Functional characterization of BRD4 in transcription elongation and termination

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

Mirjam Arnold

2022

The dissertation was prepared under the supervision of Dr. Andreas Mayer at the Max Planck Institute for Molecular Genetics in Berlin from November 2017 to March 2022.

1st reviewer: Dr. Andreas Mayer

2nd reviewer: Prof. Dr. Florian Heyd

date of defense: July 13, 2022

Selbstständigkeitserklärung

Hierdurch versichere ich, dass ich meine Dissertation selbstständig verfasst und keine anderen als die von mir angegeben Quellen und Hilfsmittel verwendet habe.

Berlin, den 11. April 2022

……

Mirjam Arnold

Acknowledgement

I'd like to sincerely thank the following people:
My supervisor Andreas Mayer, for his guidance, support and encouragement, and for the many exciting opportunities I got over the last four years;
Florian Heyd, for reviewing this thesis as second reviewer;
Annkatrin Bressin, whose ideas, skills and commitment were essential for the BET project;
Liza Mochalova, the best lab friend I could imagine, who was always there to critically discuss failed experiments and celebrate successful ones, to have lunch together outside and visit art exhibitions;
David Meierhofer, for his tremendous help with all mass spectrometry aspects of this project;
Olga Jasnovidova, from whom I learned a lot, be it about library preparation, stem cell culture or personal integrity;
Susanne Freier, for all the great help and for tolerating the chaos I sometimes caused in the lab;
Mario Rubio, for many helpful comments on the thesis draft;
Jelena Ulicevic, for ∞ fun discussions;
Thomas Kratz, for his exceptional effort and super solid work during his lab rotation;
all Mayer lab members, especially Elisabeth Altendorfer, Nicole Eischer and Martyna Gajos, for the nice working atmosphere and collegiality;
Denes Hnisz, for his help setting up Hi-ChIP;
Tuğçe Aktaş, for advice both scientific and career-related;
Wiebke Skeffington, for organizing everything IRTG-related so well;
the sequencing core, for the fast and reliable service and help with establishing new protocols;
Beata Lukaszewska-McGreal, for always taking good care of our protein samples;
Alexander Hartemink and Sneha Mitra for helpful discussions about chromatin accessibility;
lastly, my parents, Maria Kilarska-Arnold and Norbert Arnold, for supporting me in every possible way!

TABLE OF CONTENT

1 SUMMARY

The eukaryotic transcription machinery consisting of RNA polymerase II (Pol II) and a small number of minimally required initiation, elongation and termination factors is well-characterized structurally and biochemically, while functional characterization traditionally has been hampered by the limitations of the available perturbation strategies and readouts. In addition, knowledge of accessory transcription regulators is likely incomplete. Proteins of the conserved bromodomain and extra-terminal domain (BET) family have been demonstrated to function in maintaining normal transcription genome-wide by promoting elongation. However, the underlying molecular mechanism and also the specific contribution of each of the protein family members remained elusive. Moreover, it was unclear if BET proteins serve additional direct or indirect functions in transcription or transcription-coupled processes.

To address these aspects, this study combines targeted protein degradation, a new strategy to near-completely deplete a protein of interest within minutes to hours, with complementary transcriptome-, genome- and proteome-wide readouts of high temporal or spatial resolution. It provides evidence that specifically BRD4 is involved in transcription control, although BRD2 and BRD3 might have partially overlapping functions. Within 120 minutes, BRD4-selective degradation results in a global reduction of elongating Pol II from the gene body of most protein-coding and long non-coding RNA genes, as can be detected by native elongating transcript sequencing with spike-in normalization (SI-NET-seq). Concomitantly, Pol II accumulates in the promoter-proximal region, suggesting a defect in promoter-proximal pause release. Profiling elongation factors upon BRD4 depletion reveals a decrease in occupancy, suggesting an assembly defect of the active elongation complex. In particular, PAF1 binding is decreased throughout the transcribed region.

Unexpectedly, BRD4-selective degradation also induces a severe transcription termination defect at a subset of genes. Particularly, Pol II continues transcribing on average three kilobases beyond its usual termination zone, which is accompanied by inefficient cleavage of the nascent transcript at the pA site. Chromatin immunoprecipitation with sequencing (ChIP-seq) reveals a significant reduction of RNA 3’ end processing factors of the CPSF and CstF modules near the 3’ gene end as well as in the promoter-proximal and the gene body region. The failure to recruit RNA 3’ end

processing as well as PAF subunits, SPT5 and SPT6, but not CDK9, could be independently confirmed by quantitative mass spectrometry. Co-immunoprecipitation experiments indicate that the recruitment of 3' processing factors either directly depends on BRD4 or could be mediated by elongation factors.

Also, BRD4-selective degradation locally increases the binding of Pol II and SPT5 at transcribed enhancers, whereas enhancer accessibility appears to be not BRD4-dependent.

Altogether, the data establish a role of BRD4 in coordinating the recruitment of elongation and RNA 3' end processing factors during an early step of transcription.

ZUSAMMENFASSUNG

Die eukaryotische Transkriptionsmaschinerie — insbesondere RNA-Polymerase II (Pol II) und eine überschaubare Gruppe essentieller Initiation-, Elongations- und Terminationsfaktoren — ist strukturell als auch biochemisch gut charakterisiert, wohingegen die funktionelle Charakterisierung durch das Fehlen geeigneter Pertubationsmethoden und sensitiver experimenteller Analyseverfahren lange erschwert wurde. Auch sind viele nicht-essentielle Transkriptionsregulatoren wahrscheinlich noch unbekannt.
Es konnte gezeigt werden, dass die durch zwei Bromodomänen und eine sog. extra-terminale Domäne gekennzeichneten Proteine der konservierten BET-Proteinfamilie nötig sind, die normale Transkriptionsaktivität der Zelle — vermutlich durch Regulation der Elongation — aufrecht zu erhalten. Unklar sind jedoch der molekulare Mechanismus als auch die Rolle der einzelnen Proteine der Proteinfamilie. Darüber hinaus ist denkbar, dass BET-Proteine direkt oder indirekt auch an weiteren transkriptionellen oder co-transkriptionellen Prozessen beteiligt sind.

Die genannten Aspekte adressiert die vorliegende Studie einerseits durch die Verwendung eines relativ neuen Verfahrens zur gezielten und schnellen Degradierung eines Proteins — in diesem Fall BRD4 — und andererseits mit komplementären transkriptom-, genom- und proteomweiten Methoden, die eine hohe zeitliche oder räumliche Auflösung bieten. Auf diese Weise kann gezeigt werden, dass die

Transkription in hohem Maße von BRD4 abhängt, obgleich nicht auszuschließen ist, dass BRD2 oder BRD3 eine ähnliche Funktion haben. So belegt SI-NET-seq, eine spike-in-gestützten Methode zur quantitativen Analyse naszenter RNA unter nativen Bedingungen, dass die gezielte Degradierung von BRD4 innerhalb von 120 Minuten zur globalen Verringerung transkribierender Pol II am Genkörper der meisten proteinkodierenden und langen nicht-kodierenden RNA-Gene führt. Gleichzeitig akkumuliert transkribierende Pol II in der promotor-proximalen Region, was auf einen Defekt am Übergang von der promotor-proximalen Pause zur produktiven Elongation hindeutet. Einen Hinweis auf eine Funktion von BRD4 bei der Assemblierung des Elongationskomplexes liefert die Beobachtung, dass die Chromatinbindung von Elongationsfaktoren nach BRD4-Degradierung verringert ist. Insbesondere PAF1 zeigt eine Abnahme in allen Genbereichen.

Unerwarteterweise führt die gezielte Degradierung von BRD4 bei einer großen Gruppe von Genen auch zu einem Terminationsdefekt — im Durchschnitt erfolgt die Termination drei Kilobasen später als unter Kontrollbedingungen — und vermindert die Effizient, mit der die naszente RNA an der Polyadenylierungsstelle enzymatisch geschnitten wird. Tatsächlich binden CPSF- und CstF-Faktoren der 3'-RNA-Prozessierungsmaschinerie nach BRD4-Degradierung signifikant weniger am 3'-Ende der Gene, aber auch im promotor-proximalen Bereich und am Genkörper, wie mittels Chromatinimmunpräzipitation (ChIP) gezeigt werden kann. Eine signifikante Abnahme von 3'-Prozessierungsfaktoren sowie von PAF-Untereinheiten, SPT5 und SPT6, nicht aber von CDK9, ist auch durch quantitative Massenspektrometrie nachweisbar. Zudem legen Co-Immunpräzipitationsexperimente nahe, dass 3‘-RNA-Prozessierungsfaktoren entweder direkt durch BRD4 oder indirekt über Elongationsfaktoren rekrutiert werden können.

Außerdem korreliert die BRD4-spezifische Degradierung mit der lokalen Zunahme von Pol II und SPT5 an transkribierten Enhancer-Regionen; die Zugänglichkeit des Chromatins gegenüber Tn5-Transposase erscheint hingegen nicht BRD4-abhänging zu sein.

Zusammenfassend legen unsere Daten nahe, dass BRD4 zu einem relativ frühen Zeitpunkt des Transkriptionszyklus' die Rekrutierung von Elongations-, aber auch von RNA 3'-Prozessierungsfaktoren koordiniert.

2 INTRODUCTION

2.1 THE PHASES OF RNA POLYMERASE II TRANSCRIPTION

RNA polymerase II (Pol II) is the 12-subunit enzyme complex that produces protein-coding mRNA, long non-coding RNA (lncRNA), certain primary micro RNA (pri-miRNA) [1] as well as enhancer RNA (eRNA). The transcriptional activity of Pol II is thus a critical determinant of the transcriptome and ultimately the proteome, which underlie cell identity and function. The process of transcription can be subdivided into the phases of initiation, elongation and termination [2].

2.1.1 TRANSCRIPTION INITIATION

Transcription initiation usually occurs in—and is directed by—the promoter region, located at -50 to +50 bp relative to the transcription start site (TSS), but also more distally at enhancers. Promoter regions are characterized by an open chromatin configuration, as characterized, for instance, by nuclease hypersensitivity. Chromatin accessibility is a prerequisite for recognition by sequence-specific transcription factors (TFs) and especially for the assembly of the pre-initiation complex (PIC) [3–5]. How chromatin accessibility is initially achieved in cells is not fully understood. Pioneering factors can bind to closed chromatin and open it, either passively by creating an environment that facilitates transcription activation by additional TFs, or by actively recruiting chromatin remodelers and histone chaperones [6]. Accessibility depends both on nucleosome density and turnover [7]. In line with the openness of the chromatin and high levels of TF binding, promoters are enriched in distal chromatin interactions [8].

Histone marks associated with the flanking nucleosomes at promoters and active enhancers are H3K4me1/2/3, H3K79me2, H4K5ac, H3K9ac and H3K27ac, but it remains to be proven whether these histone marks are causally involved in transcription [9]. While H3K4me3 has been considered a feature of only promoters, it is now known to be present also at highly transcribed enhancers [10]. Promoter architecture differs with regard to the functional class of the associated gene: promoters of genes fulfilling roles in terminally differentiated, specialized cells are usually sharp and focused on a single TSS [11] whereas housekeeping and developmental genes harbor broad promoters with multiple

closely spaced TSSs [12]. Most promoters are CpG-rich, while the opposite is true for enhancers [13,14].

Sequence-conserved core promoter elements have been identified. Whether they efficiently recruit general transcription factors (GTFs) depends on the similarity of their sequence to the consensus and the relative spacing between the motifs. The TATA box is an asymmetrical AT-rich element of the sequences TATAAA or TATATA [15] that is located approximately 30 bp upstream of the TSS in 10 to 20% of the mammalian genes [16], particularly those involved in tissue-specific functions [17]. The initiator (INR; sequence YYA(+1)NT/AYY) located approximately -3 to +5 bp around the TSS [18] is considered necessary and sufficient to specify the exact position of transcription initiation [19]. It is present in approximately 46% of all human core promoters and, therefore, it is the most frequent core promoter element [20]. Moreover, two infrequent sequence elements (BRE^{u}; BRE^{d}) that are recognized by TFIIB have been identified upstream and downstream of the TATA box [21]. Polarity of the structure and synergy of binding with TBP might ensure the correct orientation and ordered assembly of the PIC. Lastly, the downstream core promoter element (DPE) lies in the region of +28 to +32 bp in TATA box-deficient genes, where it serves as an alternative binding site for TFIID [22].

According to the established model that agrees with both biochemical and structural studies [23], the assembly of the PIC is initiated by the binding of TBP of the general transcription factor TFIID to the TATA box [24], followed by binding of the non-essential TFIIA that stabilizes TBP [25]. Next, TFIIB is recruited and interacts with Pol II near the RNA exit channel and the active center [26] and with TBP and the BREs, which helps melting the DNA helix at the promoter and stabilizing the clamp domain to promote RNA synthesis [27]. Since the TFIIB-containing PIC can accommodate only a 5 nt long DNA–RNA hybrid, synthesis of >10 nt in the presence of TFIIB was proposed to lead to abortive initiation, possibly functioning as a checkpoint [28]. Usually, however, Pol II dissociates from TFIIB and escapes the promoter [26]. The association of TFIIF promotes retention of TFIIB at the PIC [29] but also has a role in transcription start site selection [30]. Lastly, open PIC formation is completed by the binding of TFIIE and TFIIH. TFIIE stimulates the kinase and ATPase activity of TFIIH [31] which is required to translocate the downstream DNA towards Pol II, resulting in promoter melting and entry of the template DNA into the active center of Pol II [32]. All GTFs except certain TAF proteins are essential for transcription *in vivo* [33].

Another key factor in initiation is Mediator, a conserved complex of 26 core subunits forming a “molecular bridge” between TFs and Pol II, specifically its C-terminal domain (CTD) [34]. Mediator interacts with and partially regulates the activity of GTFs, including TFIIB [35], TFIID [36] and TFIIH [37]. Also, a role in post-initiation regulation has been suggested [38,39].

Initiation involves the ordered phosphorylation of the C-terminal domain (CTD) of the RPB1 subunit of Pol II, which consists of 52 repeats of the sequence tyrosine–serine–proline–threonine–serine–proline–serine. While TFIIE, TFIIF and TFIID interact with the hypo-phosphorylated CTD [40], a conserved module consisting of CDK7 of TFIIH, cyclin H and MAT1 phosphorylates the CTD at the Serine 5 and 7 residues [41], which promotes promoter escape [42] as well as 5’ RNA capping [43,44] in a Mediator-dependent manner [45]. In addition, by phosphorylating CDK9 and CDK12, CDK7 functions as a CDK-activating kinase (CAK) [46].

In mammalian cells, the sense TSS is usually accompanied by an antisense TSS located, on average, 110 bp upstream [10], which gives rise to a usually unstable antisense transcript [47,48]. ChIP-exo data for TBP and TFIIB have revealed that divergent antisense transcription involves the assembly of an independent PIC within the same nucleosome-depleted region [10,49]. Degradation of these antisense transcripts by the nuclear exosome [50] depends on polyadenylation sites located close to the TSS and the relative depletion of U1 small nuclear ribonucleoprotein (snRNP) binding sites [51,52]. However, also instances in which another mRNA [53] and lncRNA [54] are divergently transcribed have been described. The ability of a promoter to induce antisense transcription appears to be encoded in its sequence and is conserved between mammalian species [53].

2.1.2 EARLY AND LATE ELONGATION

Historically, recruitment of Pol II and transcription initiation were considered the main regulatory steps in Pol II transcription [56]. This view was challenged by the discovery of widespread transcriptional pausing in the promoter-proximal region in *Drosophila* and mammalian cells [57–63].

Generally, Pol II pausing is thought to provide a “window of opportunity” for the integration of regulatory signals [64–66], such as general and context-specific transcription factors acting in *trans*, *cis*-regulatory elements or epigenetic modifications. In this sense, the duration of a state correlates with the likelihood of receiving a specific signal. There is extensive evidence for kinetic or temporal coupling of elongation with splicing [65,67] as well as RNA 3’ end processing [68,69]. Promoter-proximal pausing specifically might serve as a checkpoint that ensures full maturation of the elongation machinery before transcription is resumed [64,70].

In addition, Pol II pausing is also involved in spatial coupling. The phosphorylated C-terminal domain of Pol II can serve as a “landing pad” for the recruitment and proper arrangement of factors involved for instance in RNA processing [65]. A similar function has been ascribed to co-activators like Mediator during initiation or BRD4 during early elongation [71].

Besides, promoter-proximal pausing was shown to counteract nucleosome assembly near the TSS, thereby preventing the formation of a repressive chromatin structure [72,73], and to facilitate rapid and synchronous transcription induction upon a stimulus [74,75].

2.1.2.1 Discovery of promoter-proximal pausing

Early evidence for promoter-proximal pausing came from observations at the adult β-globin locus in *Gallus gallus domesticus*, whose 5’ portion was predominantly bound by transcribing Pol II in mature but not in immature erythrocytes [76]. Sarkosyl, which is required for paused Pol II to resume transcription but dispensable during late elongation, was used for nuclei isolation. The authors proposed that Pol II can initiate but not elongate unless blockage was relieved “for instance by modification of the chromatin structure or by the release of a particular protein from the DNA template” [76]. Similar observations were made in *Drosophila*: While Pol II in uninduced cells predominantly crosslinked to the region of -12 to +65 bp from the TSS of the *Hsp70* gene, it was detectable along the entire gene upon heat shock [77]. Another early study linked promoter-proximal pausing at heat shock genes and 5’ capping [78]. First evidence for promoter-proximal pausing in human cells came from studies on the *c-MYC* locus [79,80], where paused Pol II could be mapped to the +30 bp position downstream of the TSS.

Later, genome-wide approaches, such as chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) or global run-on sequencing (GRO-seq), extended these observations from single stimulus-responsive to a majority of genes [47,57,81–83].

2.1.2.2 PHYSICAL BASIS OF POL II PAUSING

Generally, three models of Pol II pausing, which are not mutually exclusive but nevertheless might be predominant at different loci, have been described [84]. Notably, the interaction and barrier models could be seen as special cases of the kinetic model.

According to the kinetic model, features of Pol II, *e.g.,* its early elongation rate and processivity, or the template and RNA product, *e.g.,* the thermodynamic stability of the double-stranded DNA template or the RNA–DNA hybrid, are determinants of Pol II pausing. The elongation rate is equivalent to the velocity of Pol II and is usually given in kb per minute, while processivity describes the propensity of Pol II to repeatedly undergo nucleotide addition cycles rather than to terminate prematurely [86]. While Brownian motion makes Pol II alternate between a pre-translocation and a post-translocation state (in which the active site at the +1 bp position harbors an unpaired deoxyribonucleotide), the forward movement of Pol II is governed by the thermodynamically favored incorporation of a nucleoside triphosphate (NTP) into the nascent RNA [85]. NTP incorporation restores the pre-translocation state, thereby resetting the translocation cycle. Transient pausing between two cycles is intrinsic to the process [86] but can be extended by encountering *trans*-acting factors, such as elongation factors [87,88] or nucleosomes [85]. Also, *cis*-elements, such as the CG-rich “pause button” in *Drosophila* [89] or similarly GC-rich motifs followed by AT-rich regions in human cells [90–92] can slow down or pause Pol II. A likely reason is that melting of a DNA duplex high in GC content requires relatively more energy. Surprisingly, also the opposite effect has been described, such that the formation of a stable structure of the nascent RNA might counteract pausing [93]. Furthermore, Pol II can undergo more stable pausing, such as transcriptional arrest, which requires TFIIS to be resolved [94,95].

Apart from discrete pausing events, also the elongation rate appears to be regulated, as it ranges from 0.5 kb/min to >5 kb/min at different genes and different gene regions [58,96,97].

According to the interaction model of Pol II pausing, DNA- or RNA-binding factors interacting with Pol II physically prevent its efficient elongation, for example, by tethering it to the promoter-proximal region. An early example is the Sigma factor σ70 in *Escherichia coli*, which causes RNA polymerase pausing by recognizing sequence elements and binding to the non-template strand within the transcription bubble of the *Lambda* late promoter [98,99]. In mammalian cells, this model is in agreement with the function of GTFs or NELF and DSIF [100,101].

An example of the barrier model [84] might be nucleosome-induced pausing, which is distinct from promoter-proximal pausing in the strict sense [86]. Here, the +1 nucleosome, which has a median distance of 214 bp to the TSS in human cells [102], acts as a barrier for Pol II [73,103]. Likewise, nucleosomes along the gene body region impede elongation, although the pausing pattern differs from that at the +1 nucleosome [104]. The biophysical basis of this type of pausing is the stability of the DNA–nucleosome interaction, which is maximal at the entrance (where Pol II stalls at the +1 nucleosome) and before the dyad (where stalling occurs at gene body nucleosomes) [105–107]. Histone chaperones and chromatin remodelers, such as FACT [108–110] and CHD1 [111] act to loosen these interactions and thereby facilitate passage through the nucleosome without disrupting its octameric structure [112,113]. Furthermore, core elongation factors like DSIF [107], the PAF1 complex and SPT6 [114] are required to transcribe through nucleosomes.
Similarly, also CTCF and YY1 have been shown to impede Pol II elongation [59].

2.1.2.3 THE MOLECULAR MECHANISM OF PROMOTER-PROXIMAL PAUSING

After initiation and promoter escape, Pol II enters the phase of early transcription and transcribes 50 to 100 nt but then pauses. Pausing involves the binding of the negative elongation factor NELF and the DRB sensitivity-inducing factor DSIF consisting of SPT4 and SPT5 (Figure 1A) [61,84]. Whether both factors are required only to maintain the pause or also for its initial induction *in vivo* is still under debate [70].

SPT5 firmly interacts with multiple domains of Pol II. Its NGN–SPT4 and KOW1–L1 domains form a DNA clamp contacting the non-template strand, while KOWx–4 and KOW5 form an RNA clamp located near the RNA exit channel [87,115,116]. Interaction with both nucleic acids appears to stabilize DSIF binding at the elongation complex, while specifically the DNA clamp ensures a closed conformation of the cleft of Pol II and correct

positioning of the non-template strand [115,117]. By stabilizing the RNA in its correct position, the RNA clamp possibly prevents R-loop formation and promotes pausing [115]. Functional analyses have revealed a dual role of SPT5 in transcription that comprises a negative function during promoter-proximal pausing and a positive function in productive elongation [118]. The overlay of the structure of Pol II bound to DSIF or the initiation complex revealed possible sterical clashes with TFIIA, TFIIB, TFIIE and TFIIF, demonstrating that promoter escape needs to happen prior to DSIF binding [115].

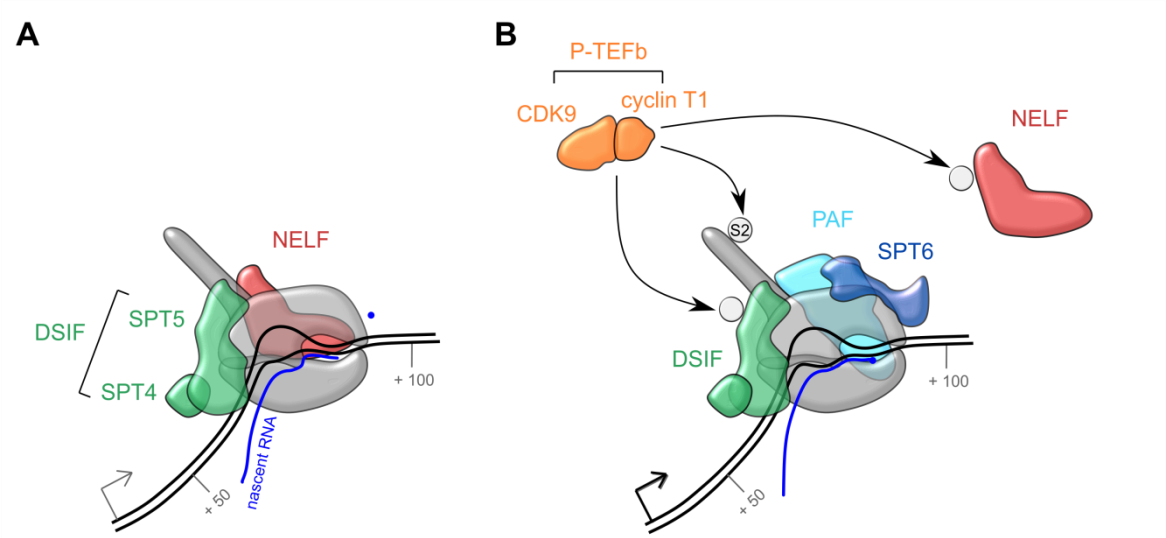


Figure 1: Schematic representation of the events during early elongation. (A) Promoter-proximal pausing, which occurs on average 50 to 100 bp downstream of the TSS, is stabilized by DSIF consisting of SPT4 and SPT5 and the Negative Elongation Factor (NELF) complex. In this pause state, the DNA–RNA hybrid in the active center of Pol II adopts a conformation that transiently precludes the extension of the nascent transcript. **(B)** The transition into productive elongation is mediated by the Positive Transcription Elongation Factor b (P-TEFb), a heterodimer of CDK9 and cyclin T1, which phosphorylates SPT5, NELF and Ser2 of the Pol II CTD. While phosphorylation is critical for the elongation-promoting activity of SPT5, it leads to the dissociation of NELF from Pol II. Subsequently, additional elongation factors, such as the PAF complex and SPT6, are recruited to the released Pol II. (Schemes are based on Noe Gonzales et al. [86]).

Once DSIF is bound, also NELF can bind and contribute to pausing [100,116]. NELF folds into three lobes, one of which (the NELF-A–NELF-C lobe) contacts the funnel and trigger loop domains of Pol II, limiting their mobility. Interestingly, binding of NELF and TFIIS to the funnel region of Pol II appears to be mutually exclusive indicating that TFIIS cannot act on NELF-bound Pol II to resolve stalling [116]. Upon association with Pol II, the DNA–RNA hybrid is tilted, with the DNA adopting a pre-translocation and the RNA adopting a post-translocation conformation, so that the NTP substrate cannot be bound [116]. Also, restriction of the movement of the trigger loop likely excludes NTP from the active site.

Knockdown of NELF in *Drosophila* cells leads to a reduction of Pol II in the promoter-proximal region of most genes which, however, is not globally accompanied by an increase at the gene body, suggesting that NELF is involved in a rate-limiting step only at a subset of genes [73,82,119]. Recent work in mammalian cells found promoter-proximal pausing to depend on NELF, but uncovered a second pause site located upstream of the dyad of the +1 nucleosome, where Pol II can accumulate even upon targeted degradation of NELF-C [120]. Interestingly, simultaneous NELF degradation and CDK9 inhibition results in the accumulation at the second pause site, underlining the role of NELF in pausing at the first pause site [120].

2.1.2.4 Regulation of promoter-proximal pause release and elongation

Pause release is mediated by the positive transcription factor P-TEFb, which consists of the kinase subunit CDK9 and cyclin T1 (Figure 1B) [121,122], but can also associate with additional factors to form the super-elongation complex (SEC) [123]. Binding by the 7SK non-coding RNA and HEXIM1/2 inhibits its kinase activity [124–126] and chromatin association [127,128]. Dissociation from HEXIM1/2 and 7SK was proposed to require cyclin T1 acetylation by the acetyltransferase P300 [129] and BRD4 [130], while recruitment to the chromatin probably relies on transcriptional co-regulators, such as Mediator [38], or context-dependent transcription factors, such as c-MYC [57,131]. The view that BRD4 recruits P-TEFb [132–134] could not be confirmed by more recent work which demonstrated that P-TEFb occupancy was unaffected by the ablation of BET proteins [135–137].

P-TEFb phosphorylates the C-terminal-repeat region of SPT5 [138,139], NELF [140] and the C-terminal domain of RPB1 at serine 2. Additionally, the elongation factors SPT6 and PAF1 are recruited to the active elongation complex, which is at least partially phosphorylation-dependent [87]. The formation of the active elongation complex leads to loosening of the DNA clamp, possibly facilitating the unwinding of the upstream DNA. At the same time, the cleft of Pol II remains closed, which prevents the dissociation of the DNA non-template strand and ensures processivity [141].

In mammalian cells, phosphorylated serine 2 (p-Ser2) occupancy peaks immediately downstream of the TSS, corresponding to a population of CDK9-phosphorylated polymerases in the process of resuming transcription elongation, and gradually increases along the gene body towards the termination zone, probably due to phosphorylation by

CDK12 [142,143]. Initially considered necessary for pause release, it might in fact be dispensable [144]. The characterization of the relationship between p-Ser2 and pause release is complicated by the lack of specificity of many widely-used CDK9 inhibitors [145] and the requirement for CDK9-mediated phosphorylation on elongation factors.

Phosphorylation of NELF at a flexible region of NELF-A destabilizes its association with Pol II [87,146], resulting in its dissociation from the transcription machinery.
In contrast, the RNA clamp of DSIF assumes a more open but still stable conformation upon phosphorylation [87]. SPT5 harbors a C-terminal repeat region (CTR) that consists of five repeats of the consensus sequence glycine-serine-arginine/glutamine-threonine-proline [139] and with regard to its high content of potential phosphorylation sites resembles the CTD of Pol II. Thr4 of the CTR motif has been shown to be the preferred substrate for CDK9 and is required for the transcription-promoting activity of DSIF [139]. CTR phosphorylation also promotes the recruitment of PAF1 [147], although this has been questioned by one study [148]. Additional interactions, for instance, with SPT6, the FACT complex or TFIIS have been identified but might be independent of the phosphorylation state [149]. SPT5 acts genome-wide and remains bound to Pol II throughout elongation [57,73].

Depletion of the Spt5 yeast homolog results in an accumulation of Pol II over the first 500 bp from the TSS and a relative decrease further downstream, suggesting a role in an elongation checkpoint [148]. Similarly, SPT5 in mammalian cells seems important to maintain processivity within the +15 to +20 kb window [150]. Targeted degradation of SPT5 results in a pronounced reduction of Pol II in the promoter-proximal region that is partially due to degradation of RPB1, and a less consistent reduction at the gene body that reflects a decrease in Pol II processivity [151,152]. SPT5 is also present at a subset of enhancers, regulates eRNA production and thereby, via recruitment of Mediator and cohesin or by promoting promoter–enhancer contacts, contributes to the activation of the target gene [151,153].

The PAF complex consists of the five subunits PAF1, CDC73, CTR9, WDR61, LEO1 and the loosely associated RTF1, which show extensive interactions with the unphosphorylated, Ser2- and Ser5-phosphorylated forms of Pol II [154]. CTR9 is the central subunit that connects the foot and funnel domains of Pol II [87,88], while PAF1 and LEO1 are in contact with RPB2. After association with Pol II in the promoter-proximal region, it remains associated until after the pA site [155]. CDK9 has been suggested to facilitate PAF recruitment, which was shown to help the recruitment of CDK12 to Pol II [156]. In turn, the

association of elongation factors, such as SPT5, SPT6 or the FACT complex, with Pol II does not depend on PAF [157]. Functional analyses of PAF1 have yielded inconsistent results: Yu et al. [156] showed that knockdown of PAF1 results in a global reduction of stable mRNA and Ser2 phosphorylation and an increase in promoter-proximal pausing, suggesting that PAF1 is a positive regulator of pause release. In contrast, an increase in nascent and mature transcripts and also of p-Ser2 was seen upon shRNA-mediated PAF1 depletion by F. X. Chen et al. [158], although some of the presented data seem ambiguous. Interestingly, some studies showed that pause release in part is mediated by enhancers activated by PAF [159,160]. Moreover, PAF depletion has been shown to cause a reduction in processivity and elongation rate [102,157].

SPT6 interacts with the RPB4–RPB7 stalk but also with the CTD linker of RPB1 rather than the heptad repeats [161]. Like PAF, SPT6 does not contact the active site of Pol II, suggesting an allosteric mode of action [87,88]. SPT6 is a replication-independent histone chaperone that facilitates transcription through nucleosomes and after passage of Pol II restores the normal chromatin structure [162,163]. However, SPT6 has also been shown to promote elongation on histone-free DNA *in vitro*, suggesting an additional function that is independent of its chaperone activity [164]. Recently, two studies employing targeted SPT6 degradation found an increase in Pol II near the 5' gene end and drastic decrease in elongation rate and processivity [102,165]. Biochemical studies showed that the PAF complex and SPT6 together, but none of them alone can stimulate elongation in the presence of active P-TEFb and DSIF, indicating cooperativity between PAF and SPT6 [87]. Long-term depletion upregulates cryptic intragenic and antisense transcription, suggesting an indirect role in transcription initiation [165,166]. Lastly, also a function of SPT6 in termination has been proposed [165,167].

In an *in vitro* extension assay, pausing was destabilized in the presence of both NELF and the PAF complex as well as active P-TEFb [87]. This is in accordance with structural data showing that NELF and PAF (in particular CTR9 and WDR61) contact the foot and protrusion domains of Pol II [87,116]. Thus, the simultaneous binding of both factors is sterically disfavored. Moreover, in contrast to the paused elongation complex, the active elongation complex can accommodate an untilted DNA–RNA hybrid and thus allows fluctuating between the pre- and the post-translocation step [87,116]. Both findings exemplify how compositional changes are mechanistically linked to pause release.

Instead of entering the phase of productive, late elongation, Pol II also might terminate prematurely [168–171].

2.1.3 RNA 3' END PROCESSING AND TRANSCRIPTION TERMINATION

The processes at the 3' gene end are rudimentarily understood compared to those at the 5' end. This is despite the severe consequences that dysregulation might have for the cell: improper termination can cause transcription interference with downstream-located genes [172–174]. For instance, in the gene-dense genome of *S. cerevisiae*, transcription usually continues for no more than 150 nt downstream of the pA site, suggesting tight regulation [175]. But even in human cells, the median distance to the termination site is only 3.3 kb [176]. Additionally, termination is a prerequisite for the initiation of a new round of transcription. From an RNA-centric point of view, processing of the 3' end of a transcript is important to define the 3' UTR properly and to ensure stability and proper targeting of the transcript, for instance, to the cytoplasm [177].

2.1.3.1 SEQUENCE ELEMENTS AND CORE COMPLEXES IN RNA 3' END PROCESSING AND TERMINATION

Transcription termination is intimately linked with RNA 3' end processing, which predominantly occurs co-transcriptionally [178,179]. The molecular machinery is a highly dynamic complex of >15 core proteins with largely specialized functions, which temporarily associate with >80 additional factors [180] (Figure 2A).

Transcription termination requires the recognition of a functional polyadenylation signal (PAS). While cleavage of the nascent transcript *per se* was considered dispensable [181,182], newer studies questioned this view [183,184]. Moreover, there is consensus that the PAS provides signals for the assembly of the 3' end processing machinery [178,185]. The central conserved sequence motif at the 3' end of protein-coding and most long non-coding transcripts is the hexameric polyadenylation signal of the sequence AAUAAA (present in approximately 69% of the pre-mRNAs) [186] or AUUAAA (in 14% of pre-mRNAs) (Figure 2B). A U- or GU-rich downstream signal element (DSE) is present in approximately 80% of PAS-containing pre-mRNAs, while the motif UGUA (in <50% of pre-mRNA) [187] or UAUA is found upstream of the PAS [188].

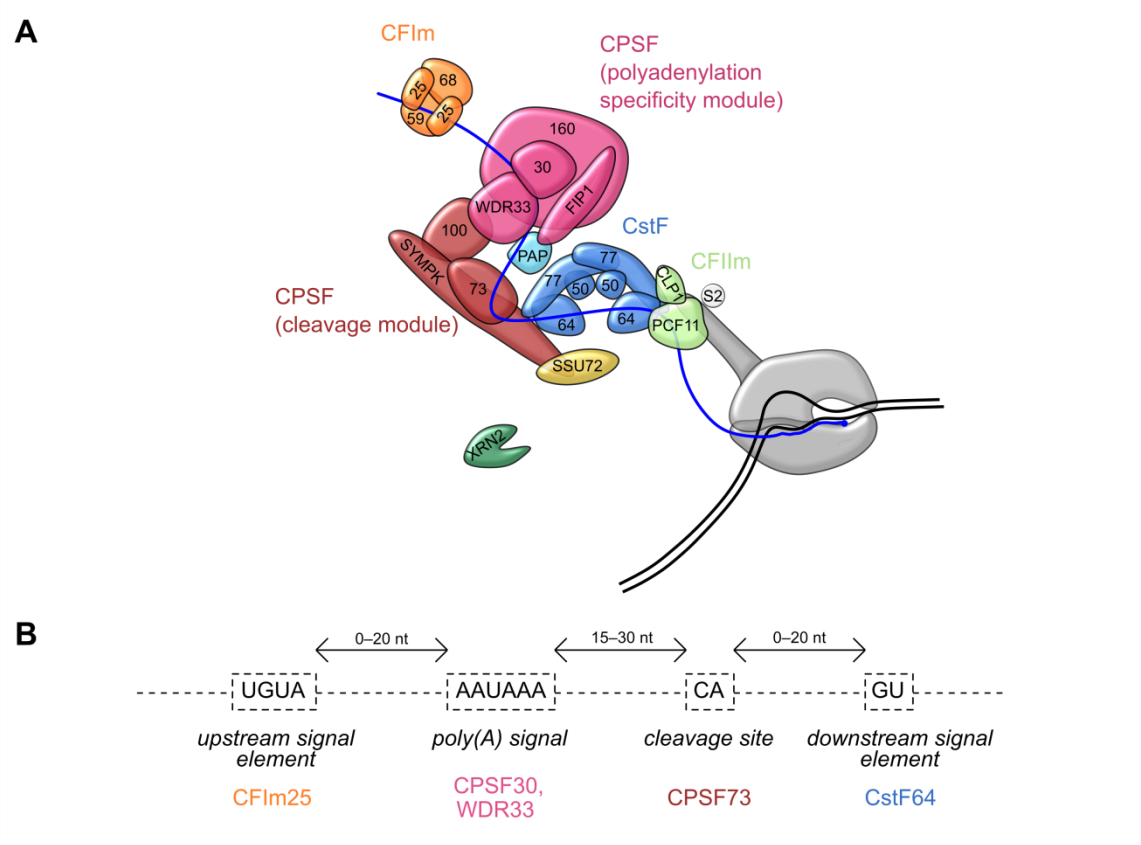


Figure 2: Schematic representation of the RNA 3' end processing machinery. (A) The RNA 3' end processing machinery is composed of four multiprotein complexes. Of central importance is the Cleavage and Polyadenylation Specificity Factor (CPSF), which consists of two distinct modules functioning in recognition of the polyadenylation signal (pink) and endonucleolytic cleavage of the nascent transcript (red). Another important complex is the Cleavage stimulation Factor (CstF; blue). Structural and biochemical studies suggested that the RNA 3' processing machinery is relatively dynamic and undergoes drastic conformational changes. **(B)** The 3' end of the nascent transcript harbors different and differently conserved sequence elements, which engage with specific subunits of the CPSF, CstF and CFIm complexes. (Schemes are based on Zhang et al. [189] and Laishram [190].

As was revealed by cryo-EM, the mammalian Cleavage and Polyadenylation Specificity Factor (CPSF) consists of a polyadenylation specificity module composed of CPSF160, WDR33, CPSF30 and FIP1, and a cleavage module formed by CPSF100, CPSF73 and symplekin [189] (Figure 2A). The PAS is recognized by the zinc finger-containing CPSF30 and by WDR33 [191,192]. FIP1, which binds CPSF30, recruits the poly(A) polymerase PAP1 but also regulates its activity [193]. CPSF160 functions as a scaffold of this overall rather static module. Interestingly, CPSF30 but not CPSF160 links CPSF and the body of Pol II in the absence of CstF, possibly to facilitate scanning for a functional polyadenylation signal site [194]. CPSF100 is a flexible tether between the two CPSF modules. The distance between the catalytic core and the PAS-binding region in currently available structures

suggests that CPSF100 can undergo remarkable rearrangements [189]. The endonuclease CPSF73 contains a metallo-β-lactamase domain and a β-CASP domain, which coordinate two zinc ions important for catalytic activity [195], and a C-terminal domain. A catalytically inactive β-CASP domain is also part of CPSF100. Cleavage by CPSF73 usually occurs 15 to 30 nt downstream of the PAS after a short CA motif and generates a free 3' hydroxyl end, to which approximately 250 adenosine nucleotides are added by PAP1 to stabilize the transcript [196,197]. Symplekin associates with the CPSF and CstF modules [198] and is dispensable for the assembly of the CPSF module as well as for recognition of the polyadenylation signal *in vitro* [199]. Moreover, symplekin associates with Pol II via the PAF complex [200] and stimulates the phosphatase activity of the Pol II CTD phosphatase Ssu72 in yeast [201].

The Cleavage stimulation Factor (CstF) overall is very flexible, which so far prevents its structural characterization at high resolution [189,202]. It consists of two copies of CstF77, CstF64 and CstF55. The symmetry of the CstF64 dimer might be necessary to bind the downstream signal element (DSE) [189]. CstF77 binds CPSF160 [203,204], while CstF50 binds the Pol II CTD [205]. CstF64 was shown to interact with the exonuclease XRN2 [206].

Cleavage Factor I (CIm) consisting of CFIm68, CFIm59 and two copies of CFIm25 was proposed to regulate alternative polyadenylation. CFIm25 specifically binds to the USE [207,208], whereas CFIm68 contains an arginine–serine repeat (RS) domain that mediates the interaction with FIP1 [187]. It has been shown that knockdown of CFIm68 and CFIm25 but not of CFIm59 favors the usage of proximal polyadenylation sites [209].

