P-TEFb-dependent pathways in cancer

Xinwen Zhu

Department of Pharmacology and Therapeutics
McGill University, Montréal, QC, Canada

September 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Abstract

Transcriptional regulators that play crucial roles in cancer progression have gained prominence as anticancer drug targets. Inhibitors of positive transcription elongation factor b (P-TEFb), a factor essential for promoter proximal pause release, have proven effective against a number of cancer types, including mixed lineage leukemia driven by translocations of the histone methyltransferase MLL. These translocations result in chimeric proteins that have been implicated in P-TEFb recruitment to target cancer-driving genes. P-TEFb is a kinase with multiple targets, including RNA polymerase II as well as two factors associated with the paused polymerase, negative elongation factor (NELF) and DRB sensitivity-inducing factor (DSIF). To determine the functional significance of P-TEFb-dependent pathways and elucidate the mechanisms underlying sensitivity to P-TEFb inhibitors in MLL-rearrangement associated leukemias, I characterized the responses to transcriptional inhibitor treatment of two cancer cell lines, THP-1, which has a MLL-rearrangement, and K562, which has unaltered MLL. I found that THP-1 exhibits reduced viability and greater loss of P-TEFb activity upon inhibition of P-TEFb. While consistent with the idea that MLL-rearrangement associated leukemias have increased dependency on P-TEFb, this work also raises the possibility that MLL-fusion proteins might globally impair P-TEFb, contrary to current models. I also took steps towards creating cell lines with phosphoacceptor site mutations in the P-TEFb target Spt5 using CRISPR/Cas9. My efforts revealed that double-stranded repair templates, which are commonly used in CRISPR/Cas9-mediated gene replacement, are non-specifically inserted into the genome at a high frequency. Overall, this work advances our understanding of the importance of P-TEFb-dependent pathways in cancer biology.

Résumé

Les régulateurs de la transcription jouent un rôle crucial dans la progression du cancer et sont, par conséquent, devenus des cibles de médicaments anticancéreux intéressantes. Les inhibiteurs du facteur positif d'élongation de la transcription (P-TEFb), un facteur essentiel pour la sortie de la pause proximale au promoteur, se sont révélés efficaces contre plusieurs types de cancers, incluant la leucémie aiguë de phénotype mixte entraînée par des translocations de la méthyltransférase MLL. Ces translocations forment des protéines chimériques qui sont impliquées dans le recrutement de P-TEFb aux gènes affectant ce cancer. P-TEFb est une kinase avec plusieurs cibles, incluant l'ARN polymérase II et deux facteurs associés à la polymérase en pause, soit le facteur négatif d'élongation (NELF) et le facteur inductible à la sensibilité au DRB (DSIF). Pour déterminer l'importance fonctionnelle des mécanismes régulés par P-TEFb et élucider les mécanismes sous-jacents à la sensibilité aux inhibiteurs de P-TEFb dans les leucémies avec réarrangement de MLL, j'ai caractérisé les réponses aux inhibiteurs transcriptionnels de deux lignées cellulaires cancéreuses, soit THP-1, qui possède un réarrangement de MLL, et K562, qui ne possède pas de réarrangement. J'ai découvert que THP-1 présente une viabilité réduite et une perte plus importante de l'activité de P-TEFb lors de traitement avec les inhibiteurs de P-TEFb. Bien qu'ils conforment à l'idée que les leucémies avec réarrangement de MLL dépendent davantage sur le P-TEFb, mes résultats soulèvent également la possibilité que les protéines de fusion MLL pourraient globalement altérer P-TEFb, contrairement aux modèles actuels. J'ai aussi tenté de créer des lignées cellulaires avec des mutations aux sites phosphoaccepteurs de Spt5, une cible de P-TEFb, en utilisant la technique CRISPR/Cas9. Mes efforts ont révélé que les molécules d'ADN à double brin, qui sont couramment utilisés comme modèles de réparation lors de l'essai CRISPR/Cas9, sont insérés dans le génome de manière non-spécifique à une fréquence élevée. Pris ensemble, mes résultats améliorent notre compréhension de l'importance des mécanismes régulés par P-TEFb dans la biologie du cancer.

Table of Contents

Acknowledgements

First and foremost, I would like to thank my supervisors, Dr. Terry Hébert and Dr. Jason Tanny, for being admirable role models and for giving me their support and guidance throughout this project. Thank you for your patience, and for allowing me room to explore and grow while pointing me in the right direction. I wouldn't be where I am today without your mentorship and positive examples. I would like to give thanks to my advisor, Dr. Lisa Münter, and to my committee members, Dr. François Robert and Dr. Dan Bernard, for their feedback and helpful suggestions. I must also thank Dr. Paul Clarke for being an excellent educator. All of the aforementioned individuals have played pivotal roles in my development of scientific thinking and communication skills which I am certain will serve me well in the future.

I would also like to thank all the members of the Hébert and Tanny labs, past and present, for helping create a supportive and positive work environment. I could not have progressed as quickly without my labmates. In particular, I would like to acknowledge Viviane Pagé for helping translate my abstract to French. I would also like to acknowledge Julien Leconte, the manager of the Flow Cytometry platform, for helping me set up my flow cytometry assay.

Last but certainly not least, I would like to express gratitude to my friends and family members who have been there for me throughout this journey. I'd especially like to give a very special shout-out to the incognito group. Thank you to everyone who invested the time and effort to check in with me and keep me on track. I wish you all the very best.

List of Figures

List of Abbreviations

Brd4	Bromodomain-containing protein 4
BET	Bromodomain and Extra-terminal domain
ChIP	Chromatin Immunoprecipitation
Cdk9	Cyclin-dependent kinase 9
CI	Confidence Interval
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTD	C-terminal domain
CTR	C-terminal region
dNTPs	Deoxynucleotides
DRB	5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole
DSIF	DRB sensitivity-inducing factor

dsDNA	Double-stranded DNA
DTT	Dithiothreitol
EC_{50}	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
eRNA	Enhancer RNA
FBS	Fetal bovine serum
HDR	Homology directed repair
HEK 293	Human embryonic kidney 293
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HEXIM	Hexamethylene bis-acetamide inducible
HIV	Human immunodeficiency virus
Hsp70	Heat shock protein 70
IGEPAL	Octylphenoxypolyethoxyethanol
NELF	Negative elongation factor
NHEJ	Non-homologous end joining
MLL	Mixed lineage leukemia
PAFc	Polymerase II–associated factor (PAF) complex
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
Pol II	RNA Polymerase II
PP1	Protein phosphatase 1
P-TEFb	Positive transcription elongation factor b
RNF20/40	Ring finger protein 20/40
RPMI-1640	Roswell Park Memorial Institute 1640 Medium
RT-qPCR	Quantitative reverse transcription PCR
SEC	Super elongation complex
SEM	Standard error of the mean
sgRNA	Single guide RNA
snRNP	Small nuclear ribonucleoprotein
siRNA	Small interfering RNA
ssODN	Single-stranded oligodeoxynucleotides

Chapter 1: General Introduction

1.1 Control of Gene Transcription

A gene is a stretch of an organism's genome that ultimately codes for a functional product, such as a protein or RNA molecule (1). The traditional study of genetics and the rising field of genomics emphasize genotype: biological variation arises because organisms possess different gene variants, which code for products with different functionalities.

However, while many individual differences and diseases can be explained by hypo- or hyper-functional gene products, phenotype, or a set of observable biological characteristics, is ultimately determined by the spatiotemporal manner in which gene products are synthesized.

The manifold determinants of gene transcription can be classified as either: 1) genetic information, 2) histone modifications and epigenetic changes, and 3) previously synthesized gene products. These three broad classifications capture various elements that work together to influence how and when genes are transcribed.

Less than 1% of the human genome is comprised of genes (2). The remainder mostly consists of repetitive elements and transposons, but also contains functionally important sequences that guide transcription. For example, short regions of DNA known as enhancers help recruit transcriptional complexes to spatially-proximal genes and thus promote their transcription. Although enhancer sequences are genetically encoded, their activity has been shown to be highly cell-context dependent (3). Genetic information is therefore not sufficient to explain how transcription varies throughout development and in response to environmental cues.

In eukaryotic cells, DNA is organized into chromatin, a complex of DNA and packaging molecules. The structural unit of chromatin is the nucleosome, which consists of DNA wound around protein octamers made up of two of each core histone protein: H2A, H2B, H3, and H4 (4). The histone proteins are subject to diverse post-translational modifications on their N-terminal and C-terminal tails such as acetylation, methylation, phosphorylation, and ubiquitination. One function of histone modifications is to alter the structure of chromatin and its accessibility to the transcription machinery (5). For instance, histone acetylation neutralizes positive charges on the histone tail, thereby weakening the intermolecular interactions between

the histones and the negatively charged DNA backbone, providing greater access for the polymerases; such acetylation marks are generally associated with increased transcription. Furthermore, histone modifications serve as binding sites for other factors; acetylated histones are recognized by bromodomain-containing proteins (6).

Last but certainly not least, gene expression is determined by levels of existing RNAs and proteins. Transcription initiation requires the recognition of the gene promoter by specific transcription factor complexes, some of which form in response to extracellular cues; different combinations of available transcription factors can regulate distinct gene expression patterns (7). The aforementioned enhancers are traditionally thought to exert their effects through bound activator proteins, which interact with general transcription factors at the promoters of regulated genes. Furthermore, recent evidence suggests that many enhancers are actively transcribed and that the resulting eRNAs mediate enhancer function, potentially through the facilitating of DNA looping, which brings enhancers in close spatial proximity to regulated promoters (3). Moreover, the histone code itself is established, interpreted, and erased by specific proteins (8). Histone acetylation is catalyzed by histone acetyltransferases and such marks are removed by histone deacetylases (6). Furthermore, bromodomain-containing proteins have been implicated in the recruitment of the transcriptional machinery (9). Finally, the gene products of core transcriptional regulatory circuitries maintain their own expression through positive autoregulatory feedback loops (10).

The physical process by which a RNA molecule is synthesized from a gene, transcription, is a complex multi-step process. Traditionally, the transcription cycle has been divided into three stages: 1) initiation, where the RNA polymerase is recruited to the promoter of a gene, 2) elongation, where the RNA polymerase synthesizes a new RNA molecule using the coding region of the gene as a template, and 3) termination, where the transcriptional complex dissociates and is recycled to begin anew (11). Recent findings have revealed that in the case of eukaryotic protein-coding genes, another major step, pausing, occurs between initiation and productive elongation. Each of these steps can be regulated to control gene expression.

RNA polymerase II (Pol II), the polymerase that transcribes all protein-coding genes in eukaryotes, is incapable of recognizing promoter sequences and initiating transcription on its own. The general transcription factors TFIIB, TFIID, TFIIE, TFIIF and TFIIH form the pre-initiation complex by recruitment and activation of Pol II. Furthermore, activator proteins bound

at enhancers interact with Pol II via the Mediator complex (12). TFIIH functions as a helicase, allowing Pol II to begin mRNA synthesis. Once the nascent mRNA transcript is long enough to fill the RNA exit channel, the complex stabilizes (13). The transition from initiation to elongation through pausing and the maintenance of steady productive elongation will be discussed in greater detail below.

Once the transcription complex reaches the terminator at the 3' end of the gene, the cleavage and polyadenylation factor complex induces release of the mRNA transcript and slows Pol II elongation. Pol II is finally released from the DNA template by the exonuclease Xrn2, which degrades RNA synthesized past the termination site (14).

1.2 Promoter-Proximal Polymerase Pausing

While control of initiation was long thought to be the most important regulatory step in transcription, interest in the regulatory mechanisms surrounding transcription elongation is increasing with emerging evidence that the process is more complicated than previously assumed (15).

At many eukaryotic genes, Pol II stalls approximately 50 base pairs from the transcription start site (TSS), a phenomenon termed promoter-proximal polymerase pausing (16). In the early 1980s, a transcriptional run-on assay of the β-globin gene in adult hen erythrocytes, which consisted of pulse-labelling nascent RNA synthesized in isolated nuclei and hybridizing it to the original gene, suggested that transcribing RNA polymerases clustered at the 5' end of the gene (17). This early hint that a crucial step in transcriptional regulation occurred in the transition between initiation and elongation was soon followed by others. In the following decade, transcriptionally-engaged RNA polymerase was found to accumulate at the promoter-proximal ends of several genes, including *hsp70* (18) and *c-myc* (19).

Despite these observations suggesting that the rate-limiting step in the expression of certain genes is downstream of initiation, it was not until modern genome-wide experimental tools became available that pausing was recognized as a more pervasive phenomenon (15). Genome-wide chromatin immunoprecipitation (ChIP) and nascent RNA-sequencing studies

revealed promoter-proximal pausing at a large proportion of genes (20, 21). High resolution approaches confirmed the widespread nature of promoter-proximal pausing, and also enabled the appreciation of polymerase pausing within the gene body (22).

Two key regulators required to maintain the paused state are negative elongation factor (NELF) and DRB sensitivity-inducing factor (DSIF), which directly interact with the paused polymerase (23). Release from this paused state requires the activity of positive transcription elongation factor b (P-TEFb), which is comprised of cyclin-dependent kinase 9 (Cdk9) and an associated cyclin (either T1, T2a, T2b, or K) (24). P-TEFb phosphorylates the NELF-A and NELF-E subunits of NELF, the Spt5 subunit of DSIF, and the C-terminal domain (CTD) of Pol II (25). Following these phosphorylation events, NELF dissociates from the transcriptional machinery (26), DSIF remains bound but is converted from a transcriptional repressor to an activator (27), and Pol II proceeds into productive transcription elongation (16).

The paused polymerase is not a stable fixture; one recent study revealed a high rate of polymerase turnover at promoters with pausing, suggesting that paused polymerases are most often removed by premature termination (28). Nevertheless, promoter-proximally paused Pol II appears to be more stably bound than Pol II at the pre-initiation complex, and paused Pol II is believed to prevent new initiation, so that only one polymerase complex is found at a promoter at a given time (29). It has been suggested that paused polymerases maintain chromatin in an active state by preventing nucleosome assembly, thus facilitating future transcription (30).

P-TEFb-mediated release of promoter-proximal polymerase pausing has been found to be the rate-limiting step in the expression of many genes in certain contexts (31, 32). P-TEFb activity is also crucial for the regulation of primary response genes, genes that do not require *de novo* protein synthesis to be expressed in response to signalling events (33).

1.3 Positive Transcription Elongation Factor b

1.3.1 Regulation

P-TEFb activity is regulated at various levels by several mechanisms. It has been proposed that multiple P-TEFb molecules must be recruited to the paused transcription complex by different mechanisms to act on various targets (25). The kinase can be recruited to chromatin as part of the super elongation complex (SEC) (16). In the context of HIV infection, the HIV Tat protein is thought to induce viral transactivation by recruiting the P-TEFb-containing SEC to the HIV promoter through interactions with cyclin T1, the Cdk9 partner cyclin, and the viral TAR RNA (34, 35). P-TEFb also interacts with gene-specific transcription factors such as NF-κB to promote transcription in response to stimuli (36). Furthermore, the Med26 subunit of the general transcription coactivator Mediator complex has putative docking sites for P-TEFb and other SEC components (37). BRD4, a BET (Bromodomain and Extra-Terminal domain) family protein that recognizes acetylated histones at enhancers, is also thought to be a major factor in recruiting P-TEFb to chromatin. BRD4 has been shown to compete with Tat for P-TEFb binding (38). Furthermore, BRD4 is thought to interact with JMJD6, another protein capable of recruiting and activating P-TEFb (39). However, in a recent study where BRD4 was depleted by inducing its degradation, acute loss of BRD4 precipitated a global decrease in elongation that did not affect Cdk9 localization, implying that BRD4 has a general role in regulating transcription elongation that does not involve P-TEFb recruitment (40).

P-TEFb activation is also controlled by association with the 7SK snRNP, a ribonucleoprotein complex containing the 7SK RNA and three RNA-binding proteins: PIP7S/LARP7, BCDIN3/MePCE, and either HEXIM1 or HEXIM2 (41, 42). In HeLa cell extracts, approximately half of all P-TEFb is sequestered by 7SK under unstimulated conditions (43), while in Jurkat cells, upwards of 75% of P-TEFb has been found in the 7SK bound form (44). This pool of 7SK-bound P-TEFb is inactive, but can be mobilized in response to stimuli. Treatment with a Cdk9 inhibitor such as DRB, seliciclib, or flavopiridol precipitates release of P-TEFb from the 7SK complex that can be detected within an hour (44). Interestingly, treatment with JQ1, a bromodomain inhibitor which prevents Brd4 binding to acetylated lysines, also releases P-TEFb from 7SK, with levels of free P-TEFb peaking an hour after stimulation and returning to basal by six hours (45). The transience of the dissociation of P-TEFb from 7SK,

despite the continuous presence of the inhibitor, indicates the presence of negative feedback control mechanisms. HIV Tat is capable of activating P-TEFb by competing with HEXIM1 for binding (34).

Like other cyclin-dependent kinases, Cdk9 must be phosphorylated at its T-loop to be enzymatically active (46). Interestingly, phosphorylation of the T-loop at Thr-186 is also necessary for P-TEFb to interact with 7SK RNA (47). Following 7SK binding, the complex is recognized by HEXIM (47).

For P-TEFb to be released from HEXIM, calcium ion (Ca^{2+})–calmodulin–protein phosphatase 2B (PP2B) must bind the 7SK snRNP and induce a conformational change that allows protein phosphatase 1α (PP1α) to dephosphorylate the T-loop at Thr-186 (48). The free unphosphorylated P-TEFb can then be recruited to chromatin, where it must be activated by autophosphorylation of the T-loop before it can promote transcriptional elongation (24).

