatti, A. Green, R. Dhir, Z. Wang, Conditional deletion of Eaf1 induces

- murine prostatic intraepithelial neoplasia in mice, *Neoplasia* 21 (2019) 752–764.
- [359] J. Ai, L.E. Pascal, K.J. O'Malley, J.A. Dar, S. Isharwal, Z. Qiao, B. Ren, L.H. Rigatti, R. Dhir, W. Xiao, J.B. Nelson, Z. Wang, Concomitant loss of EAF2/U19 and Pten synergistically promotes prostate carcinogenesis in the mouse model, *Oncogene* 33 (2014) 2286–2294.
- [360] F. Xiao, J.S. Zhang, J.Y. Zhao, D. Wu, Regulation of Eaf2 in mouse lens cells apoptosis induced by ultraviolet radiation, *Int J Ophthalmol* 5 (2012) 570–575.
- [361] G. Fournier, O. Cabaud, E. Josselin, A. Chaix, J. Adelaide, D. Isnardon, A. Restouin, R. Castellano, P. Dubreuil, M. Chaffanet, D. Birnbaum, M. Lopez, Loss of AF6/afadin, a marker of poor outcome in breast cancer, induces cell migration, invasiveness and tumor growth, *Oncogene* 30 (2011) 3862–3874.
- [362] Y. Xu, R. Chang, Z. Peng, Y. Wang, W. Ji, J. Guo, L. Song, C. Dai, W. Wei, Y. Wu, X. Wan, C. Shao, L. Zhan, Loss of polarity protein AF6 promotes pancreatic cancer metastasis by inducing snail expression, *Nat. Commun.* 6 (2015) 7184.
- [363] H. Yang, L. Jin, X. Sun, A thirteengene set efficiently predicts the prognosis of glioblastoma, *Mol. Med. Rep.* 19 (2019) 1613–1621.
- [364] Y.M. Lin, K. Ono, S. Satoh, H. Ishiguro, M. Fujita, N. Miwa, T. Tanaka, T. Tsunoda, K.C. Yang, Y. Nakamura, Y. Furukawa, Identification of AF17 as a downstream gene of the beta-catenin/T-cell factor pathway and its involvement in colorectal carcinogenesis, *Cancer Res.* 61 (2001) 6345–6349.
- [365] K. Chen, J. Quan, J. Yang, Z. Chen, The potential markers of endocrine resistance among HR + /HER2+ breast cancer patients, *Clin. Transl. Oncol.* 22 (2020) 576–584.
- [366] C.H. Huang, A. Lujambio, J. Zuber, D.F. Tschaharganeh, M.G. Doran, M.J. Evans, T. Kitzing, N. Zhu, E. de Stanchina, C.L. Sawyers, S.A. Armstrong, J.S. Lewis, C.J. Sherr, S.W. Lowe, CDK9-mediated transcription elongation is required for MYC addiction in hepatocellular carcinoma, *Genes Dev.* 28 (2014) 1800–1814.
- [367] G. De Falco, C. Bellan, A. D'Amuri, G. Angeloni, E. Leucci, A. Giordano, L. Leoncini, Cdk9 regulates neural differentiation and its expression correlates with the differentiation grade of neuroblastoma and PNET tumors, *Cancer Biol Ther* 4 (2005) 277–281.
- [368] C. Moiola, P. De Luca, K. Gardner, E. Vazquez, A. De Siervi, Cyclin T1 over-expression induces malignant transformation and tumor growth, *Cell Cycle* 9 (2010) 3119–3126.
- [369] K. Storch, N. Cordes, The impact of CDK9 on radiosensitivity, DNA damage repair and cell cycling of HNSCC cancer cells, *Int. J. Oncol.* 48 (2016) 191–198.
- [370] Y. Zhao, L. Wang, S. Ren, L. Wang, P.R. Blackburn, M.S. McNulty, X. Gao, M. Qiao, R.L. Vessella, M. Kohli, J. Zhang, R.J. Karnes, D.J. Tindall, Y. Kim, R. MacLeod, S.C. Ekker, T. Kang, Y. Sun, H. Huang, Activation of P-TEFb by androgen receptor-regulated enhancer RNAs in castration-resistant prostate Cancer, *Cell Rep.* 15 (2016) 599–610.
- [371] L.C. Franco, F. Morales, S. Boffo, A. Giordano, CDK9: a key player in cancer and other diseases, *J. Cell. Biochem.* 119 (2018) 1273–1284.

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

Koushik Ghosh,¹ Ming Tang,² Nidhi Kumari,¹ Arijit Nandy,¹ Subham Basu,¹ Dheerendra Pratap Mall,¹ Kunal Rai,² and Debabrata Biswas^{1,3,*}

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

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

³Lead Contact

*Correspondence: 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 α -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- κ 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.