Cleavage Factor II (CFIIm) consists of CLP1 and PCF11. PCF11 interacts with nascent RNA and (via a CTD-interacting domain, CID) [210] with Ser2-phosphorylated Pol II. *In vitro* studies using human cell extracts revealed that PCF11 is required for timely termination in a PAS-dependent manner [211]. Because it is substoichiometric in human cells and only binds to a subset of genes, it was suggested to be an accessory rather than a core component of the termination machinery [212].

Replication-dependent histone pre-mRNAs are processed by a specific set of proteins but also require the CPSF cleavage module including CPSF73 [213,214], CstF64 [202,215] and symplekin [216].

Some studies indicated that factors of the 3’ RNA processing machinery are bound in the 5’ region of exemplary genes in yeast [217–219] and human cells [220–223] and travel with the transcription machinery to reach the 3’ gene end [194,224,225]. Also, looping between promoter regions and the termination zone has been described in yeast [226–228] and for selected genes in human cells [229,230] but not on a global scale.

2.1.3.2 REGULATORY PRINCIPLES OF 3’ RNA PROCESSING AND TERMINATION

Two not mutually exclusive models were envisaged to explain how recognition of the PAS eventually leads to the dismantling of the elongation complex. According to the torpedo model, the endonucleolytic cleavage of the nascent RNA by CPSF73 creates an uncapped and thus unprotected 5’ phosphatidyl end that serves as an entry point for the 5’–3’ nuclease XRN2 [231]. By successively degrading the transcript produced downstream of the gene locus, XRN2 follows Pol II and upon reaching it disrupts its contacts with the nascent RNA and the DNA template. The same mechanism was proposed for the function of the yeast termination factor Rat1 [232,233]. Importantly, Rat1 depletion or mutation of its nuclease domain is not sufficient to abolish transcription termination entirely but rather delays it [232,234], which led to the idea that additional proteins, for instance helicase Sen1 (senataxin in human), are needed [68,235]. Likewise, incomplete knockdown of XRN2 in human cells does not substantially affect termination [222] while the targeted degradation has moderate effects [184,236]. Notably, since to date no other 5’–3’ exonucleases have been identified in human cells, the incomplete defect suggests that the termination depends not solely on the torpedo model.

The allosteric model proposes that transcription of the PAS and binding of the 3’ RNA processing machinery induce destabilizing conformational changes of the elongation complex, resulting in its dissociation from the DNA template [237]. The exact cause of the conformational changes is still enigmatic. However, dephosphorylation of SPT5 by PP1/PPP1R10, which associate with the RNA 3’ processing machinery [180], recently has been shown to function as an allosteric switch that decelerates Pol II, thereby promoting termination via XRN2 [238] (see below). Additionally, SCAF4 and SCAF8 act as anti-terminator proteins, which by dissociation from Pol II cause allosteric changes and thereby promote termination after the correct pA site [239]. Also, PCF11 has been implicated in the allosteric model [212,240].

According to the current view, the termination process comprises aspects of both models [211,236,241–243]. Evidence for this unified model comes from the observation that, first, targeted depletion of CPSF73 induces a remarkably more pronounced readthrough than depletion of XRN2 [236,244], and, second, that simultaneous depletion of CPSF73 and XRN2 causes a more severe termination defect then depletion of CPSF73 alone [236]. If the torpedo model was entirely correct, one would expect the perturbation of CPSF73 and XRN2 to essentially have the same effects. This suggests that CPSF73 has a function beyond creating an entry site for the exonuclease and possibly affects Pol II also allosterically. Of note, targeted depletion consistently caused clearer transcription phenotypes in these experiments than knockdown, underlining the importance of rapid perturbation strategies.

In line with the concept of kinetic coupling, slowing or pausing of Pol II was detected downstream of the pA site at selected genes and on a genome-wide scale [222,224,225,245–247]. Furthermore, termination of both slow and fast Pol II mutants was found to be dysregulated [68,69] suggesting that an optimal elongation rate might aid the assembly of 3' processing factors inducing allosteric changes or help XRN2 catch up to Pol II. Mechanistically, what leads to slowing or pausing at the 3' gene end is not well understood yet. It has been proposed that CPSF initially binds the body of Pol II with its CPSF30 subunit, while simultaneously scanning the nascent RNA for the presence of the PAS. Recognition of the motif could induce pausing of the transcription complex, recruitment of CstF and the transfer of CPSF from the Pol II body to its CTD [194]. Moreover, the heterodimer of PPP1R10 and the phosphatase PP1 was shown to dephosphorylate SPT5 in a PAS-dependent manner [238,248]. Loss of SPT5 phosphorylation correlates with a reduction in Pol II elongation rate downstream of the pA site and is required for timely transcription termination, as was found by ChIP-seq in transcriptionally synchronized cells [238].

Also elongation factors are directly involved 3' RNA processing. It has been shown that PAF stimulates polyadenylation in *S. cerevisiae* [200] and that chromatin association of the yeast homolog of CPSF160, CFT1, in the gene body region depends on the PAF complex, specifically on CDC73 and RTF1 [249]. The so far only study conducted in human cells found that CDC73 depletion impaired the recruitment of CPSF30, CPSF73, CstF64 and CstF77 to selected loci, although neither spatial distribution along the gene nor aspects determining gene specificity could be resolved [250].

SPT5, itself an interactor of the PAF complex in yeast, also interacts with the yeast CPSF100 homolog CFT2, RNA14 and RNA15 (homologs of CstF77 and CstF64, respectively), poly(A) polymerase PAP1 and the yeast-specific REF2 [149] and recruits them to the pA site and downstream region [251]. Interestingly, recent studies in human cells demonstrated that targeted degradation of SPT5 results in readthrough transcription [151,165].

Moreover, the CTD of Pol II was implicated in regulating RNA 3' processing and termination. p-Ser2 is an interesting candidate phospho-residue due to its CDK12-dependent increase towards the 3' gene end and particularly its peak at the pA site [142,252]. Also, it is known to directly interact with PCF11 in *S. saccharomyces* [253,254]. Replacement of Ser2 by alanine prevents PCF11 recruitment and efficient 3' end RNA cleavage, although the reduced elongation rate of the mutated Pol II may be a confounding factor [255]. Reduction of p-Ser2 levels by depletion of CDK12 reduces chromatin association of CstF77 in human cells [225].

Another relevant phospho-form in this context is phosphorylated tyrosine 1 (p-Tyr1) which peaks not only at promoters and enhancers but also at the pA site in mammalian cells. Partial replacement of Tyr1 by phosphorylation-incompetent phenylalanine in the C-terminal CTD repeats results in readthrough transcription of tens to hundreds of kilobases in sense direction and a downstream shift of the 3' end pausing site, as well as milder termination defects in antisense direction [256]. Surprisingly, MS analysis of the interactome of the mutated CTD revealed a highly significant reduction of interactions with Integrator and Mediator subunits but not of CPSF and CFIm subunits or XRN2 [256]. Altogether, this indicates that p-Tyr1 is possibly required for steps subsequent to cleavage and polyadenylation of the pre-mRNA.

Phoshorylated threonine 4 (p-Thr4) shows a maximum downstream of the pA site of protein-coding genes but also along the gene body region of lncRNA genes in mammalian cells [257,258] and therefore has been linked to transcription termination. Substitution of Thr4 by valine reduces the binding of CPSF100 and SLBP to 3' end of replication-dependent histone genes, which results in reduced histone mRNA and a dramatic reduction in cell viability [259]. In contrast, mRNA levels of other protein-coding genes are not affected [259]. Lastly, Thr4 phosphorylation downstream of the pA site is reduced upon targeted degradation of CPSF73 [184].

2.1.4 CO-TRANSCRIPTIONAL SPLICING

Splicing is crucial to obtain functional mRNA molecules but also to increase the coding potential of the genome [260]. Definition of exon–intron borders involves the conserved dinucleotide sequences GU and AG at the 5' and 3' splice site, respectively, a branch point sequence (BPS) 18 to 40 nt upstream of the 3' splice site [261] and a polypyrimidine tract downstream of the BPS [262]. Moreover, regulatory so-called splicing enhancers and silencers have been described [263].The simplicity of the *cis*-regulatory elements is in stark contrast to the complexity of the *trans*-acting splicing machinery that in the canonical (U2-dependent) case consists of a core of U1, U2, U5 and U4/U6 snRNPs and >200 accessory proteins [264,265]. Uridine-rich small nuclear ribonucleoproteins (snRNPs) consist of a non-coding RNA and a protein portion both of which are crucial for spliceosomal function.

The step-wise assembly of the splicing machinery has been extensively characterized both biochemically and *in cellulo* [266]. Base pairing of the snRNAs with each other and the nascent RNA is central to the coordinated assembly of the splicing machinery. The splicing cycle is initiated by the binding of the U1 snRNP to the 5' splice site and non-snRNP factors (SF1, U2AF) to the BPS and polypyrimidine tract, followed by the binding of the U2 snRNP to the BPS, resulting in the A complex. Recruitment of the pre-assembled U4/U6.U5 tri-snRNP and loss of the U1 snRNP results in the formation of the B complex, which undergoes substantial conformational rearrangements and dissociates from the U4 snRNP, leading to the $B_{activated}$ complex. Upon catalytic activation by additional factors, *e.g.*, the PRP19 complex (also known as nineteen complex, NTC) [267,268], the complex is competent to perform the first of two transesterification reactions: the 2' hydroxyl group of the BPS adenosine carries out a nucleophilic attack on the phosphate group of the 5' splice site guanosine, resulting in the formation of an intronic lariat structure. The thus-formed C complex catalyzes the second transesterification reaction, *i.e.,* the attack of the reactive 3' hydroxyl group of the upstream exon on the nucleotide downstream of the 3' splice site guanoside, so that the intron is cleaved away and the exons are covalently joined. Lastly, the spliceosome is released, disassembled and recycled.

Splicing is thought to happen largely, but not strictly co-transcriptionally [269]. Coordination of splicing and transcription is thought to increase efficiency and reliability of the splicing process but has also been shown to provide feedback on transcription [270–273]. Different

subunits of the splicing machinery physically interact with GTFs and elongation factors or with chromatin remodelers and modifiers that indirectly influence transcription [270]. For instance, PHF5A, a protein associated with the U2 snRNP, stabilizes the PAF complex and thereby promotes transcription elongation [274]. Also, P-TEFb has been shown to recruit TAT-SF1 [275]. In addition, intact U1 snRNP, U2 snRNP and SRRM1 are required for efficient RNA 3' processing [276,277].

Pol II itself, for instance via the phosphorylated RPB1 C-terminal domain (CTD), impacts spliceosomal assembly and activity [278]. Large-scale proteomics analyses in yeast and mammalian cells found interactions predominantly between spliceosomal proteins—especially of the U1 snRNP—and phosphorylated serine 5 (p-Ser5) [167,279]. Ser5-phosphorylated RPB1 is also globally associated with splicing intermediates [167]. Moreover, Ser2 appears to be involved in splicing, as the replacement by alanine prevents recruitment of U2AF65 and the U2 snRNP [255]. Besides, phosphorylation-independent interactions are known, for instance, between Pol II and FUS, which promotes the interaction with U1 snRNP [280].

Transcription coordinates splicing also by providing a limited "window of opportunity" for splice site recognition, affecting both constitutive and alternative splicing. At a given time during transcription, only a certain portion of the nascent transcript is available, which restricts which splice sites are presented to the spliceosome and also which splicing enhancers or silencers are active [281]. Splice site selection is further modulated by elongation velocity [67]. In a simplified view, slow elongation favors the use of upstream splice sites and thus the inclusion of alternative exons, while the selective advantage is lost when elongation is fast [282]. Moreover, genome-wide studies detected Pol II pausing at exon–intron boundaries and an overall higher Pol II density at alternative retained than skipped exons [59,283]. However, other studies demonstrated that splicing requires an "optimal" elongation velocity and often happens minutes after a splice site was transcribed [284], suggesting more fine-grained regulation [67].

2.2 The Bromodomain and extra-terminal domain (BET) family of proteins

The family of BET proteins consists of the ubiquitously expressed BRD2, BRD3 and BRD4 [285,286] and the predominantly testis-specific BRDT [287]. BRD2, BRD3 and BRD4

show a similar binding pattern across the genome, with most binding occurring in the promoter and promoter-proximal region and at intronic and intergenic regions likely reflecting enhancers [137,288]. While BRD4 predominantly binds to promoters, a larger proportion of BRD2 peaks is found at introns and intergenic regions. Almost all BRD4 binding sites are co-occupied by BRD2, but only 40% are also bound by BRD3 [137].

Chromatin binding is dictated by the unique domain structure that allows recognition of acetylated histones but also of sequence-specific, acetylated or non-acetylated TFs [289,290]. These interactions underlie BRD4's role in both general and context-dependent Pol II transcription.

2.2.1 DOMAIN STRUCTURE

BET proteins are characterized by the presence of two N-terminal bromodomains (BD) and an extra-terminal (ET) domain (Figure 3). In addition, the long isoform of BRD4 (152 kDa) and BRDT carry a C-terminal motif [130,291,292], which in case of BRD4 is connected to the N-terminal part by an extensive, proline- and glutamine-rich intrinsically disordered region (IDR). The short isoform BRD4-S(a) (80 kDa) and the still poorly characterized BRD4-S(b) (88 kDa) [293,294] lack the C-terminal domain and most of the IDR. With 70%, 61% and 73% for BD1, BD2 and ET, respectively, the amino acid sequence is highly conserved between the paralogs (Eischer et al., 2022; accepted for publication).

Structural information is available for the BD1 domain of human BRD2 [295], BRD3 ([296]; PDB: 6qju; PDB: 6u4a), BRD4 [296,297] and the BD2 domain of BRD2 [298], BRD3 ([296]; PDB: 5A7C) and BRD4 (for example [299]. Each BD consists of a left-handed bundle of the long α-helices αZ, αA, αB and αC and the connecting loops ZA and BC, which form a deep hydrophobic cleft. Acetyl binding is primarily mediated by conserved tryptophan, tyrosine, asparagine and methionine residues in or near the loop regions [300]. A hydrophobic, electroneutral cavity surrounded by a collar of negatively charged residues is located in the center of the cleft [296]. The first 3D structure of an ET domain became available for BRD4 [301], followed by structures of the ET domains of BRD2 and BRD3. The structure of the ET domain comprises three α-helices α1, α2 and α3 that are connected by loop regions and form a hydrophobic core. α2 assumes an antiparallel orientation relative to α1 and α3, which results in a unique, stable conformation of the α2–α3 loop that might be involved in protein–protein interactions. 3D structural information of other parts of BET

proteins or entire BET proteins is currently not available, although α-helical regions upstream of BD1 and downstream of the ET domain of BRD2 and BRD3, as well as in the CTD region of BRD4 has been predicted [302] (AlphaFold: AF-P25440-F1, AF-Q15059-F1, AF-O60885-F1).

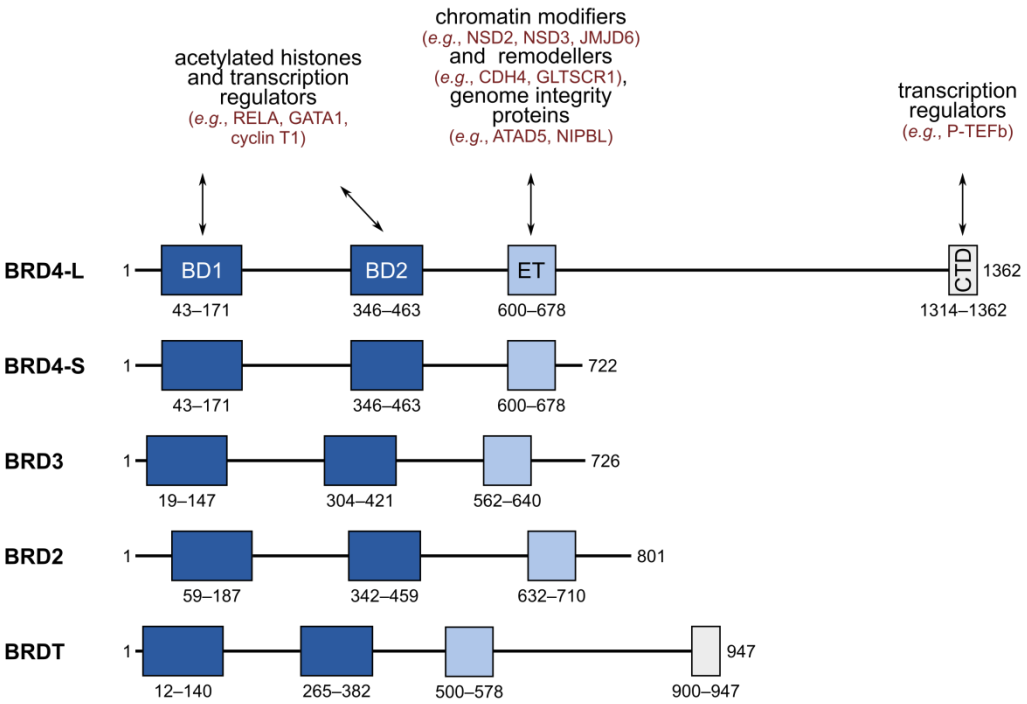


Figure 3: Illustration of the domain structure of BET proteins. The domains were annotated based on a sequence alignment done with Clustal O and the comparison of available structures, particularly the position of α-helices therein. Proteins were drawn in scale. A second short BRD4 isoform, BRD4-S(b), and BRD3-R, are not shown. A modified version of the figure was included in Eischer et al. (2022; accepted for publication).

The bromodomains bind to the ε-amino-acetyl groups of lysine residues of acetylated histones [285,296,303] and TFs [304–307]. Primarily histones H3 and H4, but to a lesser extent also H2B and H2A are bound [296,308]. Closely spaced acetylated lysines can cooperatively bind to the same bromodomain, thereby increasing the avidity of the interaction [309,310]. Among the TFs which in their acetylated form interact with BET proteins are GATA1 [307], RelA [304] and Twist [306]. BET bromodomains were also found to bind acetylated cyclin T1 [130,132,133]. The ET domain interacts with chromatin-modifying proteins, such as the histone lysine methyltransferase NSD3 and the arginine demethylase and lysine hydroxylase JMJD6 [311]. Moreover, interactions were also described with subunits of ATP-dependent chromatin remodelers, including proteins of the BAF and the NuRD complexes [311,312], and with the cohesin agonist NIPBL [313]. Lastly, the C-terminal domain of BRD4 and BRDT binds CDK9 [310] but not CDK7 or CDK8 [132].

2.2.2 BRD4 AS A REGULATOR OF POL II TRANSCRIPTION

Evidence for a role of BRD4 in transcription initiation is relatively sparse and at least in part overlaps with the evidence of its function mediated by enhancers [136], which are usually seen as regulators of transcription initiation at nearby genes. Additionally, enhancers resemble promoters in that they are sites of transcription, which is evident by enhancer RNA (eRNA) production [314]. However, enhancers are emerging as regulators also of early elongation [315,316].

BRD4 is a well-established co-factor of the Mediator complex [132,317] and co-localizes with it genome-wide at sites marked by H3K27ac, *i.e.*, at promoters and active enhancers [318–320]. JQ1 inhibitor treatment reduces Mediator binding to the chromatin in a locus-specific manner, while knockdown of MED12 or MED23 phenocopies the effect of BET inhibition on known target genes of BRD4 that only partially overlap with super-enhancers (SEs) [318]. Chromatin association of Mediator might depend on the phosphorylation state of BRD4 [321]. Of note, the kinase subunits of Mediator, CDK8 and CDK19, seem to counteract the activating function of BRD4 [322,323], which is in line with a repressive effect of CDK8 on the association of Mediator and Pol II. Whether BRD4 is required to recruit GTFs to the chromatin, is yet unknown [324,325]. Also, an increase in the initiation-associated CTD p-Ser5 was seen at the 5' gene end upon BRD4 degradation, while gene body occupancy was decreased [325].

Much of the early work on BRD4 was centered on its interactions with P-TEFb, implicating a function in early elongation. In human cells, P-TEFb is sequestered in an inactive complex with HEXIM1- and 7SK snRNA [326,327] or is present in an active form that is BRD4-bound [132,133]. c-MYC recruits P-TEFb to a small group of genes [57,131], but which factors recruit P-TEFb globally and in a sequence-independent manner remains unclear. P-TEFb activity is directly correlated with BRD4 levels *in vitro* and *in cellulo* and was suggested to depend on BRD4 for proper localization to the promoter-proximal region [132]. This fostered the idea that BRD4 is globally required to recruit P-TEFb and thereby promotes transcription elongation [134]. This is in line with the finding that the C-terminal domain of BRD4 can act as a transactivator and recruit P-TEFb to a GAL4 promoter of a reporter and induce luciferase expression [291]. A later proteome-wide analysis confirmed the association of BET proteins with P-TEFb and other proteins of the transcription-activating SEC complex, as well as the PAF complex, and also demonstrated the sensitivity of this association for the BD inhibitors that mimic an acetyl moiety [328].

Furthermore, the BD-specific inhibition leads to the dissociation of BRD4, accompanied by the loss of CDK9 and also Mediator, from TSS-distal regions and particularly enhancers [319]. Nevertheless, while these studies strengthened the molecular link between BRD4 and P-TEFb, they did not provide evidence for a requirement for BRD4 in recruiting P-TEFb, in part due to a lack of spatial or temporal resolution of the used methods.

Understanding of BET protein function was improved by the development of BET-specific proteolysis targetic chimera (PROTAC) degraders, *e.g.* dBET6. In contrast to JQ1 treatment which most strongly affects a small group of super-enhancers highly enriched in context-specific TFs, degrader treatment shows no obvious bias, but leads to an overall more pronounced reduction of the transcriptional output [135]. pan-BET degradation reduces transcribing Pol II throughout the transcribed region but especially at the gene body, indicating a defect in pause release [135,329,330]. Additionally, according to some studies, induced BET degradation reduces p-Ser2 levels [329] and phenocopies the effect of CDK9 inhibition by NVP-2 but does not decrease CDK9 or cyclin T1 occupancy, strongly arguing against an effect on P-TEFb recruitment [135]. Moreover, upon BRD4-selective degradation, no decrease in p-Ser2 is observed [331] or it is restricted to the gene body region [325]. Similarly, targeted degradation of neither BRD2, BRD3, nor BRD4 in DLD-1 cells results in the loss of CDK9 from the chromatin or affects p-Ser2 levels [137].

The current view is that BRD4 regulates CDK9 activity rather than its recruitment: it has been suggested that the BD2 of BRD4 binds acetylated cyclin T1 even when part of the inactive complex, while specifically the interaction of the C-terminal domain of BRD4 makes HEXIM1 and 7SK dissociate[130], possibly by inducing a conformational change of 7SK [332], or dependent on the proteolytic activity of the BRD4 interactor JMJD6 [333]. Moreover, a study found that JMJD6, which forms a complex with BRD4, destabilizes the inhibitory CDK9-containing complex by demethylating the 7SK cap structure and also $H4R3me2_{symmetrical}$, which is recognized by 7SK [334].

It was also proposed that BRD4 is an atypical kinase that can phosphorylate the serine 2 residue of the RPB1 CTD, but not serine 5 and serine 7 or GTFs, *in vitro* and *in vivo* [335]. In view of its substrate specificity and resistance to many commonly used wide-spectrum kinase inhibitors, the putative activity of BRD4 seems to be very distinct from that of other transcription-associated kinases, including CDK9 [335]. Kinase activity could be mapped to kinase subdomains localized in its N-terminus. Recently, BRD4 was shown to also phosphorylate and thereby destabilize c-MYC [336].

2.2.3 BRD4 AND ENHANCER FUNCTION

Enhancers are bidirectionally transcribed regulatory elements which, upon activation, affect the transcription of target genes and not least play a pivotal role in cell differentiation and disease. Enhancers act via different mechanisms, for instance, by facilitating the recruitment of tissue-specific TFs and GTFs [337] or by altering the chromatin structure at the target gene [338]. Also, enhancer-mediated effects on early and late elongation have been demonstrated for selected enhancer–promoter pairs [334,339,340].

The role of BRD4 in transcription regulation cannot be readily separated from its function at enhancers. BRD4 is enriched at enhancer and super-enhancer regions genome-wide [137,316,318,341], mainly due to binding of the bromodomains to H3K27ac-marked chromatin, or binding to TFs or eRNA [342]. At enhancers as well as promoters, BRD4 co-localizes with sequence-specific TFs in various cellular systems, suggesting a mechanism for the context-dependent recruitment of BRD4. In acute myeloid leukemia (AML), which is known to depend on BRD4, focal BRD4 binding in nucleosome-depleted regions correlates with binding of PU.1, ERG, MYB, C/EBPα or -β, etc. while the edges of the BRD4-bound region correlate with acetylated nucleosomes [305]. Overexpression of these TFs in an unrelated cellular system is sufficient to alter the binding pattern of BRD4 [305]. Mechanistically, TFs recruit BRD4 via the histone acetyltransferases p300 and CBP acting as co-activators, or directly [304,305,343,344]. Also, the histone methyltransferases KMT2C and KMT2D seem to be involved [324]. Overall, this provides an example of how acetylation-dependent and -independent modes of BRD4 recruitment can cooperate.

Most evidence suggests that BRD4 regulates enhancers on the transcription level. eRNA production correlates with BRD4 occupancy or activity [136,342,344]. For instance, JQ1 treatment and BRD4 knockout reduce eRNA production and the binding of Pol II but not of P-TEFb at enhancers [136]. Moreover, BET inhibition affects transcription of enhancer target genes [324,345]. Also, a—possibly not general—mechanism by which enhancers co-occupied by BRD4 and the arginine demethylase JMJD6 can release P-TEFb from its inhibitory complex and induce pause release at target genes has been proposed [334].

JQ1 treatment preferentially depletes BET proteins from super-enhancers, *i.e.*, clusters of enhancers highly enriched in Pol II, co-factors, master TFs and chromatin regulators, which often control disease-critical genes [346,347]. BRD4 levels at a subset of super-enhancers are particularly sensitive to JQ1 treatment [318,345], which through reduction of

MED1 and P-TEFb binding leads to a strong repressive effect on SE-regulated genes, such as *c-MYC* [319]. Notably, one study aiming at the prediction of BET inhibitor sensitivity could not confirm this SE-based model [325].

Whether BRD4 is involved in forming promoter–enhancer contacts is unclear. Based on knockdown experiments, BRD4 appears dispensable for looping between so-called anti-pause enhancers and promoters [334]. Similarly, short-term BET inhibition, long-term degradation and hexane-1,6-diol treatment, which dissolves condensates of BRD4 and MED1, disrupt transcription at known promoter–enhancer pairs but not the physical interaction [341]. This led to the suggestion that CTCF and cohesin rather than BET proteins stabilize short-range DNA–DNA contacts but are not sufficient for transcription, whereas the weak interactions of BRD4 mediate transcription activation but not DNA folding. Contrarily, a more recent study found changes in long-range interactions upon rapid BRD4-selective degradation, possibly due to a failure of loop extrusion [313]. This view appears to be in accordance with the co-localization of BRD4 with the cohesin loading factor NIPBL and the cohesin subunits RAD21, SMC1A and SMC3 at the chromatin [348].

2.3 METHODS TO DISSECT GLOBAL NASCENT TRANSCRIPTION

Dissecting transcription to uncover its mechanistic principles is technically challenging, as it requires specific readouts. For instance, total RNA-seq captures a combination of transcriptional and post-transcriptional effects and thus provides only an imperfect view of transcription dynamics. It is not capable of monitoring pausing, abortive transcription and co-transcriptional processing. Additionally, RNA secondary and tertiary structures but also covalent modifications and interactions with RNA-binding proteins affect RNA stability within the cell, resulting in a bias in RNA-seq. Differences in stability are especially relevant in the context of non-coding RNA species like promoter upstream transcripts (PROMTS) [176,349,350] or eRNA. Also, while Pol II transcribes a vaster number of genomic sites than Pol I and Pol III [351], the latter contribute by far more RNA, leading to dilution effects [352,353].

Besides, several sequencing-based assays exist to profile Pol II position and abundance, although not all of them rely on the capacity of Pol II to synthesize RNA. These include chromatin immunoprecipitation (ChIP)- and run-on-based approaches as well as NET-

seq. Additional methods, *e.g.*, permanganate footprinting, are not discussed in the following for the sake of conciseness.

ChIP relies on the antibody-based enrichment of the DNA that is bound by a protein of interest *in cellulo*. Combined either with microarrays or high-throughput sequencing, it can be used to profile transcribing Pol II genome-wide [354,355]. For instance, ChIP-on-chip provided early evidence for the local accumulation of Pol II at the 5' end also of inactive genes [356]. Depending on the availability of suitable antibodies, ChIP can be used to profile CTD phospho-forms [82,357] or to test for co-occupancy of, for instance, Pol II and modified histones on the same DNA molecule [358]. Spike-in-based normalization, which is the basis for quantitative comparisons between different conditions, became amenable with the development of ChIP-Rx [359].
The problem of limited spatial resolution was overcome by ChIP-exo [360] and ChIP-nexus [361], which offer near-single nucleotide resolution. Disadvantages that remain are the lack of strand-specificity, which precludes differentiating sense and antisense transcription from the same promoter, and the inability to distinguish transiently paused from stalled or backtracked Pol II.

In contrast, run-on methods require Pol II to resume transcription *in vitro* [47,362]. In brief, transcription is halted first by chilling the cells on ice and then by nuclei isolation, permeabilization and the removal of free nucleotides. Subsequently, NTPs labeled with 5-bromouridine 5'-triphosphate (brUTP) or biotin are added in a GRO-seq or PRO-seq experiment, respectively [47,362]. By removing pausing factors [363], treatment of the isolated chromatin with the anionic detergent sarkosyl enables Pol II to resume transcription. Incorporation of brUTP allows for the enrichment of the RNA molecule using antibodies, while biotin-11-NTP incorporation terminates RNA chain extension. Thus, while GRO-seq narrows down the localization of Pol II, PRO-seq provides single-nucleotide resolution. In contrast to ChIP-seq, run-on methods do not require formaldehyde crosslinking or depend on specific antibodies, and particularly PRO-seq provides a high signal-to-noise ratio[362]. Useful variants such as PRO-cap to map TSSs [362] or TV-PRO-seq, which allows for an estimation of pausing durations [364], have been developed. A disadvantage, however, is that transcription needs to be restarted *in vitro*, *i.e.*, under highly artificial and also error-prone conditions [365].

NET-seq, like run-on methods, enriches for elongation-competent Pol II but in addition also captures stalled polymerases. The native ternary complex composed of Pol II,

nascent RNA and DNA template is highly stable [366] and can be quantitatively purified using mild detergents and urea [59]. Cell fractionation is performed in the presence of the potent Pol II inhibitor α-amanitin, which prevents unintended run-on transcription and therefore is essential to maintaining single-nucleotide resolution [367]. Likewise, the library preparation aims to preserve the identity of the last nucleotide that was incorporated into each nascent RNA molecule in the cell. To enable quantitative comparisons between samples and to make the procedure robust against experimental variation, whole mouse or *Drosophila* cells can be spiked in prior to cell fractionation [368]. The library preparation procedure has recently been updated to increase robustness and decrease hands-on time [369] (Figure 4), while mispriming artifacts were reduced both experimentally and computationally [92].

In contrast to run-on methods, NET-seq can be combined with an IP step, allowing for profiling specifically of CTD-phosphorylated polymerases (mNET-seq) [222]. Variants of the NET-seq protocol have been successfully combined with long-read sequencing [284,368,370,371]. With NET-CAGE, a protocol to profile 5' ends of nascent transcripts is available [372]. As in case of run-on techniques, low mapping efficiency of reads shorter than 20 bp precludes the analysis of the earliest phase of transcription. Importantly, NET-seq but also PRO-seq can be used for *de novo* enhancer detection and activity analysis [373] (Bressin et al., 2022; manuscript in preparation).

A downside of this approach is that mature chromatin-associated RNAs, *e.g.*, XIST or spliceosomal RNAs, are usually also present in the libraries. However, different strategies have been employed to minimize these contaminations, such as the addition of the ionic detergent Empigen BB during cell fractionation [374] or subtractive hybridization of most abundant contaminants [375]. Also, a protocol that additionally involves a 4sU labeling step to increase library complexity but also deplete mature transcripts has been developed (Bressin et al., 2022; manuscript in preparation).

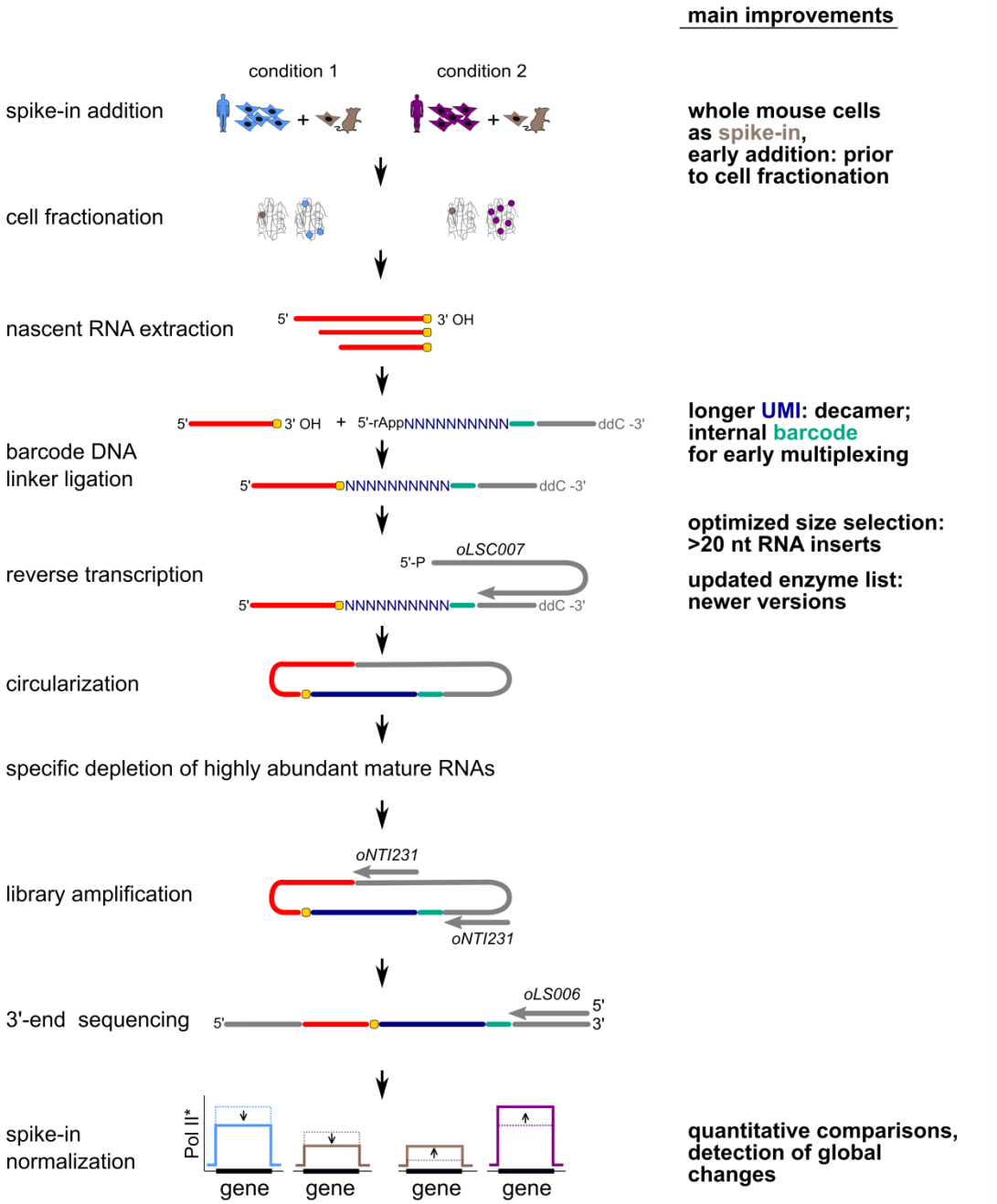


Figure 4: Schematic outline of SI-NET-seq. Changes as compared to the original protocol published by Mayer and Churchman [375] are indicated on the right side. The figure was adapted from Arnold et al. [368] and Jasnovidova et al. [369].

2.4 TARGETED DEGRADATION AS A STRATEGY TO PROBE PROTEIN FUNCTION

Depleting a specific protein and analyzing the resulting phenotype to infer the protein's function is a key idea in the field of functional genomics. For some time, this was achieved either by irreversibly mutating or deleting the coding region, for instance, using CRISPR/Cas9, or by RNA interference. Another promising approach, which in case of

BET proteins was a breakthrough, is to target the protein itself chemically. For instance, the cell-permeable small-molecule inhibitor JQ1 binds to the bromodomains present in all BET proteins, resulting in their partial dissociation from the chromatin within hours [319].

Yet another improvement was the development of small-molecule degraders named PROTACs that depend on the cellular ubiquitin–proteasome system. Generally, PROTACs are bifunctional molecules that non-covalently bind to both the target protein and a cullin–RING–ubiquitin ligase (CRL) complex, thereby enforcing spatial proximity[376]. To date, mainly the E3 ligase substrate recognition proteins cereblon (CRLN) [376,377] and VHL [378] as well as the plant-derived F-box protein OsTIR1 [379], which are both nucleoplasmic and cytoplasmic [380,381], have been utilized. Unlike inhibitors, degraders are not required to bind to an active site or interaction interface to be functional and therefore vastly increase the catalog of targetable proteins [382].

Targeted protein degradation occurs rapidly, resulting in a near-complete removal of the targeted protein within minutes to hours, which is crucial to avoid indirect or adaptive effects in downstream assays [135,368]. Additionally, essential and developmentally regulated proteins become amenable to depletion [383]. Targeted degradation is a holistic approach in that the target protein is not only blocked with regard to some of its functions but removed entirely. Because a degrader molecule can participate in multiple subsequent degradation cycles, relatively low concentrations, mostly in the nanomolar range, are sufficient [368]. Therefore, off-target toxicity can often be mitigated better than in case of inhibitors. Additionally, in contrast to genome editing strategies, the effect of PROTACs is reversible.

Binding to the target protein can either involve a naturally occurring or an engineered so-called degron, for instance, a dTAG [331] (Figure 5), AID tag [384], or Halo tag [385] integrated at the respective endogenous locus. While naturally occurring degrons are advantageous in cellular systems that cannot be easily manipulated genetically, engineered degrons offer flexibility in target choice. A potent pan-BET degrader that was also used in this study is dBET6. Biochemically, dBET6 was derived by combining JQ1 and the phthalimide thalidomide, which binds CRBN, in one molecule [376], followed by optimization of the linker to increase membrane-permeability, potency and efficacy [135]. Thus, dBET6 acts through the BET bromodomains and leads to the specific degradation of all BET proteins.

Given the structural similarity of the BET proteins, dissection of their individual functions requires the insertion of a degron tag. The dTAG, which was utilized to deplete both BRD4 isoforms in this study, consists of a mammalian FKBP12 protein in which phenylalanine 36 was replaced by valine to create a pocket that enables highly specific binding to the small molecule ligand AP1867 [386]. Consequently, a chimeric molecule of AP1903 and thalidomide, termed dTAG7, can induce degradation of any protein fused to $FKBP12^{F36V}$ but not wild-type FKBP12 [331].

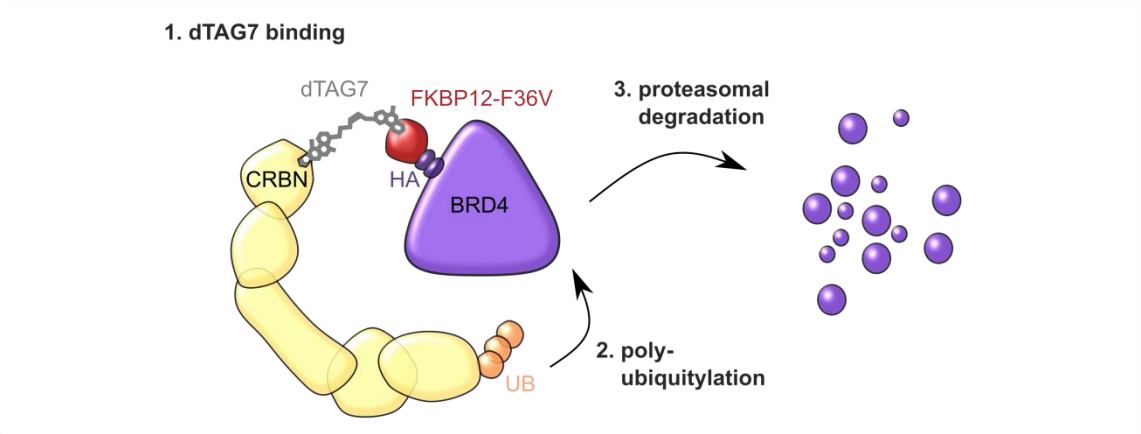


Figure 5: Schematic depiction of the BRD4-selective degradation strategy. Binding of the small molecule degrader dTAG7 to the dTAG, a mutated FKBP12 protein, which was inserted at the N-terminus of BRD4, and cereblon induces dimerization. Because cereblon is the substrate recognition subunit of an E3 ligase complex, dimerization stimulates polyubiquitylation of the dTAGged BRD4, followed by its degradation via the proteasome.