1.3.2 The P-TEFb cyclin

Early studies of P-TEFb's ability to bind Tat and promote HIV transactivation identified cyclin T1 (originally named cyclin T) as the binding partner of Cdk9 (35). Subsequent studies revealed that two alternatively spliced forms of cyclin T2 (T2a and T2b) are also capable of activating Cdk9, although they do not bind HIV Tat (24, 49). Finally, cyclin K was found to form functional P-TEFb by association with Cdk9 (50). However, another study revealed that FLAG-tagged cyclin K mainly associates with Cdk12 and Cdk13 and has no detectable ability to co-immunoprecipitate Cdk9 (51).

Recently, a histidine-rich domain of cyclin T1 has been shown to promote P-TEFb phosphorylation of the Pol II CTD but not autophosphorylation. This intrinsically disordered region is important for recruiting the CTD into phase separated droplets where it can be optimally phosphorylated (52).

1.3.3 P-TEFb-dependent pathways

As previously mentioned, active P-TEFb has a number of phosphorylation targets within the promoter-proximal paused polymerase complex that promote transcriptional elongation once phosphorylated.

1.3.3.1 RNA polymerase II

The CTD domain of Rpb1, the largest subunit of Pol II, consists of multiple repeats of the heptapeptide sequence YSPTSPS; the number of repeats is positively correlated with the complexity of the organism, with 26-27 repeats found in yeast and 52 repeats in mammals, suggesting expansion over evolutionary time (53). The reversible phosphorylation of various residues in this heptapeptide repeat coordinates the many steps of mRNA synthesis by favouring the recruitment or dissociation of different elements of the transcriptional and pre-mRNA processing machinery (54).

During transcription initiation, a hypophosphorylated form of Rpb1 is incorporated into the pre-initiation complex and phosphorylated at Ser-5 and Ser-7 by Cdk7, a component of TFIIH (55). Since depletion of P-TEFb results in a loss of actively elongating Pol II, which is hyperphosphorylated at Ser-2 of the CTD repeat, Cdk9 was originally thought to be a Ser-2 kinase. This idea was supported by *in vitro* kinase experiments performed with affinity purified FLAG-tagged P-TEFb and GST-tagged CTD (56). Furthermore, inhibition of Cdk9 leads to a loss of Ser-2 phosphorylation in cell based assays (57). However, other recent *in vitro* experiments identify Ser-5 and not Ser-2 as the substrate for Cdk9, suggesting that P-TEFb phosphorylates Rpb1 in the same position as Cdk7 (58). According to this model, P-TEFb activity toward other substrates, such as NELF and DSIF, remains necessary for Pol II to enter productive elongation in its Ser-2 phosphorylated form, but the Rpb1 Ser-2 phosphorylation event is mainly attributed to kinases such as Cdk12 or Cdk13 (59). Genome-wide studies of Pol II phosphoisoform occupancy in yeast have revealed that Ser-2 phosphorylation levels increase throughout the gene body, peaking near the 3' end of the gene (60, 61). The accumulated evidence is consistent with the notion that most Ser-2 phosphorylation is an indirect effect of Cdk9 activity, which permits pause release, allowing Pol II to enter the downstream gene regions where Ser-2 phosphorylation occurs (62). This release from promoter-proximal pausing seems to be a consequence of P-TEFb phosphorylation of NELF and the Spt5 subunit of DSIF.

1.3.3.2 Negative Elongation Factor

NELF is a complex of four proteins: NELF-A, NELF-B, either NELF-C or NELF-D, and NELF-E (63). NELF does not bind strongly to either Pol II or DSIF alone, but has stable interactions with a preformed DSIF/Pol II complex. In addition, NELF-E has RNA-binding

activity, which seems to be necessary for its repressive effect on transcription (64). Interestingly, although NELF induces pausing, *in vivo* loss of NELF leads to downregulation of NELF-mediated genes, suggesting that NELF activity sustains gene expression (65). This observation is consistent with the idea that one of the functions of transcriptional pausing is to facilitate future rounds of transcription through nucleosome displacement (30), although it may also point to the existence of unexplored functions of NELF.

1.3.3.3 Spt5

As mentioned above, NELF phosphorylation by P-TEFb leads to the dissociation of the complex from the polymerase. In contrast, the exact role of Spt5 phosphorylation is less clear. Spt5 is the larger subunit of the DSIF complex, a heterodimer of Spt5 and Spt4. The C-terminal region (CTR) of the protein contains seven iterations of the consensus repeat GS(R/Q)TP, which can be phosphorylated at the threonine by P-TEFb (66). This phosphorylation converts Spt5 from a transcriptional repressor to an activator (27). Interestingly, the CTR motif has recently been found to not be necessary for the establishment of promoter-proximal polymerase pausing; in an *in vitro* system, truncated Spt5 lacking the CTR was capable of establishing pausing (66). Furthermore, while complete knockout of Spt5 is lethal and zebrafish with CTR-truncated Spt5 exhibit impaired embryogenesis (67), the CTR is not required for cell proliferation (68). Dephosphorylation of the Spt5 CTR plays a role in the transition from elongation to termination. In the fission yeast *S. pombe*, Spt5 is dephosphorylated by Dis2, a PP1 homolog that is part of the cleavage and polyadenylation factor complex. In addition to directly phosphorylating Spt5, Cdk9 activity maintains phospho-Spt5 levels by inhibitory phosphorylation of Dis2. Inactivating mutations in Dis2 lead to transcription past the termination site, suggesting that Spt5 must be dephosphorylated for efficient termination (69).

Spt5 CTR phosphorylation has also been implicated in the recruitment of Rtf1, a factor that is functionally related to the polymerase II–associated factor (PAF) complex (70).

1.3.3.4 Polymerase II–associated factor complex

As its name implies, the PAFc was originally identified as a complex that co-purified with Pol II in budding yeast. The complex is conserved in higher eukaryotes; aside from Rtf1, the human PAFc is comprised of Paf1, Cdc73, Ctr9, Leo1, and Ski8 (71). Interestingly, the PAFc

has been found to have contrasting effects on transcriptional elongation in different experimental systems. In HCT116 or MCF7 cells, loss of Paf1 leads to an increase in the gene body to promoter distribution of Pol II, most notably at highly paused genes, and to increased SEC recruitment, suggesting that the PAFc maintains pausing and prevents interactions between the SEC and Pol II (72, 73). On the other hand, a study done with THP-1, a mixed lineage leukemia cell line harboring a fusion of MLL and AF9, a component of SEC, concluded that PAFc plays a cooperative role in promoter proximal pause release and Ser-2 phosphorylation of the Pol II CTD (74). Other experiments done with HeLa cells support the idea that PAFc participates in the recruitment of SEC to Pol II at transcribed genes (75).

The PAFc is required for the co-transcriptional monoubiquitination of histone 2B (H2Bub1), a histone modification with important functions in transcriptional elongation which occurs at lysine 123 in *S. cerevisiae* and at lysine 120 in humans (70, 71). Experiments performed with yeast have established that the role of the PAFc in this process is to link the E2 ubiquitin-conjugating enzyme and the E3 ubiquitin ligase to the elongating complex (76, 77). H2Bub1 is a prerequisite for other histone modifications, namely H3K4 and H3K79 methylation (78, 79).

The relationship between P-TEFb and H2Bub1 is not unidirectional. Disruption of P-TEFb-mediated Spt5 phosphorylation has been shown to reduce H2Bub1 in the fission yeast *Schizosaccharomyces pombe* (80), presumably through a decrease in Rtf1 recruitment. In humans, H2Bub1 is also downstream of Ser-2 phosphorylation of the Pol II CTD, which recruits a factor that associates with and stabilizes RNF20/40, the E3 ligases that mediate H2Bub1 (81, 82). Furthermore, Cdk9 has been shown to directly phosphorylate and activate UBE2A, the E2 ubiquitin-conjugating enzyme responsible for H2Bub1 in humans (83). P-TEFb activity therefore positively regulates the co-transcriptional deposition of H2Bub1 marks. In turn, H2Bub1 promotes further P-TEFb activity. Fission yeast strains with an ubiquitination-defective H2B mutant or lacking the ubiquitination machinery exhibit decreased Cdk9 recruitment and decreased Spt5 phosphorylation. This positive feedback loop is thought to balance antagonistic effects of pathways downstream of P-TEFb and H2Bub1 (80). Positive feedback also links P-TEFb and RNF20 in mammalian cells, although the role of Spt5 phosphorylation has not been determined (84).

1.3.3.5 mRNA processing

Another important function of P-TEFb activity is co-ordination of co-transcriptional mRNA processing events. The promoter-proximally paused polymerase has been found to co-localize with mRNA capping enzymes (85), which can bind to and may be recruited by DSIF (86, 87). These capping enzymes appear to relieve NELF and DSIF-mediated pausing in *in vitro* transcription assays (86), though it is unclear if they promote P-TEFb activity *in vivo*.

Factors for 5' capping, splicing, and 3' processing are all thought to be recruited by both direct interactions with the mRNA and interactions with phosphoisoforms of the Pol II CTD (85, 88), and Ser-2 phosphorylation is essential for cell survival and efficient mRNA processing (89).

Human PAFc also has putative roles in mRNA surveillance, the quality control mechanisms that remove improperly synthesized or processed mRNA. The Ski8 subunit of the complex, which is not found in the yeast PAFc, is also part of SKI, a complex thought to mediate 3'-5' mRNA degradation in conjunction with the exosome (90).

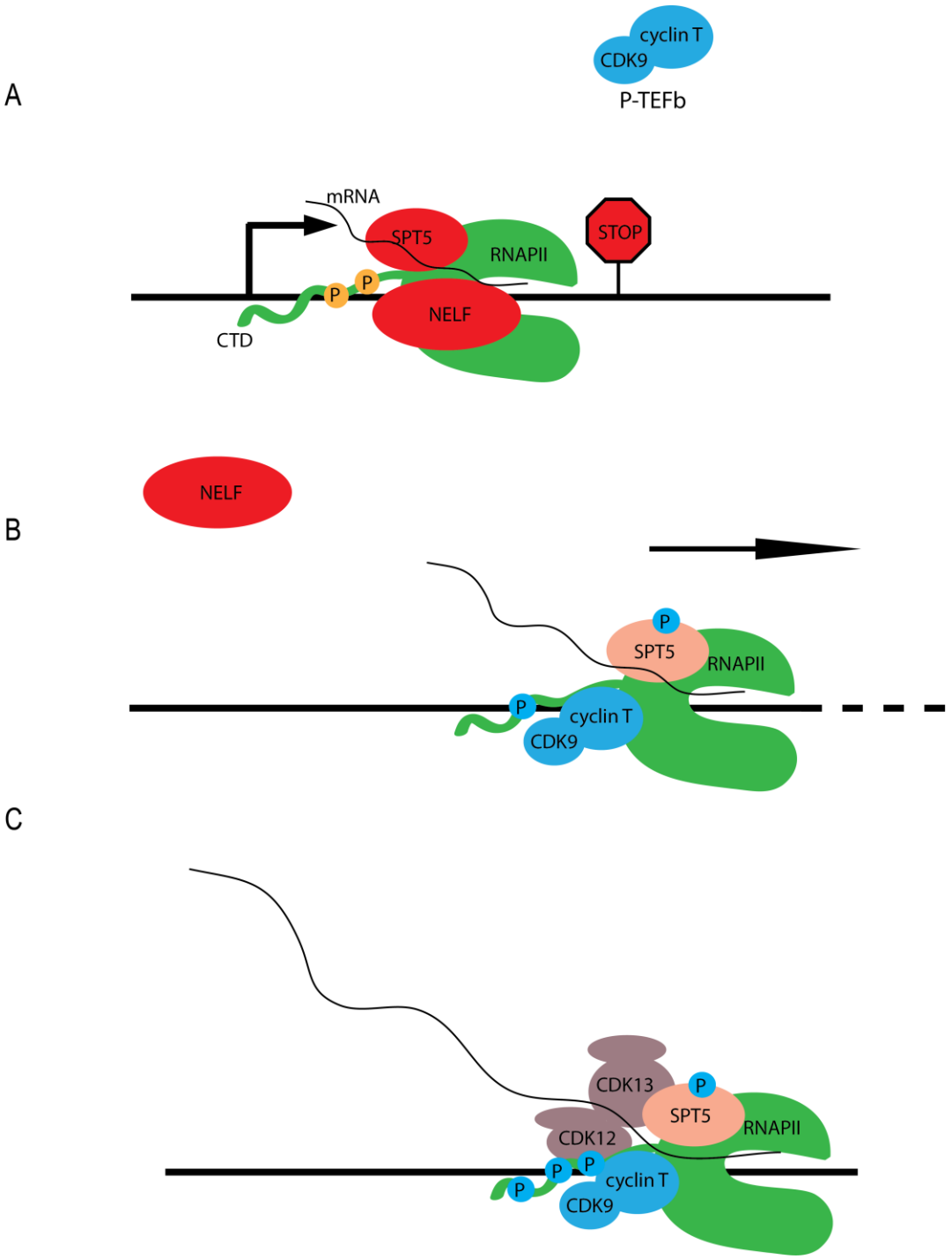

Figure 1.1 Control of transcription elongation by P-TEFb. A) The polymerase in complex with DSIF (Spt4 and Spt5) and NELF remains paused until P-TEFb is recruited. B) P-TEFb activity leads to the dissociation of NELF, the conversion of Spt5 from a transcriptional repressor to an activator, and release of the paused polymerase. C) Ser-2 phosphorylation of the Rpb1 CTD is further catalyzed by Cdk12 and Cdk13.

1.4 Transcriptional Dysregulation in Cancer

Cancers are diseases of aberrant gene expression. While different types of cancer vary in their etiologies and prognoses, all evolve towards malignancy by progressively acquiring characteristics necessary for sustained proliferation and survival (91). This evolution is driven by the accumulation of mutations which disrupt normal cell function and dysregulate transcriptional programs. For instance, genetic fusions of the histone methyltransferase MLL and components of a transcriptional elongation complex lead to the abnormal recruitment of transcriptional elongation machinery to MLL binding sites; the resulting increase in MLL target gene expression is thought to cause mixed lineage leukemia (92). Mixed lineage leukemia occurs predominantly in infants and manifests as either acute myeloid leukemia or acute lymphoid leukemia, but is characterized by leukemic blast cells that express characteristics of both the myeloid and the lymphoid lineages, sometimes switching between them (93). Such ambiguity in cell identity is symptomatic of dysregulated transcriptional programs.

Cell identity is established and maintained by core transcriptional regulatory circuitries, comprised of networks of transcription factors that promote their own expression as well as the expression of other cell-type specific genes (10). In order to become cancerous, a cell must escape its determined fate by disrupting this core regulatory circuitry and activating alternative transcriptional programs that facilitate cell growth or metastases. One of the most commonly upregulated oncogenes, *c-myc*, codes for a transcription factor, the Myc protein, that globally amplifies gene expression (94). Myc is therefore a key participant in pathways leading to cancer progression.

Unfortunately, key oncogenes such as Myc and other transcription factors are not always tractable drug targets. Although Myc is an attractive target for cancer therapy, attempts to directly inhibit the protein have not yielded great success (95). One promising alternative strategy is to exploit transcriptional dependencies which arise in the cancer cells themselves to indirectly suppress oncogenic activity (95, 96). Because aberrant transcription is crucial to the progression of cancer, cancer cells tend to become more sensitive to disruption of certain elements of the global transcriptional machinery than non-cancerous cells. Transcription regulators therefore present interesting potential therapeutic targets for anticancer strategies.

1.5 Transcriptional regulators as anticancer targets

A recurring theme in current anticancer pharmacology is the development of targeted therapies that specifically interfere with proteins driving oncogenic processes (97). Myc, the best known oncogenic transcription factor, activates its target genes (including many cell-cycle regulators important for cell proliferation) by binding to E-box elements near their promoters. Recently Myc has also been found to act as a universal amplifier of gene expression; overexpressed Myc binds many non-E-box promoters that have already recruited Pol II, and its binding at a gene is positively correlated with expression of that gene (94). Mechanistically, transcription amplification by Myc is thought to involve promoter-proximal pause release (98). In the context of mixed lineage leukemia, MLL-fusion proteins drive increased expression of Myc, which promotes oncogenesis through gene-specific and global effects on gene expression.

Many regulators of transcription play important roles in cancer progression and have attracted attention as drug targets. For instance, small molecules targeting nuclear hormone receptors, which act as transcription factors upon ligand binding, have seen great clinical success (99). p53, one of the most commonly inactivated tumour suppressors, is also a transcription factor; by activating the expression of effectors such as p21 and certain miRNAs, p53 triggers a transcriptional program that ultimately shuts down expression of genes required for cancer proliferation (100). p53 is therefore another promising drug target; a compound that reactivates mutant p53 has been observed to inhibit xenograft tumour growth (101).

In contrast to targeted therapies, which usually aim to selectively block activation of certain transcriptional programs and the resulting synthesis of specific gene products involved in oncogenic transformation, global inhibitors of transcription are also under investigation. While disruption of basal transcriptional machinery is expected to be toxic to all living cells to a certain degree, cancerous cells are more sensitive because they require high levels of RNA synthesis (102). Furthermore, since the mRNA transcripts of many oncogenes have short half-lives, it has been suggested that global transcriptional inhibition preferentially decreases oncogene expression (103).

1.5.1 P-TEFb in Cancer

Small molecule inhibitors of P-TEFb or Brd4 have shown promise as anticancer agents in recent preclinical and clinical trials. In preclinical trials, Cdk9 inhibitors such as flavopiridol and dinaciclib have been effective against acute myeloid leukemia (104, 105) and chronic lymphocytic leukemia (106). The best studied Cdk9 inhibitor, flavopiridol, achieved complete response rates between 40% and 70% in phase I and phase II clinical trials against acute myeloid leukemia when used in combination with standard treatment drugs, cytarabine and mitoxantrone (FLAM) (107). As of May 2014, flavopiridol has been granted orphan drug status for acute myeloid leukemia by the U.S. Food and Drug Administration, a designation intended to encourage the development of drugs for the treatment of rare medical conditions (108). BET inhibitors have been investigated as treatment options in acute myeloid leukemia, NUT midline carcinoma, triple negative breast cancer, and various other cancers (109). An RNA interference screen has shown that Brd4 is indispensable for the maintenance of acute myeloid leukemia (110). Furthermore, Brd4 inhibitors have demonstrated potential as part of combinatorial therapies; they can induce defects in homology-directed DNA repair by decreasing expression of a crucial DNA double-stand break resection protein, CtIP, and thus confer sensitivity to PARP inhibitors, which selectively kill cells with deficient homology-directed repair (111, 112).