Multivalent Role of Human TFIID in Recruiting Elongation Components at the Promoter-Proximal Region for Transcriptional Control

Dipika Yadav,¹ Koushik Ghosh,¹ Subham Basu,¹ Robert G. Roeder,² and Debabrata Biswas^{1,3,*}

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

²Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10065, USA

³Lead Contact

*Correspondence: dbiswas@iicb.res.in

<https://doi.org/10.1016/j.celrep.2019.01.012>

SUMMARY

Despite substantial progress in our understanding of the players involved and the regulatory mechanisms controlling the initiation and elongation steps of transcription, little is known about the recruitment of elongation factors at promoter-proximal regions for the initiation-to-elongation transition. Here, we show evidence that human TFIID, which initiates pre-initiation complex (PIC) assembly, contributes to regulating the recruitment of super-elongation complex (SEC) components at the promoter-proximal region through interactions among selective TAF and SEC components. *In vitro* direct interactions, coupled with cell-based assays, identified an important poly-Ser domain within SEC components that are involved in their interaction with TFIID. DNA template-based recruitment assays, using purified components, further show a direct role for poly-Ser domain-dependent TFIID interaction in recruiting SEC components on target DNA. Consistently, ChIP and RNA analyses have shown the importance of this mechanism in TFIID-dependent SEC recruitment and target gene expression within mammalian cells.

INTRODUCTION

The initiation step of RNA polymerase II (Pol II hereafter)-driven eukaryotic transcription begins with activator-dependent recruitment of TATA-binding protein (TBP)-containing TFIID complex (TFIID hereafter) at the promoter region. Human TFIID contains ~14 other TBP-associated factors (TAFs) that help in recruiting TFIID and other general transcription factors (GTFs) at the promoter region for assisting pre-initiation complex (PIC) assembly. Although the regulation of transcription initiation is a key step for the expression of the majority of genes, studies from the last several years have shown that a substantial number of eukaryotic genes are also regulated at the promoter-proximal pausing step (Adelman and Lis, 2012; Henriques et al., 2013, 2018; Williams et al., 2015).

Soon after promoter clearance, Pol II is subjected to pausing through the action of two negative regulatory complexes—5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor (DSIF) and negative elongation factor (NELF), as well as Pol II-associated factor 1 (PAF1) (Chen et al., 2015, 2017; Yamaguchi et al., 2013; Zhou et al., 2012). The majority of the Pol II-mediated pausing events happen after the synthesis of 20–50 bases of nascent RNA (Yamaguchi et al., 2013; Zhou et al., 2012). It is widely believed that the phosphorylation activity of positive transcription elongation factor b (P-TEFb) complex helps in relieving the paused Pol II to facilitate its entry into productive elongation. Human P-TEFb complex is a heterodimer consisting of cyclinT1/2 and cyclin-dependent kinase 9 (CDK9) (Peng et al., 1998; Price, 2000). Upon recruitment, the P-TEFb complex phosphorylates the DSIF and NELF complexes along with the C-terminal domain (CTD) of Pol II. Upon its phosphorylation, the NELF complex loses its ability to interact with the nascent RNA as well as the inhibitory DSIF complex and thereby relieves the paused Pol II to enter into productive elongation (Yamaguchi et al., 2013).

The majority of the human P-TEFb resides in an inactive complex, 7SK snRNP, in association with Hexim1, LARP1, MePCE2, and 7SK small nuclear RNA (snRNA) (He et al., 2008; Li et al., 2005; Yang et al., 2001). The transcriptionally active P-TEFb complex has been shown to associate with Brd4 and components of the super-elongation complex (SEC). Human SEC was first reported as an ~1.5 MDa complex containing multiple elongation factors such as AF9, AF9 family protein ENL, AF4, AFF4, ELL, and P-TEFb complex (Biswas et al., 2011; He et al., 2010; Lin et al., 2010; Yokoyama et al., 2010). Genome-wide association studies have demonstrated a strong presence of the SEC components at the promoter-proximal region (Lin et al., 2011). A further study has shown a role for the SEC in regulating inducible gene expression (Takahashi et al., 2011). Although studies have shown a role for c-Myc as well as histone acetylation at H3K27 in SEC recruitment for transcriptional activation (Rahl et al., 2010; Zhang et al., 2017), the role of transcription initiation factors in the SEC recruitment processes, and thus assisting in the transition from initiation to the elongation step, is poorly known.