2.5 Motivation of this study

Several studies have reported interesting transcriptional phenotypes upon BET protein or BRD4 perturbation but could not pinpoint the molecular process that was disrupted. One possible explanation is that BRD4, the BET protein predominantly implicated in transcription, acts pleiotropically and, for instance, affects both transcription initiation and elongation. In light of its prominent protein–protein interaction domains [304,305,311] but also the intrinsically disordered C-terminal domain, which has been shown to phase-separate and thereby engage with proteins of similar biophysical properties [71], this appears very plausible. Moreover, widely used inhibitors like JQ1 affect not only BRD4 but a whole family of different, at least to some extent, functionally specialized BET protein species, which might result in a complex phenotype.

Another possible explanation, however, is that conventional perturbation strategies, for instance, extended inhibitor treatments, could elicit indirect or unspecific effects. Such effects, which often arise on a longer time scale, can obscure the actual phenotype caused by a treatment. Also, some standard readouts might lack sensitivity and thereby hamper the detection of direct effects.

In light of this, we defined our main questions as the following:

1. Is BRD4 the BET protein that predominantly regulates transcription?
2. What is the specific mechanistic role of BET proteins (or BRD4) in transcription elongation?
3. Does BRD4 have additional functions, for instance, in other phases of transcription or co-transcriptional processing?

To address these questions, we made use of recently developed methods that can overcome the challenges mentioned above: we rapidly and near-completely depleted the endogenous BET proteins or selectively BRD4 using PROTAC degraders. To determine the immediate consequences, nascent transcription was quantified at high temporal and spatial resolution using SI-NET-seq. Lastly, complementary approaches, such as mass spectrometry and ChIP-Rx, were employed to reach mechanistic conclusions.

3 MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 EQUIPMENT AND REAGENTS

Table 1: Instruments.

name	source
Bioruptor plus sonicator	Diagenode
centrifuge 5424S	Eppendorf
centrifuge 5810R	Eppendorf
Countess II cell counter	Thermo Fisher Scientific
Dionex Ultimate 3000 LC system	Thermo Fisher Scientific
E220evolution sonicator	Covaris
EVE cell counter	NanoEntek
Evos XL Core microscope	Thermo Fisher Scientific
Heracell vios 160i CO_2 incubator	Thermo Fisher Scientific
mini tube rotator	Boekel Scientific
Nanodrop One UV-Vis spectrometer	Thermo Fisher Scientific
NovaSeq 6000 sequencer	Illumina
Nucleofector 2b	Lonza
Odyssey CLx infragred imager	Licor
Q-Exactive HF Orbitrap mass spectrometer	Thermo Fisher Scientific
QuantStudio 7 thermocycler	Applied Biosystems
Qubit 3.0 fluorometer	Invitrogen
rotator with vortex	Neolab
S220 sonicator	Covaris
Safe Imager 2.0	Invitrogen
TapeStation 1450 electrophoresis device	Agilent
TC20 cell counter	Biorad
Trans-Blot mini electrophoretic transfer chamber	BioRad
vapo.protect Mastercycler pro S	Eppendorf
vortex K5250	IKA
XCell SureLock mini gel electrophoresis system	Invitrogen

Table 2: Degraders and inhibitors.

name	catalog number	source
dBET6	HY-112588	MedChemExpress
dTAG7	HY-123941	MedChemExpress
JQ1	HY-13030	MedChemExpress
α-amanitin	A2263	Sigma-Aldrich

Table 3: General chemicals, buffers and consumables.

name	source
“Complete” protease inhibitor cocktail	Roche
15% and 10% TBE–urea gels	Thermo Fisher Scientific
8% TBE gels	Thermo Fisher Scientific
AMPure XP beads	Beckman Coulter
benzonase, ≥90% purity	Merck Millipore
biotin-14-dATP	Jena Bioscience

DNA ladder, 20 bp	Takara
Dynabeads Protein G	Thermo Fisher Scientific
Dynabeads streptavidin C1	Thermo Fisher Scientific
GlycoBlue coprecipitant	Invitrogen
NEBNext Fragmentation Solution	NEB
Novex TBE–urea sample buffer (2X)	Thermo Fisher Scientific
NuPage 4–12% Bis-Tris acrylamide gel	Thermo Fisher Scientific
PhosStop phosphatase inhibitor cocktail	Roche
Pierce Protein A/G magnetic beads beads	Thermo Fisher Scientific
PowerUp SYBR green master mix	Thermo Fisher Scientific
Precision Plus Protein Kaleidoscope ladder	BioRad
Qiazol	Qiagen
Spin-X centrifuge tube filter, 0.22 µM, cellulose acetate	Corning
Superase-In RNase inhibitor	Thermo Fisher Scientific
SYBR Gold nucleic acid stain	Thermo Fisher Scientific
T20 blocking solution	Thermo Fisher Scientific
Trizol	Invitrogen

Table 4: Kits.

name	catalog number	source
ChIP DNA Clean & Concentrator	D5205	Zymo
DNA Clean & Concentrator-5	D4013	Zymo
DNeasy Blood & Tissue kit	69504	Qiagen
Gel and PCR Clean-up kit	740609.250	Machery-Nagel
miRNeasy micro	217084	Qiagen
miRNeasy mini	217004	Qiagen
NEBNext Multiplex Oligos for Illumina	E7645S	NEB
NEBNext Ultra II DNA Library Prep Kit for Illumina	E7645S	NEB
NucleoBond Xtra midi EF kit	740420.50	Machery-Nagel
Nucleofector Kit V	VCA-1003	Lonza
NucleoSpin Plasmid mini kit	740588.250	Machery-Nagel
ProteoExtract Protein Precipitation kit	539180	Merck Millipore
Qubit DNA HS kit	Q32851	
RNA Clean & Concentrator-5	R1016	Zymo
SILAC Protein Quantitation Kit (Trypsin), RPMI 1640	A33973	Thermo Fisher Scientific
Tagment DNA Enzyme and Buffer kit	20034197	Illumina
TapeStation DNA D1000 kit	5067-5582/3	Agilent
ZR small-RNA PAGE Recovery kit	R1070	Zymo

Table 5: Enzymes.

name	catalog number	source
CircLigase	CL4111K	Lucigen
Klenow fragment	M0210S	NEB
MboI	R0147M	NEB
Phusion High Fidelity polymerase	M0530S	NEB
Proteinase K	3719.2	Carl Roth
RNase A	10753721	Thermo Fisher Scientific
SuperScript III reverse transcriptase	18080044	Invitrogen
T4 DNA ligase	M0202L	NEB
Tn5 transposase	20034197	Illumina
truncated T4 RNA ligase 2	M0239S	NEB

3.1.2 BUFFERS

Table 6: Buffers and solutions for immunoblotting.

SDS loading buffer (2X)	100 mM Tris-HCl (pH 7.0), 4% (w/v) SDS, 20% (v/v) glycerol, spatula tip Orange G in H_2O;
TBS–T (1X)	20 mM Tris, 150 mM NaCl, 1% (v/v) Triton X-100 in H_2O;
blocking buffer	5% (w/v) in TBS-T;
transfer buffer	25 Mm Tris base, 190 mM glycine, 20% (v/v) methanol in H_2O;

Table 7: Buffers for gDNA extraction.

Bradley lysis buffer	10 mM Tris-HCl, 10 mM EDTA, 0.5 % SDS, 10 mM NaCl, 1 mg/ml Proteinase K in H_2O;

Table 8: Buffer and solutions for SI-NET-seq.

Cytoplasmic lysis buffer	0.15% (v/v) NP-40, 10 mM Tiris-HCl (pH 7.0), 150 mM NaCl, 25 µM α-amanitin, 10 U Superase-In, 1X protease inhibitor mix in H_2O;
sucrose buffer	10 mM Tris–HCl (pH 7.0), 150 mM NaCl, 25% (w/v) sucrose, 25 µM α-amanitin, 20 U Superase-In, 1X protease inhibitor mix in H_2O;
nuclei wash buffer	0.1% (v/v) Triton X-100, 1 mM EDTA, 25 µM α-amanitin, 40 U Superase-In, 1X protease inhibitor mix in PBS;
glycerol buffer	20 mM Tris–HCl (pH 8.0), 75 mM NaCl, 0.5 mM EDTA, 50% (v/v) glycerol, 0.85 mM DTT, 25 µM α-amanitin, 10 U Superase-In, 1X protease inhibitor mix in H_2O;
nuclei lysis buffer	1% (v/v) NP-40, 20 mM HEPES (pH 7.5), 300 mM NaCl, 1 M urea, 0.2 mM EDTA, 1 mM DTT, 25 µM α-amanitin, 10 U Superase-In, 1X protease inhibitor mix in H_2O;
depletion oligo mix	200 µM of each biotinylated oligo in 10 mM Tris-HCl (pH 8.0) H_2O;
equilibration buffer (2X)	5 mM Tris-HCl (pH 7.0), 2 M NaCl, 1 mM EDTA, 0.2% (vol/vol) Triton X-100 in H_2O;
soaking buffer	10 mM Tris-HCl (pH 8.0), 300 mM NaCl, 0.97 mM EDTA in H_2O;
DNA loading buffer (6X)	6 g sucrose, 30 mg Orange G dissolved in nuclease-free H_2O, increased volume to 25 ml;

Table 9: Buffers and solution used for ChIP-RX.

blocking buffer	0.5% BSA fraction V in PBS;
lysis buffer I	50 mM HEPES, 85 mM KCl, 0.5% (v/v) NP-40, 1X protease inhibitor mix, 1X phosphatase inhibitor mix in H_2O;
lysis buffer II	10 mM Tris (pH 8.0), 150 mM NaCl, 1 mM EDTA, 1% (v/v) NP-40, 0.1% (w/v) SDS, 0.1% (w/v) sodium deoxycholate, 1X protease inhibitor mix, 1X phosphatase inhibitor mix in H_2O;
washing buffer I	20 mM Tris (pH 8.0), 150 mM NaCl, 2 mM EDTA, 1% (v/v) Triton X-100, 0.1% (v/v) SDS, 1X protease inhibitor mix, 1X phosphatase inhibitor mix in H_2O;
washing buffer II	20 mM Tris (pH 8.0), 500 mM NaCl, 2 mM EDTA, 1% (v/v) Triton X-100, 0.1% (v/v) SDS, 1X protease inhibitor mix, 1X phosphatase inhibitor mix in H_2O;
washing buffer III	10 mM Tris (pH 8.0), 250 mM LiCl, 1 mM EDTA, 1% (v/v) NP-40, 1% (v/v) sodium deoxycholate in H_2O;
TE	1 mM EDTA, 10 mM Tris (pH 8.0) in H_2O;
elution buffer	100 mM $NaHCO_3$, 1% (w/v) SDS in H_2O;

Table 10: Buffers and solutions used for ATAC-seq.

digitonin stock solution (100X ≙ 1% (w/v))	11.2 mg powder dissolved in 560 µl DMSO, diluted with 560 µl nuclease-free H_20;
resuspension buffer (RSB; 1X)	10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$ in nuclease-free H_2O;
RSB–TND	1X RSB, 0.1% (v/v) NP-40, 0.1% Tween-20 (v/v), 0.01% (w/v) digitonin;
RSB–T	1X RSB, 0.1% Tween-20;

Table 11: Buffers and solutions for Hi-ChIP.

Hi-C lysis buffer	10 mM Tris (pH 8.0), 10 mM NaCl, 0..2% (v/v) NP-40, 1X protease inhibitor mix in H_2O;
sonication buffer	50 mM HEPES (pH 7.5), 140 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% (v/v) Triton X-100, 0.1% (w/v) sodium deoxycholate, 0.1% (v/v) SDS, 1X protease inhibitor mix in H_2O;
sonication buffer, high salt	50 mM HEPES (pH 7.5), 500 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% (v/v) Triton X-100, 0.1% (w/v) sodium deoxycholate, 0.1% (v/v) SDS, 1X protease inhibitor mix in H_2O;
LiCl wash buffer	20 mM Tris (pH 8.0), 1 mM EDTA, 250 mM LiCl, 0.5% (v/v) NP-40, 0.5% (w/v) Na deoxycholate, 1X protease inhibitor mix in H_2O;

TE with NaCl	10 mM Tris (pH 8.0), 1 mM EDTA, 50 mM NaCl in H_2O;
elution buffer	50 mM Tris (pH 8.0), 10 mM EDTA, 1% (w/v) SDS in H_2O;
Tween-20 wash buffer	5 mM Tris (pH 7.5), 0.5 mM EDTA, 1 M NaCl, 0.05% (v/v) Tween-20 in H_2O;
2X biotin binding buffer	10 mM Tris (pH 7.5), 1 mM EDTA, 2 M NaCl in H_2O;

Table 12: Buffer for whole cell proteome analysis by mass spectrometry.

guanidine hydrochloride buffer	8 M guanidine hydrochloride, 50 mM Tris-HCl (pH 8.0), 1X protease inhibitor mix in H_2O;

Table 13: Buffers for native immunoprecipitation.

native IP buffer A	20 mM Tris HCl (pH 8.0), 50 mM NaCl, 0.5% (v/v) NP-40, 10% (v/v) glycerol; 2X protease inhibitor mix; 2X phosphatase inhibitor mix in H_2O;
native IP buffer B	20 mM Tris HCl (pH 8.0), 125 mM NaCl, 0.5% (v/v) NP-40, 10% (v/v) glycerol; 2X protease inhibitor mix; 2X phosphatase inhibitor mix in H2O;

Table 14: Buffers for immunoprecipitation with formaldehyde crosslinking and MS.

LB1	50 mM HEPES, 1 mM EDTA, 140 mM NaCl, 0.25% (v/v) Triton X-100, 0.5% NP-40, 10% (v/v) glycerol, protease inibitor mix in H_2O;
LB2	10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 200 mM NaCl, 0.5 mM EGTA, protease inibitor mix in H_2O;
LB3	10 mM Tris-HCl (pH 8.0), 1 mM EDTA, 100 mM NaCl, 0.5 mM EGTA, 0.1% (w/v) sodium deoxycholate, 0.5% (v/v) N-lauroylsarcosine in H_2O;
modified RIPA	50 mM HEPES (pH 7.6), 0.5% (v/v) NP-40, 0.4% (w/v) sodium deoxycholate, 1 mM EDTA, 300 mM LiCl, protease inibitor mix in H_2O;

3.1.3 ANTIBODIES

Table 15: Antibodies for immunoblotting.

target protein	manufacturer, catalog number
α-tubulin	Abcam, ab18251
BRD2	Bethyl, A302-583A
CASP3	Cell Signaling Technology, 9662S
GAPDH	Ambion, 326548
H2B	Santa Cruz Biotechnology, sc-10808

Table 16: Antibodies for ChIP-Rx or HiChIP.

target protein	manufacturer, catalog number, clone	amount per sample
BRD4	Bethyl, A301-985A50	8 µg
CPF30	Bethyl, A301-584A	10 µg
CPSF73	Bethyl, A301-091AM	10 µg
CstF64	Bethyl, A301-092A	10 µg
CstF77	Sigma, C0249	25 µg
CstF77	Santa Cruz Biotechnology, sc-376575	10 µg
FIP1	Bethyl, A301-461A	10 µg
H3K27ac	Abcam, ab4729	5 µg
PAF1	Abcam, ab20662	10 µg
PRPF8	Bethyl, A303-921A	10 µg
RPB1 p-Ser2	Active Motif, 61083, 3E10	8 µg
RPB1 p-Thr4	Active Motif, 61361, 6D7	20 µg
RPB1 p-Tyr1	Active Motif, 61383, 3D12	20 µg
RPB2	GeneTex, TX102535	5 µg
SF3B1	Bethyl, A300-996A	6 µg
SPT5	Santa Cruz Biotechnology, Sc-133217	12 µg
SPT6	Novus Biologicals, NB100-2582	10 µg

Table 17: Antibodies for immunoprecipitation

target protein	manufacturer, catalog number	amount per sample
BRD2	Bethyl, A300-204A	10 µg
BRD3	abcam, ab264420	2 µg
BRD3	Bethyl, A300-204A	2 µg
BRD4	Bethyl, A301-985A100	10 µg
CDC73	Bethyl, A300-170A	10 µg
GFP	Chromotek, gtd-20 (bead-coupled antibody)	25 µl
goat IgG control	Thermo Fisher Scientific, 02-6202	
hemagglutinin (HA) tag	Cell Signaling Technology, 3724, C29F4	0.67 µg
PAF1	abcam, ab20662	10 µg
PRPF8	Bethyl, A303-921A	10 µg
rabbit IgG control	Bethyl, A120-101A	
RTF1	Bethyl, A300-179A	10 µg
SF3B1	Bethyl, A300-996A	2 µg
SPT5	Santa Cruz Biotechnology, Sc-133217X	10 µg
SPT6	Novus Biologicals, NB100-2582	2 µg

3.1.4 OLIGONUCLEOTIDES

Table 18: Oligonucleotide sequences.

oligo name	sequence	task
tag-N-gF3	GACCTTACCCCTACGACGTG	genotyping
BRD4-F-new	CTTGGAGACCACAGCCAGAG	
BRD4-R-new	TTGAGCACCACTCTGAGCAG	
CMIP-F	GGCTGTGTTCAGACCATTCTTAGG	
CMIP-R	AGTCCTTAACCAGAATCTCAACCC	
MBNL1_gene_F1	AGCATGTTGGAACGTCTAGGGA	RT-qPCR-based cleavage assay
MBNL1_gene_R1	CCTGGCAAACTACAAGCTGTCCA	
MBNL1_gene_F2	AAGACGTGGTGAGGGAGGGA	
MBNL1_gene_R2	AAGAATGGTGCCCTTTGACTGA	
MBNL1_pA_F1	TTGTCTTGGGCTGCAATTTGTTTTATG	
MBNL1_pA_R1	GCAGTTGCAATCAGTTAAAACCAGGT	
MBNL1_pA_F2	GAAATTGCCCTAATAAAACTTCTCTTTCTTAAG	
MBNL1_pA_R2	CTCCCTCTCAGTGTCAAAGGATAATTAG	
MBNL1_downstream_F1	CCAACTGTTCTGGGTGCCACT	
MBNL1_downstream_R1	TCCCATTCTGTGAGCTGAGGACT	
MBNL1_downstream_F2	ACCTGGAATCACGTCGCAGC	
MBNL1_downstream_R2	CCCGACAATTCCAGTTTGCAGC	
PVT1_gene_F1	GTCCCCGTGTGATGTCCCCT	
PVT1_gene_R1	CCTCTGGCTCCTCTGTTGGGT	
PVT1_gene_F2	GCAGCATTCACGCCTGGGTT	
PVT1_gene_R2	TGGAGGAGGGAAGGGAATGGG	
PVT1_pA_F1	TCATCTCTGCCCCAACCACTCA	
PVT1_pA_R1	ACCCCAAATTTGTTACTGCCCA	
PVT1_pA_F2	CAGCAAGCACCTGTTACCTGT	
PVT1_pA_R2	CCCAAATTTGTTACTGCCCAAGG	
PVT1_downstream_F1	TAGACGCTCACCCTGCCTGT	
PVT1_downstream_R1	GGCCTCCCCACTGTGTGTCT	
PVT1_downstream_F2	TGTGCTGGGATGCTGACCTCT	
PVT1_downstream_R2	AGGGGAAAGTGGGTGGCAGT	
CDC42_gene_F1	CCTTGGGTGGTAAACTTGTGGCT	
CDC42_gene_R1	ACAGTGTGACTCAGCATTGGTGG	
CDC42_gene_F2	TGGAGCTGTGTTCTGGGACT	
CDC42_gene_R2	TCACCTAAATCTGAAGCCACAAGT	
CDC42_pA_F1	GTGTTGAACCAATGCTTTCTCATGT	
CDC42_pA_R1	TCTGAAAAGTAAGAGACGGAGGGA	
CDC42_downstream_F1	GGTTCTGTGCTGCGAATCTGA	
CDC42_downstream_R1	TGAGGCCACAGTGTACCAACCT	
CDC42_downstream_F2	TCATGTGATACCCTGGAAAGCAG	
CDC42_downstream_R2	AGGACCTATACGCCCCTGACT	
Decamer barcode DNA linker	5‘-rApp/$(N)_{10}$CTGTAGGCACCATCAAT/3‘-ddC	SI-NET-seq
reverse transcription primer oLSC007	5Phos/ATCTCGTATGCCGTCTTCTGCTTG/iSp18/CA CTCA/iSp18/TCCGACGATCATTGATGGTGCCTACAG	
indexed forward primer	AATGATACGGCGACCACCGAGATCTACAC GATCGGAAGAGCACACGTCTGAACTCCAGTCAC (Illumina TruSeq index) TCCGACGAT CATTGATGG	
universal reverse primer oNTI23	CAAGCAGAAGACGGCATACGA	
sequencing primer	TCCGACGATCATTGATGGTGCCTACAG	

oLSC006		
Ad1_505	AATGATACGGCGACCACCGAGATCTACACGTAAGGA GTCGTCGGCAGCGTCAGATGTG	ATAC-seq
Ad2_705	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCT CGTGGGCTCGGAGATGT	
Ad2_706	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCT CGTGGGCTCGGAGATGT	
Ad2_707	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCT CGTGGGCTCGGAGATGT	
Ad2_708	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTC GTGGGCTCGGAGATGT	
Nextera_504	AATGATACGGCGACCACCGAGATCTACACAGAGTAGATCGTC GGCAGCGTC	Hi-ChIP
Nextera_701	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGG GCTCGG	
Nextera_702	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGG GCTCGG	
Nextera_703	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGG CTCGG	
Nextera_704	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGG GCTCGG	
Nextera_705	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGG GCTCGG	
Nextera_706	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGG GCTCGG	

3.1.5 CELL LINES AND BACTERIA STRAINS

Table 19: Cell lines and bacteria strains.

name	identifier or reference	source
K562	CCL-243	ATCC
K562 dTAG-BRD4	Arnold et al., 2021 [368]	this study
NIH 3T3	CRL-1658	ATCC
MOLT4	CRL-1582	ATCC
HeLa S3	CCL-2.2	ATCC
HeLa 8880 LAP-CstF64	Poser et al., 2008 [387]	lab of Antony Hyman
Stellar chemically competent *E. coli*	636766	Takara

3.1.6 VECTORS

Table 20: Plasmid vectors for BRD4 tagging.

name	reference	source
pX330A-sgPITCh-sgBRD4	Nabet et al., 2008 [331]	lab of Georg Winter; based on Addgene plasmid no. 91794
pCRIS-PITChv2-dTAG-BRD4-puroR	Nabet et al., 2008 [331]	lab of Georg Winter; based on Addgene plasmid no. 91793
pCRIS-PITChv2-dTAG-BRD4-blastiR	Nabet et al., 2008 [331]	lab of Georg Winter; based on Addgene plasmid no. 91792

3.1.7 SOFTWARE, INTERNET RESOURCES AND EXTERNAL DATA SETS

Table 21: Data bases, software and packages.

name	reference	source
Bedtools v2.30.0	Quinlan et al., 2010 [388]	http://code.google.com/p/bedtools
Biomart	Smedley et al., 2009 [389]	https://www.ensembl.org/biomart/martview/
Bowtie2 v2.3.5.1, v2.4.5	Langmead et al., 2012 [390]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml http://software.broadinstitute.org/
Cancer Dependency Map	Ghandi et al., 2019 [391]	https://depmap.org/
Clustal O. v1.2.4	Sievers et al., 2011 [392]	https://www.ebi.ac.uk/Tools/msa/clustalo/
CRISPOR v4.4	Concordet and Haeussler, 2018 [393]	http://crispor.tefor.net
Cutadapt v2.4, v3.5	Martin, 2011 [394]	https://github.com/marcelm/cutadapt
deeptools v 3.5.1	Ramírez et al., 2016 [395]	http://deeptools.ie-freiburg.mpg.de
Excel 2019	Microsoft	
FastQC v0.11.9	Braham Institute	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Galaxy Europe v21.09	Afgan et al., 2018 [396]	https://usegalaxy.eu/
HiCcompare	Stansfield et al., 2018 [397]	https://github.com/dozmorovlab/HiCcompare
HiC-pro	Servant et al., 2015	https://github.com/nservant/HiC-Pro
HUGO Gene Nomenclature Committee database	Tweedie et al., 2020 [398]	https://www.genenames.org
IGV v.2.8.2, v.2.10.5	Robinson et al., 2011 [399]	https://software.broadinstitute.org/software/igv/home
Image Studio Lite v5.2	Licor	https://www.licor.com/bio/image-studio-lite/
MACS2 v2.1.2,	Zhang et al., 2008 [400]	https://pypi.org/project/MACS2/
MaxQuant v1.6.0.1	Cox and Mann, 2008 [401]	https://www.maxquant.org
NCBI Primer BLAST	Ye et al., 2012 [402]	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
PANTHER v15	Mi et al., 2019 [403]	http://www.pantherdb.org
Perseus v.1.6.2.3	Tyanova et al., 2016 [404]	https://www.maxquant.org/perseus/
Picard v.2.18.2	Broad Institute	http://broadinstitute.github.io/picard/
PolyA_DB v3.2	Wang et al., 2018 [405]	https://exon.apps.wistar.org/PolyA_DB/
Primer3	Untergasser et al., 2012 [406]	https://primer3.ut.ee/
Prism v9	Graphpad	
STRING v11.0	Szklarczyk et al., 2019 [407]	https://string-db.org/

Table 22: Publically available data.

identifier	data type	reference
ENCFF080ODW	ChIP-seq, EP300, K562	Michael Snyder, Stanford
ENCFF966BXU	ChIP-seq, C/EBP-β, K562	Michael Snyder, Stanford
ENCFF833DKE	ChIP-seq, GATA1, K562	Michael Snyder, Stanford
ENCFF793CDO	ChIP-seq, PU.1, K562	Richard Myers, HAIB
ENCFF609FYE	ChIP-seq, ETO2, K562	Michael Snyder, Stanford
ENCFF604SPG	ChIP-seq, RUNX1, K562	Richard Myers, HAIB

GSE152291	transcribed enhancers, K562	Lidschreiber et al., 2021 [408]
GSM2090427	NET-seq, JQ1, MOLT4	Winter et al., 2017 [135]
GSM2090429	NET-seq, BRD4, MOLT5	Winter et al., 2017 [135]
GSM2090425	NET-seq, DMSO, MOLT4	Winter et al., 2017 [135]
FF10824-111C5, FF10826-111C7	CAGE	Fantom5

3.2 METHODS

3.2.1 CELL CULTURE

All cells were cultured at 37°C and 5% CO_2. K562 and K562 dTAG7-BRD4 cells, which have erythroleukemic characteristics, were cultured in RPMI medium containing 10% fetal bovine serum (FBS; Biochrom FBS Superior) and 5% penicillin–streptomycin. K562 cells were counted every second day using an automated cell counter, centrifuged at 200 x g and re-seeded at a density of 5 * 10^5 cells/ml. HeLa cells were grown in DMEM medium containing 10% fetal bovine serum (FBS; Biochrom FBS Superior) and 5% penicillin-streptomycin and 1:6-diluted every two days. NIH3T3 cells, which were solely used for spike-in in this project, were grown in DMEM medium containing 10% FBS (HyClone Cosmic Calf Serum, GE Healthcare, or Iron-fortified Bovine Calf Serum, Sigma) and 5% penicillin–streptomycin. NIH3T3 cells were trypsinized, counted and plated at 2 * 10^5 cells/T75 flask every second day. For experiments, cells were usually seeded in the same way 48 h in advance.
Cells were kept in culture for no longer than four weeks. Mycoplasma tests were performed regularly.

3.2.2 CRISPR/CAS9 GENOME EDITING FOR DEGRON TAG INSERTION

3.2.2.1 DESIGN

Selective BRD4 degradation was achieved using the dTAG system [331]. The dTAG-BRD4-expressing K562 cell line was generated using CRISPR–Cas9-based genome editing following the CRIS/PITCh protocol [409], which exploits the microhomology-mediated end joining (MMEJ) DNA repair pathway. The pX330A-sgPITCh-sgBRD4 plasmid encoding a generic guide RNA for targeting the donor plasmid (sgPITCh), a BRD4-specific guide RNA (sgBRD4) as well as humanized *Streptococcus pyogenes* Cas9, and the donor plasmids pCRIS-PITChv2-dTAG-BRD4-puroR and pCRIS-PITChv2-dTAG-BRD4-blastiR

were provided by the lab Georg Winter (CeMM, Vienna) (Table 20). A schematic depiction of the genetic elements on the plasmids and the tagging strategy is summarized in Figure 6. To be able to deplete both protein isoforms, BRD4-L and BRD4-S, the degron tag was inserted at the N-terminus of the endogenous BRD4 locus. The N-terminus was targeted 13 bp downstream of the start codon by sgBRD4 of the sequence ATGTCTGCGGAGAGCGGCCCTGG. The puroR/blastiR-P2A-2HA-FKBP12^{F36V} cassette was flanked by a 5' microhomology arm of the sequence AGCATGTCTGCGGAGAGCGGC and a 3' microhomology arm of the sequence GGCCCTGGGACGAGATTGAGAA, which correspond to the 20 bases immediately upstream and downstream of the Cas9 cut site and additional bases to adjust the BRD4 reading frame.

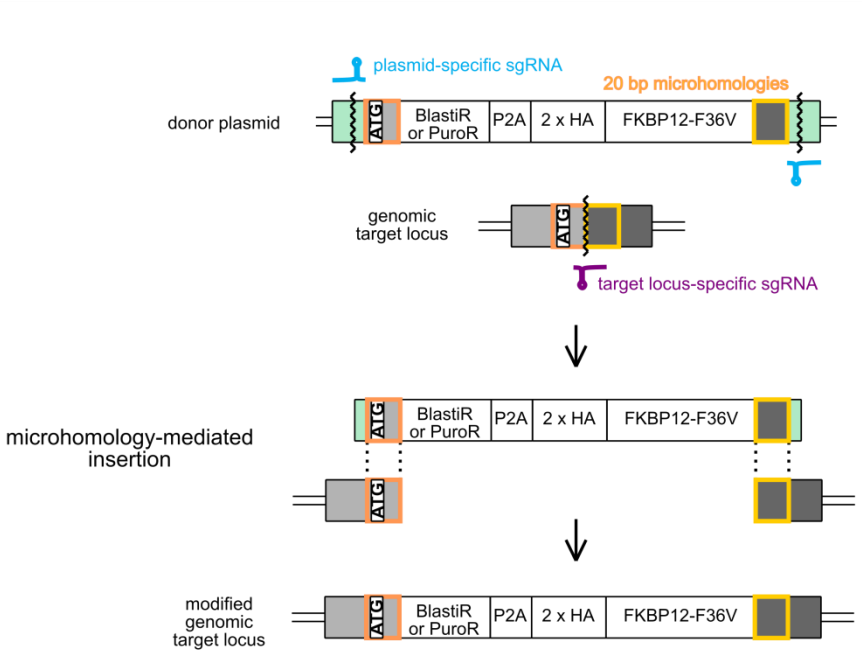


Figure 6: CRIS/PITCh (v2) strategy to insert the dTAG at the endogenous *BRD4* locus. The knock-in requires a donor plasmid encoding the construct to be integrated, which is flanked by short microhomology arms corresponding to the 20 bp upstream and downstream of the intended integration site. Additionally, sgPITCh to linearize the construct and a target gene-specific sgRNA—sgBRD4 in this case—are needed. The dTAG is accompanied by two copies of hemagglutinin to facilitate its antibody-based detection. A modified version of the scheme was also used in Arnold et al. [368].

Cleavage efficiency and specificity were predicted using CRISPOR [393]. Details of potential off-targets are summarized in Table 23. Given the high expression level of CMIP in myeloid cells and the unfavorable mismatch localization, the absence of mutations in the generated cell line was confirmed by PCR (see 3.2.2.3).

Table 23: Off-target analysis of the sgRNA used to integrate the degron tag at the *BRD4* locus.

off-target sequence	mismatch position	MIT off-target score	CFD Offtarget score	locus
AAGTCTCTGGAGAGCAGCCCAGG	.*....**.....*....	0.1766	0.5090	intron: GSAP
CTGTCAGCGGAGAGCAGTCCTGG	*....*.....*.*..	0.0323	0.4776	intergenic: RP11-286N3.1-EPB41L3
ATGGCTGAGAAGAGCAGCCCAGG	...*...*.*.....*....	0.2381	0.3792	intergenic: SMIM2-AS1-RP11-478K15.7/MIR8079
ATGTTTGTGGAGGGCAGCCCTGG	....*..*....*..*....	0.0952	0.3631	intergenic: RP11-875H7.2-WNT5A
AGGTCTGCGGGGAGCAGTCCGGG	.*.....*....*.*..	0.0296	0.3510	intron: ASPG
AGGTCTGAGGAAGGCGGCCCAGG	.*.....*...**.....	0.2724	0.3323	intergenic: RP11-756K15.2-RP11-246K15.1
CTGCCCGCGGAGGGCGGCCCCGG	*..*.*.....*.....	0.3519	0.3288	intron: ITPR3
GTGCTTGCTGAGAGCGGCCCTGG	*..**...*.....	0.8340	0.2196	intron: FAM53A
AGCACTGCAGAGAGCGGCCCCGG	.***....*.....	0.8223	0.0989	exon: CMIP
AGGGCTGGGGAGAGCGGCCTGGG	.*.*...*.....*	0.6974	0.10125	intergenic: RP11-379F4.4-RP11-379F4.8
GAGCCTGCGGAGCGCGGCCCTGG	**.*.....*.....	0.5817	0.1153	intergenic: CTD-2071N1.1-RP11-114H24.7
GTGTCTCAGGAGTGCGGCCCAGG	*.....**....*.....	0.3973	0.1341	intergenic: AL138885.1-TFAP2A
GAGTCTCCGGAGTGCGGCCCTGG	**....*.....*.....	0.3973	0.1745	exon: TSTD1/RP11-544M22.13
ATCTCTGGGGGAAGCGGCCCAGG	..*....*..**.....	0.3852	0.1500	intergenic: LINC00299-AC011747.3
GTGGCTGCGGATGGCGGCCCGGG	*..*.....**.....	0.2862	0.1568	exon: FRS3

3.2.2.2 CELL LINE GENERATION

24 h prior to transfection, K562 cells were seeded at a concentration of 2 * 10^5 cells/ml. The transfection was performed using the Nucleofector 2b device and the Amaxa Cell Line Nucleofector kit V (Lonza) following the cell line-specific recommendations of the manufacturer. In brief, 1 * 10^6 cells in 100 µl Nucleofection solution were combined with 2 µg of pX330A-sgPITCh-sgBRD4 and 1 µg of each pCRIS-PITChv2-dTAG-BRD4-puroR and pCRIS-PITChv2-dTAG-BRD4-blastiR and transfected using the program T-16. The cells were then transferred into complete RPMI medium and cultured for five days before double selection with 0.5 µg/ml puromycine and 2.5 µg/ml blasticidin S was started. To obtain a monoclonal cell line, single antibiotic-resistant cells were FACS-sorted into the wells of a 96-well plate and expanded.

3.2.2.3 CELL LINE VALIDATION

Genomic DNA (gDNA) was extracted from cells growing in the 96-well format. Therefore, cells were collected by brief centrifugation at 400 x g, followed by a PBS wash and lysis

in 50 µl Bradley lysis buffer (Table 7). After the addition of the buffer, the plate was sealed with parafilm and incubated overnight at 55°C in a humidified chamber. Next, the gDNA was precipitated by adding 150 µl of freshly prepared, ice-cold 75 mM NaCl in 100% ethanol, incubating at room temperature for 30 min and centrifugation at 3,000 x g and 4°C for 30 min. After decanting the supernatant, the DNA pellets were washed with ice-cold 70% ethanol, air-dried and re-dissolved in 200 µl nuclease-free H_2O. The DNA concentration of exemplary samples was measured by a Nanodrop spectrometer.

Because *Taq* polymerase frequently provided false-negative results, possibly because of the crude gDNA preparation, PCR genotyping was performed using Q5 DNA polymerase. Per reaction of 20 µl, 0.5 µl gDNA, 0.4 U Q5 Polymerase, 4 µl Q5 GC enhancer and primers at a final concentration of 0.5 µM were used. The following cycling scheme was found most effective: 98°C for 30 s; 30 cycles of 98°C for 10 s, 66°C for 20 s and 72°C for 1 min; 72°C for 2 min; 4°C. Combining the primers tag-N-gF3 and BRD4-R-new resulted in an amplicon of 637 bp if the tag was correctly inserted (Table 18). Zygosity of the insertion was tested by the primers BRD4-F-new and BRD4-R-new: a single product of 1,437 bp indicated homozygosity, while a product of 339 bp was seen in heterozygous and wild-type clones.

To assess the integrity of the flanking genomic regions and, in particular, determine whether InDels were generated, the 1,437 bp product was subjected to amplicon sequencing. This revealed a 10 bp deletion 2 bp upstream of the original BRD4 start codon (Figure 7). No sequence abnormalities were seen within the sequence of the construct.

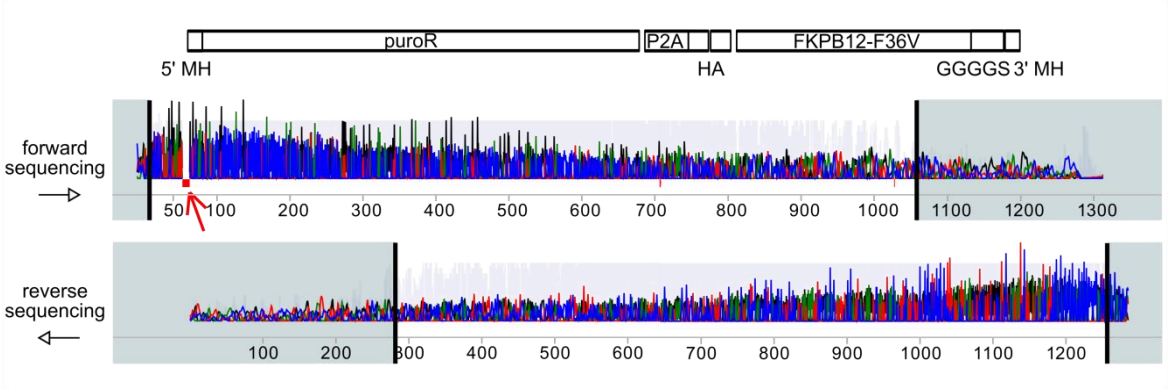


Figure 7: Electropherogram and sequence alignment of the N-terminal degron construct inserted at the *BRD4* locus in the monoclonal K562 dTAG-BRD4 cell line. The 1,437 bp product of the genotyping PCR was Sanger-sequenced in forward and reverse direction and aligned to the expected sequence. Locations of sequence mismatch are marked in red below the electropherogram. The mismatch upstream of the 5' MH (indicated by the red arrow) was confirmed to be a 10 bp deletion, while the two mismatches along the insert sequence were technical artifacts of the sequencing process.

Expression of the tagged BRD4, which contains two HA epitopes and a mass increase of 17 kDa, and absence of untagged BRD4 was confirmed by immunoblotting on whole cell lysates (see 3.2.3).

To confirm that the intracellular localization of BRD4 was not altered because of the tag, 5 * 10^6 to 1 * 10^6 cells were fractionated as described in 3.2.4.1, except that a phosphatase inhibitor cocktail was included in the buffers instead of α-amanitin and an RNase inhibitor. After fractionation, volumes of the nucleoplasm and the chromatin fractions were adjusted to the volume of the cytoplasm to facilitate direct comparisons. After volume adjustment, 2X SDS buffer was added to each fraction. Immunoblotting was carried out as described below (see 3.2.3).

To determine growth rates, modified K562 dTAG-BRD4 cells and cells of the parental K562 cell line were seeded at 3 * 10^5 cells/ml in six biological replicates and counted in duplicates after 24, 48 and 72 h using an automated cell counter. The doubling time was calculated using the equation:

$$T_{doubling} = T_{2-1}\left(\frac{ln2}{ln\frac{x_1}{x_2}}\right)$$

3.2.3 IMMUNOBLOTTING

$5 * 10^5$ to $4 * 10^6$ cells were collected by 4 min of centrifugation at 200 x g and were either snap-frozen in liquid nitrogen and stored at -80°C or immediately resuspended in 30 µl PBS supplemented with proteinase and phosphatase inhibitor mixes and 25 to 50 U benzonase and was incubated with vigorous shaking at room temperature for 20 min. 30 µl of 2X SDS buffer (Table 6) containing 10% β-mercaptoethanol and Orange G were added to the whole cell lysate, followed by denaturation at 95°C for 5 min. As the samples usually were completely homogenous and free of insoluble material at this point, no clearing by centrifugation was required. Protein concentration was not determined, as the gel loading volume was determined based on cell number, not protein content. The same procedure was followed if nuclei or chromatin were used as starting material, although the amount of benzonase was increased proportionally to the original cell number. Cytoplasm and nucleoplasm were directly mixed with 2X SDS buffer.