Myc is among the primary response genes regulated by Brd4-recruited P-TEFb, and investigation of the anticancer mechanisms of action of P-TEFb or Brd4 inhibitors has centered on Myc expression (33, 113-115). Inhibitors of BET family proteins, such as JQ1, cause a decrease in Myc expression which can be partially recovered in resistant cells through activation of a focal enhancer downstream of the Myc gene that engages the WNT signaling machinery, shifting Myc expression from a Brd4-dependent mechanism to a WNT-dependent mechanism (116). In support of the notion that downregulation of Myc is the primary mechanism by which BET inhibitors exert anti-cancer effects, higher Myc expression has been found to correlate with increased resistance to BET inhibitors in the context of mixed lineage leukemia (117).

Since JQ1 is thought to inhibit cancer cell proliferation and survival by preventing Brd4-mediated recruitment of P-TEFb to key genes such as Myc, direct inhibition of Cdk9 activity might be expected to have the same effect. However, iCdk9, a selective Cdk9 inhibitor, causes only a transient decrease in Myc expression, which is intriguingly followed by an increase above basal conditions with sustained treatment (56). Like other Cdk9 inhibitors, iCdk9 precipitates the

dissociation of P-TEFb from the 7SK snRNP. The released, active P-TEFb associates with Brd4. This complex is less sensitive to Cdk9 inhibitors and can bind at the Myc locus and increase Myc expression despite the presence of iCdk9. Thus the combination of an increase in the size of the active P-TEFb pool, along with the effect of Brd4 on CDK9 activity, leads to stimulation of Myc expression by iCdk9. As expected, the increase in Myc expression induced by iCdk9 can be blocked by JQ1; consequently, the two drugs have synergistic effects on inhibition of cancer proliferation and survival (56).

1.6 Overview and Objectives

In summary, key transcriptional regulators such as P-TEFb and its associated proteins play important roles in cancer and are emerging therapeutic targets for anticancer strategies. This thesis aims to shed light on the mechanisms underlying P-TEFb-dependent cancer progression and sensitivity to transcriptional inhibitors.

Our first objective was to determine the functional significance of P-TEFb-dependent pathways and their relevance in MLL-rearrangement driven cancers. MLL-fusion leukemias are thought to be highly sensitive to P-TEFb inhibition due to the aberrant expression of a small number of direct MLL-fusion target genes. However the notion that P-TEFb function is globally impaired in these cells has not been examined. We hypothesized that MLL rearrangements not only change gene expression patterns, but also disrupt the proper assembly and activity of transcriptional elongation complexes.

Our second objective was to establish a human cell line deficient in Spt5 phosphorylation in order to better understand the roles of this P-TEFb-mediated phosphorylation event. We hypothesized that ablation of Spt5 phosphorylation would lead to decreases in transcriptional pause release and in PAFc-related histone modifications.

Chapter 2: P-TEFb-dependent pathways in MLL-rearrangement driven cancers

2.1 Introduction

MLL, the human homolog of the *Drosophila* gene trithorax (Thx), is a large protein cleaved into two fragments that are both assembled into the same complex (93). The N-terminal fragment contains DNA-binding motifs and interacts with other DNA-binding proteins, while the C-terminal fragment contains the SET domain, which catalyzes methylation of histone H3 lysine 4 (93, 118). Similar to its fly counterpart, MLL regulates the expression of Hox genes during development (119).

Certain recurring MLL translocations at 11q23 have been found to promote leukemogenesis (120). In the majority of cases, leukemia-associated MLL-fusions consist of the N-terminus of MLL fused in frame with a partner protein (121). These fusion proteins do not have methyltransferase activity nor do they interact with the C-terminus MLL fragment (118, 120).

Aberrant transcriptional activation by the MLL-fusion protein is thought to be the main driver of leukemogenesis in mixed lineage leukemia (122). This activation is most commonly attributed to the ability of the MLL-fusion to recruit P-TEFb and other elongation factors to certain genomic loci (92, 123). Notably, the MLL partner protein is frequently a component of the SEC (93). The SEC is organized by a flexible scaffold protein, AFF4, which directly interacts with P-TEFb and either ENL or AF9. The complex also contains AF4 (AFF1) and other elongation factors, such as ELL (124). ENL and AF9 are found in distinct SECs and have been proposed to recruit the SEC to Pol II by interacting with the PAFc through their N-terminal YEATS domains (75) (Figure 2.1A). In a MLL-AF4 transformed leukemia, the YEATS domain of ENL has been shown to be necessary for cell proliferation (125). Conversely, in MLL-ENL or MLL-AF9 transformed leukemias, the YEATS domain of the fusion partner is often lost, leading to the proposal that MLL is able to substitute for the YEATS domain in targeting the SEC to chromatin (75), based on the previous observation that MLL has PAFc binding abilities which

are necessary for MLL-AF9 mediated leukemogenesis (126) (Figure 2.1B). In addition to interacting with PAFc, the YEATS domain can recognize H3K9 acetylation (127). Furthermore, the H3K79 methyltransferase Dot1L has been found to interact with MLL-AF10, and a direct fusion of MLL and Dot1L has been shown to be leukemogenic (128). However, Dot1L appears to inhibit SEC recruitment by competing with AFF4 for binding to ENL and AF9 (75). The recruitment of Dot1L and the resulting deposition of H3K79 methylation marks represent an alternate mechanism by which the transcription of MLL target genes can be activated (Figure 2.1C).

Interestingly, chromatin immunoprecipitation (ChIP) experiments have revealed that MLL-fusions regulate only a small subset of all the genes bound by wild-type MLL (129). Key target genes implicated in cancer progression include the aforementioned Hox genes and other developmental regulators (126, 129, 130).

MLL-rearranged cancers are particularly sensitive to inhibitors of P-TEFb (104) or Brd4 (131). It is unclear to what degree the effectiveness of these inhibitors is due to downregulation of *c-myc*, a Brd4-dependent primary response gene that promotes cell proliferation (56, 116), to downregulation of specific MLL-fusion target genes, or to more global effects on transcription and the rewiring of P-TEFb-dependent pathways. Rewiring could occur as the MLL-fusion disrupts existing protein interactions or enables new ones. To investigate how MLL-rearrangements might cause defects in transcriptional elongation, we treated two cancer cell lines, one with a MLL-rearrangement and one without, with transcriptional inhibitors and characterized their responses. The cell lines we selected for the present study were THP-1, an acute monocytic leukemia line with a MLL-AF9 fusion protein (132), and K562, a chronic myelocytic leukemia driven by an aberrant tyrosine kinase (133). We observed that P-TEFb inhibition reduced viability and P-TEFb activity in both cell lines, but the effects were more pronounced in THP-1 cells relative to K562 cells, raising the possibility that MLL-fusion proteins globally impair P-TEFb.

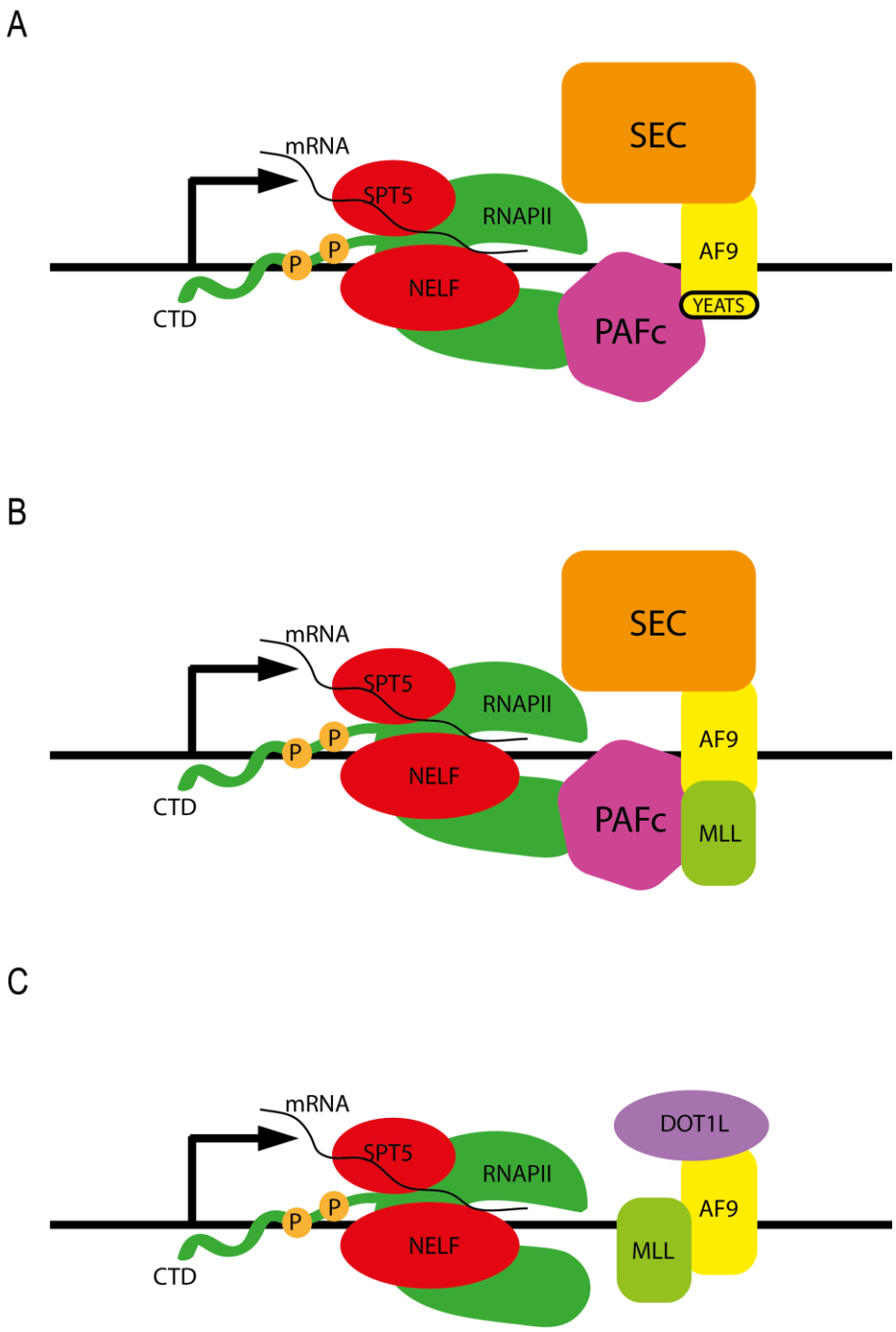

Figure 2.1 Mechanisms of transcriptional activation by MLL-AF9 fusion protein.

A) Recruitment of the SEC to Pol II via interaction between wild-type AF9 and the PAFc.

B) Recruitment of the SEC to Pol II via interaction between MLL-AF9 and the PAFc.

C) Alternate mechanisms of transcription activation. MLL is capable of directly binding to DNA, and DOT1L can compete with AFF4 for binding to AF9.

2.2 Materials and Methods

2.2.1 Cell Culture

THP-1 (ATCC® TIB-202™) cells were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum. K562 (ATCC® CCL-243™) cells were cultured in Iscove's Modified Dulbecco's Medium supplemented with 10% fetal bovine serum.

2.2.2 Transcriptional Inhibitors

THZ1 was a gift from Nathanael S. Gray (Harvard University). iCdk9 was a gift from James Sutton (Novartis). JQ1 was obtained from Cayman Chemical (Catalogue #11187-5).

2.2.3 Propridium Iodide Cell Viability Assays

THP-1 cells at a concentration of $2.0x10^6$ cells/ml or K562 cells at a concentration of $5.0x10^5$ cells/ml were cultured with the indicated concentrations of inhibitors and 0.1% DMSO for 3 or 4 days. To harvest, 1 ml of the cell suspension was centrifuged at 300 g for 5 minutes and then resuspended in 350 µl PBS. Propridium iodide (Sigma Aldrich) was added at a final concentration of 1 µg/ml. 10000 cells per sample were analyzed with the BD FACSCalibur™ flow cytometer. Cells with abnormally low or high forward-scatter and side-scatter values were filtered out, and the remaining cells were plotted according to signal measured by a fluorescence detector with a 585 nm bandpass filter. A fluorescence threshold was set based on the vehicle only samples such that cells with fluorescence levels below the threshold were considered alive and cells with fluorescence levels above the threshold were considered dead.

2.2.4 Differential Salt Extraction

Differential salt extractions were performed based on a previously described protocol (44). Briefly, $2.25x10^6$ THP-1 cells or $1.0x10^6$ K562 cells were incubated with inhibitors for the indicated time periods and collected by centrifugation at 300 g for 5 minutes. The cells were then washed once with cold PBS before being resuspended in 80 µl of low-salt buffer (10 mM KCl, 10 mM MgCl2, 10 mM HEPES-KOH (pH 7.5), 1 mM EDTA, 1 mM DTT, 0.5% IGEPAL, and protease inhibitor cocktail (Roche)) and incubated on ice for 10 min. After centrifuging the lysates at 5000 g for 5 min, the supernatants were collected and set aside as low salt extracts

(containing 7SK snRNP bound P-TEFb). The pellets were washed with 200 µl of low-salt buffer, resuspended in 80 µl of high-salt buffer (450 mM NaCl, 1.5 mM MgCl2, 20 mM HEPES (pH 7.5), 0.5 mM EDTA, 1 mM DTT, 0.5% IGEPAL, and protease inhibitor cocktail (Roche)), mixed by vortexing, and incubated on ice for 10 min. The lysates were centrifuged at 10,000 g for 5 min, and supernatants were collected and set aside as high salt extracts (containing free P-TEFb). 15 ul of each lysate was used per western blot.

2.2.5 Antibodies and Western Blotting

For western blots aiming to assess levels of Rpb1, Spt5, and H2B phosphoisoforms, cells were treated with inhibitors for the indicated time points and lysed directly in 1x Laemmli sample buffer. $2x10^5$ THP-1 cells or $5x10^4$ K562 cells were loaded per lane.

The following antibodies were used.

Target	Name: Catalog Number	Supplier
Rpb1 p-Ser2	3E10: 04-1571	SigmaMillipore
Rpb1 total	8wg16: ab817	Abcam
Spt5 p-T806	PT806	Robert Fisher (80)
Spt5 total	D-3: sc-1332117	Santa Cruz
H2Bub	Ubiquityl-Histone H2B (Lys120) (D11) XP®: 5546S	New England Biolabs
H2B total	Anti-Histone H2B (yeast) antibody: ab188271	Abcam
Cdk9	H-169: sc-8338	Santa Cruz

Densitometric analysis of western blots was performed with Image Studio™ Lite (LI-COR Biosciences).

2.2.6 RNA Isolation

For each sample, between $1x10^6$ and $2x10^6$ cells were treated with 500 nM iCdk9 or 500 nM JQ1 for the indicated time, washed twice with ice cold 1X PBS, and then resuspended in 1 ml TRI reagent (Sigma). Samples were vortexed vigorously for 30 seconds and incubated at room temperature for 5 min. 200 µl bromochloropropane was added to each sample, and mixed by vortexing. Samples were incubated for 15 min at room temperature then centrifuged at 13400g for 15 min at 4°C. 400 µl of the upper phase was transferred to a new tube, taking care not to touch the sides of the tube or disturb the lower phase. An equal volume of room

temperature isopropanol was added. Each sample was then vortexed lightly for a few seconds and incubated at room temperature for 15 min. Samples were then centrifuged at 13400g for 8 min and the supernatants were discarded. The pellets were washed twice in 75% EtOH, then air dried for 5 min. Once dry, the pellets were resuspended in 20 µl of nuclease-free water and quantified with a Nanodrop spectrophotometer.

2.2.7 cDNA Preparation

For each sample, 4 ug of RNA was treated with 0.4U DNAseI (Roche) with 1X incubation buffer (Roche) in a final volume of 20 ul for 15 min at 37°C. The DNAse was then heat inactivated at 75°C for 20 min. Half of this mixture (2 ug of RNA) was hybridized with 0.1 ng random hexamer primers in a final volume of 29 ul by incubation at 70°C for 10 min. This reaction was quenched on ice for at least 2 min before being added to the reverse transcription mix: 1x M-MLV RT buffer (Promega), 0.5 mM dNTPs, 20U recombinant RNasin (Promega), and 100U M-MLV reverse transcriptase (Promega). The reverse transcription reaction was incubated for 10 min at 23°C, then 60 min at 37°C, and finally 5 min at 70°C. This procedure was also done without reverse transcriptase for each sample to yield negative controls.

2.2.8 qPCR

Each qPCR reaction was done in a final volume of 15 µl (1X Evagreen qPCR Master Mix, 0.3 µM of each primer, and 6.6 µl of cDNA diluted 1/100). Samples were assayed in triplicate using the Corbett Gene Rotor.

The following primers were used:

c-myc F: 5'- TCCTTGCAGCTGCTTAGACGC -3'

c-myc R: 5'- GGAGGCTGCTGGTTTTCCACT -3'

U6 snRNA F: 5'- GCTTCGGCAGCACATATACTAAAAT -3' (134)

U6 snRNA R: 5'- CGCTTCACGAATTTGCGTGTCAT -3' (134)

Results were analyzed using the delta-delta Ct method.