The present study demonstrates a role for human TFIID in recruiting the SEC components at the promoter-proximal region to overcome the pausing of Pol II. Our detailed analyses further show that specific domain-dependent interactions between





Human FKBP5 Negatively Regulates Transcription through Inhibition of P-TEFb Complex Formation

Dheerendra Pratap Mall,^a Subham Basu,^a Koushik Ghosh,^a Nidhi Kumari,^a Abhishake Lahiri,^b Sandip Paul,^c  Debabrata Biswas^a

^aLaboratory of Transcription Biology, Molecular Genetics Division, CSIR–Indian Institute of Chemical Biology, Kolkata, India

^bStructural Biology and Bioinformatics Division, CSIR–Indian Institute of Chemical Biology, Kolkata, India

^cJIS Institute of Advanced Studies and Research, JIS University, Kolkata, India

ABSTRACT Although a large number of recent studies indicate strong association of FKBP5 (aka FKBP51) functions with various stress-related psychiatric disorders, the overall mechanisms are poorly understood. Beyond a few studies indicating its functions in regulating glucocorticoid receptor, and AKT signaling pathways, other functional roles (if any) are unclear. Here, we report an antiproliferative role of human FKBP5 through negative regulation of expression of proliferation-related genes. Mechanistically, we show that, owing to the same region of interaction on cyclin-dependent kinase 9 (CDK9), human FKBP5 directly competes with cyclin T1 for functional positive transcription elongation factor b (P-TEFb) complex formation. *In vitro* biochemical assays, coupled with cell-based assays, showed a strong negative effect of FKBP5 on P-TEFb-mediated phosphorylation of diverse substrates. Consistently, FKBP5 knock-down showed enhanced P-TEFb complex formation that led to increased global RNA polymerase II C-terminal domain (CTD) phosphorylation, expression of proliferation-related genes, and subsequent proliferation. Thus, our results show an important role for FKBP5 in negative regulation of P-TEFb functions within mammalian cells.

KEYWORDS CTD phosphorylation, FKBP5, P-TEFb, RNA polymerase II, transcriptional regulation

The human FKBP5 is a 51-kDa protein of the immunophilin family that binds to immunosuppressants such as rapamycin and FK506 (1, 2). The functions of this protein have largely been described in context of the HSP90 cochaperone complex (3). A recent surge in reports indicates a strong association of FKBP5 functions with several neurological diseases, including posttraumatic stress disorder (PTSD) (4–6). Furthermore, FKBP5 functions have also been correlated with multiple other diseases and processes, including type 2 diabetes, adipogenesis, and fatty acid metabolism, as well as cancers (7). In several cancers, a strong negative correlation has been observed between FKBP5 expression and severity of disease (8–11). However, little is known about the molecular mechanisms of function of this protein and its role in pathogenesis of multiple diseases.

The well-known mechanistic functions of FKBP5 have been described in the context of AKT and glucocorticoid (GC) signaling pathways (3, 7, 9, 10). FKBP5 has been shown to regulate the phosphorylation status of one of the key regulatory residues (Ser473) of AKT1 through regulating the recruitment of PH domain leucine-rich repeat phosphatase (PHLPP) (8). FKBP5 has also been shown to negatively regulate GC signaling by reducing GC binding to cognate receptors and thus reducing nuclear translocation and downstream activation of target genes (12–14). Although a plethora of literature suggests a strong connection of FKBP5 functions with neuronal abnormalities, recent studies have also indicated its role in controlling type 2 diabetes, in part through control of AKT signaling and the downstream effector protein AKT substrate 160 (AS160) (15). Besides, numerous other studies have also shown a negative correlation between

Copyright © 2022 American Society for Microbiology. All Rights Reserved.

Address correspondence to Debabrata Biswas, dbiswas@iicb.res.in.

The authors declare no conflict of interest.

Received 13 July 2021

Returned for modification 8 August 2021

Accepted 2 November 2021

Accepted manuscript posted online 15 November 2021

Published 20 January 2022



Indian Association for the Cultivation of Science
Kolkata, India

3rd IABS 2018

AN INTERDISCIPLINARY APPROACH TO BIOLOGICAL SCIENCES

February 1-3, 2018

Certificate of
Participation/Poster Presentation

This is to certify that

Subham Basu

.....
has participated in the International Conference (IABS 2018)
organised by Indian Association for the Cultivation of Science
during 1st - 3rd February, 2018 at IACS campus.

Convenor
Benu Brata Das



Bose Institute

**Indo-Japan conference on
EPIGENETICS AND HUMAN DISEASE**
(A joint seminar of Japan-India Forum for Advanced Study)
7-9 February 2018. . Bose Institute, Unified Campus, Kolkata, India

This is to certify that

Subham Basu

has participated in the

Indo-Japan conference on EPIGENETICS AND HUMAN DISEASE

held at **Bose Institute, Unified Campus, Kolkata, India**

on 7-9 February, 2018.

Siddhanta

Chairperson

Tapan K. Dutta
Convenor