Depending on the abundance of the proteins to be detected, 10 to 25 µl lysate per well and 1 µl Kaleidoscope ladder were loaded on a 4–12% gradient Bis-Tris acrylamide gel. Electrophoresis in MOPS buffer was performed at 170 V and room temperature for 1 h or was extended to 2 h to increase the resolution for larger proteins. The wet transfer onto a 0.2 µm nitrocellulose membrane was done at constant 0.03 A overnight or for 2 h at 120 V in a cold room on a magnetic stirrer. After an optional total protein staining, the membrane was blocked with 5% milk in TBS-T and then incubated overnight with a suitable primary antibody diluted in T20 blocking buffer in a cold room. Next, the membrane was washed twice with TBS-T for 10 min, incubated with an appropriate fluorophore-coupled secondary antibody diluted in 5% milk in TBS-T for 1.5 h at room temperature, washed twice with TBS-T for 15 min and once briefly with PBS.

The specific antibody-based signal was detected in the 800 nm channel, whereas the protein ladder was detected in the 700 nm channel of an Odyssey imager. Contrast or brightness of the images was not altered. Band intensity was quantified using the Image Studio Lite software. For normalization, the signal of either a house-keeping protein (GAPDH, H2B or α-tubulin) or a total protein stain was quantified per lane and divided by the highest intensity detected on the membrane to obtain a normalization factor. Then, the signal to be normalized was divided by the normalization factor calculated for the same lane.

3.2.4 NATIVE ELONGATING TRANSCRIPT SEQUENCING WITH SPIKE-IN NORMALIZATION (SI-NET-SEQ)

3.2.4.1 TREATMENT, CELL FRACTIONATION AND ISOLATION OF CHROMATIN-ASSOCIATED RNA

The SI-NET-seq protocol is based on the method published by Mayer and Churchman [375], with some modifications first introduced in Gajos et al. [92] and Jasnovidova et al. [369].
Per sample, 1.2 * 10^7 K562 dTAG-BRD4 cells at a concentration of 1 * 10^6/ml were treated with 100 nM dBET6, 500 nM dTAG7 or DMSO for 40, 50 or 120 min. After treatment, the cells were centrifuged at 200 x g and 4°C for 4 min and washed twice with ice-cold PBS to remove residual traces of the degrader or DMSO. In parallel, NIH3T3 cells were trypsinized, counted using an automated counter and divided into 2.4 * 10^6 cells aliquots. Each aliquot of NIH3T3 cells was combined with 1.2 * 10^7 K562 dTAG-BRD4 cells in PBS (ratio 1:6) so that all following steps were performed on a mix of human and mouse cells.

All buffers were made from nuclease-free reagents (Table 8), contained Superase-In RNase inhibitor, a protease inhibitor cocktail and α-amanitin to inhibit Pol II run-on elongation, and were kept on ice until use. Cells were transferred into a DNA low-bind 1.5 ml tube, centrifuged at 500 x g and 4°C for 2 min. After careful removal of the supernatant, the cell pellet was resuspended in 200 µl of cytoplasmic lysis buffer by pipetting up and down and incubated on ice for 5 min. Next, the lysate was layered onto 400 µl of sucrose buffer in a fresh 1.5 ml tube and centrifuged at 16,000 x g and 4°C for 10 min. After removing the cytoplasmic supernatant, the nuclei pellet was washed with 500 µl nuclei wash buffer and centrifuged at 1,150 x g and 4°C for 2 min. The nuclei pellet was gently but nearly completely resuspended in 200 µl glycerol buffer and transferred into a fresh 1.5 ml tube. 200 µl nuclei lysis buffer were added, followed by pulse vortexing for 45 s and 2 min incubation on ice. Chromatin became visible and was pelleted by centrifugation at 18,500 x g and 4°C for 2 min. Following removal of the nucleoplasmic supernatant, 700 µl Qiazol were added to the pellet.
Resuspension was achieved by initial decompaction with a P1000 pipette and subsequent repeated vortexing, rotation on a rotating wheel and pipetting up and down using a syringe attached to a 22G needle. The sample was either stored at -80°C for a few days at this point or immediately subjected to RNA extraction using the miRNeasy mini kit. The choice of this kit was essential to capture RNA molecules of <200 nt length

and thus detect events of promoter-proximal pausing. RNA was extracted following the manufacturer's instructions. The optional on-column DNase treatment was done. 30 µl of the eluate usually contained >10 µg chromatin-bound RNA, as was found using a NanoDrop spectrometer.

3.2.4.2 SI-NET-SEQ LIBRARY PREPARATION

Per sample, three reactions containing 1 µg RNA each were performed in parallel. Corresponding treatment–control sample pairs were processed at the same time to minimize technical variation. In contrast to the original protocol by Mayer and Churchman [375], a barcode DNA linker containing a random decamer (not hexamer) sequence functioning as a unique molecular identifier and an internal 8 nt-long index was ligated to the 3' end of each RNA (Table 18). The ligation mix consisted of 20% (v/v) PEG8000, 10% (v/v) DMSO, 1X T4 RNA ligase buffer, 200 U truncated T4 RNA ligase 2, 1 µg RNA sample and 1 µg barcode DNA linker in 20 µl and was incubated at 37°C for 3 h. 2 µl NEBNext RNA fragmentation solution containing Mg^{2+} were added per tube, followed by heating to 95°C for exactly 10 min to maximize the fraction of RNAs of 40 to 150 nt. To stop the fragmentation, 2 µl stop solution (NEB), essentially EDTA, were added. Ligated and fragmented RNA was precipitated by adding 60 µl of 2 M sodium acetate, 500 µl nuclease-free H_2O, 2 µl GlycoBlue and 750 µl isopropanol, incubation overnight at -80°C and centrifugation at 20,000 x g and 4°C for >2 h.

After careful removal of the supernatant, washing with 750 µl of 80% ethanol and drying for 8 min, each pellet was resuspended in 10 µl nuclease-free H_2O. The RNA sample was mixed with 10 µl of 2X TBE–urea buffer and denatured at 80°C for 2 min, whereas the 20 bp DNA ladder, which was also used on all following gels, was denatured for 5 min at 80°C. Electrophoresis was allowed to proceed at 200V for 65 min. After staining with 1X SYBR Gold in TBE buffer for 5 min on a shaking platform, the gel was placed on a blue light table, and RNA of 55 to 140 nt, corresponding to inserts of 20 to 105 nt, were excised with a scalpel. Contamination with free linkers of 33 nt was avoided. The gel piece was transferred into a 0.5 ml DNA low-bind tube that was pierced at the bottom and placed in a 2 ml DNA low-bind tube. Centrifugation at 20,000 x g for 4 min led to efficient homogenization of the gel piece. 200 µl nuclease-free H_2O were added, resulting in a slurry that was then incubated at 70°C and 1,400 rpm for 10 min. Next, the slurry was transferred onto a 0.22 µM Spin-X centrifuge tube filter and centrifuged at 20,000 x g for 3 min. To maximize recovery of the eluate, centrifugation was repeated twice with

different orientations of the tubes. RNA was precipitated by adding 2 µl GlycoBlue, 50 µl of 3 M sodium acetate, and 750 µl isopropanol, incubation overnight at -80°C and centrifugation, as described above. All pellets of the same sample were reconstituted in 10 µl nuclease-free H_2O.

The sample was then mixed with 1X first-strand buffer, 0.5 mM dNTPs and 0.3 µM reverse primer oLSC007 and was denatured at 80°C for 2 min. After brief cooling on ice, 1.3 µl Superase-In–DTT mix and 160 U SuperScript III were added. Reverse transcription was performed at 48°C for 30 min. To remove the RNA by alkaline hydrolysis, the sample was mixed with 1.8 µl of 1 N NaOH and incubated at 98°C for 20 min. NaOH was subsequently neutralized by 1.8 µl of 1 N HCl. The sample was mixed with 20 µl of 2X TBE–urea buffer, denatured at 95°C for 5 min and separated on a 10% (w/v) TBE–urea acrylamide gel running at 200 V for 65 min. Single-stranded cDNA of 95 to 180 nt was extracted from the gel, maintaining a range of insert sizes from 22 to 105 nt. Homogenization was done as described above. For precipitation at -20°C, 2 µl GlycoBlue, 25 µl of 5 M NaCl and 750 µl isopropanol were added.

After reconstitution, as described above, the cDNA was circularized using CircLigase according to the manufacturer's instructions, except that the concentration of $MnCl_2$ was reduced to 1.25 mM to prevent the formation of a precipitate. Next, cDNA originating from 20 highly abundant mature RNA was removed by subtractive hybridization involving biotinylated sequence-specific DNA probes. In brief, 110 µl Dynabeads streptavidin C1 were washed three times with 1X equilibration buffer, resuspended in 40 µl of 2X equilibration buffer and equilibrated at 37°C for 30 min. In parallel, circularized cDNA and 200 µM of each biotinylated probe in 2X SSC buffer, which were divided by four tubes of 10 µl each, were denatured for 90 s at 99°C and allowed to hybridize while the temperature decreased to 37°C in 0.1°C/1 s decrements in a thermocycler, followed by 15 min at 37°C. 10 µl of the washed streptavidin C1 beads were added to each tube containing the hybridization mix for another 15 min at 37°C with 1,000 rpm shaking. Then the tubes were placed on a magnetic rack, and the clear supernatant was transferred into one fresh DNA low-bind tube, mixed with 2 µl GlycoBlue, 6 µl of 5 M NaCl, 74 µl H_2O and 150 µl isopropanol and precipitated at -20°C overnight. Next, the DNA was centrifuged and washed, as before, and reconstituted in 15 µl nuclease-free H_2O.

In order to determine the number of PCR cycles required to obtain sufficient quantities of the sequencing library while minimizing the formation of uninformative PCR duplicates and concatemeric products, a test PCR run in which amplification was stopped after four, six or eight amplification cycles was performed. Each reaction of 20 µl consisted of 3.8 µl of 5X Phusion HF buffer, 0.2 mM dNTPs, 0.5 µM of each the respective indexed forward

primer and the universal reverse primer, 1.8 U Phusion DNA polymerase and 1 µl of template cDNA. The following cycling scheme was used: 98°C for 30 s; 4 to 8 cycles of 98°C for 10 s, 60°C for 10 s and 72°C for 5 s; pausing at 4°C. Orange G DNA loading dye was added to each sample as well as to the 20 bp DNA ladder, before they were run on an 8% TBE gel at 180 V for 65 min, stained in 1X SYBR Gold and visualized under blue light. Following this test, five identical reactions per sample were performed using the optimal number of cycles. After gel electrophoresis, the band of on average 182 bp consisting of the library was excised. Care was taken to avoid contamination with faster migrating amplification products originating from self-ligated DNA linkers and primer dimers. After homogenizing the gel pieces as described above, a slurry in 670 µl soaking buffer was prepared and incubated overnight in a thermo-shaker at 21°C and 1,400 rpm. Next, the eluate was separated using a centrifuge tube filter, as described above. The library DNA was precipitated with 2 µl GlycoBlue and 680 µl isopropanol at -20°C overnight, centrifuged at 20,000 x g and 4°C for >2 h, followed by washing of the pellet with 80% ethanol and reconstitution in in total 6 µl of 10 mM Tris-HCl (pH 8.0).
1 µl of the eluate was used to prepare 1:5 and 1:10 dilutions to measure DNA concentration using the Qubit dsDNA HS kit and size distribution using the BioAnalyzer HS DNA kit. Libraries were sequenced on a HiSeq4000 instrument in SR75 mode at the sequencing core facility of the MPI for Molecular Genetics.

3.2.4.3 SI-NET-SEQ DATA PROCESSING

SI-NET-seq data were processed by Annkatrin Bressin as summarized in the following. First, the adapter sequence ATCTCGTATGCCGTCTTCTGCTTG was trimmed using Cutadapt [394]. Likewise, poly-G stretches indicative of failed signal detection were removed. Next, the unique molecular identifier (UMI), *i.e.*, the ten bases adjacent to the adapter, was extracted using Starcode [410]. Reads of identical insert sequence and UMI were considered PCR duplicates and counted only once. For read mapping, a combined genome consisting of the human GRCh38.p12 and mouse GRCm38.p6, in which genes transcribed by Pol I and Pol III, and Pol II-transcribed genes of chromatin-associated non-coding RNAs were masked, was used. Reads were mapped using STAR aligner [411]. Multi-mapping reads were discarded. Mapping of not only the insert portion but also the UMI to the genome indicated incorrect binding of the RT primer not to its binding site in the linker but to a genomic region of a similar sequence. Using a custom Python script, the Pol II position was defined as the 5'-most nucleotide of the insert sequence

corresponding to the last nucleotide incorporated into the nascent transcript. Lastly, reads likely originating from splicing intermediates were removed.

Pol II occupancy changes in the promoter-proximal region (defined as the region from the TSS to TSS +300 bp) and the gene body region (TSS +300 bp to the pA site) of active, non-overlapping genes of ≥1 kb length with ≥6 reads in the respective region were determined with DEseq2 [412]. Spike-in normalization exploited the relative log expression function of DESeq2, which was used to calculate a size factor based on the content of mouse reads in the individual data set.

The readthrough index (RTI) was defined as the difference in the average termination distance (ATD) at a gene j under treatment and control conditions:

$$RTI = ATD_j(Occ_{treatment}) - ATD_j(Occ_{control})$$

The ATD is a measure of the Pol II occupancy within the extended termination zone of gene j, defined as the region between the last active pA site and the point where the Pol II signal strength falls below 1 read for a window of 5 kb. In particular, Pol II occupancy is included as a weighted distance:

$$ATD_j = \frac{\sum_{i=1}^{j} d_{pA_j}(i) \times Occ(i)}{\sum_{i=1}^{j} Occ(i)}$$

For RTI calculation, only non-overlapping genes with a termination zone of >1 kb under each condition and an RPKM of >0.01 were considered.

3.2.5 ChIP-seq with reference genome normalization (ChIP-Rx)

3.2.5.1 Treatment and chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed largely as described by Baluapuri et al. [413]. Modifications are detailed below.

4 * 10^7 K562 dTAG-BRD4 cells growing at a density of 1 * 10^6 cells/ml were treated with dTAG7 at a final concentration of 500 nM or with DMSO for 2 h, washed with PBS and crosslinked with 1% methanol-free formaldehyde for 5 min at room temperature while rotating. Glycine was added to a final concentration of 250 mM to quench unreacted formaldehyde. In parallel, murine NIH3T3 cells were crosslinked with 1% formaldehyde

for 8 min at room temperature and rinsed with 250 mM glycine, followed by detaching from the culture vessel by scraping, gentle resuspension and cell counting using an automated counting system. K562 dTAG-BRD4 and NIH3T3 cells were washed with ice-cold PBS to remove traces of formaldehyde. Per sample, exactly 4 * 10^7 K562 dTAG-BRD4 and 1 * 10^7 NIH3T3 cells were combined at a ratio of 1:5. Thus, a mixture of human and murine cells was handled in all of the following steps. After cell lysis by incubating in modified lysis buffer I (Table 9) for 20 min with rotation at 4°C, the nuclei were collected by centrifugation at 420 x g and 4°C for 10 min, resuspended in lysis buffer II and divided into two AFA fiber 1 ml tubes (Covaris). Chromatin shearing was carried out in an E220evolution sonicator for 20 min at intensity 4, duty cycle 5% and 200 cycles per burst. The temperature was kept below 7°C by the integrated cooling system. After pre-clearing the lysates by centrifugation at 16,000 x g and 4°C for 10 min, 2.5% of the sample volume was saved as input. The remainder of the sample was incubated with an appropriate amount of primary antibody (Table 16), which was coupled to 80 µl Dynabeads Protein G by incubating for >6.5 h, overnight on a rotating wheel in a cold room.

On the following day, the chromatin-bound beads were washed three times for 45 s on a rotating wheel with each of the washing buffers I, II and III (Table 9). For elution, the beads were resuspended in 100 µl of freshly prepared elution buffer and incubated at room temperature for 15 min, followed by one repetition. Subsequently, the ChIP and input samples were incubated with 20 µg RNase A for 1 h at 37°C and with 160 µg Proteinase K for 2 h at 50°C before crosslinks were reversed by 5 h incubation at 65°C in the presence of 0.84 M NaCl. DNA was purified using the ChIP DNA Clean & Concentrator kit according to the manufacturer's instructions and quantified using the Qubit dsDNA HS kit.

3.2.5.2 ChIP-Rx library preparation

Libraries were constructed from 6 to 30 ng of ChIP DNA (depending on the efficiency of the ChIP enrichment) or 30 ng input DNA using the NEBNext Ultra II DNA kit and single or dual index primers according to the manufacturer's instructions. The appropriate number of PCR cycles was chosen based on the starting amount of ChIP DNA, as recommended by the manufacturer. After purification of the amplified libraries with one volume AMPure beads, fragments of 200 to 500 bp were extracted from an 8% TBE gel, as described in 3.2.4.2. This additional size selection was necessary to completely

remove residual adapters and primer dimers as well as larger, under-fragmented DNA fragments. ChIP libraries were sequenced on an Illumina HiSeq 4000 or NovaSeq 6000 sequencing instrument in PE75 or PE100 mode, respectively, at the sequencing core facility of the MPI for Molecular Genetics. 1 to 1.5 * 10^8 reads were requested per library.

3.2.5.3 ChIP-Rx data processing

All ChIP-Rx data were processed by Annkatrin Bressin as summarized in the following. After adapter trimming etc., reads were aligned to a joined reference genome consisting of the GRCh38.p12 and GRCm38.p6 sequences using Bowtie2 [390]. After the human and mouse signal were separated, raw density was extracted using Bedtools [388]. To account for differences in sequencing depth, reads per million (RPM) values were calculated as

$$RPM = c \times \alpha$$

where c is the raw read count and α is a normalization factor that is dependent on the total number of uniquely mapped reads N and is defined as

$$\alpha = \frac{1}{N} \times 10^6$$

MACS2 [400] in "bdgcmp" mode was used for input as well as Pol II normalization. Heat maps and meta-gene plots were generated using DeepTools2 [395].

3.2.6 Assay for transposase-accessible chromatin with sequencing (ATAC-seq)

3.2.6.1 ATAC-seq library preparation

ATAC-seq was performed following the Omni-ATAC protocol by Corces et al. [414] with minor modifications.

3 * 10^6 K562 dTAG-BRD4 at a concentration of 1 * 10^6cells/ml were treated with 500 nM dTAG7 or an equivalent amount of DMSO for 2 or 6 h. As cell viability was high (>93 %), no additional steps were taken to remove dead cells. 50 µl of the suspension corresponding to 50,000 cells were transferred into a 1.5 ml tube and centrifuged for 5 min at 500 x g and 4°C. After removal of the supernatant, the cell pellet was gently but fast resuspended in 50 µl ice-cold RBS–NTD buffer by pipetting up and down three times

(Table 10) and incubated on ice for 3 min. 1 ml of ice-cold RBS–T was added, followed by inverting the tube four times and centrifuging at 500 x g, 4°C for 10 min. After removing the supernatant, the nuclei were resuspended in 25 µl TD buffer, 2.5 µl DNA TDE1 enzyme (≙ Tn5 transposase), 16.5 µl PBS, 0.5 µl of 1% (v/v) digitonin and 0.5 µl of 10X Tween-20 in a final volume of 50 µl by pipetting up and down six times and were incubated for 45 min at 37°C and 1,000 rpm in a thermo-shaker. Afterwards, the samples were placed on ice immediately, diluted with 250 µl DNA binding buffer (Zymo) and purified using the DNA Clean & Concentrator kit. DNA was eluted sequentially in 10 µl and 11 µl nuclease-free H_2O. The entire eluate was combined with 2X NEBNext Q5 master mix and 1.25 µM of each the forward and the reverse primer (Table 18). The following program was used to PCR-amplify the libraries: 72°C for 5 min, 98°C for 30 s, 8 cycles of 98°C for 10 s, 63°C for 30 s and 72°C for 1 min, pausing at 4°C. Eight cycles were chosen as a compromise between enriching for the relevant subnucleosomal and mononucleosomal fractions (which amplify more efficiently due to their small size) and limiting PCR artifacts. Lastly, after running the amplified samples on an 8% TBE gel at 180 V for 60 min and staining with SYBR Gold, fragments of 180 to 500 bp were excised and purified as described in 3.2.4.2. Thereby, dinucleosomal (ca. 550 bp without adaptors) and polynucleosomal DNA were excluded. ATAC-seq samples were sequenced in PE100 mode on a NovaSeq 6000 for 5 * 10^7 reads per sample at the sequencing facility of the MPI.

3.2.6.2 ATAC-SEQ DATA PROCESSING

All analysis steps were performed using the Galaxy Project platform [396]. Cutadapt [394] was used to trim Nextera adapter sequences (CTGTCTCTTATACACATCTCCGA GCCCACGAGAC; CTGTCTCTTATACACATCTGACGCTGCCGACGA). Read pairs with a distance of 500 to 1,000 nt and a quality score of ≥20 were mapped to the GRCh38.p12 genome assembly with the parameter "--very sensitive" using Bowtie2 [390]. Next, read pairs with a mapping quality of <30 or of mitochondrial origin, which accounted for 20 to 24% of all reads, were excluded. Picard MarkDuplicates was used to identify and discard PCR duplicates. Peaks were called using MACS2 [400] with the parameters "--shift -100" to center the signal on the 5' end of the reads, *i.e.*, the cutting site of Tn5, and "--extend 200". Metagene plots were generated using DeepTools [395]. Lastly, the data were reads per kilobase million (RPKM)-normalized using the definition

$$RPKM = c \times \frac{1}{N \times L} \times 10^9$$

where c is the raw read count, N is the total number of uniquely mapped reads and L is a fixed bin size set to 50 kb here.

3.2.7 Hi-ChIP

3.2.7.1 Hi-ChIP sample preparation

Hi-ChIP was performed following a protocol by Mumbach et al. [415] with some modifications.
5 * 10^6 K562 dTAG-BRD4 at a concentration of 5 * 10^6 cells/ml were treated with 500 nM dTAG7 or DMSO for 2 h, followed by crosslinking with 1 % methanol-free formaldehyde for 10 min at room temperature and quenching with glycine at a final concentration of 125 mM. The cells were pelleted at 400 x g and 4°C for 5 min, washed with ice-cold PBS and frozen in liquid nitrogen to facilitate lysis. Subsequently, the cell pellet was thawed on ice and lysed in 1000 µl ice-cold Hi-C lysis buffer while rotating at 4°C for 30 min (Table 11), followed by centrifugation at 2,500 x g and 4°C for 5 min. After one washing step with Hi-C lysis buffer, the nuclei pellet was gently but fast resuspended in 100 µl of 0.5% SDS in H_2O and incubated at 62°C for exactly 10 min. SDS was then quenched by incubating with 1.1% Triton X-100 at 37°C for 15 min. 50 µl NEB buffer 2 and 375 U MboI, which recognizes the sequence GATC, were added dropwise to prevent clumping. The samples were placed on a rotator at 37°C overnight and heated to 62°C for 20 min for inactivation of the enzyme on the following day.
For end repair and biotin-labeling, 1.5 µl of 10 mM dCTP, dGTP, dTTP as well as biotin-14-dATP and 50 U Klenow fragment were added per sample and incubated at 37°C for 1 h. To religate the DNA ends, the samples were incubated with 4,000 U T4 DNA ligase, 140 µl of 10X T4 ligase buffer, 125 µl of 10% Triton X-100 and 7.5 µl of 20 mg/ml BSA at room temperature for 4 h. Afterwards, the nuclei were pelleted at 2,500 x g for 5 min.

Next, each nuclei pellet was resuspended in 980 µl ChIP sonication buffer and transferred into a 1 ml AFA fiber tube. Chromatin shearing was carried out in an S220 sonicator for 6 min at intensity 4, duty cycle 5% and 200 cycles per burst. The samples were then cleared by centrifugation at 16,100 x g and 4°C for 10 min and combined with

an H3K27ac-specific antibody coupled to magnetic beads for overnight incubation on a rotator a 4°C (Table 16). Input samples were not collected. On the next day, the samples were placed on a magnetic stand to remove the unbound material and washed once with 1 ml of sonication buffer, sonication buffer with high salt content, LiCl washing buffer and TE buffer containing NaCl, while rotating for 5 min at 4°C. For elution of the chromatin, the samples were resuspended in freshly prepared elution buffer and heated to 65°C for 15 min. Next, the samples were incubated with 80 µg RNase A for 2 h at 37°C and with 200 µg Proteinase K for 45 min at 50°C, before crosslinks were reversed at 65°C for 5 h. DNA was purified using the ChIP DNA Clean & Concentrator kit. A 1:5 dilution was prepared to determine the concentration and size distribution of the DNA.

Per sample, 5 µl Streptavidin C1 beads were transferred into a fresh tube and washed with 1 ml Tween-20 wash buffer by rotating for 5 min at 4°C, followed by resuspension in 10 µl of 10X biotin binding buffer for each sample. Next, 50 ng DNA in 10 µl H_2O were combined with 10 µl of the bead suspension and incubated at room temperature for 15 min to allow for biotin–streptavidin binding while mixing by flicking the tube every 5 min. Subsequently, the samples were resuspended twice in 500 µl Tween-20 washing buffer and incubated at 55°C and 900 rpm shaking for 2 min to remove unspecific binding. Care was taken to completely remove all liquid after this step. The samples, which were still bead-bound, were resuspended in 25 µl of 2X Nextera TD buffer, supplemented with 2.5 µl Tn5 transposase and 22.5 µl H_2O and incubated for 10 min at 55°C and 900 rpm shaking. After removing the supernatant, 500 µl of 50 mM EDTA were added for a 30 min incubation at 50°C. Again, the beads were washed twice with 500 µl Tween-20 washing buffer at 55°C, and once with 500 µl of 10 mM Tris-HCl (pH 7.5).
The bead-bound sample was combined with 15 µl NEBNext Q5 master mix and 1.25 µM of the forward and the reverse Nextera primers in a 45 µl reaction and was amplified using the following cycling scheme: 72°C for 5 min, 98°C for 30 s, 8 cycles of 98°C for 10 s, 63°C for 30 s and 72°C for 1 min, pausing at 4°C. The amplified library was purified using the DNA Clean and Concentrator kit and quantified using a Qubit spectrometer and TapeStation.

Lastly, the library was purified and size-selected using AMPure beads following the manufacturer's recommendations. Specifically, the beads were added at a ratio of 0.8 and 0.5 to enrich for fragments of 200 to 750 bp length. Sequencing was done in PE50 mode on a HiSeq 4000 instrument for $1.3 * 10^7$ reads in the sequencing facility of the MPI.

3.2.7.2 Hi-ChIP data processing

Hi-ChIP data were processed and analyzed by Annkatrin Bressin. Briefly, reads were mapped to the GRCh38.p12 genome assembly using HiC-Pro [416]. Those reads that could not be readily mapped were considered chimeric and therefore were split at MboI restriction sites. Subsequently, these halves were mapped to the genome while reads without MboI sites as well as PCR duplicates were discarded. Next, the reads were divided into 10 kb bins. Using HiCcompare [397], regions showing copy number variation were discarded, and loess normalization was applied to the interaction matrix. To predict the functional state of the interaction sites, ChromHMM [417] was trained on ENCODE data of H3K27me3, H3K4me1, H3K27ac, H3K4me3, H3K79me2 and H3K36me3.

3.2.8 Quantitative native immunoprecipitation with mass spectrometry (native IP-MS) or immunoblotting

3.2.8.1 Sample preparation and analysis

4 * 10^7 K562 dTAG-BRD4 or one confluent T175 flask of HeLa 8880 LAP-CstF64 [387] of untagged HeLa S3 cells per replicate were collected and washed with ice-cold PBS. One copy of GFP is part of the N-terminal LAP tag, which facilitates affinity enrichment. The cells were incubated in 450 µl native IP buffer A containing 2X protease and phosphatase inhibitor mixes on ice for 4 min and centrifuged at 16,000 x g and 4°C for 10 min. Then, the nuclei were resuspended in 410 µl native IP buffer A containing 2X protease and phosphatase inhibitor mixes and supplemented with 1,000 U benzonase (Table 13) and incubated on ice for up to 2 h with repeated gentle mixing by pipetting. The homogenous nuclei lysate was centrifuged at 14,000 x g and 4°C for 10 min to remove insoluble material. In case of experiments with immunoblotting as readout, 10% of the supernatant was saved as input. Also, a set of replicates was processed to this step, precipitated and submitted for mass spectrometric analysis to serve as a background, for instance, in a subsequent GO term analysis. 0.8 µM EDTA and an appropriate amount of the specific antibody or isotype-matched normal IgG were added to each K562 sample (Table 17). Lysate and antibody were incubated for 4 h on a rotating wheel at 4°C, and for additional 1.5 h in the presence of 150 µg Dynabeads protein G. In case of the CstF64 experiment in HeLa cells, no EDTA was added, and the lysates of LAP-CstF64 and untagged control cells were incubated with 25 µl GFP-Trap Dynabeads for a total duration of 2 h. For

immunoblotting, the protein-bound beads were washed five times with 1 ml native IP buffer B and eluted in 60 µl of 1X SDS loading buffer at room temperature. The samples were denaturation at 95°C for 5 min and supplemented with β-mercaptoethanol to a final concentration of 5%. Bead-bound samples for mass spectrometry were washed five times with native IP buffer B and additionally three times with 1 ml of 100 mM ammonium bicarbonate and were handed over to the mass spectrometry facility of the MPI for further processing and analysis.
In brief, the bead-bound samples were digested with trypsin and desalted using Pierce C18 tips, reconstituted in 5% acetonitrile and 2% formic acid in H_2O and sonicated for 30 s. Peptides were separated by nanoflow reverse-phase liquid chromatography using a C18 resin analytical column and analyzed on an orbitrap instrument.

Five or four replicates were analyzed for each IP-MS experiment in K562 and HeLa cells, respectively.

3.2.8.2 DATA ANALYSIS

The raw data were processed using MaxQuant [401]. Amino acid sequences were searched against the human UniProtKB database, with a false discovery rate of 0.01 and a mass tolerance of 4.5 ppm and 20 ppm for precursor and fragment ions, respectively. Two missed tryptic cleavage sites were tolerated. Cysteine carbamidomethylation was set as a fixed modification, while N-terminal acetylation and methionine oxidation were set as variable modifications. Label-free quantification (LFQ) intensities were used for the following analysis performed using Perseus [404]. Only proteins identified by ≥2 common peptides were considered in case of the native IP-MS experiments, while the minimum number was set to 1 to increase sensitivity for crosslink IP-MS (see 3.2.9). After filtering for potential contaminants, proteins matching the reverse sequence database or proteins for which only modified peptides were detected, LFQ intensities were log2-transformed. Proteins identified in <70% of the replicates of at least one sample type were discarded. Missing values were imputed from a normal distribution of the measured values of each replicate using the default settings. Lastly, a two-tailed two-sample Student's t-test with Benjamini–Hochberg correction was performed. Since imputation can introduce a bias if signal intensities are relatively low, unique peptide counts were shown instead of p-values in case of the LAP-CstF64 IP experiment.

3.2.9 QUANTITATIVE IMMUNOPRECIPITATION WITH FORMALDEHYDE CROSSLINKING AND MS (CROSSLINK IP-MS)

5 * 10^7 K562 dTAG-BRD4 or untagged K562 control cells were used per replicate. Samples were prepared as described by Mohammed et al. [418].

All lysis and washing buffers used were supplemented with a protease inhibitor mix. In brief, cells were collected by centrifugation, resuspended in 10 ml RPMI without additives and crosslinked with methanol-free formaldehyde at a final concentration of 1% for 8 min at a rotating wheel at room temperature, followed by the addition of 0.125 M glycine and thorough inversion of the tube. Cells were pelleted at 200 x g and 4°C for 4 min, washed with PBS twice and snap-frozen in liquid nitrogen. After thawing on ice, the pellet was resuspended in 10 ml ice-cold LB1 (Table 14), rotated for 10 min in a cold room and centrifuged at 200 x g and 4°C for 4 min, followed by lysis with LB2 in the same way. The nuclei pellet was then resuspended in 900 µl LB3 and divided into three 1.5 ml TPX sonication tubes (Diagenode). Sonication was carried out in a Bioruptor Plus at high intensity for five cycles of 30 s shearing and 40 s pausing. Next, Triton X-100 was added to a final concentration of 1%, and the sample was centrifuged at 20,000 x g and 4°C for 10 min. In parallel, 15 µl (≙ 300 µg) Pierce Protein A/G beads were blocked with 0.5% BSA in PBS, and incubated with 0.67 µg HA antibody for 1 h at room temperature. After washing with BSA in PBS three times, the antibody-coupled beads were added either to the cleared lysate of K562 dTAG-BRD4 or of untagged control cells, which were then placed on a rotation wheel in a cold room overnight. On the next day, the protein-bound beads were washed eight times with 1 ml cold modified RIPA buffer and three times with 1 ml of 100 mM ammonium bicarbonate before they were handed over to the mass spectrometry facility of the MPI for further processing and analysis, as described above (3.2.8.2).

Five HA-BRD4 IP replicates and four control replicates were processed and analyzed.

3.2.10 QUANTITATIVE CHROMATIN-MS

3.2.10.1 SILAC LABELING, TREATMENT AND SAMPLE PREPARATION

The SILAC labeling procedure was largely done as described by Ong and Mann [419]. Prior to the experiment, K562 dTAG-BRD4 cells were cultured in sterile-filtered RPMI medium containing 10% dialyzed FBS (GIBCO), 5% penicillin–streptomycin and either $^{13}C_6{}^{15}N_2$ L-lysine-2HCl and $^{13}C_6{}^{15}N_4$ L-arginine-HCl ("heavy RPMI"; Thermo Fisher), or L-lysine and L-arginine containing only light isotopes ("light RPMI") for at least five cell division cycles. Cell proliferation was comparable in heavy and light medium (not shown). To determine incorporation efficiency, 2 x 10^6 cells grown in heavy and 2 x 10^6 cells grown in light medium were collected by centrifugation at 200 x g and 4°C for 4 min, washed with PBS and lysed in 100 µl guanidine hydrochloride buffer in the presence of 500 U benzonase in a cold room (Table 12). To aid lysis, the cells were vigorously shaken for 40 to 60 min. Then, proteins were purified using the ProteoExtract kit following the manufacturer's instructions. Dry protein pellets were submitted to the mass spectrometry core facility of the MPI for Molecular Genetics and processed, as described below. As 99% of the lysine and arginine residues in the heavy sample contained only ^{13}C and ^{15}N, aliquots of the cells were frozen for subsequent experiments.

To determine the effects of BET or BRD4-selective degradation on chromatin composition, cells were treated with 100 nM dBET6, 500 nM dTAG7 or an equivalent of DMSO for 120 or 360 min in four to five biological replicates. To exclude effects of the media, cells of two replicates were degrader-treated in heavy, cells of the other two replicates in light media. After treatment, cells were washed with ice-cold PBS and pairwise pooled, so that degrader-treated cells and corresponding control cells could be further processed in one tube, minimizing technical variation. Pooled samples were fractionated as described in 3.2.4.1, except that the only additives to the fractionation buffers were 1X protease inhibitor mix and 1X phosphatase inhibitor mix. Moreover, chromatin pellets were solubilized by the addition of 750 U benzonase, 10 min incubation on ice and 5 min agitation at room temperature. The chromatin was precipitated overnight at -20°C and purified using the ProteoExtract kit following the manufacturer's instructions. Dry protein pellets were submitted to the mass spectrometry facility of the MPI for further processing and analysis.

In brief, the samples were reconstituted in denaturation buffer, denatured at 95°C and sonicated. After quantification using a bicinchoninic acid assay (BCA), 30 µg protein

where digested with LysC and subsequently with trypsin. Peptides were desalted using Pierce C18 tips and separated into four fractions by strong cation exchange chromatography. Liquid chromatography and mass spectrometric analysis were performed as described in 3.2.8.1.

Six biological replicates per treatment condition were analyzed for the 120 min dTAG7 and 360 min dBET6 experiments, four replicates per condition were analyzed for the remaining experiments.

3.2.10.2 DATA PROCESSING

The raw data were processed using MaxQuant [401], as described above, except that $^{13}C_6{}^{15}N_4$-arginine and $^{13}C_6{}^{15}N_2$-lysine were defined to distinguish the treatment conditions. Using the Perseus software [404], the downstream analysis was performed based on normalized SILAC ratios, which were calculated based on at least two common peptides. After filtering as described in 3.2.8.2, the inverse ratio was calculated for all label swap replicates, and all ratios were log2-transformed. Significance was determined by two-tailed one-sample Student's t-test with Benjamini–Hochberg correction. k-means clustering was performed in Perseus [404].

Statistical overrepresentation of gene ontology (GO) terms linked to biological processes or cellular components was tested using PANTHER [403]. Specifically, the curated "PANTHER GO-Slim" list of cellular component terms and the complete list of terms associated with biological processes were used. The search was performed against a reference list consisting of all proteins robustly detected in at least one of the experiments. A Fisher's exact test with Bonferroni correction was applied to reduce the chance of false positive results.

3.2.11 RT-QPCR-BASED CLEAVAGE ASSAY

To determine if BET or BRD4-selective depletion decreased cleavage efficiency at the main pA site of exemplary genes, a cleavage assay similar to an assay described in Davidson et al. [225] was designed. In particular, the coordinates of the main pA site of *MBNL1* and *PVT1*, which based on our SI-NET-seq data showed readthrough

transcription, and the control gene *CDC42* were extracted from the database PolyA_DB [405]. Primer3 [406] was used to design primer pairs flanking the pA site so that an amplicon was formed only if the transcript serving as a template was not cleaved (Figure 8) (Table 18). Additionally, primer pairs binding further upstream and downstream of the annotated pA site were designed for normalization purposes. NCBI Primer BLAST [402] was used to predict primer specificity. Primer efficiency and specificity were tested using a template dilution series and by melting curve analysis, respectively.

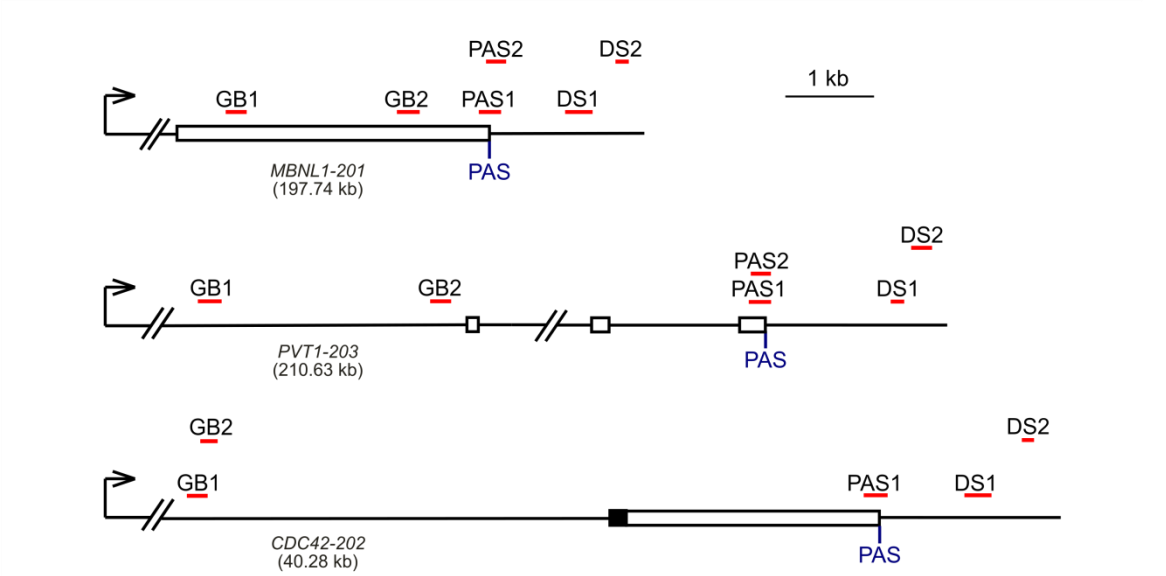


Figure 8: Positions of the RT-qPCR amplicons that were quantified in the cleavage assay. Primers were designed to bind flanking the main pA site (PAS), in the gene body region (GB) or downstream of the pA site (PAS). Amplicons (shown in red), introns, translated (black) and untranslated exons (white) near the 3' RNA end are drawn in scale. Names and lengths of the transcript isoforms that were targeted in this assay are indicated. *MBNL1* and *PVT1* were selected because of the strong termination defect seen upon dBET6 and dTAG7 treatment with SI-NET-seq, whereas *CDC42* served as a negative control. A modified version of the scheme was included in Arnold et al. [368].

For each of three biological replicates, 4 * 10^6 K562 dTAG-BRD4 cells at a concentration of 1 * 10^6 cells/ml were treated with 100 nM dBET6, 500 nM dTAG7 or an equal volume of DMSO for 120 min or were left untreated. Subsequently, the cells were fractionated as described above (3.2.4.1), except that α-amanitin was omitted from all buffers. Chromatin-associated RNA was extracted using the miRNeasy micro kit following the manufacturer's instructions, except that the duration of the on-column DNase I digest was extended to 30 min. RNA was eluted in 14 µl nuclease-free H_2O and quantified by Qubit. For reverse transcription, 0.5 µg RNA, 1 mM dNTP mix, and 75 ng random hexamer oligos in a total volume of 10 µl were incubated at 65°C for 5 min and then placed on ice.

10 µl of the reaction mix consisting of 2X SuperScript III buffer, 10 mM $MgCl_2$, 20 mM DTT, 40 U RNaseOUT RNase inhibitor and 200 U SuperScript III reverse transcriptase were added per sample, which was then incubated for 10 min at 25°C, 50 min at 50°C and 5 min at 85°C. One sample to which no reverse transcriptase was added was processed as a control for each condition, as DNA contamination could otherwise be misinterpreted as uncleaved RNA. Lastly, each sample was treated with 2 U RNase H at 37°C for 20 min.