2.2.9 Statistical Analysis

Dose-response curves were created by nonlinear regression with variable slope in GraphPad Prism 6. Statistical tests were done in MATLAB.

2.3 Results

2.3.1 THP-1 and K562 differ in their sensitivity to transcriptional inhibitors

In order to assess the cytotoxic effects of transcriptional inhibitors on THP-1 and K562, we assessed cell viability with propridium iodide after 3 or 4 days of treatment with either THZ1, a covalent Cdk7 inhibitor (135), iCdk9, a selective Cdk9 inhibitor which binds the ATP-binding pocket of the kinase (56), or JQ1, a competitive inhibitor of Brd4 binding to chromatin (136). In the case of THZ1, the K562 cell viability did not drop below 50% after 4 days even at the highest dose (200 nM), while this same dose and incubation time was nearly completely lethal for THP-1 cells (Figure 2.2A). However, the estimated EC_{50} values were similar (THP-1 cells at 4 days, EC_{50} = 27.93 nM, 95% CI = [25.60 nM, 30.48 nM]; K562 cells at 4 days, EC_{50} = 31.62 nM, 95% CI = [20.81 nM, 48.04 nM]), suggesting that both cell lines were affected at similar doses, even if the K562 cells were less sensitive during the tested time frame. This similarity is expected given that Cdk7 is a general transcription factor whose activity is necessary for transcription initiation by Pol II in all situations.

For iCdk9 and JQ1, however, notable differences were observed. With iCdk9 treatment, the dose-response curve for K562 cells was right-shifted compared to THP-1 (THP-1 at 4 days, EC_{50} = 214.0 nM, 95% CI = [175.1 nM, 261.5 nM]; K562 at 4 days, EC_{50} = 747.8 nM, 95% CI = [669.1 nM, 835.7 nM]) (Figure 2.2B). The difference was even more pronounced with JQ1 treatment, where the THP-1 cells show an estimated EC_{50} of 218.1 nM at 4 days (95% CI = [145.3 nM, 327.4 nM]), while the K562 cells exhibited no detectable decrease in cell viability even at the highest dose tested (2 μM) (Figure 2.2C). Nevertheless, in both cell lines, JQ1 potentiated the cytotoxic effects of iCdk9 (Figure 2.2D-E), confirming a previously reported synergistic effect (56).

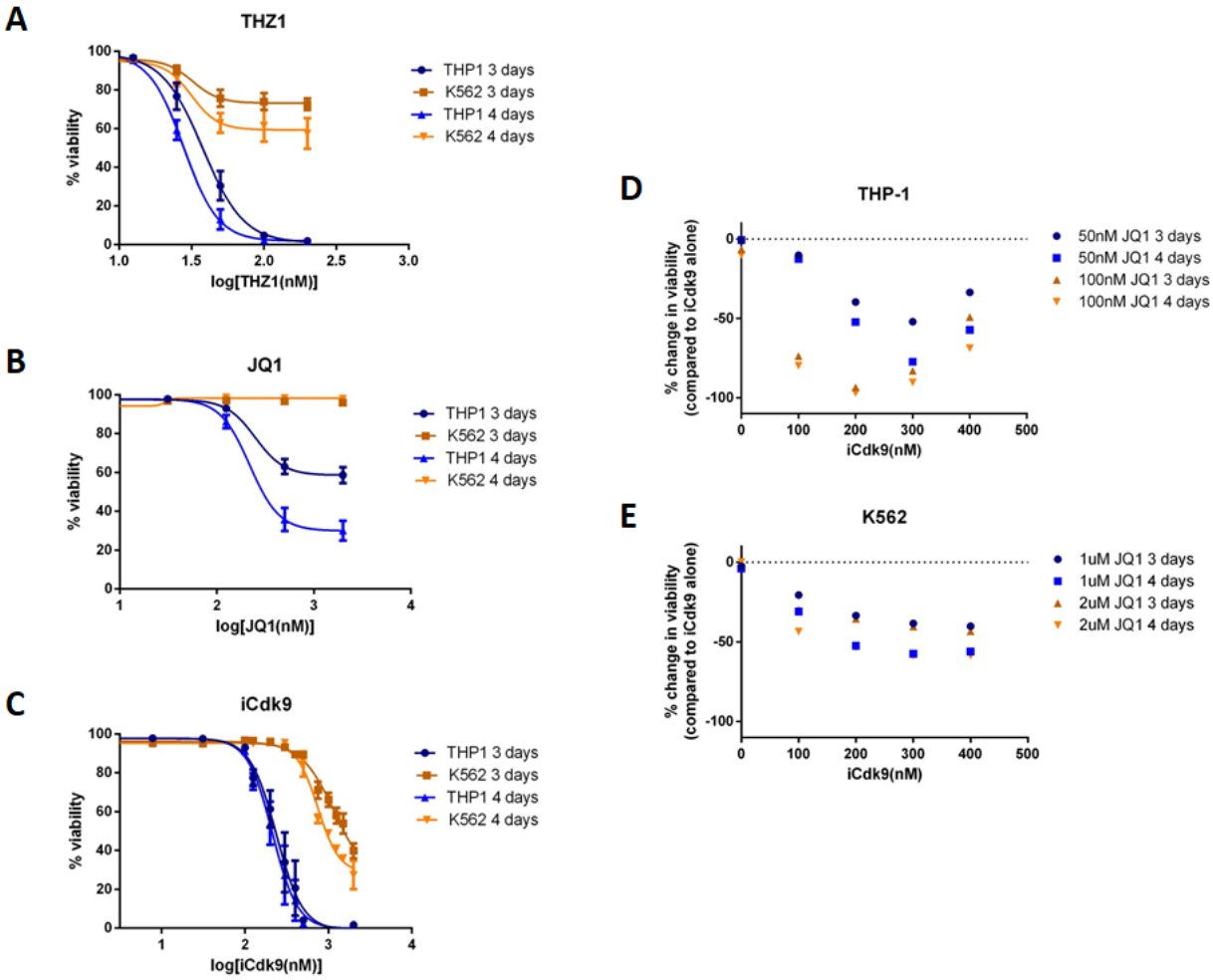

Figure 2.2 THP-1 and K562 differ in their sensitivity to transcriptional inhibitors. A-C) Dose-response curves of cell viability as measured by flow cytometry with propridium iodide. Data represent mean and SEM from between one and four experiments. D-E) Percent change in cell viability of cells treated with a given dose of JQ1 and varying doses of iCdk9 compared with iCdk9 alone. This experiment was done once.

2.3.2 THP-1 and K562 cells differ in inhibitor-induced compensatory Myc expression

Given the importance of Myc in cancer progression, we investigated the effects of iCdk9 and JQ1 on Myc expression at different time points. RT-qPCR revealed that both drugs acutely downregulated Myc in both cell lines (Figure 2.3). With iCdk9, the downregulation was more severe in THP-1 than in K562 cells; moreover, while Myc transcript levels increased above baseline by 24 hours in K562 cells, no such compensatory expression was seen in THP-1 cells. With JQ1 treatment, both cell lines were similarly affected and no compensatory expression was seen in either.

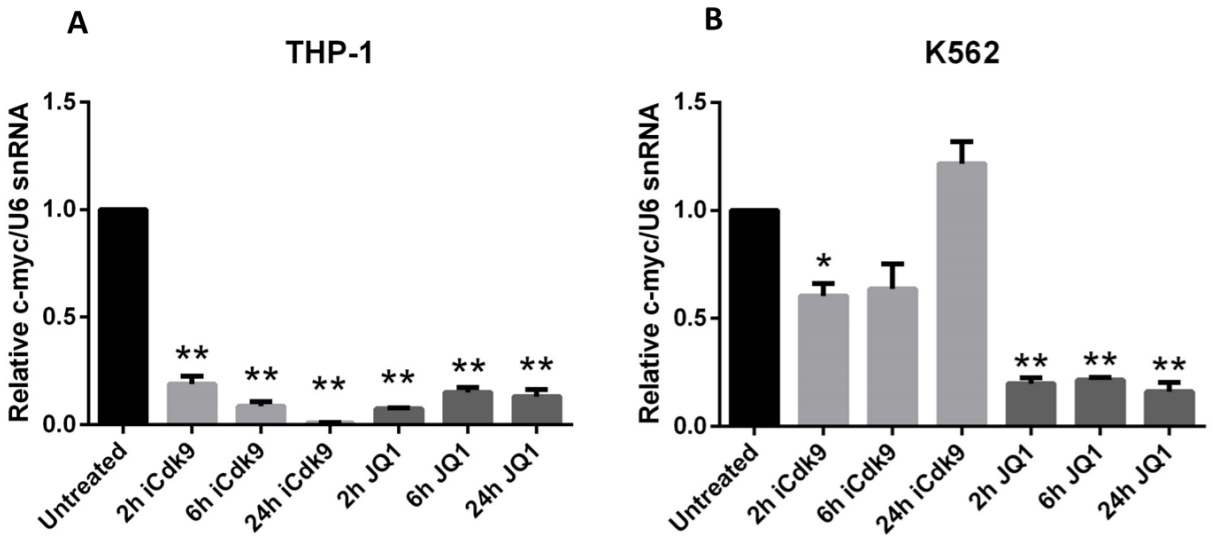

Figure 2.3 *c-myc* mRNA levels after varying lengths of inhibitor treatment. RT-qPCR of *c-myc* in THP-1 and K562 cells treated with 500 nM iCdk9 or 500 nM JQ1 for the indicated times. Results were analyzed using the delta-delta Ct method, by normalizing to the U6 snRNA. Data represent mean and SEM of three biological replicates. *p < 0.05, **p < 0.005, t-test compared to untreated.

2.3.3 THP-1 cells are more sensitive to inhibition of P-TEFb activity

To investigate how P-TEFb activity was affected by iCdk9 and JQ1, we examined levels of downstream targets of P-TEFb through western blotting. We found that Rpb1 Ser-2 phosphorylation, Spt5 CTR phosphorylation, and H2B ubiquitination all decreased in response to iCdk9 in both cell lines (Figure 2.4A-C). Interestingly, THP-1 cells appeared to be more affected at lower doses than K562 cells, suggesting that they have less overall P-TEFb activity.

Opposite trends were seen in response to JQ1 treatment (Figure 2.4D-F). While Rpb1 Ser-2 phosphorylation levels were variable and showed a downward trend in both cell types, Spt5 CTR phosphorylation and H2B ubiquitination decreased in a dose-dependent manner in K562 cells, but not in THP-1 cells. This result suggests that THP-1 cells have less Brd4-dependent P-TEFb recruitment overall, despite requiring certain Brd4-responsive genes for their growth and proliferation. This interpretation is consistent with the notion that MLL-AF9 fusion proteins mainly exert their leukemogenic effect by recruiting P-TEFb through the SEC. The fact that JQ1 caused smaller decreases in P-TEFb-dependent markers than iCdk9 is consistent with the idea that it regulates P-TEFb in a gene-specific manner. Furthermore, although Spt5 CTR phosphorylation appears to be less affected by JQ1 than by iCdk9 in both cell lines, the decrease in H2B ubiquitination in K562 cells was similar for both drugs, suggesting alternative mechanisms for regulating H2B ubiquitination that are independent of Spt5 CTR phosphorylation status.

If MLL-rearranged leukemias were sensitive to transcriptional inhibition solely due to a heavy reliance on the transcription of certain key genes, then we would not expect to see large differences in their global response to Cdk9 inhibition. Such observed differences suggest that the two cell lines have different overall levels of P-TEFb activity, or that P-TEFb pathways are differentially utilized in the two cell lines.

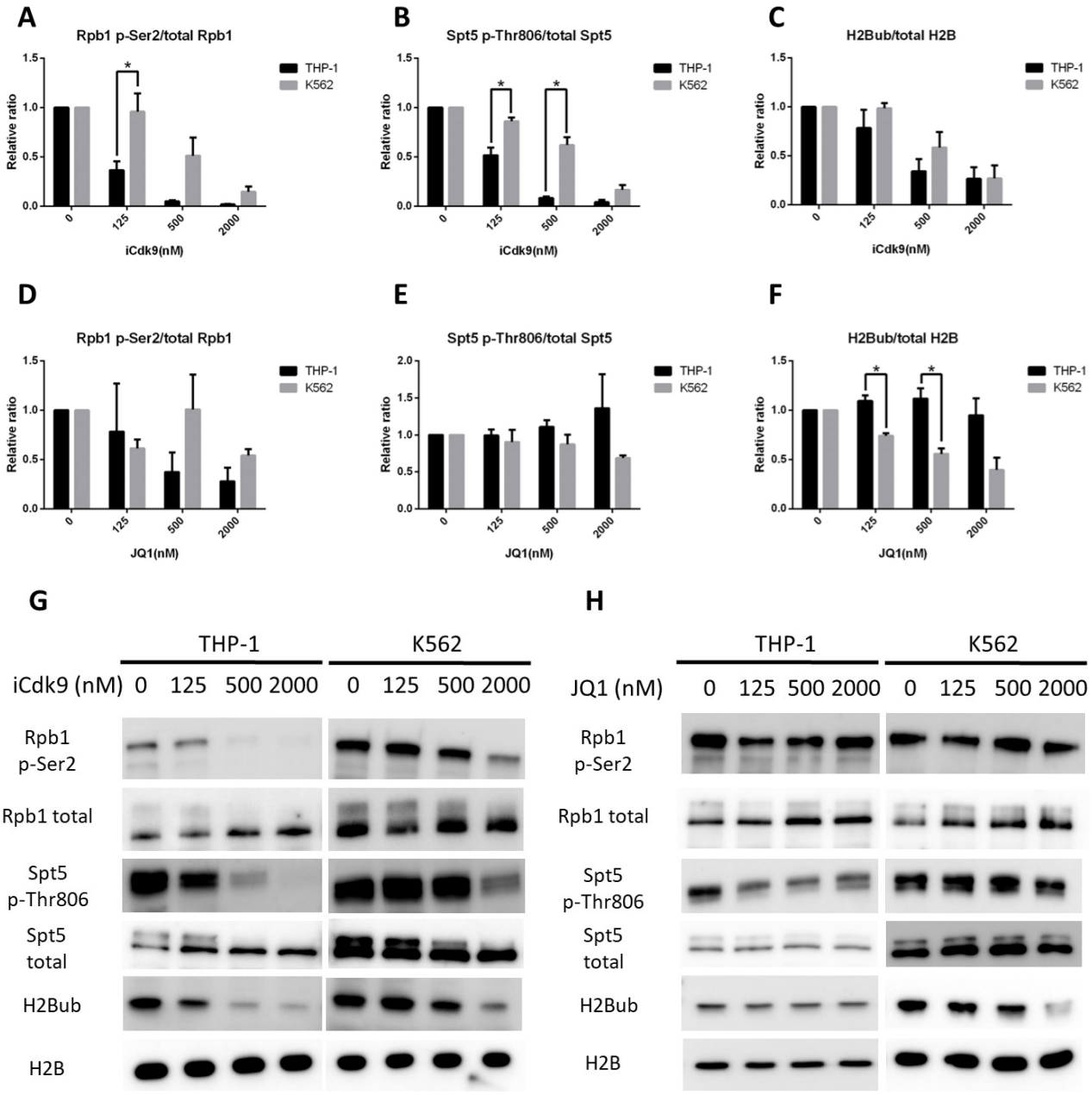

Figure 2.4 P-TEFb activity after inhibitor treatment. THP-1 and K562 cells were treated with the indicated concentrations of iCdk9 or JQ1 for 6 hours. Protein levels were determined via western blotting. A-F) Quantification of western blots, examples of which are shown in G-H. Data represent mean and SEM of three biological replicates. *$p < 0.05$, two-sample t-test. G-H) Representative western blots.

2.3.4 P-TEFb inhibition leads to an increase in free P-TEFb

The apparent decreased P-TEFb activity in the THP-1 cells could reflect a difference in P-TEFb activation. Inhibitors of P-TEFb have previously been reported to cause the release of P-TEFb from its inactive complex with 7SK snRNA. To determine what proportion of the cellular pool of P-TEFb is in the inactive 7SK-bound form as opposed to the free active chromatin-bound form and how this balance shifts with inhibitor treatment, we used a differential salt extraction method. With this method, 7SK-bound P-TEFb is extracted in a low salt buffer while the chromatin-bound form is extracted in a high salt buffer.

Both JQ1 and iCdk9 decreased the proportion of Cdk9 in the 7SK-bound form (Figure 2.5A, 2.5C). Furthermore, this decrease was dose-dependent with iCdk9 (Figure 2.5B, 2.5D). Although a greater proportion of Cdk9 is inactive in THP-1 cells in the absence of inhibitor, overall, both cell lines behaved similarly in terms of inhibitor-induced P-TEFb activation. In THP-1 cells, two bands could be resolved near the expected molecular weight for Cdk9 (Figure 2.5C, 2.5D); the upper band was quantified as it appeared to migrate and respond to inhibitor treatment similarly to the Cdk9 band from K562 cells.

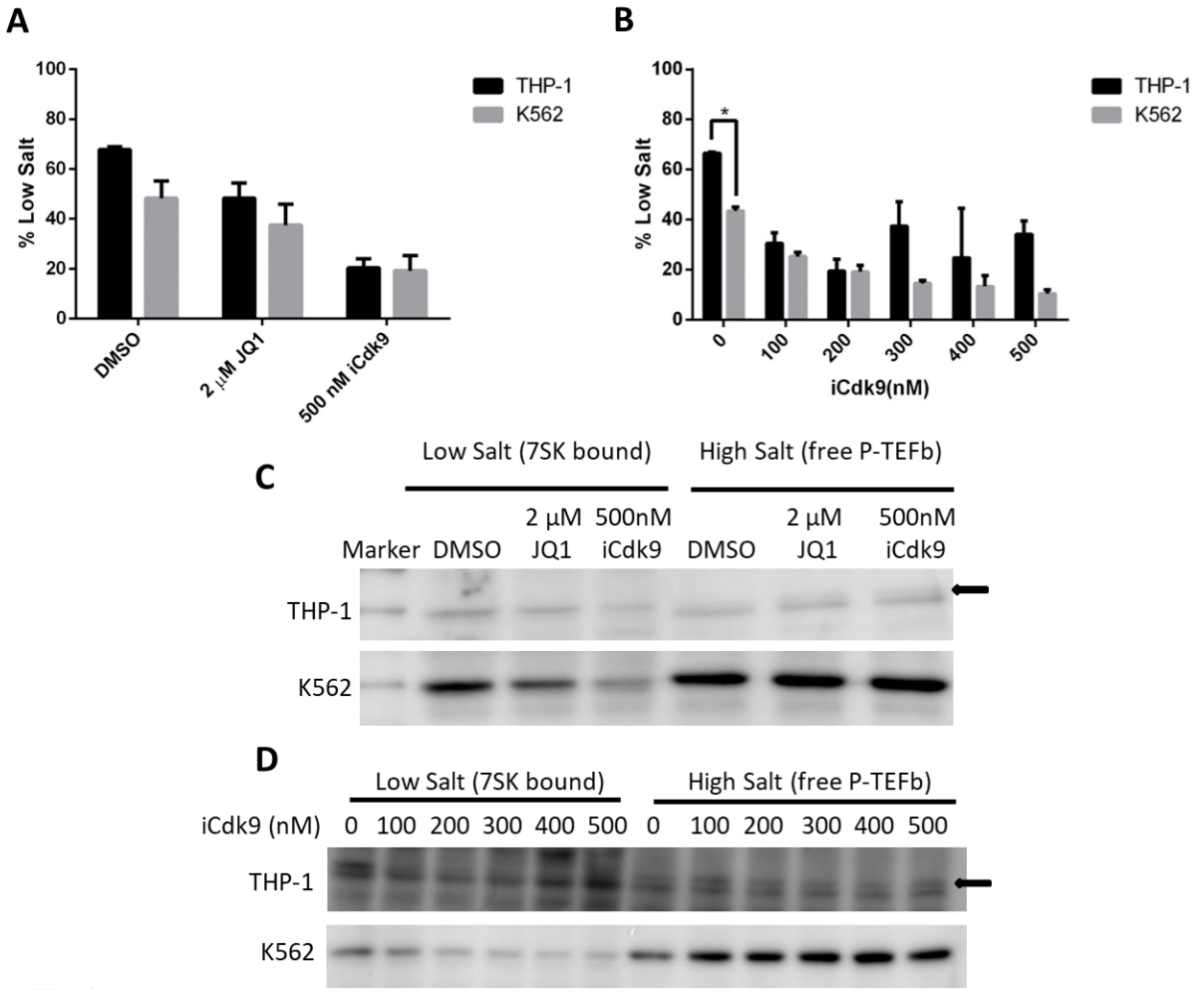

Figure 2.5 P-TEFb inhibition leads to an increase in free P-TEFb. THP-1 and K562 cells were treated with the indicated concentrations of iCdk9 or JQ1 for 1 hour. 7SK bound and free P-TEFb were separated by differential salt extraction. Cdk9 levels in each fraction were determined via western blotting. A) Quantification of western blots, example of which is shown in C. Percentage of Cdk9 found in the low salt (7SK bound) fraction after treatment with 2 μM JQ1 or 500 nM iCdk9. Data represent mean and SEM of two biological replicates. B) Quantification of western blots, example of which is shown in D. Percentage of Cdk9 found in the low salt (7SK bound) fraction after treatment with different doses of iCdk9. Mean and SEM of two biological replicates. *$p < 0.01$, two-sample t-test. C-D) Representative western blots. Arrows indicate the THP-1 bands that were quantified.

2.4 Discussion

2.4.1 Summary and interpretation of findings

In order to interrogate the mechanisms underlying the sensitivity of MLL-rearranged leukemias to transcriptional inhibitors, we examined the responses of THP-1, which has a MLL-AF9 fusion, and K562, which does not have a MLL translocation, to iCdk9 and JQ1, inhibitors of P-TEFb and Brd4 respectively. We observed that lower doses of both iCdk9 and JQ1 led to cell death in THP-1 cells compared to K562 cells. This difference in sensitivity to JQ1 has been previously reported (110). Our observations also complement an earlier study proposing that iCdk9 and JQ1 would have synergistic effects (56). This same study demonstrated that Myc expression is acutely downregulated by iCdk9 but becomes upregulated with sustained treatment in several cell lines. In our experiments, this compensatory Myc expression was observed in K562 but not in THP-1 cells.

The absence of Myc upregulation in THP-1 cells after 24 hours treatment does not necessarily indicate that THP-1 cells completely lack the regulatory circuits that upregulate Myc in other cell lines. Sustained treatment with the dosage used leads to cell death, and it is possible that the cells might be able to upregulate Myc in response to lower doses. However, it is clear that the downregulation of Myc in response to Cdk9 is greater in THP-1 than in K562 cells. The increased sensitivity of THP-1 to iCdk9 may therefore partially be explained by Myc downregulation. This raises the possibility that expression of the MLL-fusion in THP-1 cells interferes with Cdk9-dependent regulation of Myc.

In contrast, Myc downregulation under JQ1 treatment was similar at all tested times, suggesting that Myc expression differences cannot account for the differential sensitivity to JQ1. This also argues that Brd4-dependent mechanisms for Myc activation are not differentially affected in the two cell lines. It has been previously reported that certain cell lines, including K562, are able to overcome Myc downregulation by JQ1 through the activation of an alternative enhancer that can be activated independently of Brd4 (116), but this effect was not seen in our cells.

One explanation for the greater loss of cell viability in THP-1 compared to K562 upon treatment with a Cdk9 inhibitor is that THP-1 cells simply have a greater dependency on P-TEFb activity. The MLL-AF9 fusion protein has over 900 target genes in THP-1 (137), many of which

are important for maintaining an oncogenic phenotype. It is possible that a large amount of active P-TEFb is required for survival in THP-1 cells and that this accounts for their relatively increased sensitivity to iCdk9. However, while this may be true, this model would not predict global differences in how P-TEFb activity levels are affected by Cdk9 inhibition.

Strikingly, when we examined markers of P-TEFb activity, namely levels of Ser-2 phosphorylation of the Pol II CTD, Thr-806 phosphorylation of Spt5, and H2B ubiquitination, we found greater decreases in all three in THP-1 compared to K562 for the same dose of iCdk9. This discrepancy raises the possibility that THP-1 cells have less overall P-TEFb activity. Intriguingly, in K562 cells, Spt5 CTD phosphorylation is more affected by iCdk9 than by JQ1, while H2B ubiquitination is affected to a similar degree by both drugs, suggesting that JQ1 decreases H2B ubiquitination by Spt5 CTD phosphorylation-independent mechanisms. For instance, it is possible that the JQ1 effect on H2B ubiquitination instead occurs through a Mediator-dependent pathway that has previously been described in mammalian cells (138). Finally, although activation of P-TEFb in response to iCdk9 treatment seemed similar in both cell lines, a smaller proportion of the cellular pool appeared to be active in THP-1 cells under unstimulated conditions. However, due to technical difficulties with the detection of Cdk9 in THP-1 extracts, this last observation would need to be reconfirmed with the use of a more specific antibody or alternative extraction conditions.

2.4.2 Future directions

Having observed that THP-1 cells have reduced viability and reduced markers of P-TEFb activity at equal doses of iCdk9 compared to K562 cells, we would like to ask whether these differences arise because of the MLL-AF9 fusion protein. While the dominant model of MLL-rearrangement driven leukemia posits that the fusions proteins transform cells by increasing transcription of key oncogenes, it is also possible that they have more general effects on the assembly of transcriptional complexes. For instance, it is known that MLL-AF9 has PAFc binding abilities (126) which could be interfering with other PAFc interactions.

Various studies in different contexts have concluded that the PAFc acts either as a pause factor or as a positive regulator of elongation (72-74), but PAFc recruitment and activity have also been identified downstream of pause release. P-TEFb activity has been shown to promote

the recruitment of PAF and Rtf1 as well as the deposition of histone marks promoted by the PAFc, such as H2B ubiquitination (80, 139).

To better understand the functional relationship between the PAFc and P-TEFb, we are performing ChIP for Pol II and PAFc subunits in THP-1 and K562 with and without iCdk9. One intriguing possible outcome of this experiment would be that PAFc recruitment to target genes is differentially sensitive to iCdk9 in the two cell lines, perhaps pointing to a mechanistic difference in P-TEFb pathways that could explain the differential sensitivity to iCdk9 that we observe.

To corroborate our observation that a larger fraction of P-TEFb is sequestered by the 7SK snRNP in THP-1 cells and to study other P-TEFb interactions, we would like to perform co-immunoprecipitation experiments in both cell lines. By pulling down elements of the 7SK snRNP and the SEC complex, we would be able to isolate inactive from active P-TEFb.

One limitation of this study is that THP-1 and K562 cells differ by more than just a MLL-translocation. A wider panel of leukemia cell lines with or without MLL rearrangements would be need be assayed in order to relate any observed differences to MLL fusion proteins. To conclusively attribute effects to MLL fusion proteins, we would need to work with a system where the presence of MLL fusion proteins is a controllable variable.

MLL-AF9 has previously been knocked down in THP-1 cells with RNA interference using siRNA targeting the fusion point, leading to cell differentiation but no observed loss of proliferation (140). However, two earlier studies, which both downregulated MLL-AF9 in THP-1 using oligodeoxyribonucleotides antisense to the fusion junction, concluded that loss of MLL-AF9 led to increased apoptosis and decreased cell proliferation (141, 142). Decreased cell proliferation and viability presents an additional hurdle to many downstream assays, and would not be ideal for our purposes. Other studies have introduced MLL fusion proteins in a variety of expression systems (124, 143-145). Consistent with the notion of MLL fusion proteins being oncogenic, transduction of hematopoetic progenitor cells with MLL-AF9 led to their immortalization (144). On the other hand, expression of MLL-AF4 in Jurkat cells was observed to slow cell proliferation (145). Such findings could be consistent with the idea that MLL fusion proteins, despite activating oncogenes in certain cell contexts, globally decrease P-TEFb activity; in this case, they would be expected to negatively affect proliferation in cell types without physiological relevance to mixed lineage leukemia. In our lab, efforts are currently underway to

introduce MLL fusion proteins into cell lines with wild type MLL to test whether the presence of MLL fusions proteins could be sufficient to induce the observed characteristics, namely increased sensitivity to transcriptional inhibitors and reduced P-TEFb activity, and any other differences observed between THP-1 and K562. This introduction will be done with transient transfection of mammalian overexpression plasmids, but in the long term we would want a stable inducible expression system.

In this chapter, we sought to determine the functional significance of P-TEFb-dependent pathways in MLL-rearrangement driven cancers. We have shown that the MLL-rearranged THP-1 leukemia line exhibits reduced viability and greater loss of P-TEFb activity upon treatment with a P-TEFb inhibitor compared to a non-MLL-rearranged leukemia line, K562. Based on our observations, we propose that the high sensitivity of MLL-fusion leukemias to P-TEFb inhibition results from a global impairment of P-TEFb activity, contrary to the current model where the leukemias are dependent on P-TEFb activity at a small number of direct MLL-fusion target genes.

The next chapter details our efforts to establish a human cell line with phosphoacceptor site mutations in the P-TEFb target Spt5 using CRISPR/Cas9. Such a cell line would serve as a useful tool in the investigation of P-TEFb-dependent pathways requiring Spt5 phosphorylation.

Chapter 3: CRISPR/Cas9-mediated editing of phosphorylation sites in the Spt5 C-terminal region

3.1 Introduction

P-TEFb-mediated phosphorylation at the threonine residue of the GS(R/Q)TP repeat in the CTR of Spt5 has a positive effect on transcriptional elongation. A previous study demonstrating this effect in human cells relied on knocking down the endogenous gene and attempting to rescue function with the introduction of either the wild type protein or a phosphorylation-deficient mutant. Yamada *et al.* created HeLa cells lines that expressed RNAi-resistant versions of wild-type or mutant Spt5 under the control of tetracycline-repressible promoters. The cells were then transduced with lentiviruses expressing shRNA against endogenous Spt5, thus knocking down the endogenous protein without affecting the RNAi resistant forms. Characterization of these cell lines revealed a defect in pause release at certain genes, but it is still unknown whether loss of Spt5 phosphorylation has general effects on elongation and other P-TEFb-dependent pathways (27).

These previously-used experimental strategies for studying mutants are well-established and the generation of cell lines constitutively expressing the exogenous protein and shRNA against the endogenous form has a clear advantage over transient expression: avoiding variation from experiment to experiment due to the need to introduce exogenous nucleic acids for every experiment. However, certain disadvantages remain: knockdown of the endogenous gene is incomplete and the exogenous protein, being overexpressed by a constitutive promoter, may escape important regulatory processes.

These issues can be circumvented by replacing the endogenous gene with the mutant rather than expressing exogenous genes. This approach ensures that the mutant is regulated and expressed in a similar manner to the wild type gene. Replacement of endogenous Spt5 with a phosphorylation-deficient mutant has been accomplished in the *S. pombe* model system. Strains with this mutation recruit capping enzymes similarly to wild-type but exhibit growth defects at low or high temperatures (146). These mutants have less chromatin-associated Rtf1, although a small amount remains bound in a Cdk9-independent manner. Interestingly, although chromatin

association of the Tpr1 subunit of the PAFc is also decreased, binding remains Cdk9-dependent, suggesting that Cdk9 promotes PAFc recruitment through targets other than the Spt5 CTR. Furthermore, *in vitro* assays have confirmed that Rtf1 binding to Spt5 is increased by CTR phosphorylation (70).

Although Spt5 and other key transcriptional regulators are evolutionarily conserved, it is believed that the wiring of P-TEFb-dependent pathways could differ significantly between yeast and metazoans (81). More work must therefore be done to establish the roles of Spt5 phosphorylation in humans and other metazoans.

In recent years, the development of CRISPR/Cas9-mediated gene editing technologies has facilitated gene replacement efforts in a wide range of species. The Cas9 endonuclease forms a complex with a single guide RNA (sgRNA) that allows it to recognize and cleave specific DNA sequences that are complimentary to an 18-20 nucleotide sequence on the sgRNA, which can be easily engineered to confer specificity to a DNA target of interest (147, 148). The resulting double strand break can be repaired by non-homologous end joining (NHEJ), which can introduce non-specific insertions or deletions, or by homology-directed repair (HDR), which can introduce desired mutations if an appropriate repair template is supplied (149).

Several methods have been developed to improve the relative efficiency of obtaining cells edited by HDR, ranging from the use of small molecules to favour HDR over NHEJ (150, 151) to the introduction of selectable markers in the repair template (152). However, selectable markers such as fluorescent proteins can only be incorporated at either the N- or the C-terminus of a protein to minimize disruption of protein folding and function; such markers cannot be used when editing a region such as the Spt5 CTR, which is encoded across two internal exons.

Given this restriction, a marker-free coselection strategy has recently been developed for use in human cells. Since the efficiency of HDR tends to depend on cell-specific factors such as cell cycle status and the global activation of certain DNA repair pathways, HDR events are not independent: a cell that has successfully undergone HDR at one genomic locus is more likely to have successful HDR at another locus. The method in question hinges on the fact that a two amino acid substitution of the sodium potassium pump ATP1A1 confers resistance to the cardiac glycoside ouabain. By cotransfecting a sgRNA and repair template for ATP1A1 that mutate the two amino acids in question along with the CRISPR components for the gene of interest, one could enrich for successfully edited cells by selection with ouabain (153).

There has also been considerable debate over the ideal form of the repair template. For relatively small edits of a few nucleotides, single-stranded oligodeoxynucleotides (ssODN) less than 100 nucleotides in length are optimal (154). However, longer homology arms of at least 1-2 kilobase pairs are required for editing larger regions (155). The repair templates used for these larger regions have typically consisted of double-stranded DNA (dsDNA), either circular or linear. A recent study suggests that the most efficient dsDNA template for HDR is a circular plasmid containing the region to be edited along with homology arms flanked by sgRNA target sequences; this donor is linearized in the nucleus by the same Cas9/sgRNA complex that cleaves the genomic region of interest (156). Based on these recent insights, we devised and employed a CRISPR/Cas9 strategy to replace the threonines of the Spt5 CTR repeat region with alanines, in order to generate a mammalian cell line deficient in Spt5 phosphorylation. This cell line would serve as a model for interrogating P-TEFb-dependent pathways that rely on Spt5 CTR phosphorylation, and confirm whether loss of Spt5 phosphorylation would decrease transcriptional pause release and PAFc related histone modifications in a mammalian system.

3.2 Materials and Methods

3.2.1 Sequencing of Spt5 CTR Genomic Locus in THP-1 cells

Genomic DNA from THP-1 cells was extracted using a Genomic DNA Mini Kit for Blood/Cultured Cells (Geneaid). 50 ng genomic DNA was used per PCR. PCRs were done in final volumes of 50 ul with 50 ng genomic DNA, 0.2 mM dNTPs, 0.4 μM of each primer, and 2.5U Taq polymerase. PCR products were sent to Genome Quebec for Sanger sequencing. The following primers were used for amplification and sequencing:

1F: 5’- GCTTCTATCCTAGCCCCGTG -3’

1R: 5’- ACGGTCCACAGAGATGGTCT -3’

2F: 5’- AAGCAAGTGAAGGGGACGTT -3’

2R: 5’- CATAGGTCGAGGTCATGCCG -3’

3F: 5’- AGACCATCTCTGTGGACCGT -3’

3R: 5’- GCGACTGTGGTGATAATGCG -3’

4F: 5’- CTCCATCCCCAACCTCAGTA -3’

4R: 5’- GGGCAGCATAGGGAGAGAAC -3’

5F: 5’- ACCCACAGGAGGGTAGACG -3’

5R: 5’- AACAAATGGGGTGAGCAGAC -3’

1Fn: 5’- ATGTGACCAACTTCACCGTG -3’

1Rn: 5’- TGGACTCTGTGGCATCTTTC -3’

4Fi: 5’- TGGTCTGTGTTGTACCTGCA -3’

4Fo: 5’- TGGGCTGGTAGGAACCTTGT -3’

We used a combination of nested PCR and gel extraction techniques to obtain pure PCR amplicons suitable for Sanger sequencing. Primers 3F, 3R, and 4R were also used for screening.