For qPCR, which was done in technical triplicates or quadruplicates in 384 well plates in a QuantStudio 7 thermocycler, 2.5 ng cDNA, 0.2 µM of the forward and the reverse primer and 1X PowerUp SYBR green mix were included per reaction of 10 µl. The following cycling program was used: 50°C for 2 min, 95°C for 2 min, 40 cycles of 95°C for 15 s, 56°C for 15s and 72°C for 1 min.

To account for differences in nascent transcription that could be caused by the treatment and complicate quantitative comparisons across the samples, all C_t values were normalized by the C_t value at the most upstream gene body position, *i.e.*, GB1, of the respective condition. The resulting ΔC_t values were statistically analyzed by an unpaired t-test with Welch correction.

4 RESULTS

4.1 A NEW ROLE OF BRD4 IN REGULATING TRANSCRIPTION ELONGATION AND TERMINATION

4.1.1 BET PROTEINS AND BRD4 CAN BE EFFICIENTLY AND SPECIFICALLY DEPLETED USING PROTAC DEGRADERS

We treated K562 dTAG-BRD4 cells with 100 nM dBET6, a concentration that has been found efficacious also in MOLT4 cells, and observed a substantial reduction by >70% in the long isoform of BRD4 (BRD4-L) after 120 min (Figure 9A). Likewise, BRD2 and BRD3 at this time point were reduced by >75% and >85%, respectively. A time course experiment indicated a first clear reduction of BRD4 by >40% as early as 60 min after the beginning of the treatment. After 360 min of treatment, <15% of BRD2, BRD3 and BRD4 were still detectable by immunoblotting. While the initial decrease in BRD2 seemed comparable to that of the other BET proteins, higher BRD2 levels remained present in the plateau state, possibly reflecting a lower maximal capacity of the degrader against BRD2, as has also been seen for the structurally related dBET1 [420]. Overall, our observations in K562 cells were comparable with earlier studies in that dBET6 showed a similar level of activity towards each BET member [135]. No remarkable increase in protein levels was observed after overnight treatment, indicating the stability of the compound over prolonged treatment times. Also, no signs of apoptosis or cellular stress were observed neither by brightfield microscopy nor when probing for cleaved caspase 3 by immunoblotting (not shown).

Likewise, dTAG7 treatment of the K562 dTAG-BRD4 cell line at a final concentration of 500 nM proved to degrade BRD4 efficiently. In particular, >85% of BRD4-L and >90% of BRD4-S were degraded after 120 min of treatment (Figure 9B, C). Sampling in 10 min intervals revealed a decrease in BRD4 signal by >40% already after 40 min of dTAG7 treatment (Figure 9C). Notably, accurate quantification of BRD4-L was complicated by inefficient transfer onto the nitrocellulose membrane, which was evident even after extensive optimization.

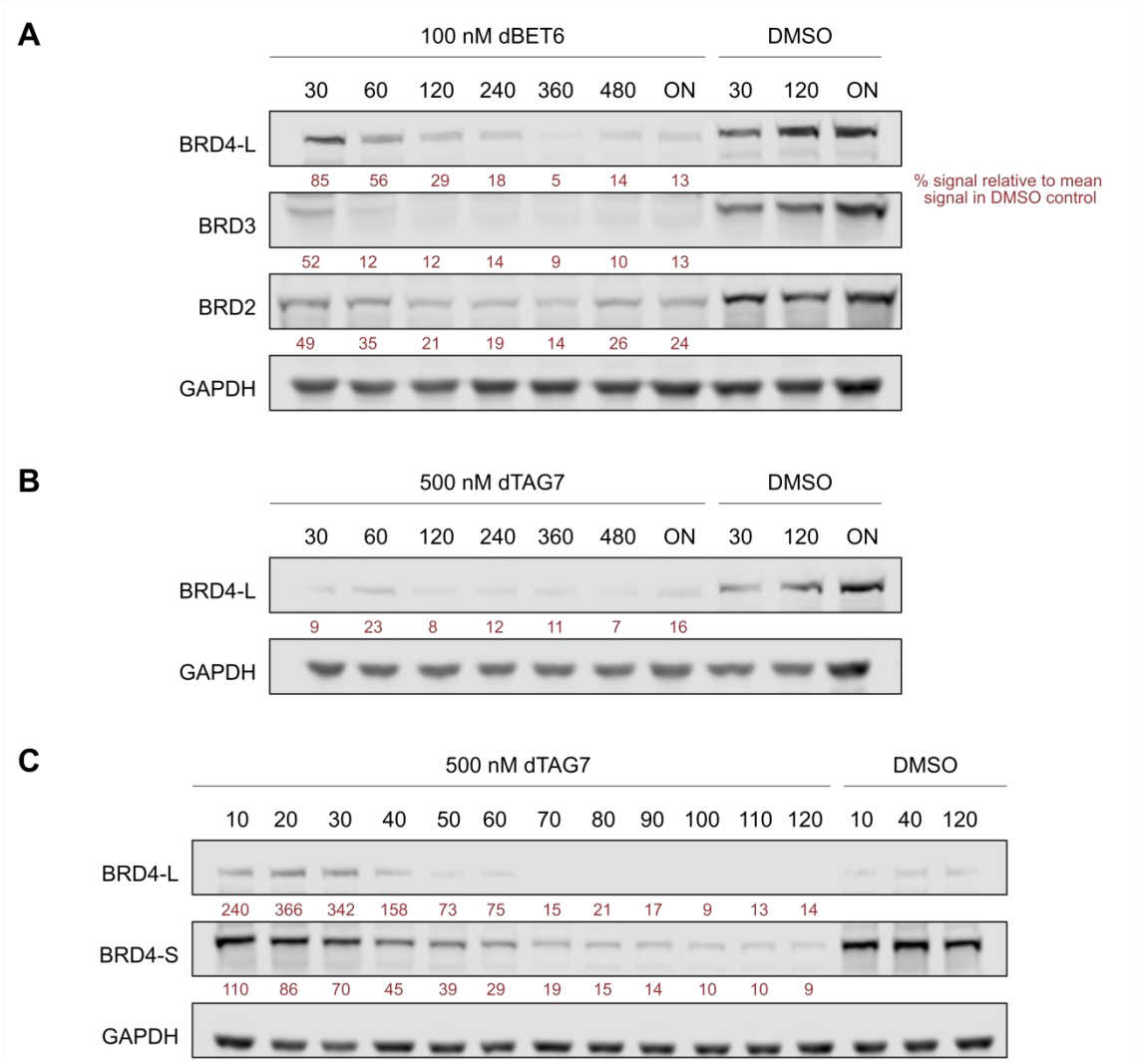


Figure 9: dBET6 and dTAG7 treatment lead to rapid degradation of the target proteins. (A) Immunoblot for BRD2, BRD3 and BRD4 upon treatment with 100 nM dBET6 for 30 min up to overnight, or under control conditions (DMSO), in K562 dTAG-BRD4 cells. dTAG-BRD4 was detected using an HA-specific antibody. GAPDH was included as a loading control and was also used to normalize the intensity values. The percentages below the blot refer to the signal intensity per lane relative to the mean signal intensity of all DMSO control lanes. **(B)** Immunoblot for BRD2, BRD3 and BRD4 upon treatment with 500 nM dTAG7 for 30 min up to overnight. Quantification was done as described under (A). **(C)** Immunoblot for the short and long isoforms of BRD4 after dTAG7 treatment. Samples were collected every 10 min. Quantification was done as described under (A).

Based on these experiments, we assumed similar degradation kinetics of BRD4 with either treatment. As a compromise between BRD4 reduction and treatment time, we chose to treat the K562 cells for 120 min in the following experiments, unless stated otherwise.

To prove the specificity of the degrader, we performed a SILAC-based whole cell MS experiment after 120 min of dTAG7 treatment (see 3.2.10.1). No significant changes were

seen for 2,814 of the robustly detected proteins ($p \geq 0.05$, corresponding to $-\log_{10} p \leq 1.30103$). In contrast, 38 and 30 proteins were up- or down-regulated, respectively. With a decrease of 35% and a p-value of 0.00556, BRD4 showed the largest significant reduction of all detected proteins (Figure 10). In contrast, the levels of BRD2 and BRD3 were not altered. The discrepancy in the remaining BRD4 levels as measured by different assays could be explained by the fact that immunoblotting only quantified BRD4 with a yet intact N-terminus, while whole cell mass spectrometry quantified peptides also of BRD4 degradation intermediates. Notably, none of the other significantly affected proteins had a known function in transcription, as was also confirmed by gene ontology analysis (not shown).

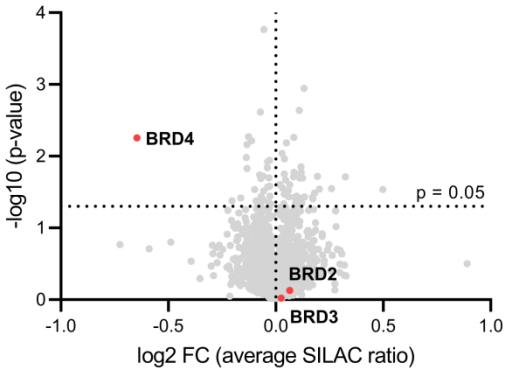


Figure 10: dTAG7 treatment specifically depletes BRD4 but does not affect BRD2 or BRD3. K562 dTAG-BRD4 cells were treated with dTAG7 for 120 min and subjected to SILAC-assisted whole cell mass spectrometry. Statistical significance of the fold-change was determined by two-tailed one-sample t-test. Four biological replicates were analyzed.

4.1.2 The degradation tag does not impact the cellular localization of BRD4 or cell proliferation

Degradation of both BRD4 isoforms required the degron tag to be inserted at the N-terminus of the protein, which on the genomic level could interfere with communication between the promoter elements of the BRD4. Additionally, despite its small size of 17 kDa, the tag could potentially affect the stability of BRD4 or its intracellular localization. Similar effects have been described by others [421]. Therefore, we fractionated monoclonal K562 dTAG-BRD4 cells as well as untagged cells of the parental K562 cell line and quantified BRD4 across the cellular compartments. Indeed, dTAG-BRD4 levels were roughly 50% lower as compared to the levels of untagged BRD4 in the parental cell line

(Figure 11A). However, the relative local distribution seemed only mildly affected, with 60% and 50% of BRD4 localized at the chromatin in the wild-type and the tagged cell line, respectively.

Despite the decrease in BRD4 levels, the proliferation capacity of K562 dTAG-BRD4 cells was not reduced (Figure 11B). Repeated cell counting in biological and technical replicates over the course of 72 h revealed a median doubling time of 24.8 and 24.7 h for K562 wild-type and K562 dTAG-BRD4 cells, respectively.

To rule out that the degron tag triggered apoptosis, we performed a dTAG7 time course experiment and probed for the cysteine protease caspase 3 and its cleavage product, a well-established apoptosis marker. However, only the full-length protein but not the fragment was detected even upon extended treatment times (Figure 11C).

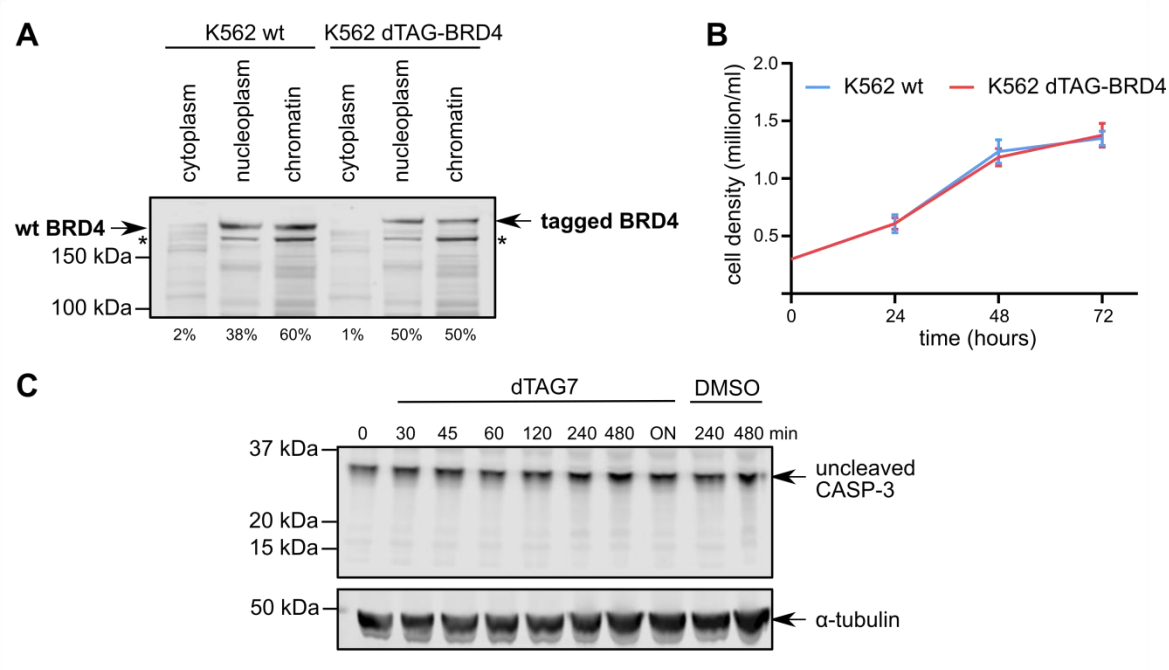


Figure 11: The dTAG neither affects intracellular localization of BRD4 nor cell viability. **(A)** Cellular distribution of untagged BRD4 in K562 wild-type (wt) and degron-tagged BRD4-L in K562 dTAG-BRD4 cells. dTAG-BRD4 runs slightly higher on the gel than wild-type BRD4, which corresponds to the increase in mass by 17 kDa. As the antibody recognized an epitope near the C-terminus of BRD4 (between the amino acids 1,312 and 1,362), the short isoform BRD4-S could not be detected. The asterisk marks an unspecific band. Percentages indicate the fraction of the BRD4 signal in the respective cellular fraction relative to the sum of the signal in cytoplasm, nucleoplasm and chromatin. **(B)** Proliferation curves of K562 wt and K562 dTAG-BRD4 cells. Error bars indicate the standard deviation. Cells were counted in six biological replicates. **(C)** Immunoblot for caspase 3. Only the full-length caspase 3 but not its cleavage product of 17 to 19 kDa, which is a marker for apoptosis, was detected, showing that even an extended degrader treatment does not elicit this adverse effect. α-tubulin is shown as a loading control.

4.1.3 BET PROTEIN DEGRADATION GLOBALLY DISRUPTS TRANSCRIPTION ELONGATION

A profound impact of dBET6 treatment on transcription has been described in MOLT4 cells [135], but the molecular mechanism remained largely unknown. To characterize the effect of BET ablation in our cellular system at high spatial and temporal resolution, we treated K562 dTAG-BRD4 cells with dBET6 or DMSO for 120 min and performed a NET-seq experiment that relied on spiked-in mouse cells for normalization (SI-NET-seq). The use of spike-ins allowed for a quantitative analysis and was considered crucial, as the previous study has indicated an effect on global transcription, *i.e.*, the assumption of a group of genes that was unaffected by the treatment and therefore could be used for normalization would not hold true.

To distinguish effects on early and late elongation, we defined the promoter-proximal region as ranging from the TSS to the TSS +300 bp. The remaining gene region until the pA site was classified as the gene body. Gene body occupancy of actively transcribing Pol II was significantly reduced at most genes, *i.e.*, at 6,655 out of 9,404 active genes. Conversely, a significant increase in gene body occupancy was seen only at 26 loci. A decrease was also evident in the promoter-proximal region of most genes. In summary, 71% of the genes included in this analysis showed a decrease in Pol II occupancy in both the promoter-proximal and the gene body region (Figure 12A). In case of 28% of the genes, the decrease was confined to the gene body, while Pol II accumulated in the promoter-proximal region.

Next, we aimed at understanding which features of a gene might underlie its dBET6 sensitivity. A GO analysis did not reveal an enrichment of specific terms linked to biological processes, indicating heterogeneity of the group of affected genes (not shown). The decrease in Pol II occupancy at the gene body was comparable for protein-coding and long non-coding RNA genes, while no significant occupancy change was seen for histone genes (one-sample Mann–Whitney U test, $p = 0.7086$) (Figure 12B). Visual inspection revealed that exemplary histone genes showed a distinct Pol II binding profile, such that Pol II binding upon treatment appeared to increase in the promoter-proximal and the gene body region (Supplemental Figure 1). Therefore, we focused on protein-coding and lncRNA genes for all following analyses.
Interestingly, with a median length of 42.8 kb, genes showing a significant decrease in gene body occupancy were significantly longer than the remaining genes, which had a

median length of 14.0 kb (Mann–Whitney U test, $p < 0.0001$) (Figure 12C). This trend was not caused solely by histone genes, although genes of <5 kb indeed showed the smallest reduction in gene body occupancy. Instead, the Pol II occupancy change was also significantly less pronounced for genes 5 to 20 kb in length than for genes longer than 20 kb.

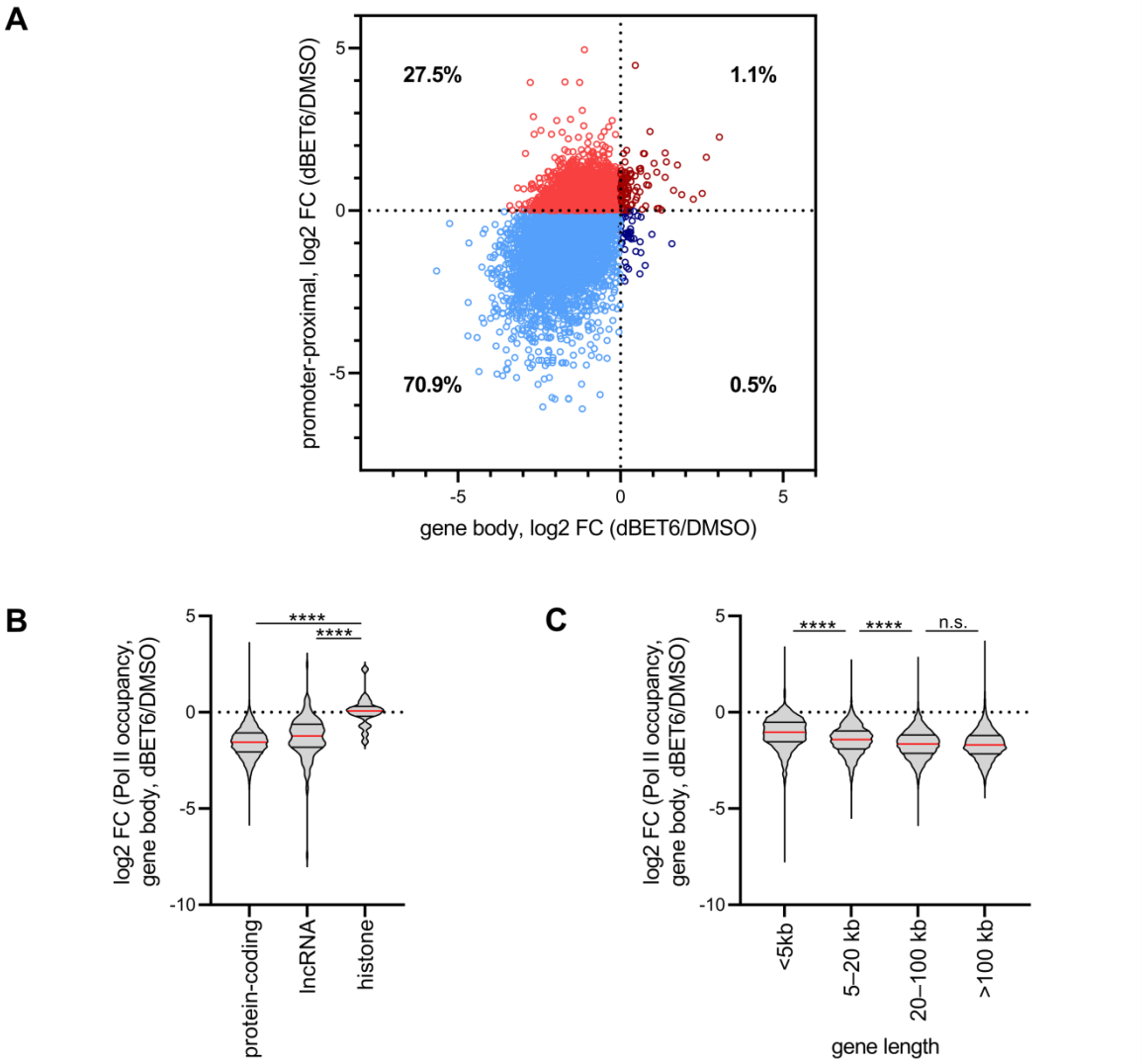


Figure 12: 120 min of dBET6 treatment leads to the global loss of transcribing Pol II from the promoter-proximal and the gene body region. (A) Comparison of the effects of dBET6 treatment on the promoter-proximal and the gene body region, as determined by SI-NET-seq. The promoter-proximal region was defined as ranging from the TSS to TSS +300 bp, the gene body region as ranging from TSS +300 bp to the main annotated pA site. Data of one representative replicate (out of two) are shown. 8,071 genes displayed Pol II binding in both regions and were therefore considered in this analysis. A reduction of Pol II in both regions was observed in case of 71% of these genes. **(B)** dBET6 treatment reduced Pol II occupancy at the gene body of protein-coding and lncRNA but not histone genes.8,735 protein-coding, 441 lncRNA and 19 histone genes were considered. **(C)** The decrease in Pol II occupancy in the gene body region correlated with gene length for genes of ≤100 kb in length. 708 genes shorter than 5 kb, 2,526 genes of 5 to 20 kb, 4,403 genes of 20 to 100 kb and 1,699 genes of >100 kb were included in the analysis. A Mann–Whitney U test was performed to determine the significance of the differences between the median values (red line). ****: $p \leq 0.0001$; n.s. (not significant): $p > 0.05$.

4.1.4 BRD4 DEGRADATION IMPAIRS THE RELEASE OF PROMOTER-PROXIMALLY PAUSED POL II

It remained unclear from the previous experiment which of the BET proteins was mainly responsible for the observed effect. Given the recent evidence for a role in transcription elongation [325], we focused on BRD4, induced its degradation by dTAG7 treatment and performed SI-NET-seq after 120 min of treatment.

For 2,834 of the 7,574 genes considered in the analysis, we observed a significant reduction of actively elongating Pol II in the gene body region, indicating that also dTAG7 treatment globally disrupts elongation. However, with a median log2 fold-change of -0.6416 corresponding to a decrease by >35%, the decrease was less severe than the decrease seen upon dBET6 treatment (median log2 fold-change = -1.546 ≙ decrease by >65%). Thus, the similarity of the effects indicated that specifically BRD4 is required for efficient transcription, although a contributory role of BRD2 and BRD3 appeared probable. As in the case of dBET6 treatment, Pol II occupancy was reduced at the gene body of protein-coding and lncRNA genes but not of histone genes (not shown).

Strikingly, for a majority of 65% of genes, we observed a decrease in Pol II occupancy at the gene body but an increase in the promoter-proximal region (Figure 13A). A decrease in both regions was only seen at 31% of the genes. Thus, while the effect of BRD4- and pan-BET degradation on Pol II occupancy at the gene body was comparable, a main difference could be seen in the promoter-proximal region of protein-coding and lncRNA genes: here, actively elongating Pol II accumulated upon dTAG7 treatment, but was depleted upon dBET6 treatment (Figure 13B). This finding suggested that Pol II in the absence of BRD4 could initiate transcription and undergo promoter-proximal pausing but could not exit from the pause state to resume elongation. In contrast, in the absence of all BET proteins, transcription appeared to be dysregulated additionally at an earlier step, possibly during initiation.

Notably, spike-in normalization was a prerequisite to detect this difference in the promoter-proximal region, as it allowed us to directly and quantitatively compare the effects across samples. Alternative measures*, e.g.*, the pausing index, defined as the ratio of Pol II occupancy in the promoter-proximal and the gene body region, could indicate a defect in pause release but would have failed to distinguish the scenarios seen upon dBET6 and dTAG7 treatment.

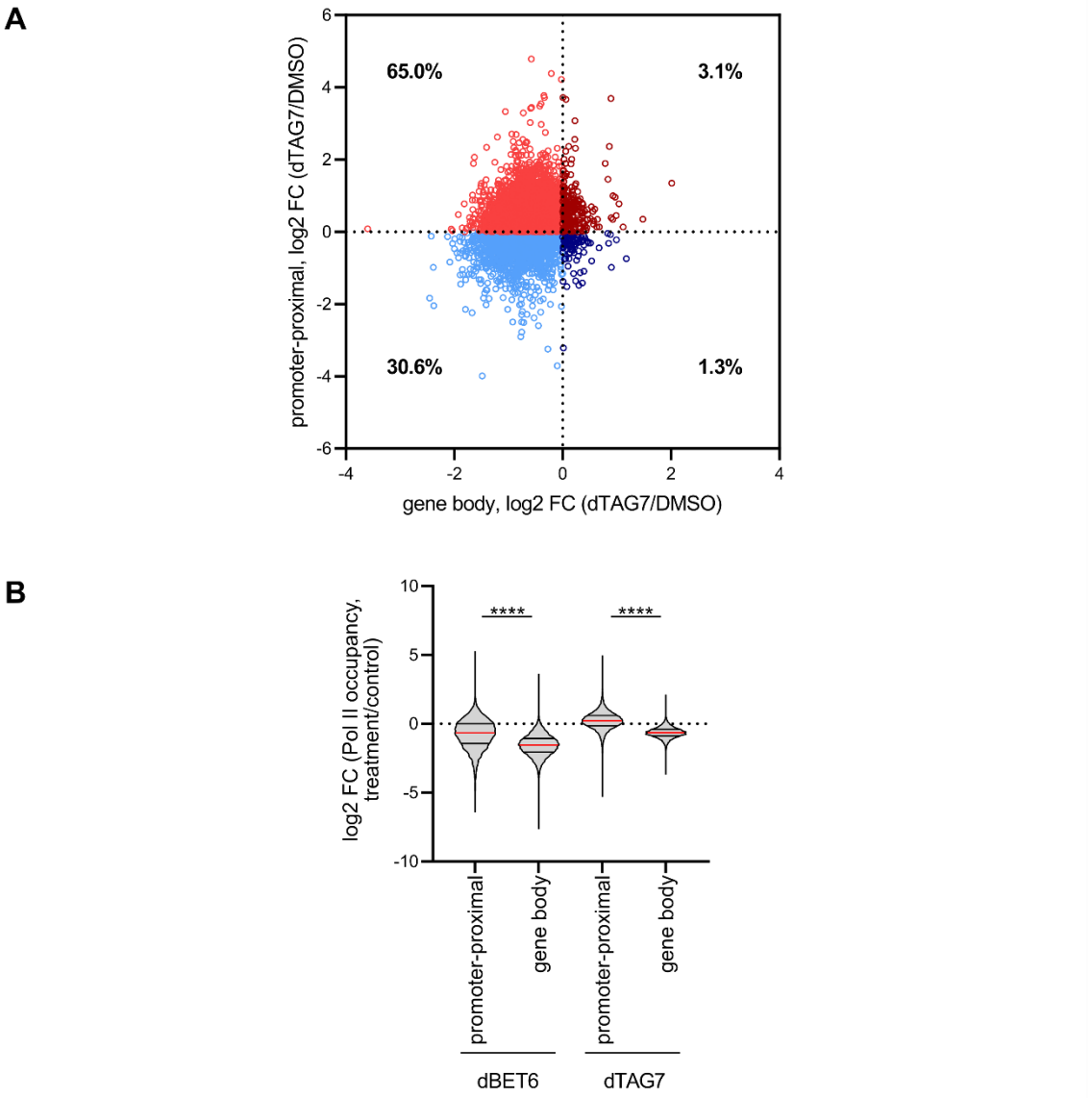


Figure 13: 120 min of dTAG7 treatment globally results in an increase in transcribing Pol II in the promoter-proximal region and a decrease at the gene body. (A) The promoter-proximal region was defined as ranging from the TSS to TSS +300 bp, the gene body region as ranging from TSS +300 bp to the pA site. Data of one representative SI-NET-seq replicate (out of two) are shown. 6,915 genes showed Pol II binding in both regions and were therefore considered in this analysis. **(B)** Comparison of the Pol II occupancy change in the promoter-proximal and the gene body region upon BET and BRD4 degradation. A Mann–Whitney U test was performed to determine the significance of the differences between the median values (red line). ****: $p \leq 0.0001$.

4.1.5 The elongation defect is already apparent after 50 min of dBET6 treatment

A time course experiment was performed to investigate if the observed reduction in elongating Pol II represented a dynamic or an equilibrium state. In particular, K562 dTAG-BRD4 cells were treated with dBET6 for 40 and 50 min and subjected to SI-NET-seq. A

significant decrease in the median Pol II occupancy in the gene body region was visible already after 40 min of BET degradation ($p < 0.0001$; one-sample Mann–Whitney U test), but was more pronounced after 50 min of treatment, indicating a highly dynamic time frame (Figure 14A). A further decrease, accompanied by a greater variability of the values, was seen towards the 120 min time point. The early onset of the elongation defect indicated that it was likely a direct consequence of BET degradation.

An moderate increase in effect strength was also visible at the gene level, with 2,670 and 2,834 genes of the >7,500 genes included in this analysis showing a significant Pol II reduction after 50 and 120 min of BET ablation, respectively (Figure 14B). Likely because no matched DMSO control sample was sequenced, so that a 60 min DMSO control data set generated on the same day had to be used for comparison, almost no genes reached significance after 40 min of treatment. A statistical overrepresentation test on the genes showing a decrease in Pol II occupancy revealed no link to specific biological processes (not shown). Also, no significant GO term association was found for the small group of upregulated genes.

Interestingly, with regard to both significance and fold-change, elongation at the c-*MYB* locus was most dramatically decreased across all tested time points. The decrease in Pol II occupancy was confirmed by visual inspection of the read distribution (Figure 14C). This observation agreed with the high sensitivity of some leukemic TFs to pan-BET inhibition or degradation [135,305].

The analysis of the distribution of elongating Pol II along the transcribed region uncovered a “wave” which likely corresponded to Pol II molecules that had entered active elongation before BET protein levels fell below a critical threshold (Supplemental Figure 2A). Comparing the position of the back edge of the transcription wave across the different time point data sets, it appeared to progressively move along the gene body (Supplemental Figure 2B). Similar waves of Pol II have been described earlier, particularly upon DRB treatment, and were indicative of impaired pause release [66].

Furthermore, the time course data showed that the elongation defect, as observed by SI-NET-seq, was not an equilibrium but became more severe over time.

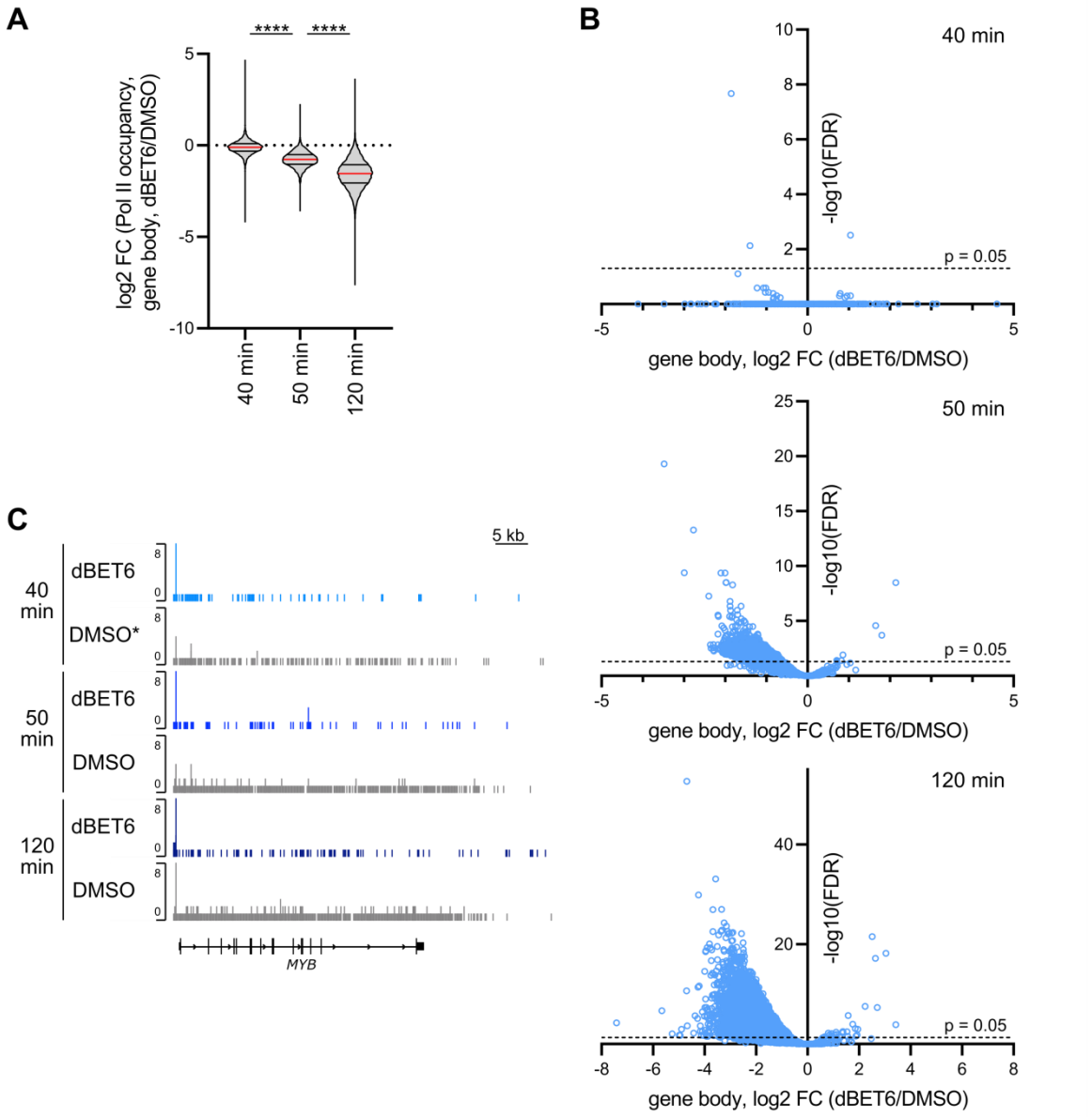


Figure 14: The elongation defect manifests already after short dBET6 treatments. **(A)** Quantification of the decrease in gene body occupancy across a dBET6 time course. 9,682, 7,761 and 9,404 genes were considered for each time point. Statistical significance of the difference between the median of the samples (red line) was determined by Mann–Whitney U test. ****: p ≤ 0.0001. The data represent one replicate out of two. **(B)** Volcano plot representation of the change in Pol II occupancy per gene after 40, 50 and 120 min of dBET6 treatment, as determined by SI-NET-seq. **(C)** Spike-in-normalized SI-NET-seq signal in reads per bp at the *MYB* locus. As no 40 min DMSO treatment was performed, the signal track of a 60 min DMSO experiment is shown. For clarity, reads mapping to the antisense strand are omitted.

4.1.6 BET PROTEIN AND BRD4-SELECTIVE DEGRADATION GLOBALLY DISRUPT TRANSCRIPTION TERMINATION

Strikingly, we also found transcription termination to be strongly affected by both treatments. SI-NET-seq revealed massive readthrough transcription upon BET and

BRD4-selective ablation, *i.e.*, Pol II upon treatment continued transcribing tens of kilobases beyond the termination zone as defined under control conditions (Figure 15A, B, C).

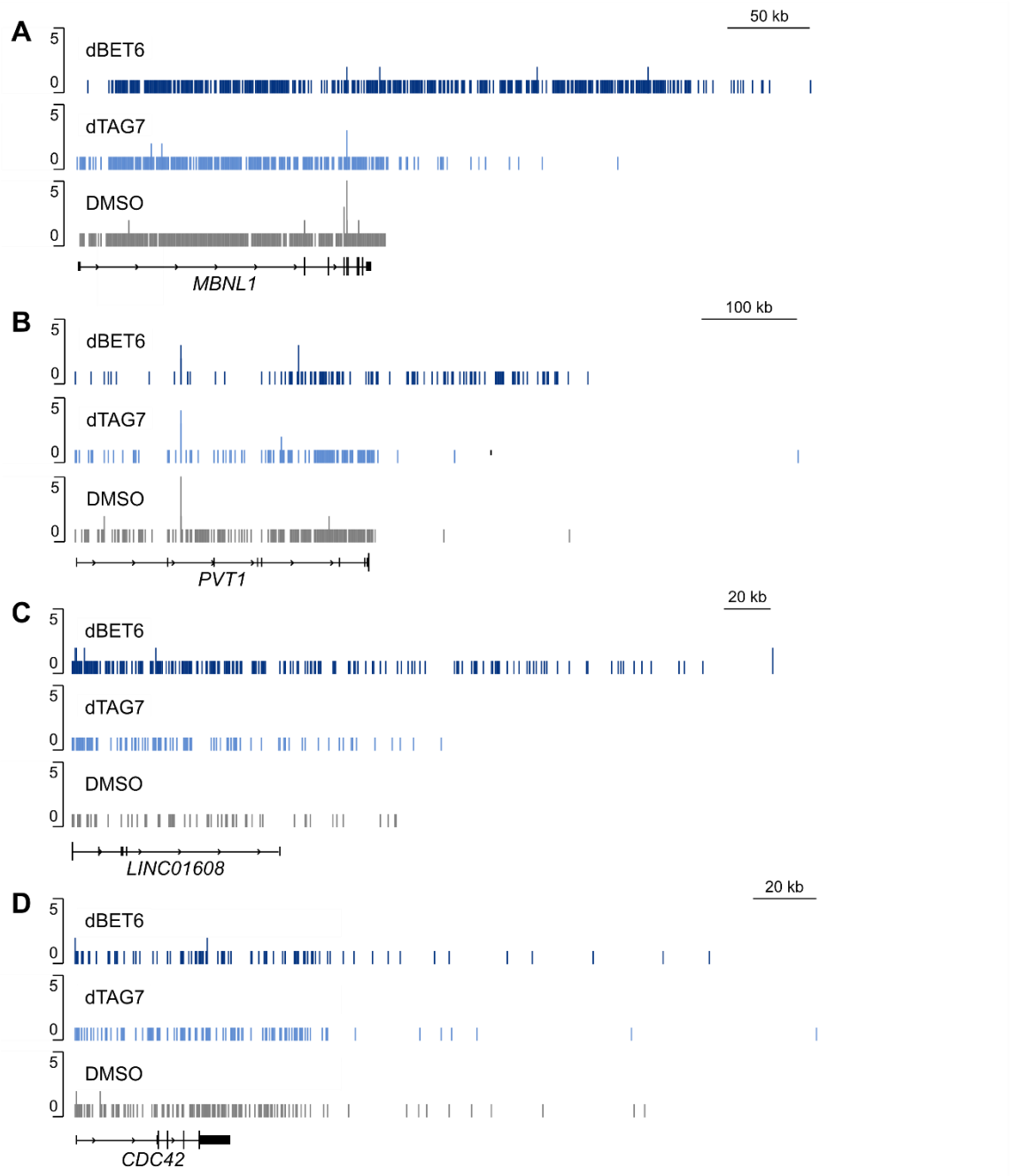


Figure 15: 120 min of treatment with dBET6 or dTAG7 results in a termination defect. (A) Spike-in-normalized SI-NET-seq signal in reads per bp at the examplary protein-coding *MBNL1* locus. **(B)**, **(C)** Spike-in-normalized SI-NET-seq signal in reads per bp at the non-coding *PVT1* and *LINC01608* loci. **(D)** NET-seq signal at the *CDC42* locus which did not display a termination defect and therefore was chosen as a negative control in the cleavage assay (see 4.1.7). All genes are transcribed from left to right in this representation. The data were visualized using the IGV genome browser [399].

To quantify the extent of readthrough transcription, we developed the readthrough index (RTI), a measure defined as the change of the average termination distance, *i.e.*, the weighted distance of Pol II from the last active pA site, under treatment and control conditions (see also 3.2.4.3). Thus, it took into account that termination occurs in a variable distance from the pA site even under control conditions [69,176].

Upon dBET6 treatment, an RTI of ≥5.0 kb was found for 2,865 genes, while the median RTI of these genes was 9.3 kb (Figure 16A). A qualitatively similar but quantitatively weaker effect was seen upon dTAG7 treatment. In particular, 1,154 genes displayed an RTI of ≥5.0 kb and a median shift of the termination zone by 7.6 kb (Figure 16B). However, the effects were well correlated at the gene level (Pearson correlation coefficient r = 0.7095). Impaired termination at overlapping sets of genes upon pan-BET and BRD4-selective degradation indicated a common underlying molecular mechanism. Moreover, the data showed that loss of BRD4 was sufficient to cause dysregulation of termination, although an additional role of BRD2 and BRD3 could not be excluded. Notably, a weak termination defect was already seen after 40 and 50 min of dBET6 treatment, arguing for a direct rather than an indirect effect (Figure 16C).