3.2.2 Single guide RNA design and Surveyor assay

Single guide RNAs targeting the Spt5 CTR were designed using the Crispr.mit tool developed by the Zhang Lab (157). We selected two candidate sgRNAs with on-target activity scores above 80 and which started with a guanosine nucleotide, a requirement for efficient transcription initiation through the U6 promoter.

These candidate sgRNAs were cloned into the eSpCas9(1.1) vector using previously established methods (158). eSpCas9(1.1) was obtained from Feng Zhang (Addgene plasmid #71814) (159). 1.5 µg of eSpCas9(1.1) with the candidate sgRNA was transfected into 4.5 x 10^5 HEK 293F cells using Lipofectamine 2000 following the manufacturer’s protocol. Three days later, the cells were harvested and genomic DNA was extracted using a Genomic DNA Mini Kit for Blood/Cultured Cells (Geneaid). The locus of interest was first amplified by PCR using the primer pair 3F and 3R. The resulting product was diluted 1/20 and 2 µl were used as the template in a subsequent round of PCR using the following nested primers:

3Fn: 5’- CCATCTCTGTGGACCGTCAG -3’

3Rn: 5’- CTGAGGTTGGGGATGGAGAG -3’

The products of this second PCR were used for the Surveyor assay (Integrated DNA Technologies) following the manufacturer’s protocol.

3.2.3 CRISPR/Cas9 Plasmids and Repair Template

eSpCas9(1.1)_No_FLAG_ATP1A1_G3_Dual_sgRNA (Addgene plasmid # 86613) and ATP1A1_plasmid_donor_RD (Addgene plasmid #86551) were obtained from Yannick Doyon.

sgRNA#2 was cloned into eSpCas9(1.1)_No_FLAG_ATP1A1_G3_Dual_sgRNA using previously established methods (158).

The m7 Spt5 repair template was synthesized by Integrated DNA Technologies. We added restriction sites and overhangs to the repair template by PCR with Taq polymerase and the following primers:

Fwd: 5'- TAAGCACTCGAGGGTACCGTGAGCCGTACA -3'

Rvs: 5'- GGTGCAAAGCTTGAATTCCGAACACCCATG -3'

The resulting PCR product was cloned into pBlueScript II SK+.

The m4 intronless Spt5 repair template was synthesized and cloned into pUC57-Amp by BioBasic.

3.2.4 Transfection, Ouabain Selection, and Screening

HEK 293F or 293 parental cells were cultured in Dulbecco's Modified Eagle's medium supplemented with 5% FBS. eSpCas9(1.1)_No_FLAG_ATP1A1_G3_Dual_sgRNA containing sgRNA#2, ATP1A1_plasmid_donor_RD, and either the m7 or the m4 intronless Spt5 repair template cloned into a circular plasmid were transfected into a half-confluent T75 flask of HEK 293F or 293 parental cells at a 1:1:3 molar ratio. 10 µl of Lipofectamine 2000 was used to transfect 10 µg of plasmid DNA according to the manufacturer's protocol. Two days after transfection, cells were trypsinized and replated in complete media containing 1 µM ouabain (Sigma). Cells were grown in ouabain-containing media for at least a week and a half, or until colonies were clearly visible and confluent at their centers. At this point, ouabain-resistant cells were separated into monoclonal lines.

To screen for successful genomic integration of the repair template, crude cell lysates were prepared. Cells were washed in Kreb's buffer (126 mM NaCl, 2.5 mM KCl, 25 mM $NaHCO_3$, 1.2 mM NaH_2PO_4, 1.2 mM $MgCl_2$, 2.5 mM $CaCl_2$), then resuspended in 10 mM Tris-HCl pH 8.8 at 2 x 10^4 cells/ul. The cells were heated at 95°C for 10 minutes, then cooled to room temperature before the addition of proteinase K at a final concentration of 0.4 mg/ml. The lysates were incubated at 56°C for 30 minutes, and then at 95°C for another 10 minutes to deactivate the proteinase K. 5 ul of this crude lysate was used for PCR with Taq Polymerase and with either the primer pair 3F and 3R to assess the presence of the repair template independent of location or 3F and 4R to assess the presence of the repair template at the specific locus of interest. PCR

products were treated with the SacI restriction enzyme and visualized by gel electrophoresis with a 1% agarose gel or sent to Genome Quebec for Sanger sequencing.

3.3 Results

3.3.1 Single guide RNA validation

We tested two candidate sgRNAs (Figure 3.1A) for the ability to induce double strand breaks at the genomic locus of interest using the Surveyor assay, which detects the presence of mismatches. These mismatches occur as a result of insertion or deletion mutations introduced by NHEJ. Such mismatches could be detected in HEK 293F cells transfected with eSpCas9 and either candidate sgRNA, but not in cells transfected with eSpCas9 alone (Figure 3.1B), indicating that both candidate sgRNAs are capable of directing Cas9-mediated cleavage of the target genomic locus. We chose to proceed with sgRNA#2.

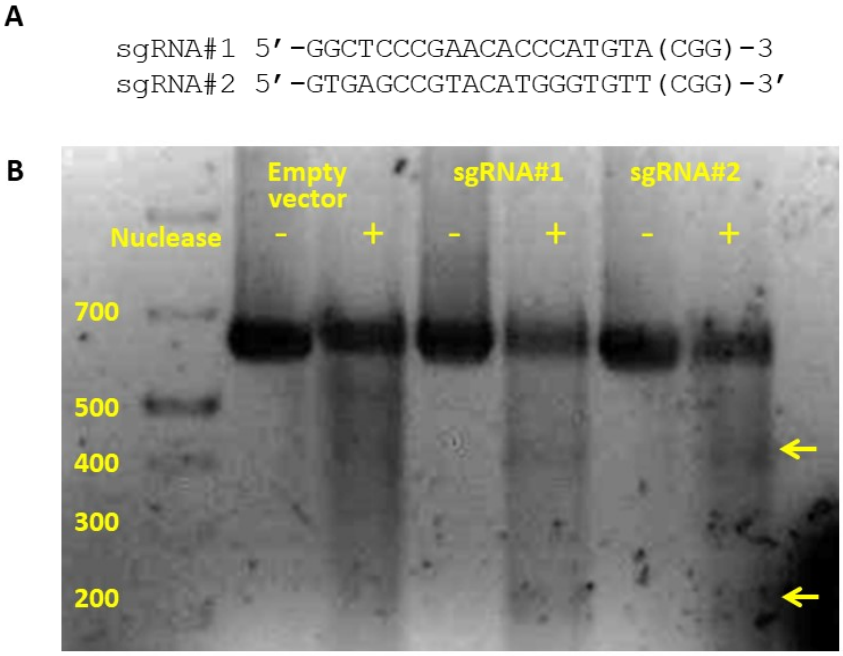

Figure 3.1 Single guide RNA design and validation. A) sgRNAs tested. B) Surveyor assay assessing the presence of NHEJ-induced mutations in HEK293F cells transfected with the candidate sgRNAs. Arrows indicate positions of expected bands resulting from mismatches at the desired location.

3.3.2 Repair Template Design

We opted for a dsDNA repair template in a circular plasmid with 750 base pair homology arms flanked by sgRNA recognition sites on either side of the CTR (Figure 3.2A). In order to determine the sequence of the homology arms and the wild-type CTR, we isolated genomic DNA from THP-1 cells and performed PCR followed by Sanger sequencing. Select point mutations were introduced to replace all seven CTR threonines (T768, T775, T784, T791, T799, T806, T814) with alanines, to disrupt the sgRNA recognition sequence, and to introduce a SacI restriction site for facilitated screening of edited cells. This repair template was named the m7 Spt5 repair template (m7).

The efficiency of HDR is known to decrease drastically with increased distance between the sgRNA cut site and the region to be edited (160). When multiple proximal point mutations are introduced with a single donor template, it has been reported that the mutations closest to the cut site are incorporated with a higher frequency (161, 162). Nevertheless, CRISPR-mediated gene replacement can be effective for long distances: the replacement of a human gene with its mouse homolog, a 2.7 kilobase edit, has been reported (163). The key factor in this success was avoiding homology in the region to be replaced. It has been proposed that homologous regions between the desired point mutations and the cut site can act as homology arms during DSB repair (162). This results in incomplete HDR where the DSB repair is only partially guided by the donor template and the point mutation is not introduced into the target genome. Such a mechanism can account for the decreased frequency of incorporation of point mutations further from the cut site.

Since the two exons that code for the Spt5 CTR are separated by an 85 base pair intron, we suspected that the frequency of incorporation of the point mutants after the intron would be decreased. We were also concerned about the potential lethality of completely abolishing Spt5 CTR phosphorylation. For this reason, we designed an alternate repair template, the m4 intronless Spt5 repair template (m4). This template lacks the intronic region and only mutates the first four threonines (T768, T775, T784, T791), but is otherwise identical to m7.

Figure 3.2 m7 Spt5 repair template design. A) Repair template sequence. B) Amino acid changes between wild-type and mutant.

3.3.3 Creation of heterozygote HEK 293F cells

We transfected HEK 293F cells with our CRISPR/Cas9 components and the m7 repair template and screened 22 ouabain-resistant clones. Screening was done by PCR amplification of the region of interest with a primer that bound outside the region overlapping the homology arms of the repair template, so that only repair template sequences incorporated in the correct genomic locus would be amplified. PCR products were then assayed for the presence of a SacI restriction site or sent for Sanger sequencing. Out of the 22 clones, four incorporated all point mutations. However, the levels of these point mutations were approximately half the levels of the wild-type allele (Figure 3.3), suggesting that these cells might be triploid at the locus on interest and that only one out of three copies of the gene was successfully edited.

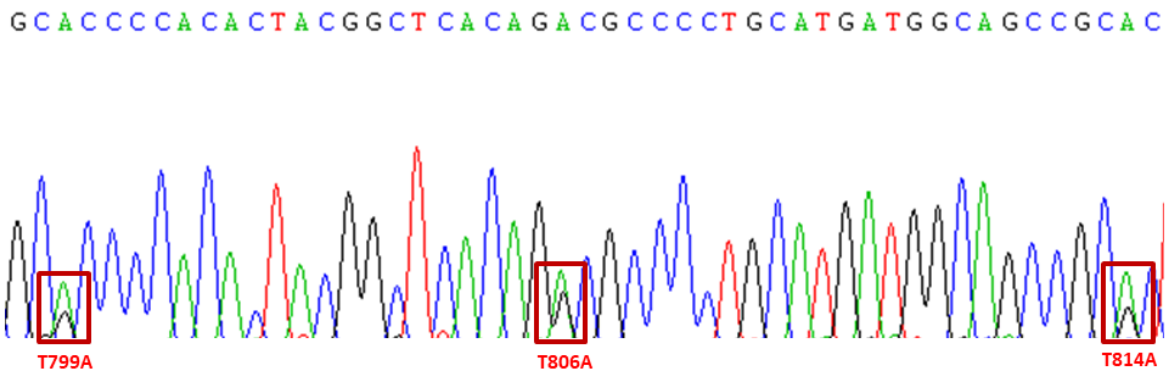

Figure 3.3 Sanger sequencing results of Spt5 CTR in HEK 293F cells treated with CRISPR/Cas9 components and the m7 Spt5 repair template.

3.3.4 Non-specific repair template insertions are widespread

Reasoning that the observed aneuploidy might be characteristic of the cell line, we introduced our CRISPR components and either the m7 or the m4 repair template in the HEK 293 parental cell line. We performed PCR on polyclonal cell lysates before and after ouabain selection with two primer pairs: 3F3R, which are both located within the homology arms, and 3F4R, where one of the primers is outside the homology arms (Figure 3.4B). While the 3F3R pair is expected to amplify both wild type and repair template DNA regardless of location, the 3F4R pair would only amplify sequences at the correct genomic locus. Treating the PCR products with SacI revealed that the vast majority of the 3F3R amplicons contained a SacI restriction site, while SacI cleavage was undetectable in the 3F4R amplicons (Figure 3.5). These results suggest that the repair template is incorporated into the genomic DNA, but not at the locus of interest.

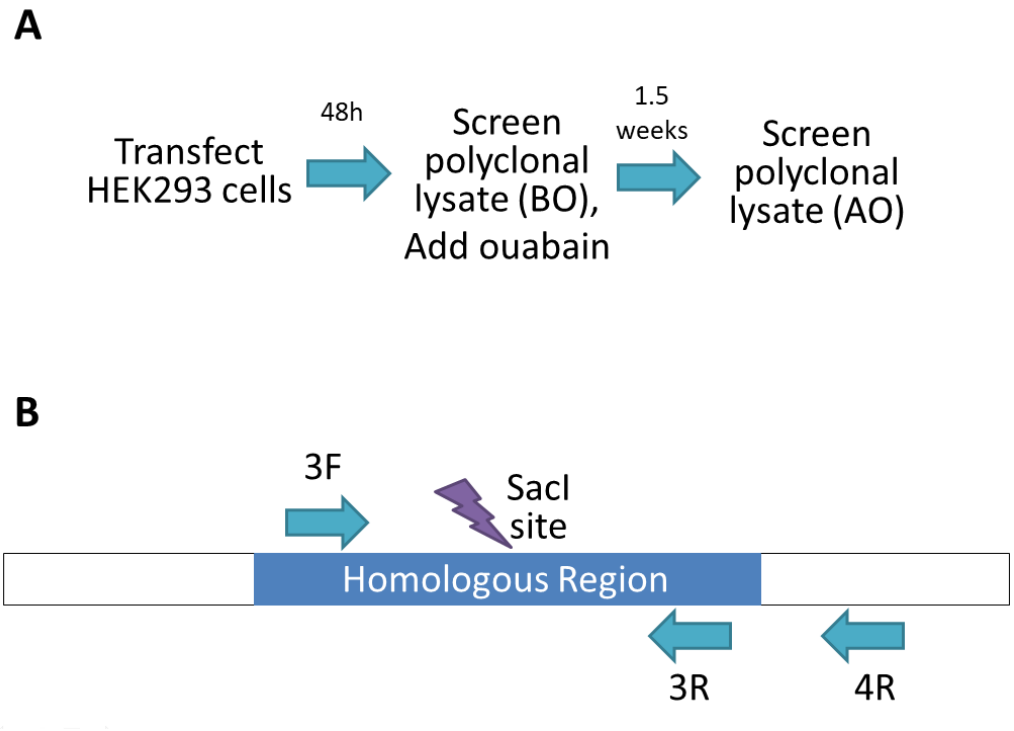

Figure 3.4 Screening strategy. A) Experimental workflow. B) Locations of PCR primer binding and SacI recognition site on the Spt5 CTR.

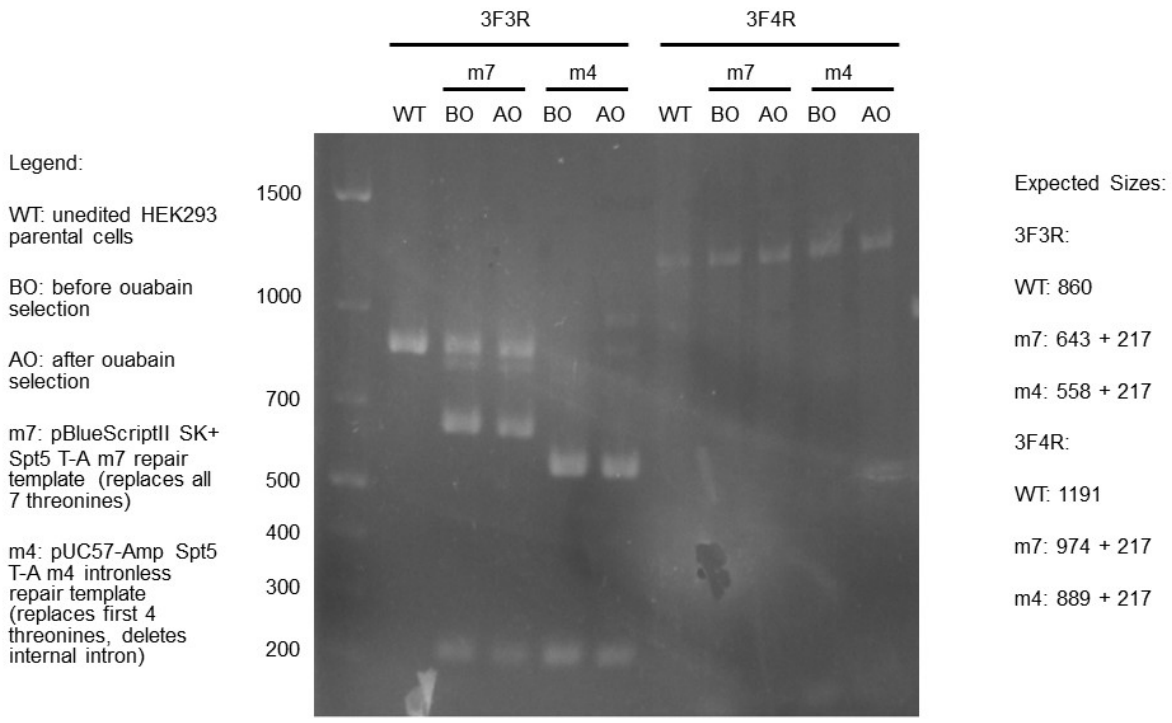

Figure 3.5 Non-specific repair template insertions are widespread. Agarose gel electrophoresis of SacI-treated PCR products from polyclonal lysates of cells treated with CRISPR/Cas9 components. Cells were harvested before (BO) or after (AO) ouabain selection.