Readthrough transcription was observed for protein-coding and long non-coding genes (Figure 15B, C), *i.e.,* genes terminating in a polyadenylation signal-dependent manner, but not for histone genes. When considering only genes with an RTI of ≥5 kb, lncRNA genes showed a significantly stronger termination defect than protein-coding genes (Figure 16D). Moreover, similarly to the elongation defect, the termination defect was correlated with gene length. The median RTI was significantly higher for genes of >100 kb than of 20 to 100 kb length irrespectively of the degrader used (Figure 16E). Also, genes of 20 to 100 kb length had a significantly higher median RTI than genes of 5 to 20 kb upon dBET6 but not dTAG7 treatment.

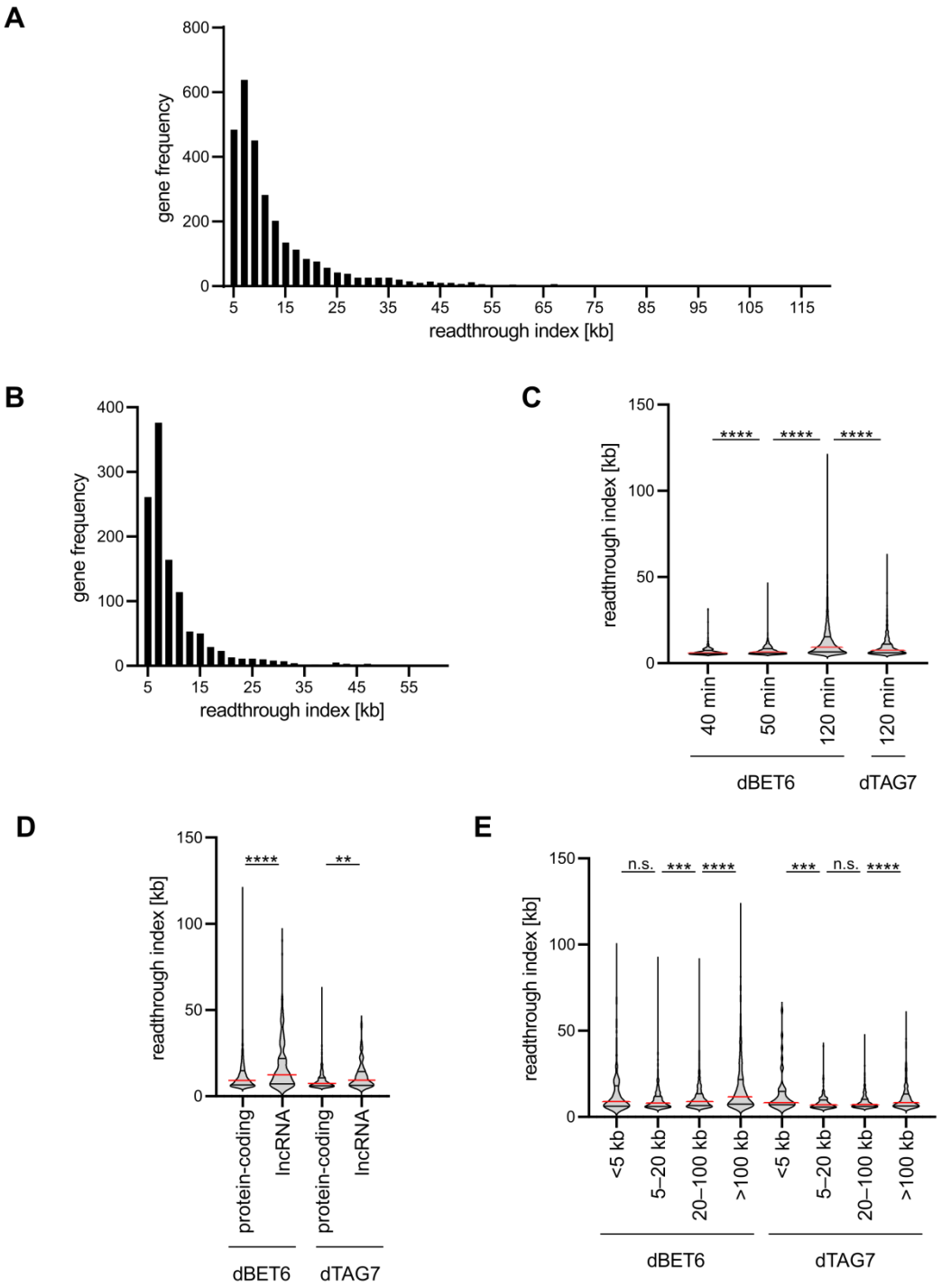


Figure 16: The readthrough index (RTI) is increased upon dBET6 and dTAG7 treatment and is associated with specific gene features. (A) Distribution of the RTI upon 120 min of dBET6. 2,865 genes with an RTI of ≥5 kb were considered. **(B)** Distribution of the RTI upon 120 min of dTAG7 treatment. 1,154 genes with an RTI of ≥5 kb were considered. **(C)** Quantitative comparison of the RTI at different timepoints of dBET6 or dTAG7 treatment. **(D)** Gene biotype-dependent differences in the RTI. Gene annotation was obtained from Ensembl. The analysis was performed based on 2,615 and 1,031 protein-coding genes as well as 117 and 94 lncRNA genes present in the dBET6 or the dTAG7 data set, respectively. **(E)** Correlation of the RTI with gene length. Statistical significance of the differences of the median values was determined by Mann-Whitney test. ****: $p \leq 0.0001$: ***: $p \leq 0.001$; **: $p \leq 0.01$; *: $p \leq 0.05$; n.s.: $p > 0.05$.

4.1.7 EFFICIENT 3' END CLEAVAGE OF THE NASCENT RNA IS IMPAIRED UPON BET AND BRD4-SELECTIVE DEGRADATION

Transcription termination is intimately coupled with 3' end processing of the nascent transcript in higher eukaryotes [422]. Given that particularly protein-coding and lncRNA genes showed BET-dependent termination defects, we decided to focus on polyadenylation signal-dependent 3' processing. According to the torpedo but not the allosteric model, cleavage of the nascent transcript by CPSF73 is immediately required to create an entry site for the exonuclease XRN2. Thus, we hypothesized that defective 3' end processing accompanied or even mechanistically underlay the termination defect.

In order to detect and quantify those transcripts that were not cleaved at the pA site, we designed a cleavage assay similar to an experiment by Davidson et al. [225]. Therefore, we used RT-qPCR primers binding shortly upstream and downstream of the main annotated pA sites of the genes *MBNL1* and *PVT1*, which showed strong readthrough transcription upon both BRD4 and BET degradation, and the control gene *CDC42* (Figure 15A, B, D). Additional primer pairs were designed to bind upstream and downstream of the pA site to determine nascent transcription levels at the gene body and readthrough activity, respectively.

Upon BET degradation, significantly more nascent transcripts were found spanning the pA site of *MBNL1* and *PVT1* compared to DMSO-treated or untreated cells (Figure 17A). In contrast, at the *CDC42* locus, no such relative increase was seen. Notably, robust signal detection at the pA site of all tested genes under control conditions possibly was due to not yet cleaved transcripts. Likewise, we performed the cleavage assay upon selective degradation of BRD4. While treatment had no significant effect on *CDC42* and *MBNL1*, higher RNA levels were found for the regions around the pA site of *PVT1* and downstream (Figure 17B). Thus, while the effect of dTAG7 treatment was somewhat weaker than the effect of dBET6, this experiment demonstrated that BRD4 is required for efficient 3' cleavage of the nascent RNA.

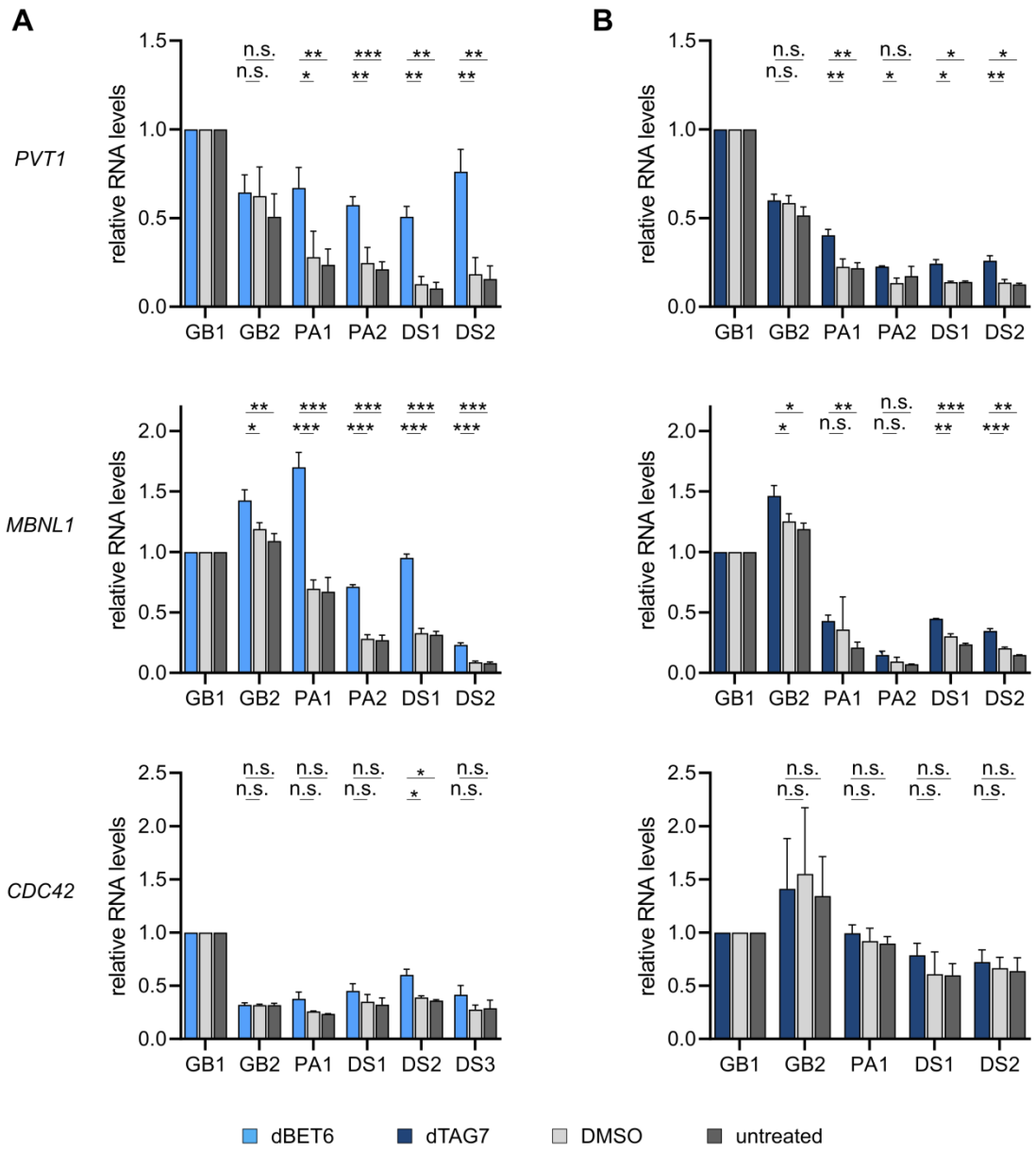


Figure 17: BET and BRD4-selective degradation result in a defect in RNA 3' end cleavage. **(A)** Quantification of the effect of dBET6 treatment (120 min) on 3' RNA cleavage at the exemplary loci of *PVT1* and *MBNL1*, which based on SI-NET-seq showed readthrough transcription, and *CDC42* that served as a negative control. Nascent RNA was enriched by cell fractionation and reverse-transcribed using random hexamer primers. The cDNA was used as the template for RT-qPCR. Primers were designed to bind flanking the main pA site (PA), upstream at the gene body region (GB) or further downstream (DS) (see Figure 8). To account for differences in transcriptional activity, all C_t values were normalized by the C_t values at the most upstream gene body position (GB1). Thus, the plots show ΔC_t values. Per condition, three biological replicates and three technical replicates were performed. Error bars indicate the standard deviation. **(B)** Quantification of the effect of dTAG7 treatment (120 min) on 3' RNA cleavage. Significance was determined by an unpaired t-test with Welch correction. ***: FDR-adjusted p-value ≤0.001; **: adjusted p-value ≤0.01; *: adjusted p-value ≤0.05; n.s. (not significant): adjusted p-value >0.05.

4.1.8 dBET6 AND dTAG7 TREATMENT RESULT IN THE DISPLACEMENT OF TRANSCRIPTION ELONGATION FACTORS FROM THE CHROMATIN

In view of the effects on different phases of transcription, we hypothesized that the recruitment of proteins involved might be disturbed in the absence of BET proteins or BRD4. To determine the effects of BET protein depletion on the chromatin proteome in an unbiased and quantitative manner, we based on the cell fractionation protocol used for NET-seq developed chromatin-MS. Cell fractionation was performed in the presence of protease and phosphatase inhibitors to preserve protein–protein interactions, followed by benzonase treatment to completely solubilize the native chromatin and by a precipitation step to remove mass spectrometry-incompatible detergents. While the destabilization of certain proteins at the chromatin under the chosen conditions could not be excluded, it was not expected to impair comparative analyses, as destabilization would affect all conditions. SILAC labeling of the cells and combining the “heavy” and “light” cells subjected to different treatments prior to cell fractionation, *i.e.*, in an early step of the procedure, minimized technical variation and improved sensitivity.

To assess the quality of the chromatin purification, we performed a “cellular component”-focused GO analysis with Bonferroni correction on all proteins detected in the dBET6 experiment. 951 of the detected proteins were associated with the term “nucleus” ($p = 3.8 * 10^{-124}$), while 726 proteins were associated with “cytoplasm” ($p = 8.21 * 10^{-21}$). Proteins with no expected function in the chromatin were also found enriched in chromatin samples prepared using different protocols [423], indicating that cytoplasmic contaminations were to be expected. Furthermore, it is increasingly acknowledged that proteins could serve functions in several cellular compartments [423,424]. Probably due to similar sedimentation properties, also 82 proteins linked to the ribosome ($1.52 * 10^{-27}$), 105 mitochondrial proteins ($p = 5.39 * 10^{-04}$) and 82 proteins of the cell envelope ($p = 1.12 * 10^{-07}$) were found. However, the group of significantly depleted proteins was not enriched in proteins classified as cytoplasmic, indicating that putative cytoplasmic contaminants only contributed to the background but not to the subset of regulated proteins that were of most interest.

Treatment with dBET6 led to a significant decrease in the chromatin-bound amounts of 146 proteins (7%) of the 2,219 proteins that were robustly detected, while the amount of 43 proteins (12%) was significantly increased. Upon dTAG7 treatment, 76 proteins (8%) out of 964 were significantly depleted from the chromatin, while 28 proteins (3%) were

significantly enriched. With regard to the fold-change, dBET6 had a more pronounced effect than dTAG7. Notably, given the short treatment time, gene expression changes were unlikely to be visible on the proteome level. Instead, a reduction in signal indicated failure to recruit a protein to or maintain its binding at the chromatin. Also, treatment-dependent changes in posttranslational modifications affecting the detectability of the protein could not be excluded.

BRD4 was decreased by 72% and thus showed the most robust significant reduction of all detected proteins after 120 min of dBET6 treatment but was not detected upon 120 min of dTAG7 treatment, probably for technical reasons. Likewise, neither BRD2 nor BRD3 was detected in either of the two experiments. In both experiments, several subunits of Pol II appeared as overall most significantly depleted upon treatment (Figure 18A, B). Strikingly, the DSIF subunit SPT5 and also SPT6 were depleted to a comparable or even greater extent than the Pol II subunits upon either treatment. Also, SPT4, which was only detected in the dBET6 experiment, showed a similar fold-change. In contrast, NELFA and NELFCD of the negative elongation factor were detected upon dBET6 treatment, but their levels did not change significantly. Among the most significantly depleted proteins in both experiments were also core subunits of the PAF complex, particularly PAF1, CDC73 and CTR9. Of the more loosely associated PAF subunits LEO1 and RTF1, which, however, are necessary for PAF associating with chromatin [249], only LEO1 was detected, and was found to be significantly depleted upon dBET6 treatment.

Given the debated hypothesis that BRD4 recruits P-TEFb, the proteins of this complex were of particular interest. While cyclin T1 was not detected in any of the experiments and HEXIM1 localization was not changed upon BET degradation, CDK9 remained undetectable upon dTAG7 treatment but was significantly enriched in the chromatin sample upon dBET6 treatment. Moreover, subunits of the super-elongation complex were not detected.

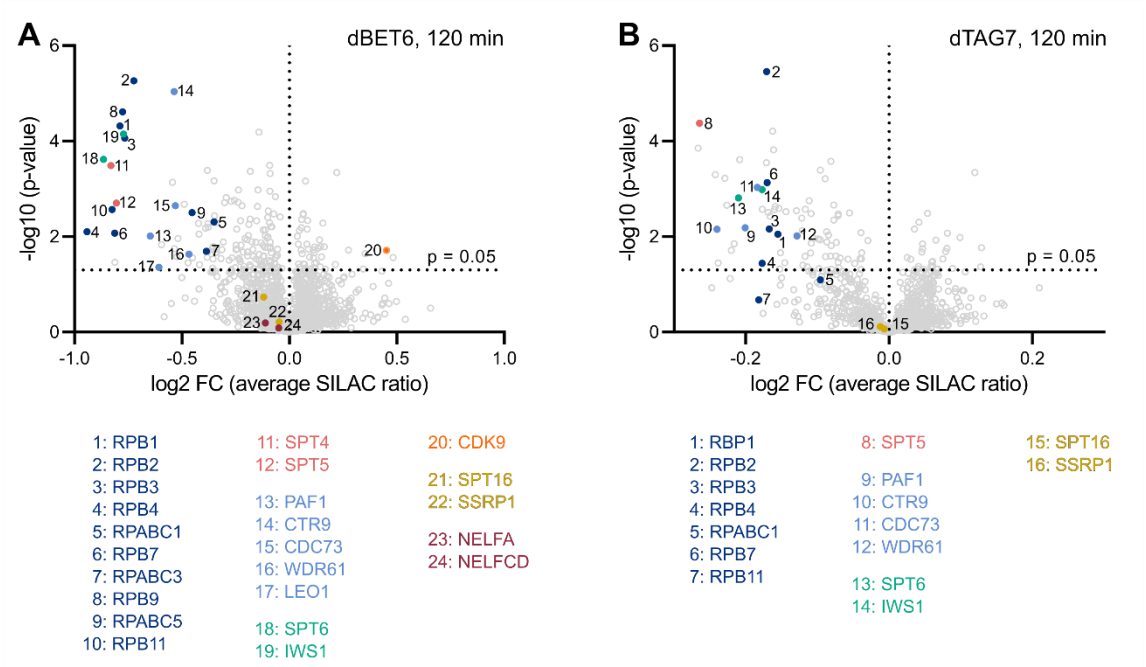


Figure 18: BET and BRD4-selective degradation displace Pol II and key elongation factors from the chromatin. (A) Effect of 120 min of dBET6 treatment on chromatin binding, as determined by quantitative (SILAC-based) chromatin-MS. Differently SILAC-labeled cells were treated with dBET6 or DMSO, combined after treatment and fractionated, followed by MS analysis. 2,219 proteins or protein groups were detected. Relevant proteins are color-coded by complex. For reasons of clarity, BRD4 (log2 FC (average SILAC ratio) = -1.83) is omitted from the graph. **(B)** Effect of 120 min of dTAG7 treatment. 964 proteins or protein groups were detected in this experiment. Significance was determined by a two-tailed one-sample t-test.

4.1.9 BET AND BRD4-SPECIFIC DEPLETION PREVENT THE RECRUITMENT OF 3' END PROCESSING FACTORS GENOME-WIDE

In view of the defect in RNA 3' end processing, we used our chromatin-MS data to assess the possibility that proteins involved in 3' end processing were not efficiently recruited. Particularly, both treatments resulted in a significant reduction of the CPSF proteins CPSF160, CPSF100, CPSF73 and CPSF30 from the chromatin (Figure 19A, B). FIP1 and WDR33 were significantly depleted only upon BRD4-selective and pan-BET degradation, respectively. CstF64 and CstF77 of the CstF complex were significantly depleted under both treatment conditions. In contrast, the reduction of CstF50 was significant only in the dBET6 experiment. Overall, the effects on subunits of the central CPSF and CstF modules were very consistent across treatments.

By contrast, few significant changes were observed for CFIm and CFIIm: CFIm25 was significantly depleted upon dTAG7 treatment but, like CFIm68 and PCF11, showed no

change upon dBET6 treatment. Likewise, XRN2 was not changed upon dTAG7 but was mildly but significantly reduced upon dBET6 treatment.

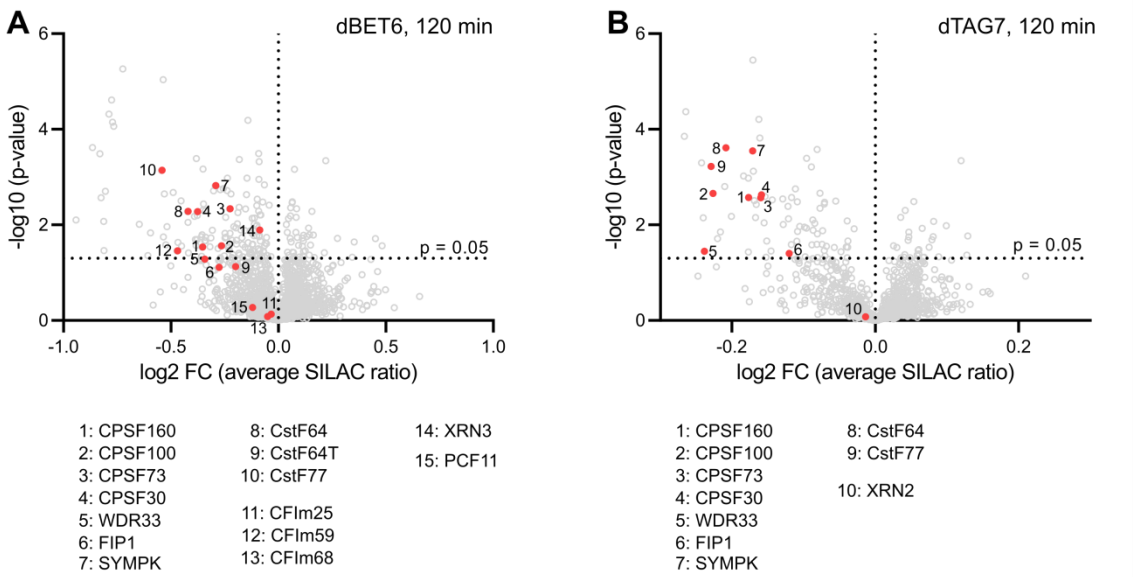


Figure 19: BET and BRD4-selective degradation reduce the chromatin association of RNA 3' end processing factors. (A) Effect of 120 min of dBET6 treatment on chromatin binding, as determined by quantitative (SILAC-based) chromatin-MS. 2,219 proteins were detected in this experiment. Proteins are grouped by complex. **(B)** Effect of 120 min of dTAG7 treatment. In total, 964 proteins were detected. Significance was determined by a two-tailed one-sample t-test.

For an unbiased view on the effects of BET and BRD4 depletion on chromatin composition, we performed a GO term analysis on the subsets of proteins that were significantly depleted or enriched upon treatment. Proteins lost from the chromatin upon dBET6 treatment were enriched in terms linked to "RNA 3' end processing" ($p = 3.55 * 10^{-3}$), "mRNA splicing, via spliceosome" ($p = 8.01 * 10^{-11}$), "transcription by RNA polymerase II" ($p = 6.31 * 10^{-3}$) but also "rRNA processing" ($p = 9.7 * 10^{-4}$). A similar set of terms was linked to the proteins depleted by dTAG7 treatment.

4.1.10 BET DEPLETION DOES NOT ALTER CHROMATIN ASSOCIATION OF GENERAL TRANSCRIPTION FACTORS AND INTEGRATOR SUBUNITS

Our chromatin-MS data sets allowed us to determine the effect of rapid BRD4 or BET depletion on various proteins known or implicated in transcription regulation. For instance, BRD4 has been suggested to activate transcription by directly recruiting GTFs,

particularly TFIID, and Mediator [324]. However, none of the subunits of TFIID, although robustly detectable at least in the dBET6 experiment, were significantly changed in amounts (Supplemental Table 1). A significant but small increase of the TFIIH subunits GTF2H2 and GTF2H3 was seen upon dBET6 treatment, while CDK7 and other subunits were not significantly affected. Interestingly, also five subunits of the Mediator complex were detected but not depleted from the chromatin upon 120 min of dBET6 treatment. This is in contrast to earlier studies which detected a reduction of MED1 from the chromatin upon extended degrader treatment [135,341].

Recently, the Integrator complex has been ascribed a function in premature termination [425] but also promotor-proximal pause release at protein-coding genes [426–428]. We, therefore, determined the effect of BET depletion on Integrator proteins and the interacting PP2A phosphatase. However, while INTS1 and the endonuclease subunit INTS11 were significantly depleted upon dBET6 treatment, several other INTS proteins as well as subunits of PP2A (PPP2R1A, PPP2R2A, PPP2CA, PPP2R5C, PPP2R5D, PPP2R5E) were detected but did not change significantly (Supplemental Table 1). Upon dTAG7 treatment, fewer INTS and PP2A proteins were detected and were not changed in amounts.

Moreover, none of the core subunits of the PP1 phosphatase (PPP1CA, PPP1CB, PPP1CC) showed a significant trend upon either of the treatments. In contrast, PPP1R10 (≙ PNUTS) was significantly depleted upon dBET6 treatment and close to the significance threshold upon dTAG7 treatment (-log10 p-value = 1.12). PP1 and PPP1R10 have been shown to catalyze the dephosphorylation of SPT5 downstream of the pA site, thereby decelerating the elongation complex and promoting termination according to a combined termination model [238].

Interestingly, RPRD1A, RPRD1B and RPRD2, which correspond to Rtt103 in *S. cerevisiae* and were implicated in promoting the dephosphorylation of Ser5 of the RPB1 CTD [429,430], were significantly depleted or close to the threshold upon dBET6 treatment, but insignificant upon dTAG7 treatment (Supplemental Table 1).

4.1.11 THE ELONGATION COMPLEX ASSEMBLES INCOMPLETELY IN THE ABSENCE OF BRD4

Chromatin-MS provided an unbiased view of the compositional changes of the chromatin proteome upon BET or BRD4-specific depletion. As effects were averaged across the cell population but also across the genome, it appeared particularly suitable to detect strong, global trends. However, the lack of spatial resolution posed a limitation to identifying potential causes. Furthermore, the detection of a protein in chromatin-MS strongly depended on its physicochemical and biomolecular properties. To overcome these limitations, we performed a ChIP-Rx experiment for the elongation factors SPT5, PAF1 and SPT6 upon 120 min of BRD4-selective degradation. Since the effect of BRD4 depletion on CTD phosphorylation was unclear, Pol II was profiled using an RPB2-specific antibody, *i.e.*, in a CTD phosphorylation-independent manner.

RPB2 showed the typical bimodal signal distribution indicative of bidirectional transcription at the TSS (Figure 20A), with a maximum in sense signal between +50 and +200 bp. An additional, weaker RPB2 accumulation was observed at the pA site and downstream. Similarly, the SPT5 signal flanked the TSS, indicating an association with sense- and antisense-transcribing Pol II (Figure 20B). Like Pol II, also SPT5 showed a narrow peak of relatively low amplitude at the pA site.
While PAF1 and SPT6 were also found at both gene ends, their profiles differed remarkably from that of SPT5 (Figure 20C, D). At the gene 5' end, PAF1 and the SPT6 signal peaked more distally from the TSS than did SPT5, *i.e.*, between +150 and +300 bp, but flattened more slowly at the gene body (Supplemental Figure 3A, B). The different position of the local maximum likely resulted from the fact that both factors associate not with paused but with released Pol II. Also, SPT6 appeared to associate only weakly with Pol II transcribing in antisense direction. At the 3' gene end, no sharp peak but instead a broad signal shoulder downstream of the pA site was seen (Supplemental Figure 3B).

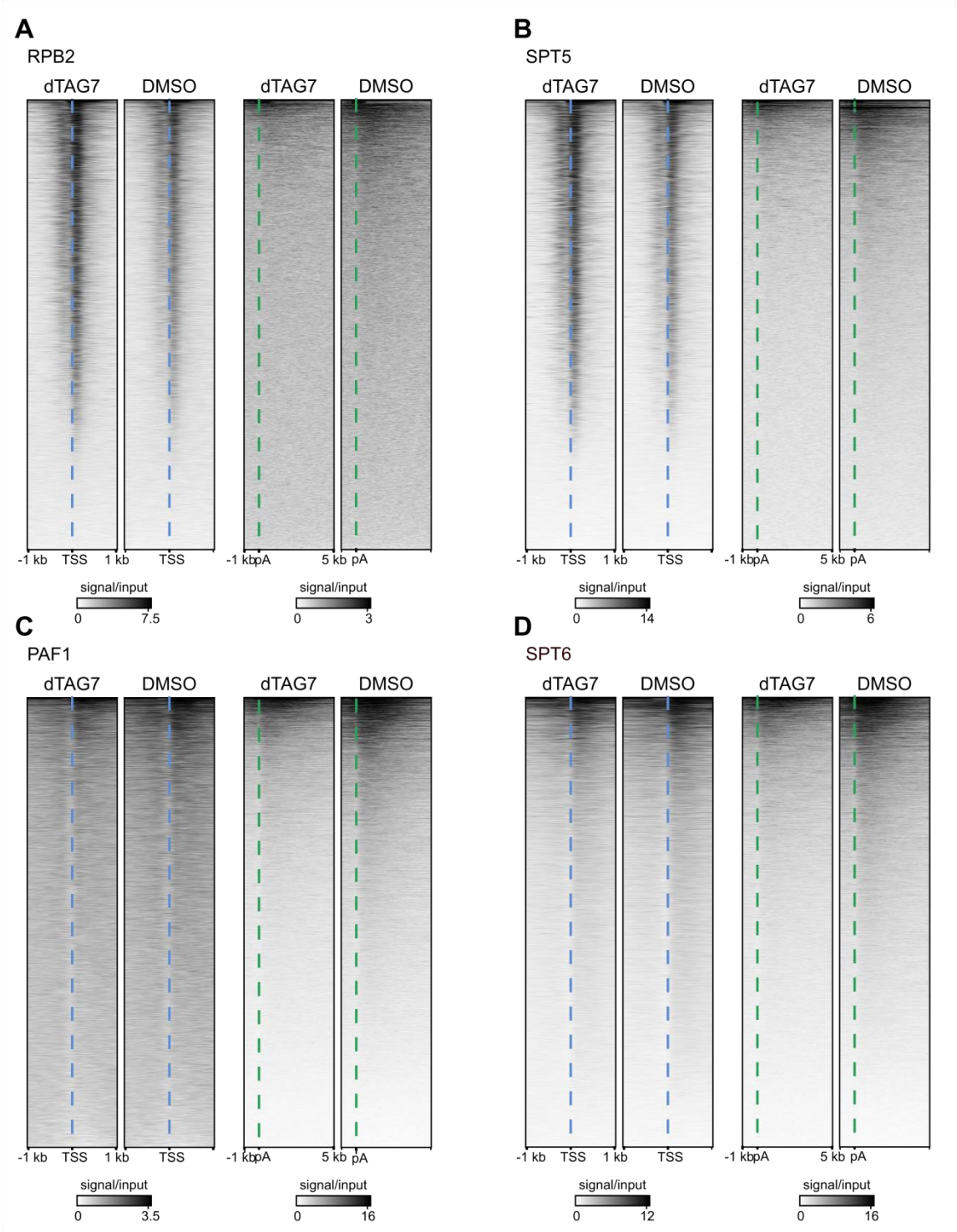


Figure 20: Selective BRD4 degradation affects binding of RPB2, SPT5, PAF and SPT6 at the 5' and 3' gene ends. ChIP-Rx was performed after 120 min of dTAG7 treatment or under control conditions. The heat maps show the input-normalized signal centered at the annotated TSS or pA site (indicated by a blue or green line, respectively) of 13,040 active genes without scaling and sorted by mean signal intensity. **(A)** Occupancy of RPB2, representative of total Pol II. Data of one replicate out of two is shown. **(B)**, **(C)**, **(D)** Occupancy of the elongation factors SPT5, PAF1 and SPT6. Four, three and one biological replicates were analyzed, respectively.

Upon BRD4-selective degradation, binding of RPB2 and SPT5 at the 5' gene end was increased, probably reflecting the increase in promoter-proximal pausing also observed by SI-NET-seq, while binding at the 3' gene end was reduced (Figure 20A, B). Notably, the RPB2 profile revealed an increase in signal in the promoter-proximal and a decrease in the gene body region (Supplemental Figure 4) and thus recapitulated the change in occupancy that was seen upon dTAG7 treatment with SI-NET-seq. This showed that the overall trend was detectable independently of the method used. PAF1 and SPT6 levels were decreased along the entire transcribed region in the absence of BRD4 (Figure 20C, D). Also, for none of the profiled proteins, a relative shift of the 5' peak position was seen (Supplemental Figure 3A).

Reduced elongation factor binding upon BRD4 depletion could be a mere consequence of the global decrease in transcription and, in this case, would only indirectly depend on BRD4. To test this possibility, we normalized the SPT5 and PAF1 signal by the signal of RPB2. This analysis revealed a highly significant global reduction of SPT5, which was more pronounced than the reduction of RPB2, in the termination zone upon BRD4 depletion (Figure 21). However, no significant difference in SPT5 binding under treatment and control conditions was found for the promoter-proximal and the gene body region. In contrast, the Pol II-normalized PAF1 levels were significantly lower upon treatment than under control conditions in all three tested regions, suggesting that PAF1 was inefficiently recruited or unstably bound to the transcription machinery in the absence of BRD4.

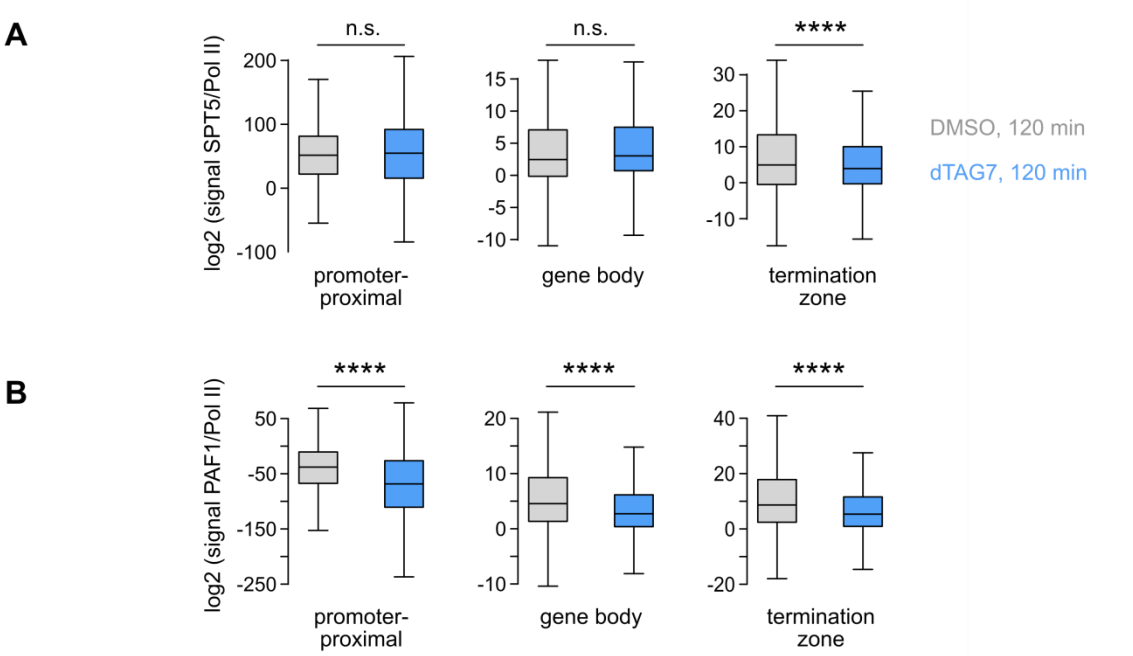


Figure 21: Upon dTAG7 treatment, SPT5 and PAF1 show a locally more pronounced decrease in chromatin binding than Pol II. (A) Quantification of the SPT5 signal normalized to RPB2 signal upon 120 min of dTAG7 treatment or under control conditions. The regions were defined as follows: promoter-proximal: TSS to TSS +300 bp; gene body: TSS +300 to pA site; termination zone: pA site to pA site +3 kb. Significance was determined by one-tailed paired Mann–Whitney U test. ****: $p \leq 0.0001$. **(B)** Quantification of PAF1 occupancy normalized to RPB2 occupancy. The analysis was performed by Annkatrin Bressin and with minor changes included in Arnold et al. [368].

4.1.12 BRD4 DEPLETION IMPAIRS THE RECRUITMENT OF CPSF AND CSTF PROTEINS TO THE CHROMATIN

In order to better understand the effect of BRD4 depletion on the 3' RNA processing machinery, we performed ChIP-Rx experiments for selected proteins of the CPSF and CstF modules, in particular CPSF73 and FIP1 as well as CstF64. Also, data sets for CPSF30 and CstF77 were generated but excluded from further analysis due to low ChIP enrichment. Notably, the lack of interaction interfaces of these proteins with DNA and the unavailability of ChIP-grade antibodies complicated the generation of high-quality ChIP-Rx profiles.

Interestingly, we found all tested factors to accumulate at both the 5' and the 3' end of active genes (Figure 22A). This binding pattern was evident for exemplary coding and non-coding genes (not shown) as well as on the meta-gene level (Figure 22B). While enrichment near the 3' end was expected given the primary function in RNA 3' end processing, the consistent strong enrichment at the 5' end was surprising. Based on the

meta-gene analysis, the CPSF proteins appeared more strongly enriched at the 5’ end than the 3’ end, while the opposite was seen for CstF64. Similar differences in the binding pattern of the two protein complexes have been described earlier for exemplary loci [224,225]. Interestingly, the signal maximum of each factor was located 150 to 300 bp downstream of the TSS, *i.e*., at a position similar to the position of the peaks of PAF1 and SPT6, but more distal to the TSS than the SPT5 peak.

Upon 120 min of dTAG7 treatment, occupancy of the 3’ end processing factors at both gene ends was reduced on a global scale (Figure 22B). Of note, the decrease in the promoter-proximal region was contrary to the effect seen for RPB2, which showed an increase in binding upon BRD4 depletion. This indicated that RNA 3’ processing factors were inefficiently recruited to the transcription machinery in the absence of BRD4.

Based on these findings, we hypothesized that RNA 3’ end processing factors were recruited already near the 5’ gene end and moved towards the 3’ gene end together with Pol II. However, because the signal-to-noise ratio appeared to be relatively low along the gene body, it was unclear from meta-gene data if the tested RNA 3’ end processing factors were also present in this region. Therefore, we compared the RPKM values of each factor upon treatment with dTAG7 or DMSO at the gene body region. Differences were significant in case of CstF64 and CPSF73 but not FIP1, showing that CstF64 and CPSF73 were indeed present at the gene body and possibly traveled with elongating Pol II and additionally bound chromatin in a BRD4-dependent manner (Figure 23).