3.4 Discussion

The objective of this project was to establish a mammalian cell line deficient in Spt5 phosphorylation that could be used as a tool for investigating the function of this post-translational modification. We designed and validated sgRNAs against the Spt5 CTR and designed dsDNA repair templates to incorporate multiple point mutations. We were successful in creating HEK 293F cells heterozygous for the desired mutations, although they may still possess two wild type alleles in addition to the mutated allele.

Without karyotypic data, we cannot disregard the possibility that two functioning copies of wild-type Spt5 are required for efficient cell proliferation, and that our gene editing attempts have selected for cells that have amplified the portion of chromosome 19 that contains the Spt5

gene. However, the observed amplification of Spt5 may simply be a feature of the cells used. Parental HEK 293 cells have been described as hypotriploid (164), and previous studies have reported amplification of chromosome 19, which also contains the adenoviral genome fragment originally used to immortalize the cells (165). HEK 293F cells are derived from the parental cell line and adapted to growth in suspension, but they have likely retained the chromosome 19 amplification, among other chromosomal abnormalities.

Based on previous studies, our replacement of the Spt5 CTR phosphorylation sites with alanines would not be expected to be lethal. In HeLa cells expressing a mutant Spt5 lacking the CTR under the control of a tetracycline-repressible promoter, complete knockout of endogenous Spt5 combined with tetracycline treatment leads to apoptosis, but this effect can be avoided in the absence of tetracycline. The CTR has thus been shown to be non-essential to cell survival, although proliferation with the truncated Spt5 was slower than with wild-type (68). Functions associated with Spt5 CTR phosphorylation include promoting elongation and recruiting Rtf1 to promote H2Bub1. If the CTR is truly dispensable to cell survival, then efficient elongation is possible without Spt5 phosphorylation. Furthermore, Cdk9 is essential in all organisms studied, so many of its functions must be mediated by pathways that do not involve the Spt5 CTR.

Unexpectedly, our repair template was observed to insert non-specifically in the genome of HEK 293 cells. Our dsDNA template is transfected in a circular plasmid but is designed to be linearized by the sgRNA-bound Cas9. The use of linear dsDNA is thought to increase CRISPR-mediated knockin efficiency (156), but it may also increase random integration. Our results complement recent demonstrations that linear dsDNA can incorporate into HEK 293 genomes even in the absence of CRISPR/Cas9 components, and that this non-specific incorporation rate increases with the addition of Cas9 and a sgRNA (166). These undesired integrations may be mitigated by using single-stranded DNA repair templates.

Alternative strategies for establishing a mammalian cell line deficient in Spt5 phosphorylation might include inserting the phosphorylation-deficient Spt5 mutant gene at an ectopic genomic locus, similarly to previous work (27). Phenotypic analyses could be conducted at this stage, since the mutant is expected to retain the ability to associate with Pol II and establish promoter-proximal pausing and should thus be dominant negative. CRISPR/Cas9 may also be used to knock out the endogenous gene. The disadvantage of such a strategy relative to

CRISPR/Cas9 knockin at the endogenous locus is the loss of endogenous control of gene expression.

While HEK 293 cells are commonly used in CRISPR/Cas9 gene editing studies because of their relative tractability, other cell types may be more appropriate for the study of Spt5 phosphorylation. For instance, THP-1 cells are known to be sensitive to P-TEFb inhibition (104), suggesting an increased reliance on pathways downstream of P-TEFb. We initially used THP-1 for the creation of the repair template with the intention of performing the gene replacement in THP-1 cells. Furthermore, despite several other reported chromosomal abnormalities, THP-1 cells seem to carry two wild-type copies of chromosome 19 (132).

In summary, we have successfully created HEK 293F cells heterozygous for the desired mutations with CRISPR/Cas9. We have also uncovered a major pitfall of the usage of linearized dsDNA repair templates, namely their tendency for non-specific genomic insertion, and suggested alternative strategies for establishing a mammalian cell line deficient in Spt5 phosphorylation. Such a cell line would be highly beneficial for studies of pathways downstream of Spt5 phosphorylation.

Chapter 4: Concluding Remarks

The focus of this thesis was the investigation of P-TEFb-dependent pathways in mixed lineage leukemia. We have shown that the MLL-rearranged THP-1 leukemia line is more sensitive to iCdk9, an inhibitor of P-TEFb activity, and to JQ1, an inhibitor of Brd4 recruitment, than a non-MLL-rearranged leukemia line, K562. With iCdk9, this increased sensitivity is also seen when assessing biomarkers of P-TEFb activity, leading us to the novel hypothesis that MLL fusion proteins, despite recruiting the SEC to key cancer genes, might generally impair P-TEFb activity. Future work will test whether the presence of a MLL fusion protein is sufficient to confer sensitivity to P-TEFb inhibitors. We will also explore the functional relationship between the PAFc and P-TEFb in both cell lines, to determine if differences in PAFc recruitment could account for the observed differential sensitivity to iCdk9.

PAFc recruitment has previously been reported to be downstream of Spt5 phosphorylation. In order to better understand the roles of P-TEFb-mediated Spt5 phosphorylation, we aimed to establish a human cell line with mutated phosphoacceptor sites in the Spt5 CTR using CRISPR/Cas9. We created a cell line heterozygous for the desired mutations but also observed non-specific genomic insertion of our linearized dsDNA repair templates. Our results suggest that alternative repair template formats may be preferable for gene replacement.

Understanding the regulation of P-TEFb and P-TEFb-dependent pathways in cancer progression could yield valuable insight into cancer biology and treatment.

References

1. Pearson H. What is a gene? Nature. 2006;441:398.
2. International Human Genome Sequencing C. Initial sequencing and analysis of the human genome. Nature. 2001;409:860.
3. Melamed P, Yosefzon Y, Rudnizky S, Pnueli L. Transcriptional enhancers: Transcription, function and flexibility. Transcription. 2016;7(1):26-31.
4. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature. 1997;389:251.
5. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Research. 2011;21(3):381-95.
6. Fuchs SM, Laribee RN, Strahl BD. Protein Modifications in Transcription Elongation. Biochimica et biophysica acta. 2009;1789(1):26-36.
7. Wolberger C. Combinatorial transcription factors. Current Opinion in Genetics & Development. 1998;8(5):552-9.
8. Jenuwein T, Allis CD. Translating the Histone Code. Science (New York, NY). 2001;293(5532):1074.
9. Fujisawa T, Filippakopoulos P. Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. Nat Rev Mol Cell Biol. 2017;18(4):246-62.
10. Saint-Andre V, Federation AJ, Lin CY, Abraham BJ, Reddy J, Lee TI, et al. Models of human core transcriptional regulatory circuitries. Genome Res. 2016;26(3):385-96.
11. Svejstrup JQ. The RNA polymerase II transcription cycle: cycling through chromatin. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression. 2004;1677(1):64-73.
12. Krishnamurthy S, Hampsey M. Eukaryotic transcription initiation. Current Biology. 2009;19(4):R153-R6.
13. Luse DS. Promoter clearance by RNA polymerase II. Biochimica et biophysica acta. 2013;1829(1):63-8.
14. Proudfoot NJ. Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. Science (New York, NY). 2016;352(6291).
15. Mayer A, Landry HM, Churchman LS. Pause & go: from the discovery of RNA polymerase pausing to its functional implications. Current opinion in cell biology. 2017;46:72-80.
16. Guo J, Price DH. RNA Polymerase II Transcription Elongation Control. Chemical Reviews. 2013;113(11):8583-603.
17. Gariglio P, Bellard M, Chambon P. Clustering of RNA polymerase B molecules in the 5' moiety of the adult beta-globin gene of hen erythrocytes. Nucleic Acids Research. 1981;9(11):2589-98.
18. Rougvie AE, Lis JT. The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell. 1988;54(6):795-804.
19. Krumm A, Meulia T, Brunvand M, Groudine M. The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. Genes & development. 1992;6(11):2201-13.
20. Pugh BF, Venters BJ. Genomic Organization of Human Transcription Initiation Complexes. PloS one. 2016;11(2):e0149339.

21. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science (New York, NY). 2008;322(5909):1845-8.
22. Mayer A, di Iulio J, Maleri S, Eser U, Vierstra J, Reynolds A, et al. Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. Cell. 2015;161(3):541-54.
23. Missra A, Gilmour DS. Interactions between DSIF (DRB sensitivity inducing factor), NELF (negative elongation factor), and the Drosophila RNA polymerase II transcription elongation complex. Proceedings of the National Academy of Sciences. 2010;107(25):11301-6.
24. Paparidis NFdS, Durvale MC, Canduri F. The emerging picture of CDK9/P-TEFb: more than 20 years of advances since PITALRE. Molecular BioSystems. 2017;13(2):246-76.
25. Lu X, Zhu X, Li Y, Liu M, Yu B, Wang Y, et al. Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. Nucleic Acids Res. 2016;44(14):6853-67.
26. Fujinaga K, Irwin D, Huang Y, Taube R, Kurosu T, Peterlin BM. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Molecular and cellular biology. 2004;24(2):787-95.
27. Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H. P-TEFb-Mediated Phosphorylation of hSpt5 C-Terminal Repeats Is Critical for Processive Transcription Elongation. Molecular cell. 2006;21(2):227-37.
28. Krebs AR, Imanci D, Hoerner L, Gaidatzis D, Burger L, Schübeler D. Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters. Molecular cell.67(3):411-22.e4.
29. Shao W, Zeitlinger J. Paused RNA polymerase II inhibits new transcriptional initiation. Nat Genet. 2017;advance online publication.
30. Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, Li L, et al. Pausing of RNA Polymerase II Disrupts DNA-Specified Nucleosome Organization to Enable Precise Gene Regulation. Cell. 2010;143(4):540-51.
31. Liu L, Xu Y, He M, Zhang M, Cui F, Lu L, et al. Transcriptional Pause Release Is a Rate-Limiting Step for Somatic Cell Reprogramming. Cell Stem Cell.15(5):574-88.
32. Zhou Q, Yik JHN. The Yin and Yang of P-TEFb Regulation: Implications for Human Immunodeficiency Virus Gene Expression and Global Control of Cell Growth and Differentiation. Microbiology and Molecular Biology Reviews. 2006;70(3):646-59.
33. Fowler T, Sen R, Roy Ananda L. Regulation of Primary Response Genes. Molecular cell. 2011;44(3):348-60.
34. Zaborowska J, Isa NF, Murphy S. P-TEFb goes viral. BioEssays. 2016;38:S75-S85.
35. Wei P, Garber ME, Fang SM, Fischer WH, Jones KA. 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. 1998;92(4):451-62.
36. Barboric M, Nissen RM, Kanazawa S, Jabrane-Ferrat N, Peterlin BM. NF-κB Binds P-TEFb to Stimulate Transcriptional Elongation by RNA Polymerase II. Molecular cell. 2001;8(2):327-37.

37. Conaway JW, Takahashi H, Parmely TJ, Tomomori-Sato C, Sato S, Conaway RC. The Mediator of RNA Polymerase II Transcription: Links to Transcription Elongation and Leukemogenesis. The FASEB Journal. 2012;26(1_supplement):227.2-.2.
38. Yang Z, Yik JHN, Chen R, He N, Jang MK, Ozato K, et al. Recruitment of P-TEFb for Stimulation of Transcriptional Elongation by the Bromodomain Protein Brd4. Molecular cell. 2005;19(4):535-45.
39. Liu W, Ma Q, Wong K, Li W, Ohgi K, Zhang J, et al. Brd4 and JMJD6-associated Anti-pause Enhancers in Regulation of Transcriptional Pause Release. Cell. 2013;155(7):1581-95.
40. Winter GE, Mayer A, Buckley DL, Erb MA, Roderick JE, Vittori S, et al. BET Bromodomain Proteins Function as Master Transcription Elongation Factors Independent of CDK9 Recruitment. Molecular cell. 2017;67(1):5-18.e9.
41. C. Quaresma AJ, Bugai A, Barboric M. Cracking the control of RNA polymerase II elongation by 7SK snRNP and P-TEFb. Nucleic Acids Research. 2016;44(16):7527-39.
42. He N, Jahchan NS, Hong E, Li Q, Bayfield MA, Maraia RJ, et al. A La-Related Protein Modulates 7SK snRNP Integrity to Suppress P-TEFb-Dependent Transcriptional Elongation and Tumorigenesis. Molecular cell. 2008;29(5):588-99.
43. Yang Z, Zhu Q, Luo K, Zhou Q. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. Nature. 2001;414:317.
44. Biglione S, Byers SA, Price JP, Nguyen VT, Bensaude O, Price DH, et al. Inhibition of HIV-1 replication by P-TEFb inhibitors DRB, seliciclib and flavopiridol correlates with release of free P-TEFb from the large, inactive form of the complex. Retrovirology. 2007;4:47.
45. Bartholomeeusen K, Xiang Y, Fujinaga K, Peterlin BM. Bromodomain and extra-terminal (BET) bromodomain inhibition activate transcription via transient release of positive transcription elongation factor b (P-TEFb) from 7SK small nuclear ribonucleoprotein. The Journal of biological chemistry. 2012;287(43):36609-16.
46. Pavletich NP. Mechanisms of cyclin-dependent kinase regulation: structures of cdks, their cyclin activators, and cip and INK4 inhibitors1,211998 Awardee, Walter J. Johnson Prize for the Encouragement of Research in the Life Sciences2Edited by P. E. Wright. Journal of Molecular Biology. 1999;287(5):821-8.
47. Chen R, Yang Z, Zhou Q. Phosphorylated positive transcription elongation factor b (P-TEFb) is tagged for inhibition through association with 7SK snRNA. The Journal of biological chemistry. 2004;279(6):4153-60.
48. Chen R, Liu M, Li H, Xue Y, Ramey WN, He N, et al. PP2B and PP1α cooperatively disrupt 7SK snRNP to release P-TEFb for transcription in response to Ca(2+) signaling. Genes & development. 2008;22(10):1356-68.
49. Peng J, Zhu Y, Milton JT, Price DH. Identification of multiple cyclin subunits of human P-TEFb. Genes & development. 1998;12(5):755-62.
50. Fu TJ, Peng J, Lee G, Price DH, Flores O. Cyclin K functions as a CDK9 regulatory subunit and participates in RNA polymerase II transcription. The Journal of biological chemistry. 1999;274(49):34527-30.
51. Blazek D, Kohoutek J, Bartholomeeusen K, Johansen E, Hulinkova P, Luo Z, et al. The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. Genes & development. 2011;25(20):2158-72.

52. Lu H, Yu D, Hansen AS, Ganguly S, Liu R, Heckert A, et al. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. Nature. 2018.
53. Corden JL. Tails of RNA polymerase II. Trends in Biochemical Sciences. 1990;15(10):383-7.
54. Hirose Y, Manley JL. RNA polymerase II and the integration of nuclear events. Genes & development. 2000;14(12):1415-29.
55. Nilson KA, Guo J, Turek ME, Brogie JE, Delaney E, Luse DS, et al. THZ1 Reveals Roles for Cdk7 in Co-transcriptional Capping and Pausing. Molecular cell. 2015;59(4):576-87.
56. Lu H, Xue Y, Yu GK, Arias C, Lin J, Fong S, et al. Compensatory induction of MYC expression by sustained CDK9 inhibition via a BRD4-dependent mechanism. eLife. 2015;4:e06535.
57. Ni Z, Schwartz BE, Werner J, Suarez J-R, Lis JT. Coordination of Transcription, RNA Processing, and Surveillance by P-TEFb Kinase on Heat Shock Genes. Molecular cell. 2004;13(1):55-65.
58. Czudnochowski N, Bosken CA, Geyer M. Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. Nat Commun. 2012;3:842.
59. Bartkowiak B, Liu P, Phatnani HP, Fuda NJ, Cooper JJ, Price DH, et al. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. Genes & development. 2010;24(20):2303-16.
60. Mayer A, Lidschreiber M, Siebert M, Leike K, Söding J, Cramer P. Uniform transitions of the general RNA polymerase II transcription complex. Nature Structural &Amp; Molecular Biology. 2010;17:1272.
61. Bataille AR, Jeronimo C, Jacques PE, Laramee L, Fortin ME, Forest A, et al. A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. Molecular cell. 2012;45(2):158-70.
62. Bartkowiak B, Greenleaf AL. Phosphorylation of RNAPII. Transcription. 2011;2(3):115-9.
63. Narita T, Yamaguchi Y, Yano K, Sugimoto S, Chanarat S, Wada T, et al. Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. Molecular and cellular biology. 2003;23(6):1863-73.
64. Yamaguchi Y, Inukai N, Narita T, Wada T, Handa H. Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. Molecular and cellular biology. 2002;22(9):2918-27.
65. Sun J, Li R. Human Negative Elongation Factor Activates Transcription and Regulates Alternative Transcription Initiation. Journal of Biological Chemistry. 2010;285(9):6443-52.
66. Qiu Y, Gilmour DS. Identification of Regions in the Spt5 Subunit of DSIF That Are Involved in Promoter Proximal Pausing. The Journal of biological chemistry. 2017.
67. Chen H, Contreras X, Yamaguchi Y, Handa H, Peterlin BM, Guo S. Repression of RNA Polymerase II Elongation In Vivo Is Critically Dependent on the C-Terminus of Spt5. PloS one. 2009;4(9):e6918.
68. Komori T, Inukai N, Yamada T, Yamaguchi Y, Handa H. Role of human transcription elongation factor DSIF in the suppression of senescence and apoptosis. Genes to Cells. 2009;14(3):343-54.