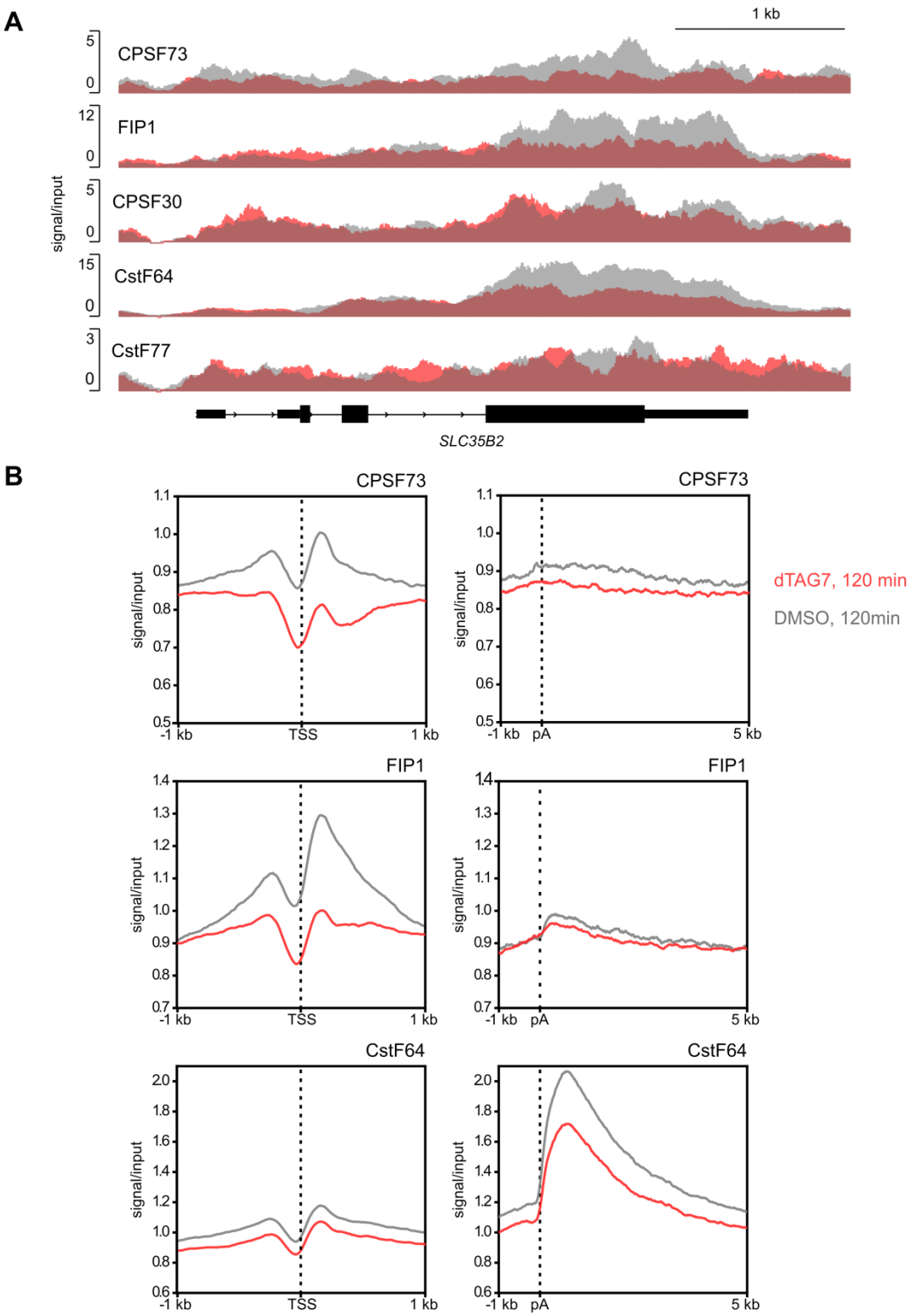


Figure 22: BRD4 ablation reduces binding of CPSF and CstF at both the 5' and the 3' gene end. (A) Input-normalized ChIP-Rx signal of selected proteins of the CPSF and CstF complexes at the *SLC35B2* locus. For direct comparison, the signal tracks of the dTAG7-treaded sample (red) and the DMSO control (grey) are overlaid. Data of one representative biological replicate (out of two) are shown. **(B)** Meta-gene representation of the input-normalized occupancy centered at the TSS and the pA site (indicated by dotted lines) upon dTAG7 treatment (red) and under control conditions (grey).

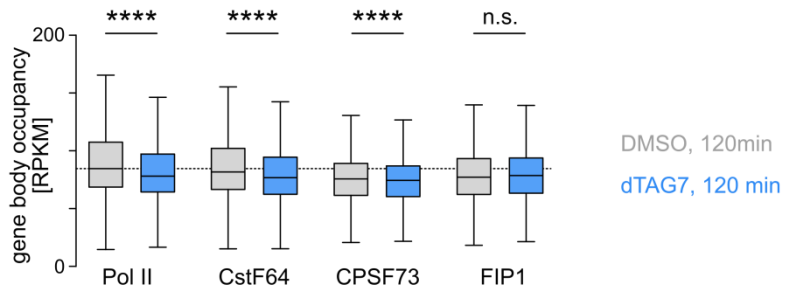


Figure 23: The gene body occupancy of CstF64 and CPSF73, but not FIP1, is significantly decreased upon BRD4 degradation. The gene body region was defined as ranging from TSS +300 bp to the pA site. Statistical significance of the difference in protein binding between the treatment and control condition was determined by one-tailed paired Mann–Whitney U test. The analysis was performed by Annkatrin Bressin and with minor changes included in Arnold et al. [368].

4.1.13 BRD4, LIKE ELONGATION AND RNA 3' END PROCESSING FACTORS, BINDS IN THE PROMOTER-PROXIMAL REGION

While co-localization at the chromatinized genome is no prerequisite to assume a functional relationship, it would offer a parsimonious explanation of how BRD4 could promote the recruitment of the 3' RNA processing machinery. Therefore, to determine the binding regions of BRD4 in our cell system, we performed ChIP-Rx for BRD4 upon 40 min and 120 min of dTAG7 treatment and under control conditions (DMSO). As no HA tag-expressing NIH3T3 cell line was readily available as spike-in standards, a BRD4-specific antibody was used.

A genome-wide analysis revealed that BRD4 binding under control conditions occurred predominantly at the 5' gene end and only to a minimal extent along the gene body or near the 3' end (Figure 24A). Particularly, binding was evident immediately upstream and downstream of the annotated TSS, suggesting an association with Pol II transcribing in sense and antisense direction (Figure 24B). Binding near the promoter was also confirmed by visual inspection of the input-normalized signal at single gene loci (Supplemental Figure 5). Thus, the genomic distribution suggested that the interaction of BRD4 with factors of the 3' processing machinery would most likely occur in the promoter-proximal region.

Unexpectedly, BRD4 levels were reduced to a similar extent upon 40 and 120 min of dTAG7 treatment (Figure 24A, B), which appeared to conflict with the gradual decrease in BRD4 signal seen by whole cell immunoblotting. In accordance with a role in enhancer regulation, BRD4 binding was also detected at enhancer centers and was decreased upon 40 and 120 min of dTAG7 treatment (Figure 24C).

Peak calling and assigning the peaks based on their relative position confirmed that BRD4 under control conditions was bound primarily near the promoter but also in intronic and intergenic regions (Figure 24D). Specifically, 59% and 11% of the 44,013 peaks were found within a 1 kb or a 1 to 3 kb distance to the TSS of annotated genes, respectively. 13% of the peaks were located at intergenic sites—presumably intergenic enhancers. Of the intragenic regions distal to the promoter, introns contained 16% of the peaks, while exons contained only <1%. Treatment reduced the number of BRD4 peaks from on average 41,966 to 3,041 and 4,667 peaks after 40 and 120 min, respectively.

To test if BRD4 occupancy could, to some extent, explain the differences in BET sensitivity of elongation and termination that were observed, we determined the mean input-normalized signal of BRD4 under control conditions for different Ensembl gene biotypes as well as for different gene length classes. Meta-gene plots revealed a higher BRD4 occupancy at protein-coding than at long non-coding or histone genes (Figure 25A). No clear difference was apparent between long non-coding and histone genes, although only the former displayed an elongation defect upon BRD4 degradation. Also, the mean BRD4 occupancy was very similar irrespective of gene length (Figure 25B). In all cases, most of the signal was found near the 5' gene end.

Lastly, we asked how the effect of BET degradation compared to that of BET inhibition. Reanalysis of publically available BRD4 ChIP-seq data collected after 120 min of JQ1 treatment in MOLT4 cells [135] revealed a reduction of >40% at the TSS, while BRD4 was almost completely lost after 120 min of dBET6 treatment (Figure 25C). In view of this, stronger phenotypes should be expected upon BET degradation.

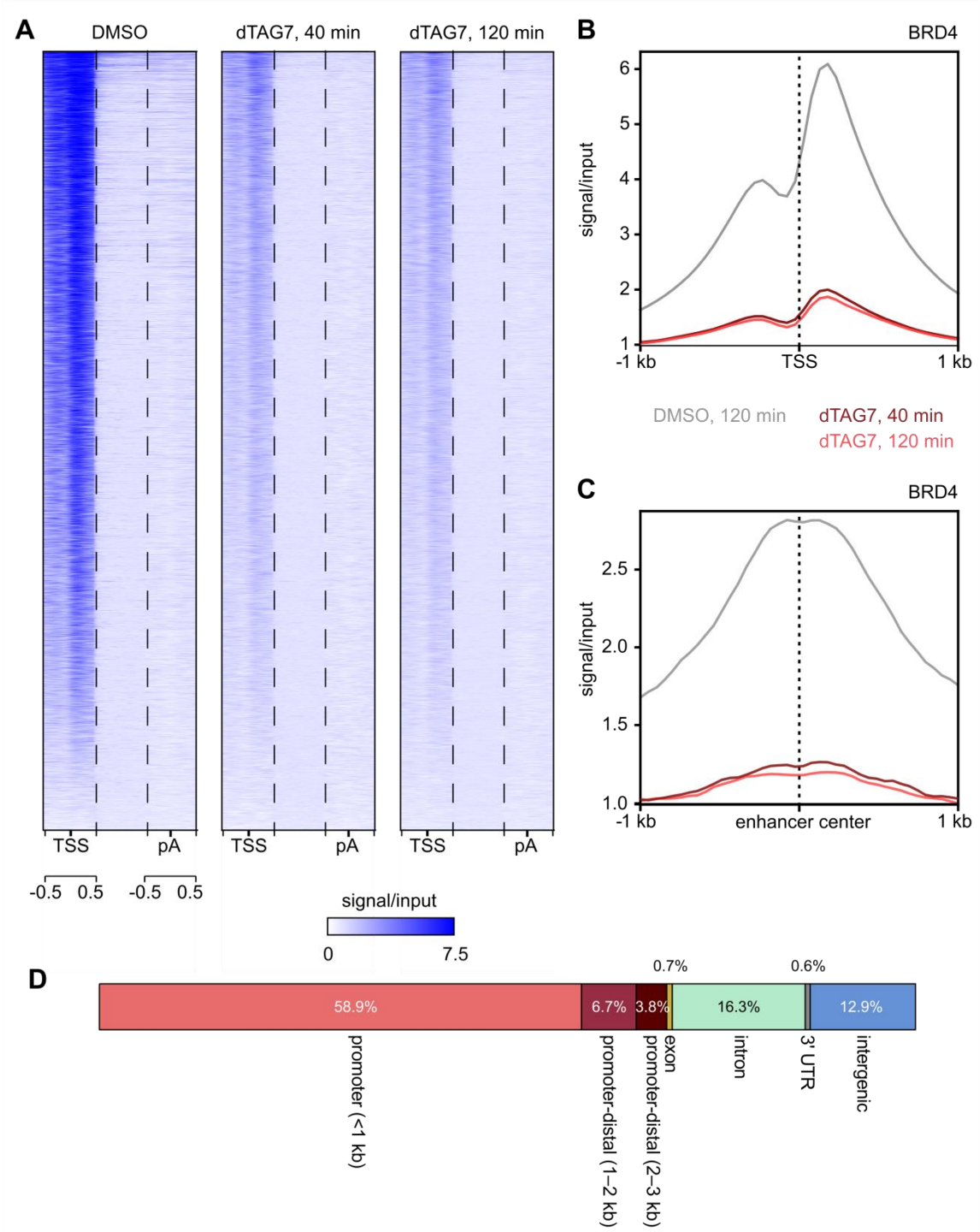


Figure 24: BRD4 predominantly binds the 5' gene end as well as enhancers and is rapidly lost upon dTAG7 treatment. (A) Heat map representation of the input-normalized BRD4 signal under dTAG7 treatment (40 min, 120 min) or control conditions (120 min). 13,040 coding and non-coding genes were considered. The signal is centered at the TSS and the pA site. Scaling is only applied to the gene body defined as TSS +500 bp to pA site -500 bp. Genes are ranked by descending mean signal intensity. **(B)** Meta-gene representation of the BRD4 occupancy centered at the TSS (dotted line) upon dTAG7 treatment (red) and under control conditions (grey). The same set of genes was used as in (A). **(C)** Meta-gene representation of the BRD4 occupancy aligned at the center of 3,228 enhancer regions in K562 (dotted line), as identified by Lidschreiber et al. [408]. **(D)** Annotation of BRD4 peaks identified under control conditions relative to annotated gene regions. The bioconductor package ChIPseeker was used [431].

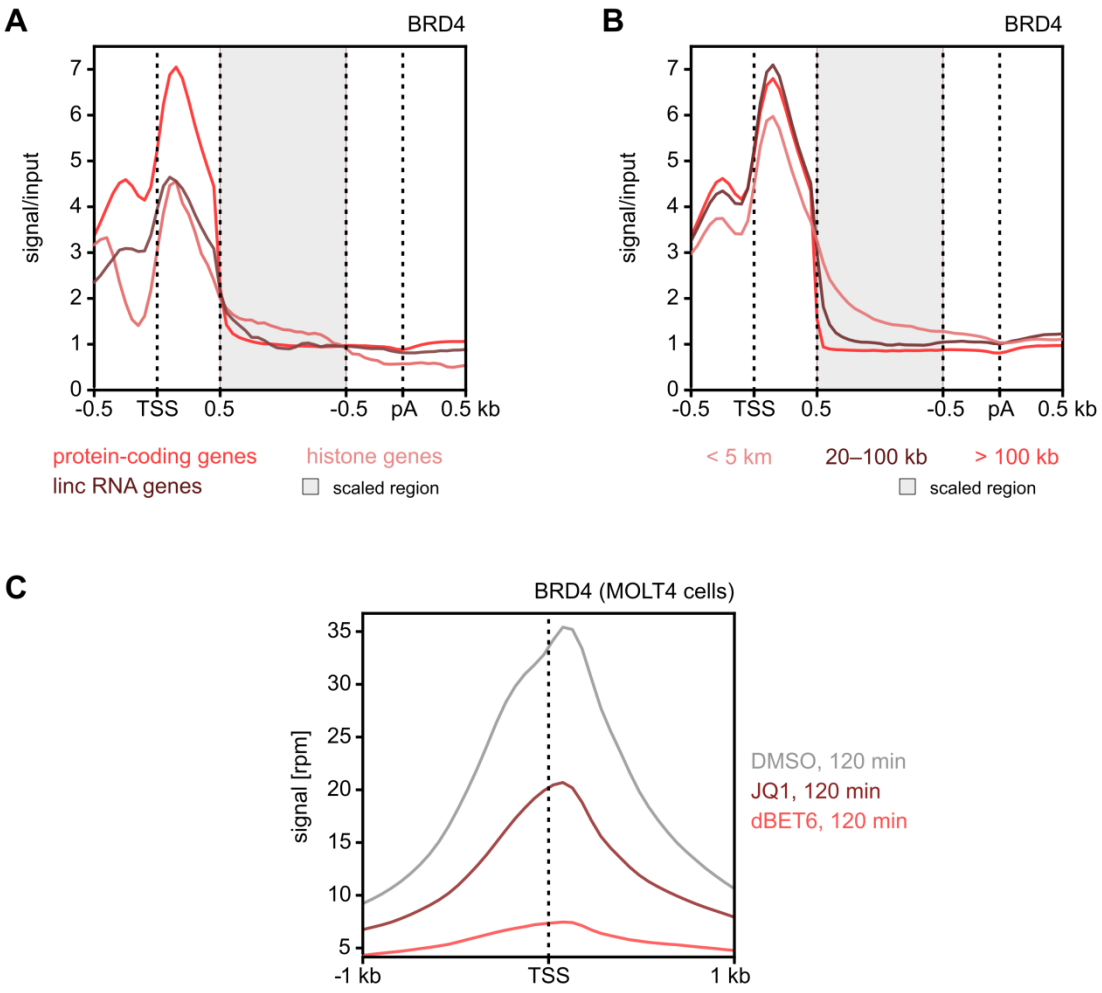


Figure 25: BRD4 binding levels only partially explain differences in the severity of the elongation defect. (A) Meta-gene representation of the BRD4 occupancy for different gene biotypes. The input-normalized signal is centered at the TSS and the pA site. Scaling is only applied to the gene body defined as TSS +500 bp to pA site -500 bp (grey-shaded). Biotype annotations were obtained from Ensembl Biomart. 6,369 protein-coding genes (excluding histone genes), 336 lincRNA and 55 histone genes were included. Other non-intergenic lncRNA were excluded to avoid artifacts due to overlapping regions. **(B)** Meta-gene representation of the BRD4 occupancy for different gene length classes. The signal is centered and scaled as described for (A). Only protein-coding and lincRNA genes were included. 1,389, 3,179 and 6,165 genes were included in the different categories. **(C)** Meta-gene representation of the BRD4 occupancy centered at the TSS upon JQ1 treatment (120 min) or under control conditions in MOLT4 cells. The ChIP-Rx data were generated by Winter et al. [135].

4.1.14 BRD4 CO-PRECIPITATES WITH FACTORS OF THE TRANSCRIPTION MACHINERY UNDER NATIVE CONDITIONS

While direct physical interactions between proteins are challenging to prove and require specialized techniques, such as crosslinking mass spectrometry (CL-MS) [432]; or a yeast-two-hybrid assay [433], immunoprecipitation followed by mass spectrometry (IP-MS) or

SDS-PAGE can give valuable hints about the spatial proximity of proteins or protein complexes within the cell. To test if BRD4 itself could be involved in the recruitment of elongation but also RNA 3' processing factors, we performed a native IP-MS experiment in K562 dTAG-BRD4 cells. The sample preparation included treatment with benzonase nuclease to disrupt interactions mediated by RNA or DNA. An antibody targeted at the C-terminal domain of BRD4 was used.

Of the 966 proteins or protein groups that were robustly detected, 650 were significantly enriched by immunoprecipitation of BRD4 (Figure 26). Among these, there were known interactors of BRD4, including NSD3, CHD4 and GLTSCR1 [311]. The analysis moreover revealed a significant enrichment of core components of the paused and the activated elongation complex. In particular, BRD4 was found to interact with the two central subunits of Pol II, RPB1 and RPB2, as well as RPABC1 (≙ RPB5) and RPB9, which form part of the lower and upper jaw, respectively (Figure 26). Also, all subunits of the DSIF complex (SPT4, SPT5) and the NELF complex (NELFA, NELFB, NELFCD, NELFE) complex were significantly enriched. Interestingly, both CDK9 and cyclin T1 but not HEXIM1 were detected, indicating that BRD4 interacted with CDK9 in its active but not in its HEXIM1-inhibited form, which is in accordance with the literature [132]. Furthermore, PAF1, CDC73 and LEO1 of the PAF complex, but neither the core component CTR9 nor WDR61 were among the interactors. RTF1 was close to the significance threshold ($p = 0.05615$). Lastly, also SPT6 and its associated factor IWS1 were enriched.

Interestingly, several proteins involved in 3' end RNA processing or transcription termination were significantly enriched by IP of BRD4 (Figure 26). These included CstF64 of the CstF module, WDR33 and symplekin of the CPSF module, CFIm25 of CFIIm and PCF11 of CFIm. The most significant enrichment was seen for the endonuclease XRN2. Notably, despite an extensive nuclease treatment to exclude interactions mediated by nucleic acids, three of the detected factors contain dedicated RNA-binding domains or even use RNA as a substrate.

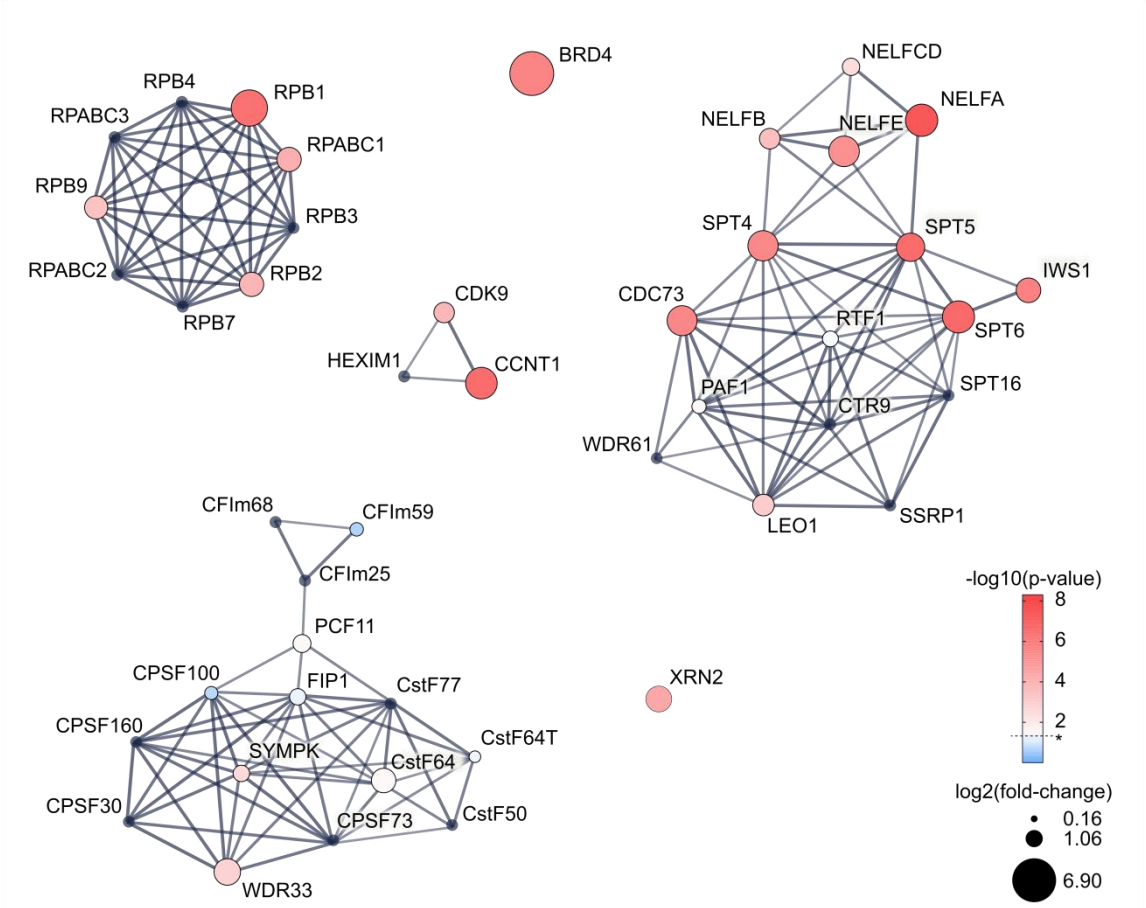


Figure 26: Native IP-MS reveals interactions of BRD4 with Pol II subunits, elongation factors and RNA 3' processing and termination factors. The experiment was performed in five biological replicates. Networks were constructed based on high-confidence physical interactions as annotated in the STRING database [407]. Significance of the enrichment was calculated based on LFQ intensities. Proteins detected as BRD4 interactors by native IP-MS are highlighted, with red shading indicating significant enrichment and the size of the circle representing the enrichment. Proteins that were not detected or did not pass the filtering criteria are shown in grey-blue.

4.1.15 Formaldehyde-assisted IP-MS confirms the interaction of BRD4 with the transcription machinery

Native immunoprecipitation critically depends on assay conditions, including detergent content and salt strength, to preserve protein–protein interactions formed under the physiological conditions inside the cell. In contrast, IP strategies involving a short crosslinking step chemically stabilize those interactions. Thus, more stringent lysis and washing conditions can be applied to reduce unspecific background. This benefits mainly, but not exclusively, the detection of transient, substoichiometric interactions [434]. Assuming that interactions of transcriptional regulators are usually transient, we, in addition to the native approach described in 4.1.14, chose an IP-MS protocol that involved a short

crosslinking step with formaldehyde prior to cell lysis. Additionally, nucleic acids were only sheared by sonication but not removed enzymatically. Lastly, a high-affinity antibody against the HA tag, which was part of the dTAG construct inserted at the BRD4 locus, was used to enrich BRD4 more efficiently.

Following this approach, 552 proteins were robustly detected, 231 of which were significantly enriched by HA-specific IP. Strikingly, when ranking the interactors by fold-change, the known BRD4 interactors NSD3, CHD4, ATAD5, ZMYND8 [311] and cyclin T1 were found among the top ten proteins (not shown), indicating that the protocol could uncover specific interactions. To increase sensitivity and thereby the number of interactors, we repeated the analysis with a threshold of only one common peptide required for protein quantification. By default, relative label-free quantification (LFQ) requires at least two peptide species to be present in each sample [435]. Changing this parameter increased the number of robustly detected and of significantly enriched proteins to 896 and 379, respectively.

IP-MS with crosslinking revealed interactions with RPB1, RPB2 and RPB3 (Figure 27). No other subunits of Pol II were significant, suggesting that despite crosslinking, individual subunits rather than entire multi-subunit complexes were enriched. While elongation factors overall were among the most strongly enriched proteins, individual proteins were either not detected or not significantly enriched. In particular, NELFA, NELFB and NELFE as well as SPT5, but not NELFCD and SPT4, were found to be significant interactors of BRD4. Likewise, neither SPT6 nor IWS1 was detected with this approach, although they were well-detectable and significantly enriched in the native IP experiment. Interestingly, while CDC73 and PAF1 could be confirmed as BRD4 interactors, enrichment of the dissociable subunit RTF1 was significant only following crosslinking. WDR61 and CTR9 were still not detected. In contrast, SPT16 and SSRP1, which were missing in the native IP data set, were found to interact with BRD4 in this experiment.

Detection of proteins involved in 3' RNA processing was partially improved. Under these assay conditions, CstF77 was significantly enriched, while CstF64 was not detected. CPSF160 and FIP1 now appeared as interactors, while the enrichment of CPSF73 improved but was still not significant. The significant enrichment of symplekin, a factor suggested to couple transcription and polyadenylation [200,201] was confirmed. Unexpectedly, neither PCF11 nor XRN2 was detected with this approach.

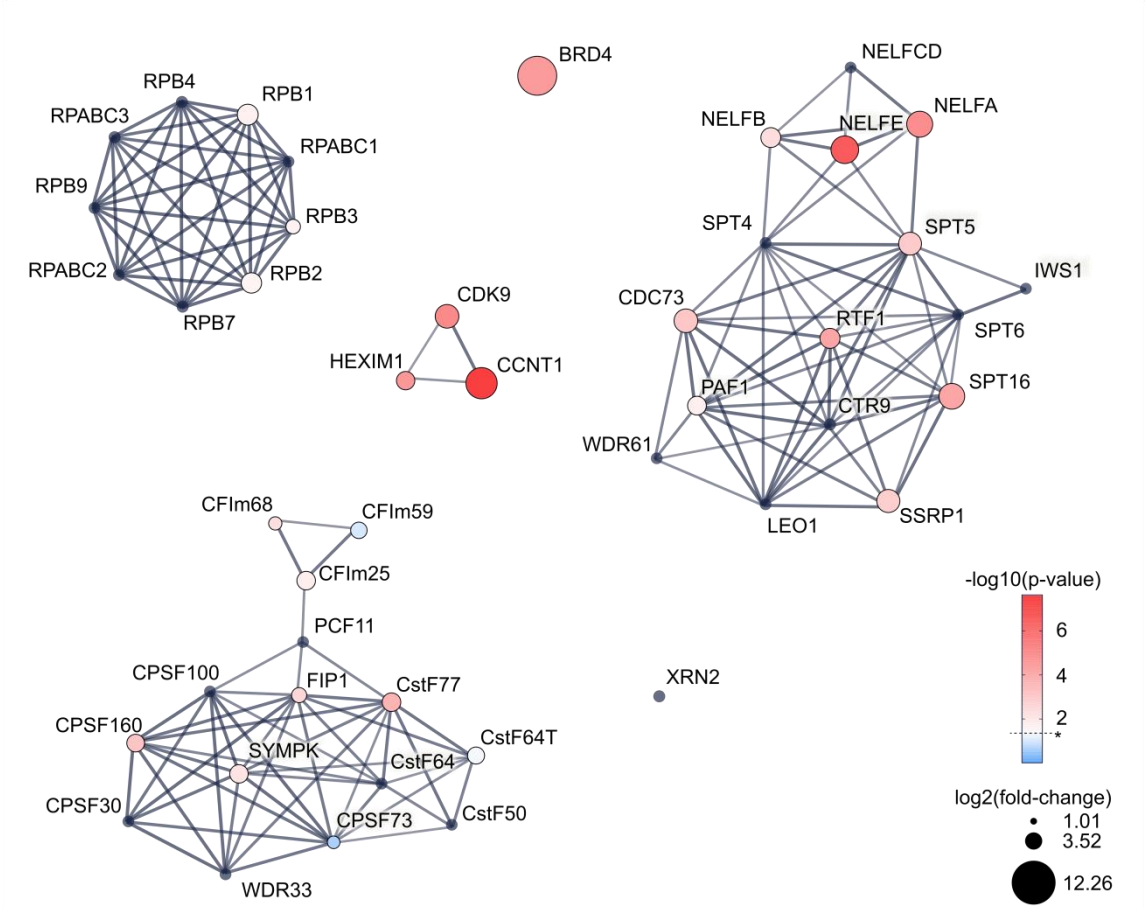


Figure 27: Formaldehyde-assisted IP-MS reveals additional interactions of BRD4 with RNA 3' cleavage and termination factors. To improve protein enrichment, K562 dTAG-BRD4 cells were crosslinked prior to lysis, and an HA-specific antibody was used for IP. Additionally, the threshold for protein quantification was lowered to one peptide (see 3.2.9). Five biological replicates were analyzed. The networks were constructed based on high-confidence physical interactions as annotated in the STRING database [407]. Significance of the enrichment was calculated based on LFQ intensities. Proteins detected as BRD4 interactors by IP-MS with crosslinking are highlighted, with red shading indicating significant enrichment and the size of the circle representing the enrichment. Proteins that were not detected or did not pass the filtering criteria are shown in grey-blue.

4.1.16 PAF INTERACTS WITH RNA 3' END PROCESSING FACTORS

Both IP-MS experiments for BRD4 revealed physical interaction with individual RNA 3' end processing factors, supporting a direct role of BRD4 in facilitating their chromatin binding. However, also elongation factors—particularly, PAF and SPT5—have been implicated in 3' RNA processing mainly in yeast [200,249,250], which offered an alternative hypothesis. To test whether the reduction of PAF observed upon BRD4 depletion could be causal for the recruitment defect of RNA 3' processing factors, we performed a native IP for PAF1, CDC73 and RTF1 of the PAF complex, followed by immunoblotting. Lysates

were treated with benzonase to increase solubility and eliminate RNA- or DNA-mediated interactions. This experiment confirmed interactions with BRD4 but also revealed interactions with the CPSF and CstF modules in our cellular system (Figure 28A). Of the tested factors, only CstF77 co-precipitated with all tested subunits, although the signal was most intense for CDC73. CPSF73, FIP1 and CPSF30 predominantly interacted with CDC73.

Since its binding profile near the pA site resembled that of RNA 3' processing factors, we also performed a native IP experiment for SPT6. Unexpectedly, only CstF77 but none of the tested CPSF proteins or XRN2 precipitated with SPT6 (Figure 28B). However, this experiment could recapitulate the interaction of SPT6 and BRD4, indicating suitable assay conditions overall (Figure 28C).

Based on these experiments, particularly the CstF module appeared to functionally link transcription elongation and termination. To validate this hypothesis, we made use of a HeLa cell line "8880" overexpressing a CstF64 protein that contained an N-terminal localization and affinity purification (LAP) tag [387] and determined the CstF64 interactome by IP-MS. As expected, we found an enrichment of unique peptides of CstF55, CstF64τ and CstF77 (Figure 28D). CstF64 itself appeared as relatively abundant also in the control, possibly due to cross-contamination between the samples during MS analysis. Moreover, unique peptides of all CPSF proteins, CFIm proteins, XRN2 and PCF11 were detected as interactors of the tagged CstF64 but not in the control, indicating that the tag did not interfere with the normal complex formation of CstF64. Interestingly, this experiment also revealed interactions of CstF64 with PAF1, CDC73, LEO1 and CTR9, but not RTF1, and also with SPT4, SPT5 and SPT6, supporting the idea that these factors potentially have a role in 3' RNA processing in human cells. Unexpectedly, BRD4 was not detected in any of the samples.

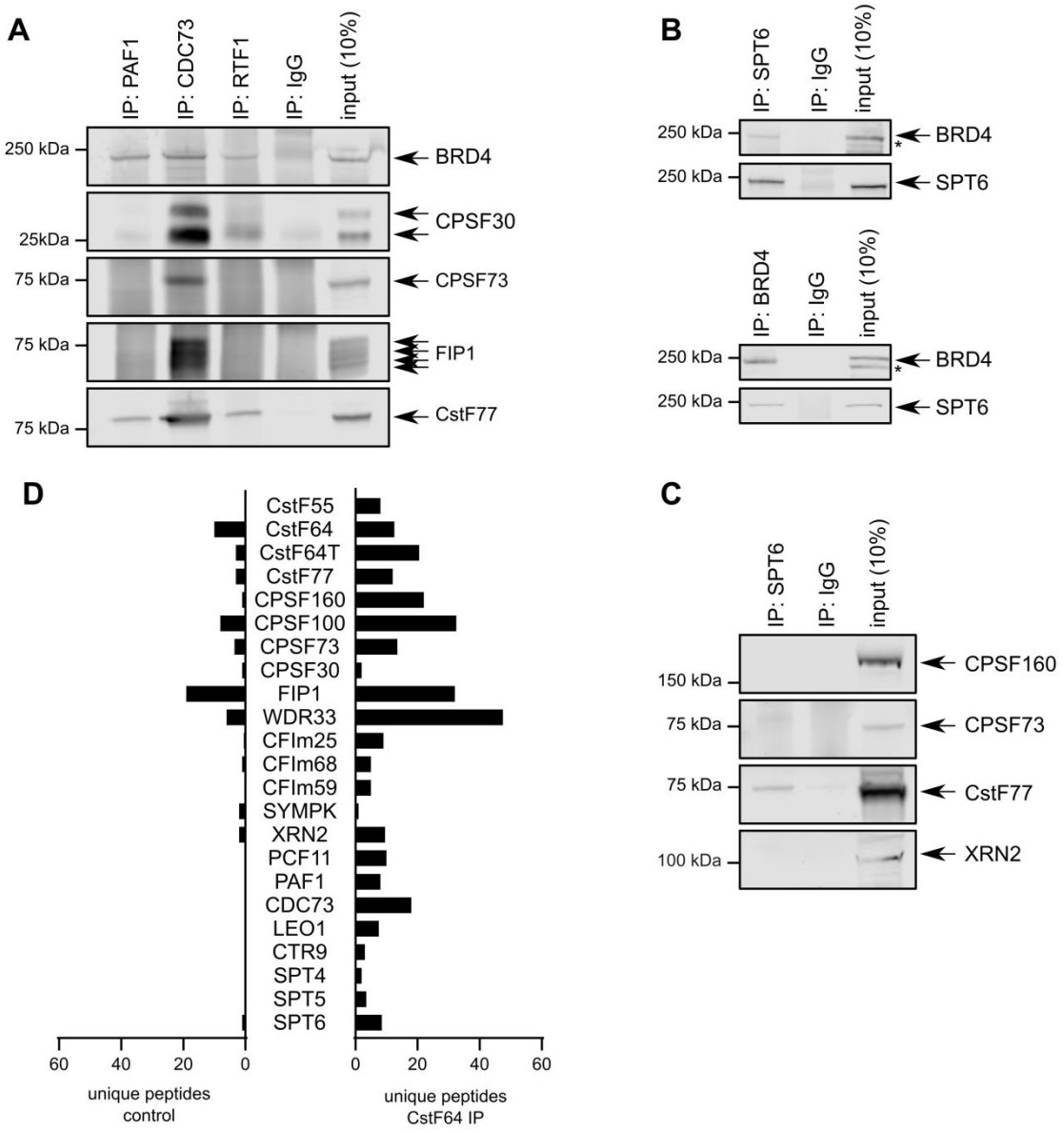


Figure 28: PAF interacts with RNA 3' end processing factors under native conditions. (A) Interactions of PAF subunits with BRD4 and selected RNA 3' end processing factors, as found by native IP and immunoblotting. **(B)** Confirmation of the interaction between SPT6 and BRD4 by native IP and immunoblotting. **(C)** Interactions of SPT6 with selected RNA 3' end processing factors. The asterisk marks an unspecific band. **(D)** Interactions of CstF64 with elongation and RNA 3' end processing factors that were identified by native IP for GFP-tagged CstF64 in HeLa cells and MS analysis. As a control, GFP-non-expressing cells were incubated with the GFP-specific antibody.

To test whether the association of PAF with RNA 3' processing factors required BRD4, we performed a native IP for CDC73 in K562 dTAG-BRD4 cells that were pre-treated with dTAG7 for 120 min. To rule out that BRD4 degradation exerted different effects on the protein populations bound to the chromatin or soluble in the nucleoplasm, either a whole nuclei lysate (prepared using the standard IP lysis protocol) or a lysate consisting solely

of chromatin was used. Immunoblotting revealed a robust decrease in CDC73 levels upon dTAG7 treatment that was particularly pronounced in the chromatin-based sample, which was in accordance with the reduced detection of elongation factors by chromatin-MS (Figure 29). Irrespectively of the treatment, immunoprecipitation from nuclear lysates efficiently enriched 3' RNA processing factors. Except for CPSF160, which was reduced by 27% upon dTAG7 treatment compared to the DMSO control, no decrease in protein amounts was seen in the chromatin lysate-based immunoprecipitation sample. However, amounts of CPSF160 were also decreased by 35% in the dTAG7-treated input, confounding the interpretation.

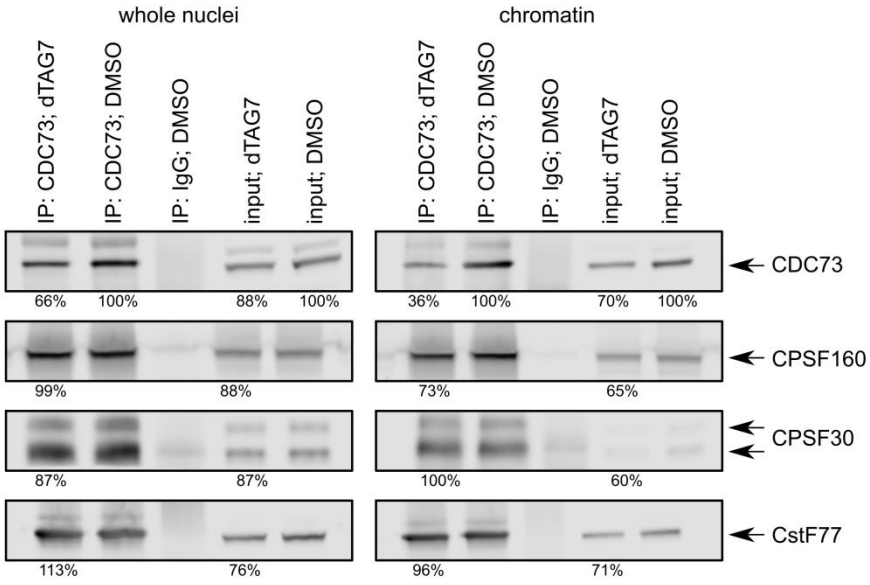


Figure 29: The interaction of CDC73 and RNA 3' processing factors does not depend on BRD4. After treating K562 dTAG-BRD4 cells with dTAG7 for 120 min, immunoprecipitation was performed either on a whole-nuclei (left) or a chromatin lysate (right). Band intensities upon dTAG7 treatment were quantified relative to the signal seen under DMSO control conditions for both IP samples and input controls.

4.1.17 Pol II CTD phosphorylation in the promoter-proximal region is increased upon BRD4 degradation

The C-terminal domain of RPB1 has been ascribed a major role in coordinating transcription and transcription-coupled processes. In particular, it was suggested to serve as a "landing pad" for the recruitment of factors involved in transcription and co-transcriptional processing [40,436]. To test whether CTD phosphorylation was disturbed upon BRD4 degradation and therefore could underlie the observed recruitment defects, we used ChIP-Rx to profile p-Ser2, p-Tyr1 and p-Thr4 upon 120 min of dTAG7 treatment.

Overall, the characteristic shape of the distribution of each Pol II phospho-form was maintained upon BRD4 degradation [257]. Under control and treatment conditions, p-Ser2 and p-Thr4 showed local signal maxima upstream and downstream of the TSS as well as flanking the pA site and a steady increase in signal along the gene body. In contrast, p-Tyr1 peaked primarily at the 5' gene end (Figure 30).

As expected with regard to its role in pause release, Ser2-phosphorylated Pol II was detected downstream of the TSS, with a peak within the first 150 nt of the transcribed region. p-Ser2 additionally accumulated in a broad peak immediately downstream of the pA site. Importantly, an increase in p-Ser2 signal in the promoter-proximal region upon dTAG7 treatment indicated that CDK9 activity was not diminished in the absence of BRD4 which is in accordance with Winter et al. [135]. However, a decrease in signal was seen downstream of the pA site upon dTAG7 treatment.

The input-normalized p-Tyr1 signal predominantly peaked immediately downstream of the TSS but additionally upstream, agreeing with a proposed role in antisense transcription [437]. Upon treatment, the signal amplitude was increased, indicating that promoter-proximally paused Pol II was phosphorylated at the Tyr1 residue. Albeit very weakly, another signal peak was seen exactly at the pA site (not shown). While more p-Tyr1 was detected at the pA site upon dTAG7 treatment, the signal fell below the level in the DMSO sample further downstream.

Thr4-phosphorylated Pol II accumulated downstream of the TSS, putatively in the region of promoter-proximal pausing. As expected with regard to its proposed role in termination [184,258], signal of similar intensity was also detected downstream of the pA site. Overall, the pattern of p-Thr4 resembled that of p-Ser2. Similar to p-Ser2 and p-Tyr1, p-Thr4 was increased at the 5' gene end but decreased at the 3' gene end in the absence of BRD4.