69. Parua PK, Booth GT, Sansó M, Benjamin B, Tanny JC, Lis JT, et al. A Cdk9–PP1 switch regulates the elongation–termination transition of RNA polymerase II. Nature. 2018.
70. Mbogning J, Nagy S, Pagé V, Schwer B, Shuman S, Fisher RP, et al. The PAF Complex and Prf1/Rtf1 Delineate Distinct Cdk9-Dependent Pathways Regulating Transcription Elongation in Fission Yeast. PLoS genetics. 2013;9(12):e1004029.
71. Tomson BN, Arndt KM. The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms. 2013;1829(1):116-26.
72. Chen FX, Woodfin AR, Gardini A, Rickels RA, Marshall SA, Smith ER, et al. PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. Cell. 2015;162(5):1003-15.
73. Chen FX, Xie P, Collings CK, Cao K, Aoi Y, Marshall SA, et al. PAF1 regulation of promoter-proximal pause release via enhancer activation. Science (New York, NY). 2017;357(6357):1294-8.
74. Yu M, Yang W, Ni T, Tang Z, Nakadai T, Zhu J, et al. RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. Science (New York, NY). 2015;350(6266):1383-6.
75. He N, Chan CK, Sobhian B, Chou S, Xue Y, Liu M, et al. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(36):E636-45.
76. Wood A, Schneider J, Dover J, Johnston M, Shilatifard A. The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. The Journal of biological chemistry. 2003;278(37):34739-42.
77. Xiao T, Kao C-F, Krogan NJ, Sun Z-W, Greenblatt JF, Osley MA, et al. Histone H2B Ubiquitylation Is Associated with Elongating RNA Polymerase II. Molecular and cellular biology. 2005;25(2):637-51.
78. Sun Z-W, Allis CD. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature. 2002;418:104.
79. Nakanishi S, Lee JS, Gardner KE, Gardner JM, Takahashi YH, Chandrasekharan MB, et al. Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. J Cell Biol. 2009;186(3):371-7.
80. Sanso M, Lee KM, Viladevall L, Jacques PE, Page V, Nagy S, et al. A positive feedback loop links opposing functions of P-TEFb/Cdk9 and histone H2B ubiquitylation to regulate transcript elongation in fission yeast. PLoS genetics. 2012;8(8):e1002822.
81. Johnsen SA. CDK9 and H2B Monoubiquitination: A Well-Choreographed Dance. PLoS genetics. 2012;8(8):e1002860.
82. Zhang F, Yu X. WAC, a Functional Partner of RNF20/40, Regulates Histone H2B Ubiquitination and Gene Transcription. Molecular cell. 2011;41(4):384-97.
83. Shchebet A, Karpiuk O, Kremmer E, Eick D, Johnsen SA. Phosphorylation by cyclin-dependent kinase-9 controls ubiquitin-conjugating enzyme-2A function. Cell cycle (Georgetown, Tex). 2012;11(11):2122-7.
84. Wu L, Li L, Zhou B, Qin Z, Dou Y. H2B ubiquitylation promotes RNA Pol II processivity via PAF1 and pTEFb. Molecular cell. 2014;54(6):920-31.

85. Glover-Cutter K, Kim S, Espinosa J, Bentley DL. RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nature structural & molecular biology. 2008;15(1):71-8.
86. Mandal SS, Chu C, Wada T, Handa H, Shatkin AJ, Reinberg D. Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(20):7572-7.
87. Pei Y, Shuman S. Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. The Journal of biological chemistry. 2002;277(22):19639-48.
88. Hsin JP, Manley JL. The RNA polymerase II CTD coordinates transcription and RNA processing. Genes & development. 2012;26(19):2119-37.
89. Hsin JP, Xiang K, Manley JL. Function and control of RNA polymerase II C-terminal domain phosphorylation in vertebrate transcription and RNA processing. Molecular and cellular biology. 2014;34(13):2488-98.
90. Zhu B, Mandal SS, Pham AD, Zheng Y, Erdjument-Bromage H, Batra SK, et al. The human PAF complex coordinates transcription with events downstream of RNA synthesis. Genes & development. 2005;19(14):1668-73.
91. Hanahan D, Weinberg Robert A. Hallmarks of Cancer: The Next Generation. Cell.144(5):646-74.
92. Mueller D, Garcia-Cuellar MP, Bach C, Buhl S, Maethner E, Slany RK. Misguided transcriptional elongation causes mixed lineage leukemia. PLoS biology. 2009;7(11):e1000249.
93. Slany RK. The molecular biology of mixed lineage leukemia. Haematologica. 2009;94(7):984-93.
94. Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, et al. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell. 2012;151(1):68-79.
95. Koh CM, Sabò A, Guccione E. Targeting MYC in cancer therapy: RNA processing offers new opportunities. Bioessays. 2016;38(3):266-75.
96. Bradner JE, Hnisz D, Young RA. Transcriptional Addiction in Cancer. Cell. 2017;168(4):629-43.
97. Tsimberidou A-M. TARGETED THERAPY IN CANCER. Cancer chemotherapy and pharmacology. 2015;76(6):1113-32.
98. Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, et al. c-Myc regulates transcriptional pause release. Cell. 2010;141(3):432-45.
99. Bhagwat AS, Vakoc CR. Targeting Transcription Factors in Cancer. Trends in Cancer. 2015;1(1):53-65.
100. Sullivan KD, Galbraith MD, Andrysik Z, Espinosa JM. Mechanisms of transcriptional regulation by p53. Cell Death Differ. 2018;25(1):133-43.
101. Yu X, Vazquez A, Levine Arnold J, Carpizo Darren R. Allele-Specific p53 Mutant Reactivation. Cancer cell. 2012;21(5):614-25.
102. Villicaña C, Cruz G, Zurita M. The basal transcription machinery as a target for cancer therapy. Cancer Cell International. 2014;14:18-.
103. Stellrecht CM, Chen LS. Transcription Inhibition as a Therapeutic Target for Cancer. Cancers. 2011;3(4):4170-90.

104. Baker A, Gregory GP, Verbrugge I, Kats L, Hilton JJ, Vidacs E, et al. The CDK9 Inhibitor Dinaciclib Exerts Potent Apoptotic and Antitumor Effects in Preclinical Models of MLL-Rearranged Acute Myeloid Leukemia. Cancer research. 2016;76(5):1158-69.
105. Boffo S, Damato A, Alfano L, Giordano A. CDK9 inhibitors in acute myeloid leukemia. Journal of Experimental & Clinical Cancer Research. 2018;37(1):36.
106. Chen R, Keating MJ, Gandhi V, Plunkett W. Transcription inhibition by flavopiridol: mechanism of chronic lymphocytic leukemia cell death. Blood. 2005;106(7):2513-9.
107. Zeidner JF, Karp JE. Clinical activity of alvocidib (flavopiridol) in acute myeloid leukemia. Leukemia Research. 2015;39(12):1312-8.
108. Orphan Drug Status for Alvocidib. Oncology Times. 2014;36(10):91.
109. Andrieu G, Belkina AC, Denis GV. Clinical trials for BET inhibitors run ahead of the science. Drug Discovery Today: Technologies. 2016;19:45-50.
110. Zuber J, Shi J, Wang E, Rappaport AR, Herrmann H, Sison EA, et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature. 2011;478(7370):524-8.
111. Sun C, Yin J, Fang Y, Chen J, Jeong KJ, Chen X, et al. BRD4 Inhibition Is Synthetic Lethal with PARP Inhibitors through the Induction of Homologous Recombination Deficiency. Cancer cell. 2018;33(3):401-16.e8.
112. Gravells P, Grant E, Smith KM, James DI, Bryant HE. Specific killing of DNA damage-response deficient cells with inhibitors of poly(ADP-ribose) glycohydrolase. DNA Repair. 2017;52:81-91.
113. Hargreaves DC, Horng T, Medzhitov R. Control of Inducible Gene Expression by Signal-Dependent Transcriptional Elongation. Cell. 2009;138(1):129-45.
114. Fowler T, Ghatak P, Price DH, Conaway R, Conaway J, Chiang CM, et al. Regulation of MYC expression and differential JQ1 sensitivity in cancer cells. PloS one. 2014;9(1):e87003.
115. Mertz JA, Conery AR, Bryant BM, Sandy P, Balasubramanian S, Mele DA, et al. Targeting MYC dependence in cancer by inhibiting BET bromodomains. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(40):16669-74.
116. Rathert P, Roth M, Neumann T, Muerdter F, Roe J-S, Muhar M, et al. Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. Nature. 2015;525(7570):543-7.
117. Chen L, Sun Y, Wang J, Jiang H, Muntean AG. Differential regulation of the c-Myc/Lin28 axis discriminates subclasses of rearranged MLL leukemia. Oncotarget. 2016;7(18):25208-23.
118. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, et al. MLL targets SET domain methyltransferase activity to Hox gene promoters. Molecular cell. 2002;10(5):1107-17.
119. Yu BD, Hess JL, Horning SE, Brown GA, Korsmeyer SJ. Altered Hox expression and segmental identity in Mll-mutant mice. Nature. 1995;378(6556):505-8.
120. Slany RK. When epigenetics kills: MLL fusion proteins in leukemia. Hematological Oncology. 2005;23(1):1-9.
121. Meyer C, Kowarz E, Hofmann J, Renneville A, Zuna J, Trka J, et al. New insights to the MLL recombinome of acute leukemias. Leukemia. 2009;23:1490.

122. Yokoyama A. Transcriptional activation by MLL fusion proteins in leukemogenesis. Experimental hematology. 2017;46:21-30.
123. Benedikt A, Baltruschat S, Scholz B, Bursen A, Arrey TN, Meyer B, et al. The leukemogenic AF4-MLL fusion protein causes P-TEFb kinase activation and altered epigenetic signatures. Leukemia. 2011;25(1):135-44.
124. Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, Florens L, et al. AFF4, a Component of the ELL/P-TEFb Elongation Complex and a Shared Subunit of MLL Chimeras, Can Link Transcription Elongation to Leukemia. Molecular cell. 2010;37(3):429-37.
125. Erb MA, Scott TG, Li BE, Xie H, Paulk J, Seo H-S, et al. Transcription control by the ENL YEATS domain in acute leukaemia. Nature. 2017;543:270.
126. Muntean AG, Tan J, Sitwala K, Huang Y, Bronstein J, Connelly JA, et al. The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. Cancer cell. 2010;17(6):609-21.
127. Li Y, Wen H, Xi Y, Tanaka K, Wang H, Peng D, et al. AF9 YEATS Domain Links Histone Acetylation to DOT1L-Mediated H3K79 Methylation. Cell. 2014;159(3):558-71.
128. Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, et al. hDOT1L Links Histone Methylation to Leukemogenesis. Cell. 2005;121(2):167-78.
129. Wang QF, Wu G, Mi S, He F, Wu J, Dong J, et al. MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. Blood. 2011;117(25):6895-905.
130. Chen CW, Armstrong SA. Targeting DOT1L and HOX gene expression in MLL-rearranged leukemia and beyond. Experimental hematology. 2015;43(8):673-84.
131. Gilan O, Lam EY, Becher I, Lugo D, Cannizzaro E, Joberty G, et al. Functional interdependence of BRD4 and DOT1L in MLL leukemia. Nature structural & molecular biology. 2016;23(7):673-81.
132. Odero MD, Zeleznik-Le NJ, Chinwalla V, Rowley JD. Cytogenetic and molecular analysis of the acute monocytic leukemia cell line THP-1 with an MLL-AF9 translocation. Genes, Chromosomes and Cancer. 2000;29(4):333-8.
133. Grosveld G, Verwoerd T, van Agthoven T, de Klein A, Ramachandran KL, Heisterkamp N, et al. The chronic myelocytic cell line K562 contains a breakpoint in bcr and produces a chimeric bcr/c-abl transcript. Molecular and cellular biology. 1986;6(2):607-16.
134. Yang L-h, Wang S-l, Tang L-l, Liu B, Ye W-l, Wang L-l, et al. Universal Stem-Loop Primer Method for Screening and Quantification of MicroRNA. PloS one. 2014;9(12):e115293.
135. Kwiatkowski N, Zhang T, Rahl PB, Abraham BJ, Reddy J, Ficarro SB, et al. Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. Nature. 2014;511(7511):616-20.
136. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, et al. Selective inhibition of BET bromodomains. Nature. 2010;468(7327):1067-73.
137. Prange KHM, Mandoli A, Kuznetsova T, Wang SY, Sotoca AM, Marneth AE, et al. MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. Oncogene. 2017;36(23):3346-56.

138. Yao X, Tang Z, Fu X, Yin J, Liang Y, Li C, et al. The Mediator subunit MED23 couples H2B mono-ubiquitination to transcriptional control and cell fate determination. The EMBO Journal. 2015;34(23):2885-902.
139. Liu Y, Warfield L, Zhang C, Luo J, Allen J, Lang WH, et al. Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. Molecular and cellular biology. 2009;29(17):4852-63.
140. Fleischmann KK, Pagel P, Schmid I, Roscher AA. RNAi-mediated silencing of MLL-AF9 reveals leukemia-associated downstream targets and processes. Molecular Cancer. 2014;13:27-.
141. Kawagoe H, Kawagoe R, Sano K. Targeted down-regulation of MLL-AF9 with antisense oligodeoxyribonucleotide reduces the expression of the HOXA7 and -A10 genes and induces apoptosis in a human leukemia cell line, THP-12001. 1743-9 p.
142. Pession A, Martino V, Tonelli R, Beltramini C, Locatelli F, Biserni G, et al. MLL-AF9 oncogene expression affects cell growth but not terminal differentiation and is downregulated during monocyte-macrophage maturation in AML-M5 THP-1 cells. Oncogene. 2003;22(54):8671-6.
143. Skucha A, Ebner J, Schmollerl J, Roth M, Eder T, Cesar-Razquin A, et al. MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity. Nat Commun. 2018;9(1):1983.
144. Muntean AG, Giannola D, Udager AM, Hess JL. The PHD fingers of MLL block MLL fusion protein–mediated transformation. Blood. 2008;112(12):4690.
145. Xia Z-B, Popovic R, Chen J, Theisler C, Stuart T, Santillan DA, et al. The MLL fusion gene, MLL-AF4, regulates cyclin-dependent kinase inhibitor CDKN1B (p27(kip1)) expression. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(39):14028-33.
146. Schneider S, Pei Y, Shuman S, Schwer B. Separable Functions of the Fission Yeast Spt5 Carboxyl-Terminal Domain (CTD) in Capping Enzyme Binding and Transcription Elongation Overlap with Those of the RNA Polymerase II CTD. Molecular and cellular biology. 2010;30(10):2353-64.
147. Bolukbasi MF, Gupta A, Wolfe SA. Creating and evaluating accurate CRISPR-Cas9 scalpels for genomic surgery. Nature methods. 2016;13(1):41-50.
148. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nature biotechnology. 2014;32(3):279-84.
149. Graham DB, Root DE. Resources for the design of CRISPR gene editing experiments. Genome Biology. 2015;16:260.
150. Pawelczak KS, Gavande NS, VanderVere-Carozza PS, Turchi JJ. Modulating DNA Repair Pathways to Improve Precision Genome Engineering. ACS Chemical Biology. 2018;13(2):389-96.
151. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotech. 2015;33(5):543-8.
152. Merkle FT, Neuhausser WM, Santos D, Valen E, Gagnon JA, Maas K, et al. Efficient CRISPR-Cas9-mediated generation of knock-in human pluripotent stem cells lacking undesired mutations at the targeted locus. Cell reports. 2015;11(6):875-83.

153. Agudelo D, Duringer A, Bozoyan L, Huard CC, Carter S, Loehr J, et al. Marker-free coselection for CRISPR-driven genome editing in human cells. Nat Meth. 2017;advance online publication.
154. Yang L, Guell M, Byrne S, Yang JL, De Los Angeles A, Mali P, et al. Optimization of scarless human stem cell genome editing. Nucleic Acids Research. 2013;41(19):9049-61.
155. Song F, Stieger K. Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks. Molecular Therapy - Nucleic Acids. 2017;7:53-60.
156. Zhang J-P, Li X-L, Li G-H, Chen W, Arakaki C, Botimer GD, et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. Genome Biology. 2017;18(1):35.
157. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nature Biotechnology. 2013;31:827.
158. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Science (New York, NY). 2013;339.
159. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. Science (New York, NY). 2016;351(6268):84-8.
160. Inui M, Miyado M, Igarashi M, Tamano M, Kubo A, Yamashita S, et al. Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. 2014;4:5396.
161. Ménoret S, De Cian A, Tesson L, Remy S, Usal C, Boulé J-B, et al. Homology-directed repair in rodent zygotes using Cas9 and TALEN engineered proteins. Scientific Reports. 2015;5:14410.
162. Wang K, Tang X, Liu Y, Xie Z, Zou X, Li M, et al. Efficient Generation of Orthologous Point Mutations in Pigs via CRISPR-assisted ssODN-mediated Homology-directed Repair. Molecular Therapy Nucleic Acids. 2016;5(11):e396.
163. Byrne SM, Ortiz L, Mali P, Aach J, Church GM. Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. Nucleic Acids Research. 2015;43(3):e21-e.
164. Bylund L, Kytölä S, Lui WO, Larsson C, Weber G. Analysis of the cytogenetic stability of the human embryonal kidney cell line 293 by cytogenetic and STR profiling approaches. Cytogenetic and Genome Research. 2004;106(1):28-32.
165. Lin Y-C, Boone M, Meuris L, Lemmens I, Van Roy N, Soete A, et al. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. Nature Communications. 2014;5:4767.
166. Takara Bio USA. Highly specific gene knockins of long sequences using CRISPR/Cas9 and a single-stranded DNA donor template. 2017. Accessed 05/29/2018. Available from: http://www.clontech.com/US/Products/Genome_Editing/CRISPR_Cas9/Technical_Notes/Long_ssDNA_for_gene_knockins.