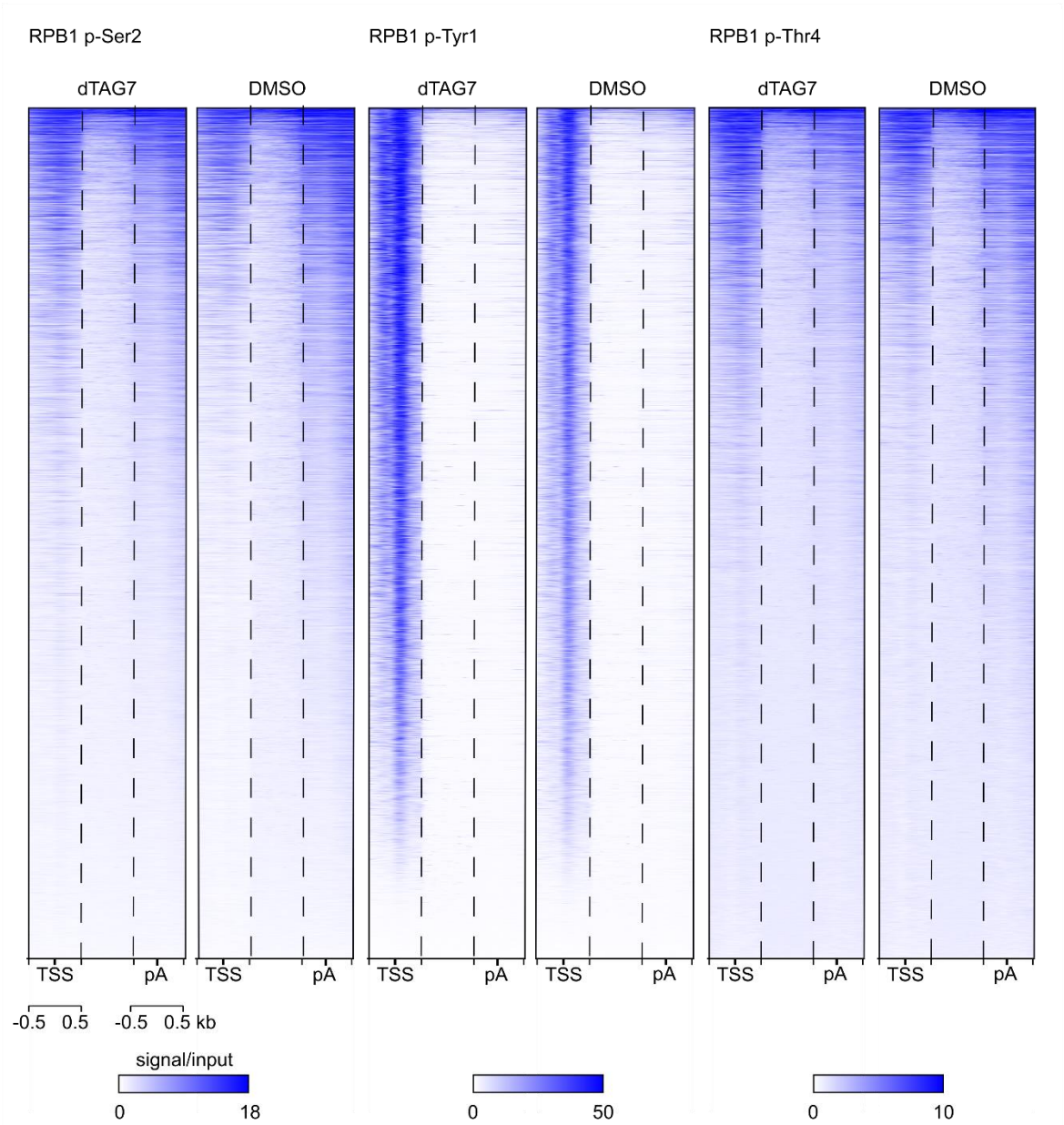


Figure 30: Targeted BRD4 degradation results in increased phosphorylation of Tyr1, Ser2 and Thr4 of the RPB1 C-terminal domain (CTD) in the promoter-proximal region. ChIP-Rx was performed after 120 min of dTAG7 treatment. 13,040 coding and non-coding genes were considered. The input-normalized signal is centered at the TSS and the pA site. Scaling is only applied to the gene body defined as TSS +500 bp to pA site -500 bp. Genes are ranked by descending mean signal intensity.

To determine whether the increase in signal seen upon dTAG7 treatment near the TSS was a consequence of the increase in promoter-proximally paused Pol II or indicated an actual change in the phosphorylation state of the RPB1 CTD, we normalized the signal by RPB2. Normalized p-Ser2 levels of both conditions were very similar in the promoter-proximal region, but further downstream increased more steeply in the dTAG7- than the DMSO-treated sample (Figure 31). Conversely, less p-Ser2 was found upstream but

especially within 5 kb downstream of the pA site upon dTAG7 treatment. dTAG7 treatment also resulted in a more pronounced increase in p-Tyr1 than in RPB2 signal in the promoter-proximal and the gene body region. Interestingly, p-Tyr1 levels were lower under treatment conditions within 3 kb downstream of the pA but beyond that region decreased more slowly than in the control. In fact, higher phosphorylation levels were maintained for at least 40 kb downstream of the pA site upon treatment (not shown). p-Thr4 was excluded from this analysis as the two biological replicates yielded inconsistent results. Overall, this analysis confirmed that CTD phosphorylation at the Ser2 and Tyr1 residues was not negatively affected by BRD4 ablation in the promoter-proximal region, but showed a characteristic decrease within a relatively narrow window downstream of the pA site.

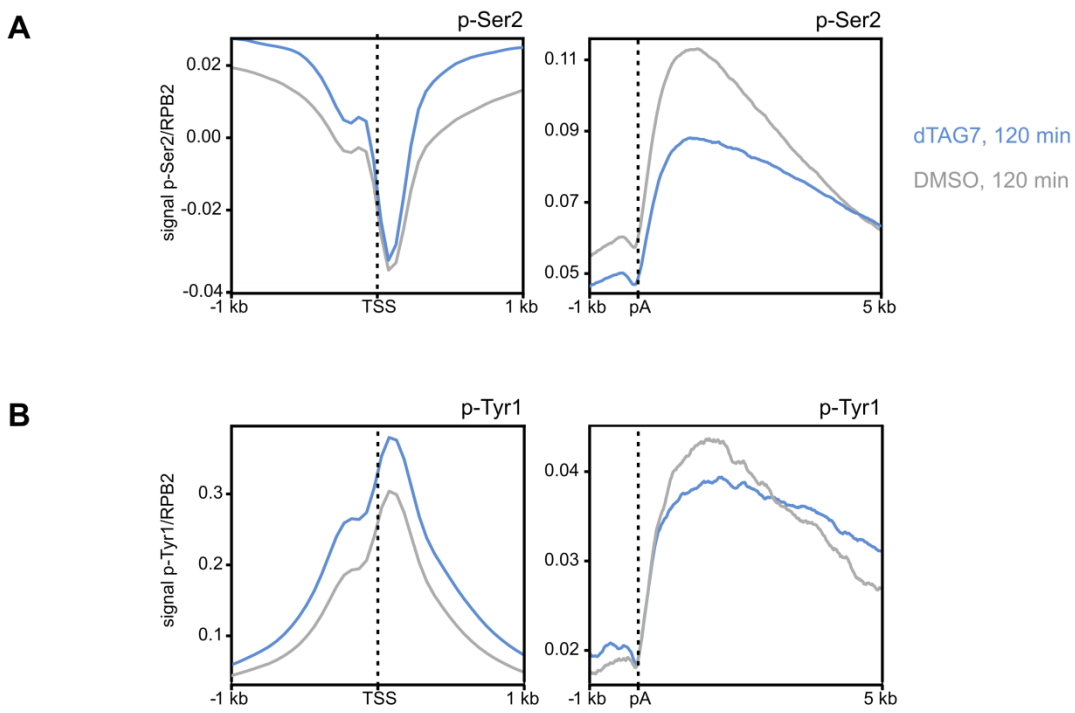


Figure 31: The increase in phosphorylated Ser2 and Tyr1 in the promoter-proximal region seen upon dTAG7 treatment does not merely reflect the increase in promoter-proximal pausing. To account for general changes in Pol II occupancy, the CTD p-Ser2 and p-Tyr1 signal (in RPM) were normalized by the RPB2 signal under the respective condition. **(A)** Meta-gene representation of the normalized p-Ser2 signal upon treatment with dTAG7 (120 min; blue) or DMSO (grey) centered at the TSS or pA site (indicated by dotted lines). **(B)** Meta-gene representation of the normalized p-Tyr1 signal centered at the TSS or pA site. The analysis was performed by Mario Rubio.

4.2 BET PROTEINS AS REGULATORS OF CO-TRANSCRIPTIONAL SPLICING

4.2.1 THE EFFECTS OF BET DEPLETION ON THE CHROMATIN PROTEOME ARE PARTIALLY STABLE OVER 360 MIN

Chromatin-MS performed after 120 min of dTAG7 or dBET6 treatment emerged as a powerful strategy to identify early changes in chromatin composition that were possibly a direct consequence of the ablation of BRD4 or BET proteins and might underlie the defects seen on the nascent RNA level. To determine whether these changes were stable over time or transient, which could be an indicator of compensatory processes, we analyzed the chromatin proteome upon 360 min of pan-BET or BRD4-selective degradation. Upon 360 min of dBET6 treatment, 424 proteins (24%) were significantly depleted from the chromatin, while 213 proteins (12%) were significantly increased (Figure 32A). Upon an extended dTAG7 treatment, amounts of only 78 proteins (8%) and 42 proteins (4%) were decreased or increased, respectively (Figure 32B). The overlap between proteins affected by 120 or 360 min of the same treatment was relatively small, showing that longer treatment had a quantitatively but also qualitatively distinct effect on the chromatin proteome. The proteins significantly depleted upon dBET6 treatment on average were decreased by 17%, those depleted upon dTAG7 treatment by only 8%.

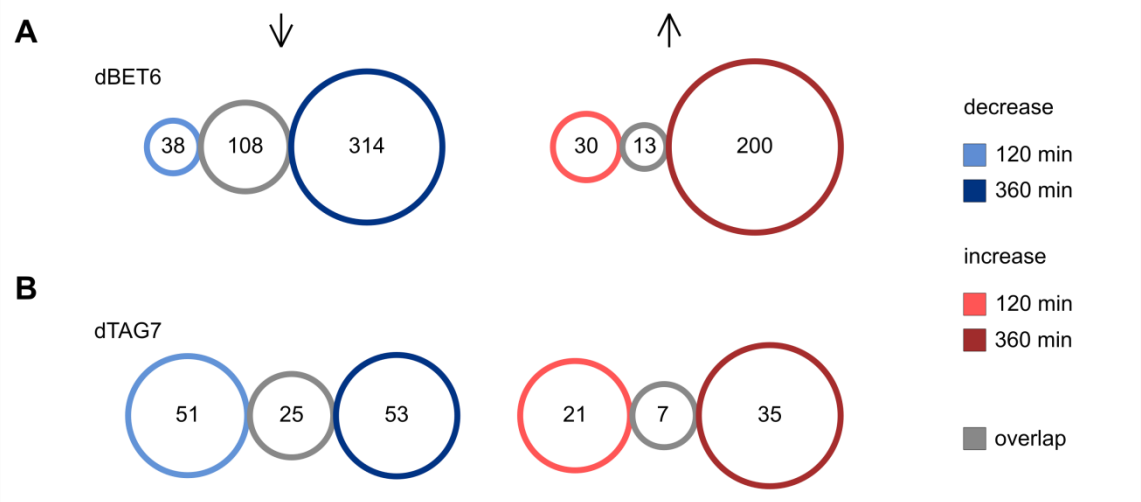


Figure 32: Extending the degrader treatment from 120 to 360 min leads to more severe changes in the chromatin proteome. The Venn-like diagram illustrates the number of proteins significantly increased (red) or decreased (blue) after 120 min (bright color) or 360 min (dark color), or at both time points (grey). Chromatin of K562 dTAG-BRD4 cells treated with dBET6 (A) or dTAG7 (B) for 120 or 360 min was isolated and analyzed by mass spectrometry. Four to six biological replicates were prepared per condition. In total, 2,219 and 1,856 proteins or protein groups were robustly detected in the two dBET6 experiments, while 964 and 967 proteins were identified in the dTAG7 experiments. Statistical significance ($p \leq 0.05$) was determined by two-tailed two-sample t-test.

To eliminate the detection bias of the individual experiments, we extracted a set of 500 proteins that were detected at both time points upon each of the treatments and performed k-means clustering based on the fold-change (Figure 33A). Interestingly, although the fold-change overall was lower upon BRD4-specific depletion, most proteins displayed similar trends upon both treatments.

Only cluster I contained proteins that were stabilized at the chromatin upon treatment. This increase was visible after 120 min of BET or BRD4-selective ablation and further increased upon prolonged treatment (Figure 33B). Interestingly, a GO term analysis revealed a significant enrichment of functions related to “protein import into the nucleus” ($p = 1.64 * 10^{-4}$), “mRNA export from the nucleus” ($p = 1.64 * 10^{-4}$) but also “nucleus organization” ($p = 5.28 * 10^{-3}$) and “chromosome organization” ($p = 8.8 * 10^{-3}$). This might be indicative of compensatory processes induced by the cell to attenuate the effects of the treatment.

At the other end of the spectrum, proteins in cluster IV showed a highly significant destabilization that was more apparent at the late time point. Again, this was commonly observed for both treatments. This group of proteins was significantly linked to “maturation of LSU-rRNA” ($p = 1.77 * 10^{-5}$), “maturation of SSU-rRNA” ($p = 2.56 * 10^{-2}$) and “mRNA splicing, via spliceosome” ($p = 2.16 * 10^{-2}$).

A similar, albeit weaker effect was seen for the proteins in cluster III. The decrease towards the 360 min time point was highly significant for dBET6. In contrast, a still significant but weaker decrease in chromatin binding upon 360 min of treatment was observed for dTAG7, possibly suggesting compensation by BRD2 or BRD3. A GO term analysis on this set of proteins indicated an enrichment of terms associated with “spliceosomal snRNP assembly” ($p = 9.69 * 10^{-3}$), “mRNA splicing, via spliceosome” ($p = 8.62 * 10^{-9}$) and “maturation of SSU-rRNA” ($p = 2.73 * 10^{-5}$).

The most pronounced reduction across all treatments and time points was observed for cluster V. In case of dBET6, the reduction was on average constant after 120 and 360 min of treatment, while it was mildly but significantly stronger after 120 min than after 360 min of dTAG7 treatment. At least over the course of this experiment, chromatin displacement of these proteins could not be rescued. Proteins of this group were specifically enriched for transcription-related GO terms, in particular “transcription elongation from RNA polymerase II promoter” ($p = 1.26 * 10^{-4}$), “mRNA polyadenylation” ($p = 1.34 * 10^{-3}$) and “mRNA cleavage” ($p = 1.40 * 10^{-2}$).

Lastly, the localization of the proteins of cluster II was not affected by either treatment. This group of proteins was linked to the GO term “regulation of RNA metabolic process” ($p = 2.20 * 10^{-2}$).

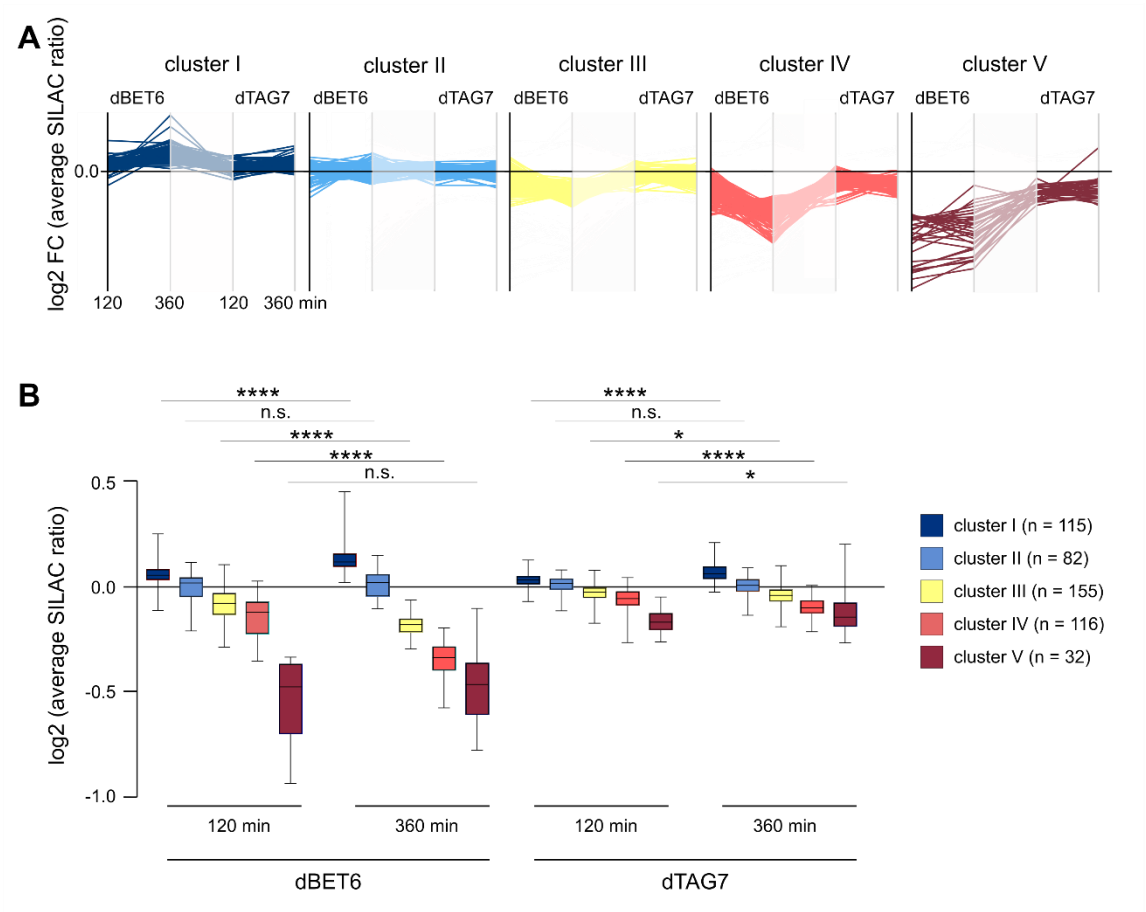


Figure 33: Proteins can be clustered based on the effect of degrader treatment on their chromatin association. (A) Chromatin of K562 dTAG-BRD4 cells treated with dTAG7 or dBET6 for 120 or 360 min in four to six replicates was isolated and analyzed by mass spectrometry. 500 proteins detected in all four experiments were extracted and k-means-clustered based on the average SILAC ratio using the Perseus software [404]. **(B)** Box plot quantification of the mean SILAC ratio of each cluster. Differences in the mean values of each cluster after 120 and 360 min of treatment were tested for statistical significance with an unpaired t-test with Welch correction. ****: $p \leq 0.0001$: ***: $p \leq 0.001$; **: $p \leq 0.01$; *: $p \leq 0.05$; n.s.: $p > 0.05$.

4.2.2 Chromatin binding of BRD2 and BRD3 increases in the absence of BRD4

We hypothesized that compensation for the BRD4 loss could be accompanied by an increase in BRD2 or BRD3 levels at the chromatin. Since BET proteins largely escaped detection in the chromatin-MS experiment, we repeated the 120 min treatment with dTAG7 or DMSO but analyzed all three cellular fractions by immunoblotting.

Quantification of BRD4 confirmed that the dTAG7 treatment was effective, as the total amount was decreased to 9% of the amount seen in the DMSO control after 120 min of treatment (Figure 34A). With 76% of the signal found in the chromatin fraction, BRD2 was the predominant BET protein associated with chromatin under control conditions, while the main part of BRD4 (57%) and BRD3 (71%) were located in the nucleoplasm. This was in contrast to a recent study that found more BRD3 and BRD4 than BRD2 bound to chromatin [288]. The relative distribution of BRD2 remained largely constant upon 120 min of dTAG7 treatment, whereas the amount of BRD3 at the chromatin decreased slightly from 21% to 12%.

To test if compensation occurs on an extended time scale, K562 cells were dTAG7-treated for 24 h, followed by cell fractionation. Cell numbers of control and treated cells were similar, indicating that loss of BRD4 did not compromise cell proliferation or survival. However, acidification of the medium upon DMSO and dTAG7 but not upon dBET6 treatment indicated an effect of specifically the latter treatment on the energy metabolisms, particularly on anaerobic respiration (not shown). Unchanged amounts at p-Ser2 and PCNA demonstrated that neither transcription nor DNA replication collapsed completely (Supplemental Figure 6). BRD4 levels were reduced to 5%, confirming the stability of the degrader compound over longer periods. After this extended treatment, relative BRD2 levels at the chromatin increased from 57% to 72%, while the fraction of BRD3 increased from 51% to 74% (Figure 34B). A possible interpretation thereof might be that the nucleoplasmic, unbound fraction of BRD2 and BRD3 acted as a reservoir of proteins and upon BRD4 ablation were re-localized to the chromatin.

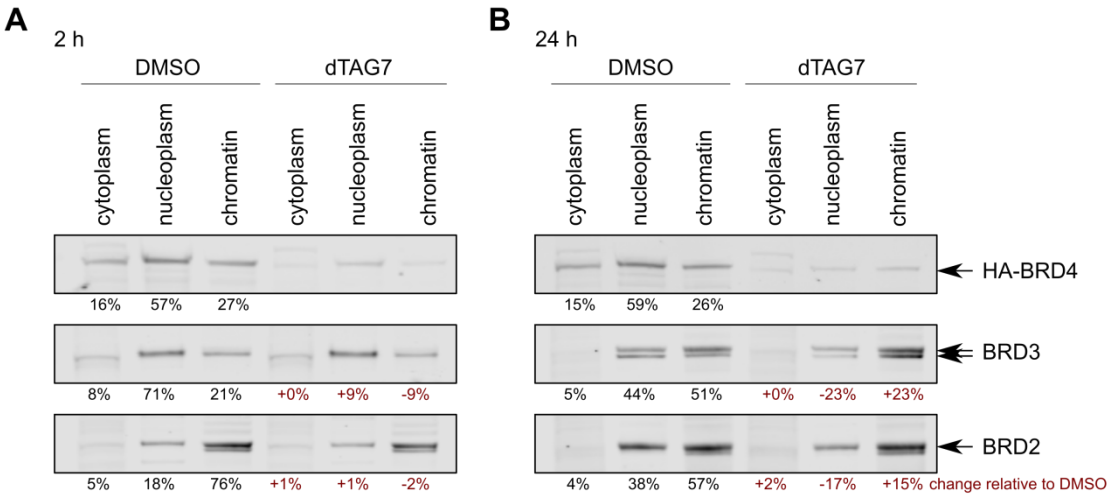


Figure 34: The chromatin-bound fraction of BRD2 and BRD3 is not increased upon 2 h but upon 24 h of dTAG7 treatment. K562 dTAG-BRD4 cells were treated with dTAG7 for 2 h (A) or 24 h (B), followed by cell fractionation and immunoblotting. To account for differences in loading, the signal intensity for the respective BET protein species was normalized by the signal of GAPDH. The relative distribution of the respective BET protein between cytoplasm, nucleoplasm and chromatin under control conditions (black) and the percental change in signal upon treatment (red) as compared to the corresponding DMSO control are indicated.

4.2.3 BRD2, BRD3 AND BRD4 SHARE INTERACTIONS WITH SEVERAL PROTEINS OF THE TRANSCRIPTION MACHINERY

Our initial finding of dTAG7 treatment recapitulating the effects of dBET6 on transcription elongation and termination suggested that BRD4 was the BET protein central to transcription regulation. On the other hand, differences in the severity of the defects suggested a contribution of BRD2 or BRD3. To assess whether redundancy between BRD4, BRD3 and BRD2 was probable, we based on native IP-MS experiments performed a comparative analysis of their interactomes.

574, 681 and 650 significant interactors were identified for BRD2, BRD3 and BRD4, respectively, indicating that the IP conditions were appropriate. Interestingly, all three BET proteins significantly interacted with one another (Figure 35A). To rule out that this was due to cross-reactivity of the antibodies, we compared the number of detected unique peptides and found that the respective target protein of the antibody used for immunoprecipitation showed the most efficient enrichment.

Surprisingly, not only BRD4 but also BRD2 and BRD3 were found to be associated with the Pol II elongation machinery. Each BET protein efficiently enriched the two largest

subunits of Pol II, while additional smaller subunits (RPABC1, RPB9) were only found in the BRD4 IP experiment (Figure 35B). BRD4 but also BRD2 interacted with both SPT4 and SPT5. While BRD3 was found to interact only with SPT5, the absence of SPT4 might be related to the lower number of expected tryptic peptides. Although the entire NELF complex was detected only by BRD4 IP, NELFA and NELFE were also identified, albeit with lower significance, by IP of BRD2 and BRD3. IPs for all BET proteins enriched CDK9 and cyclin T1, although the enrichment of CDK9 by BRD3 was not significant, confirming that the C-terminal domain of BRD4 is not required for interaction with P-TEFb. Of the subunits of the CDK9-containing super-elongation complex (SEC), ELL, AFF1 and AFF4 interacted with BRD4, but not with BRD2 or BRD3. Unexpectedly, PAF1, CDC73 and LEO1 were significantly enriched by BRD3 and BRD4 IP, whereas BRD2 interacted only with CDC73 and LEO1. Also, SPT6 and IWS1 were identified as interactors of all BET proteins. Lastly, only BRD2 appeared to interact with CDK12.

Given the possible initiation defect seen upon dBET6 treatment, we additionally focused on GTFs and the Mediator complex. In accordance with their known co-localization at enhancers and promoters, a remarkable majority of 25 Mediator subunits, including CDK8, CCNC, MED12 and MED13 of the kinase module, were enriched by BRD4 IP, while only sporadic interactions were found for BRD2 and BRD3 (Figure 35C). Additionally, BRD4 and BRD3 but not BRD2 emerged as interactors of almost all subunits of TFIID. Moreover, only BRD4 interacted with both TFIIF subunits and GTF2H1. Based on these results, BRD4 rather than the other BET proteins seemed to be involved in transcription initiation.

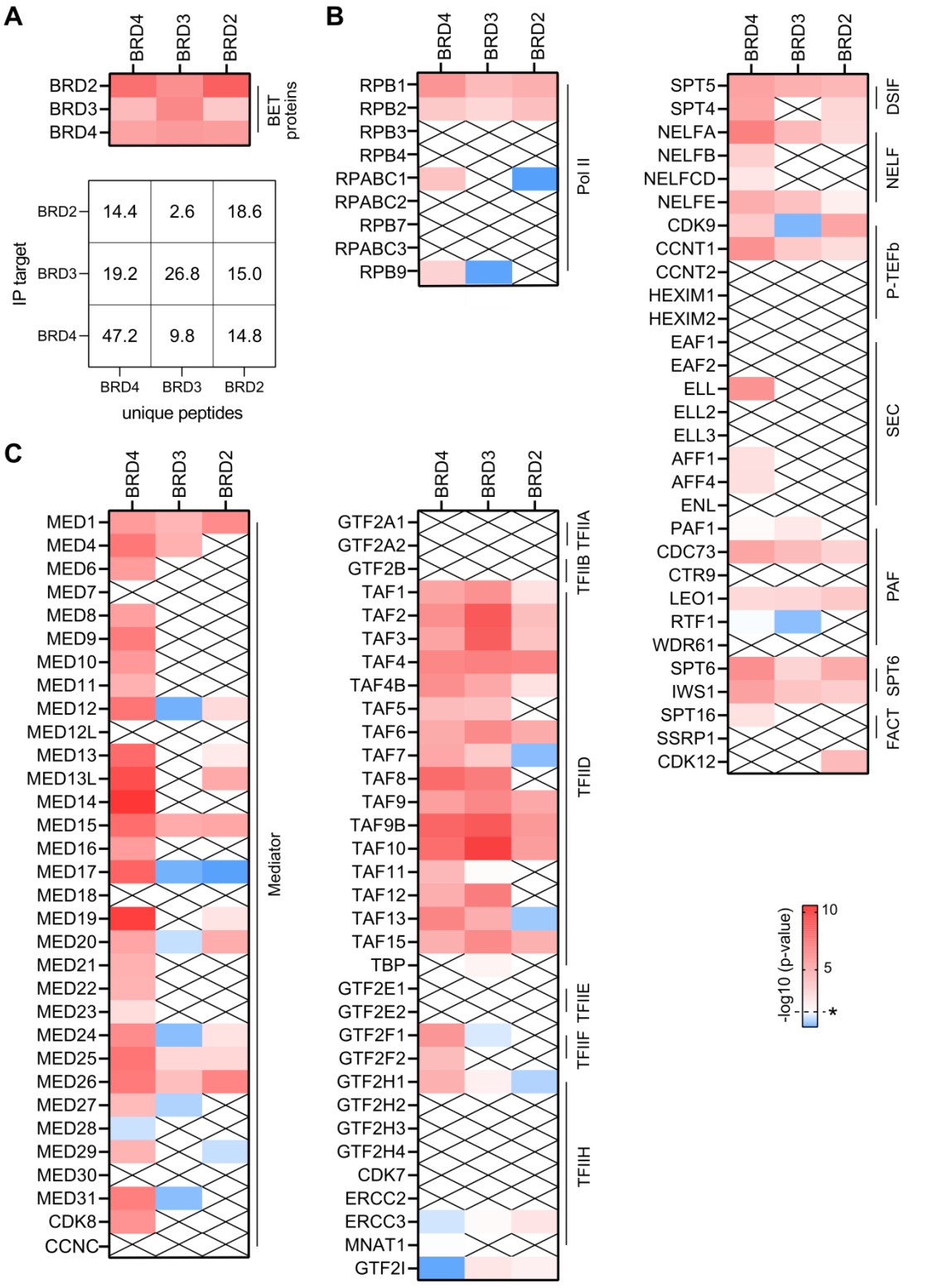


Figure 35: BRD4, but also BRD2 and BRD3, interact with the key proteins involved in transcription initiation and elongation. (A) Interactions between BET proteins. Native IP-MS of the endogenous BRD2, BRD3 and BRD4 (see also 4.1.14) was performed in five biological replicates. Significance of the IP enrichment (*: $p \leq 0.05$) was determined by two-tailed two-sample t-test and is indicated by color. The mean counts of unique peptides detected in each IP experiment are listed in the table below. **(B, C)** Heat map representation of the interactions between BRD2, BRD3 and BRD4 with Pol II subunits, initiation and elongation factors. Proteins not detected or filtered out are indicated by “x”.

Lastly, we found interactions of the CstF and CPSF modules of the RNA 3' end processing machinery to occur also with BRD2 and BRD3. In particular, all CstF subunits and most CPSF subunits except CPSF160 and WDR33 were significantly enriched by BRD2 and BRD3 IP, but, except for CstF64, WDR33 and symplekin, were undetectable or insignificant in the BRD4 IP (Figure 36A). Despite no effect of dTAG7 and dBET6 on chromatin localization of CFIm and CFIIm, we could identify interactions of both with BET proteins. In particular, CFIm59 was found to interact with BRD2, while CFIm25 was enriched by immunoprecipitation of BRD4 and BRD3. Surprisingly, XRN2 appeared as a highly significant interactor of all three BET proteins.

Although not primarily involved in the termination of protein-coding and long non-coding RNA genes, also the Integrator complex was included in this analysis. The almost entire complex was significantly enriched by BRD4 IP, while fewer subunits were part of the BRD3 and the BRD2 interactomes (Figure 36B).

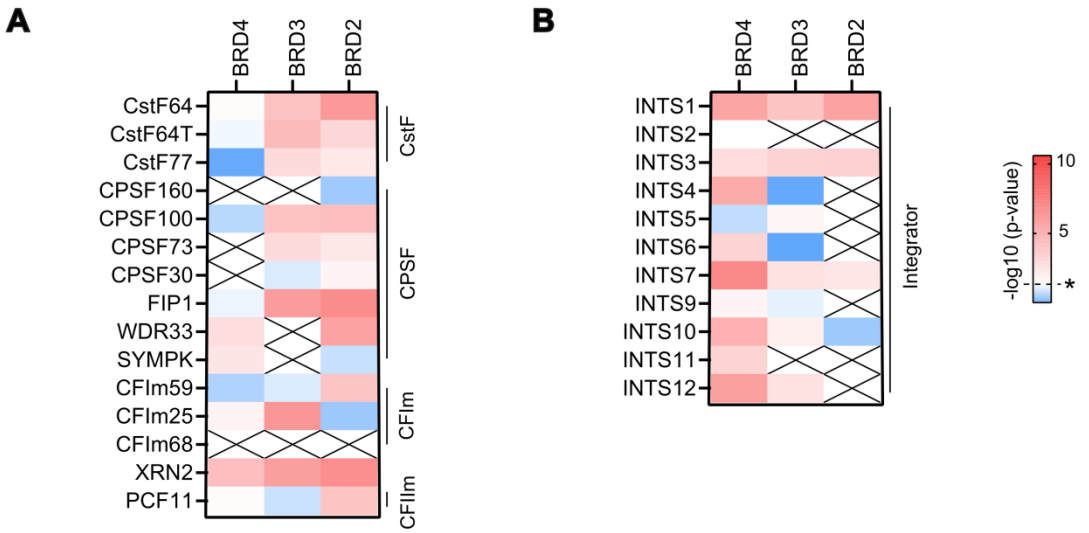


Figure 36: BRD2, BRD3 and BRD4 interact with RNA 3' end processing and termination factors (A) and with subunits of the Integrator complex (B). Native IP-MS using antibodies targeted BRD2, BRD3 and BRD4 was performed in five biological replicates. The significance of the enrichment by IP (*: $p \leq 0.05$) was determined by two-tailed two-sample t-test and is indicated by color. Proteins not detected or filtered out are indicated by "x".

4.2.4 CHROMATIN ASSOCIATION OF SPLICING FACTORS IS PARTIALLY BRD4-DEPENDENT

Based on the chromatin-MS clustering approach, chromatin association particularly of splicing factors appeared to depend on BET proteins or specifically BRD4 and was also subject to compensation processes. Therefore, we systematically assessed the effect of a 120 min treatment with each of the degraders on individual splicing factors. 154 of the 237 known or putative splicing factors that we considered in our analysis were robustly detected in ≥70% of the dTAG7 replicates, 31 of which were significantly ($p \leq 0.05$) depleted from the chromatin. In the dBET6 experiment, 187 splicing factors were detected, while 45 were significantly depleted from the chromatin. When considering all robustly detected (*i.e.*, not only significantly affected) splicing factors, the log2 fold-changes of the SILAC ratios upon dTAG7 and dBET6 were well correlated (Pearson correlation coefficient $r = 0.7131$).

Proteins of most spliceosomal snRNPs were affected by dTAG7 treatment (Figure 37). A majority of U1 and the U2 snRNP proteins were lost from the chromatin, although the reduction of not all of these proteins reached significance. SNRPA, SF3A3, SF3B1 and SF3B2 but none of the smaller SF3B proteins were significantly reduced. Besides its role in splicing, U1 snRNP is known to affect pre-mRNA polyadenylation in a process termed telescripting [438]. The SF3B module has an important function in branch site recognition and thus in formation of the A complex [439]. Moreover, two of the structurally essential Sm proteins were significantly depleted.

Strikingly, all of the detected U5 snRNP subunits except PRPF6 were also significantly depleted. In contrast, none of the U4/U6 di-snRNP-specific proteins showed a significant decrease at the chromatin. Given the concerted recruitment of the U4, U5 and U6-associated proteins as a tri-snRNP, and the coordinated binding of the U5 and U6 snRNPs throughout the activated and catalytically active phases of splicing, this finding was unexpected and might point to a distinct role of the U5 snRNP. Additionally, it also argued against the global reduction in transcription as a cause of the lower splicing factor quantities. Also proteins of the PRP19 complex and many of its associated proteins were significantly depleted. PRP19 is stably integrated into the assembling spliceosome upon formation of the activated spliceosome ($B_{activated}$) [440], in which also the U5 snRNP has a central role. Among the significantly depleted proteins was also ISY1, which together with CPSF73 controls biogenesis of certain miRNAs [441].

The detectability of non-snRNP proteins present first or only in the A, B, $B_{activated}$ and C complexes overall was lower than that of the spliceosomal subunits, possibly pointing to a limitation of chromatin-MS in detecting more dynamic protein associations. Of the B complex-specific proteins, only MFAP1, a known interactor of PRPF38A, which was close to significance ($p = 0.056645$) [442], was significantly depleted. Interestingly, the decrease in CWC27 and CWC22, which form a heterodimer in the $B_{activated}$ complex and based on recent X-ray crystallography structures dissociate from the spliceosome at the transition to the C complex [443], was highly significant.

While only the change of SAP18 was statistically significant, most proteins of the exon junction complex (EJC) or the THO and ASAP complexes were more abundant in the chromatin fraction upon treatment. Only MAGOHB was significantly depleted, with its known interactors MAGOH, RBM8A and EIF4A3 showing a similar but insignificant trend.

Most splicing regulators of the SR protein and hnRNP families did not change significantly upon treatment. Exceptions from this were hnRNPA0 and hnRNPM, which were significantly depleted, and hnRNPL, which showed a mild increase at the chromatin. Overall, a majority of proteins of these categories showed a tendency towards an increase in amount.

Lastly, CCAR2 was highly significantly depleted upon BRD4-selective degradation, while its interacting protein ZNF326 showed a significant increase. The DBIRD complex formed by both proteins is known to interact with Pol II and nascent RNA-containing RNPs [444].

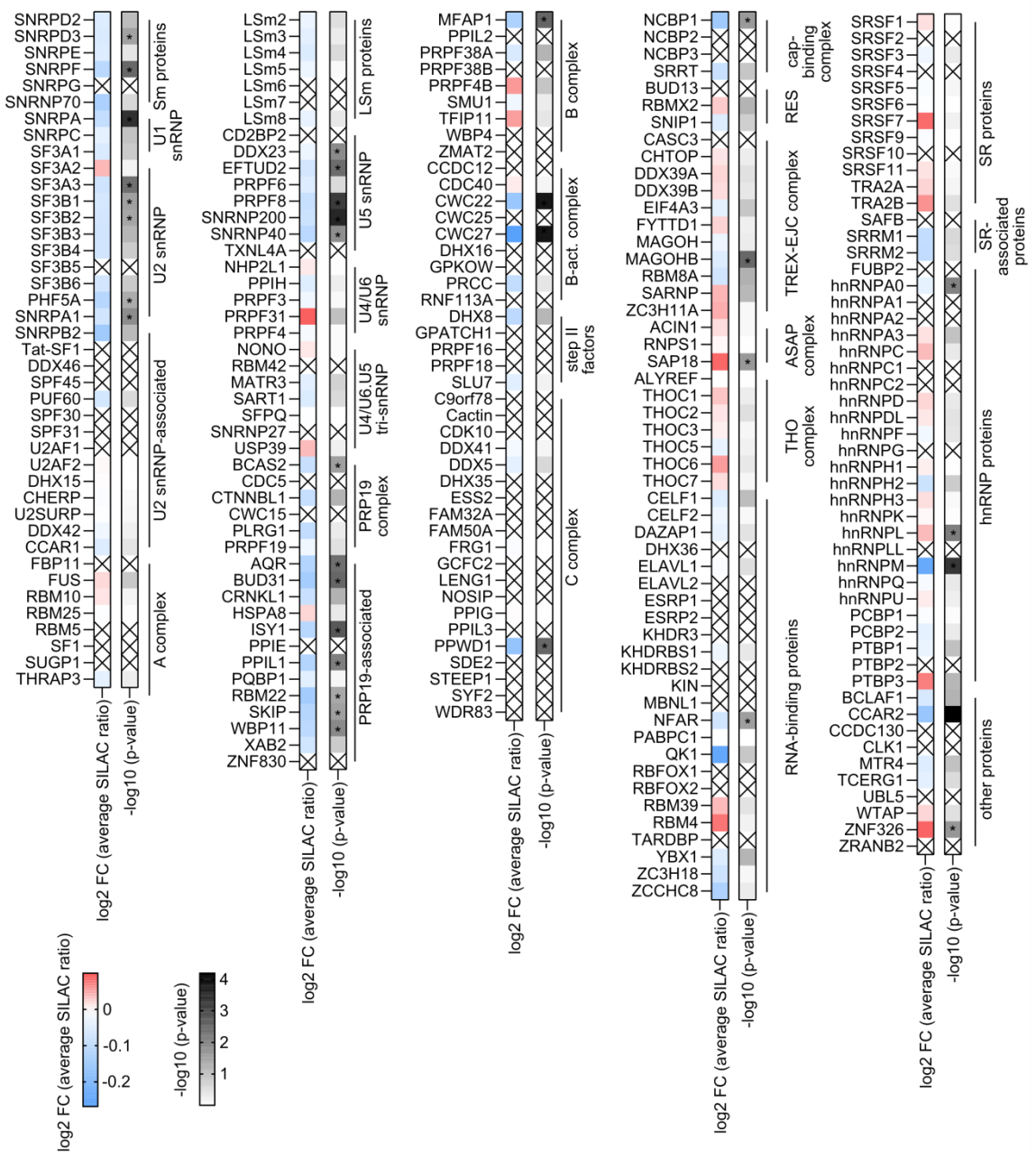


Figure 37: Splicing factors are partially depleted from the chromatin upon 120 min of dTAG7 treatment. Chromatin of K562 dTAG-BRD4 cells treated with dTAG7 in six biological replicates was isolated and analyzed by mass spectrometry. Statistical significance ($p \leq 0.05$) was calculated by two-tailed one-sample t-test and is indicated by the gray scale and "*". An increase or decrease in the chromatin fraction upon treatment is indicated by shades of red or blue, respectively. Proteins not detected or filtered out are marked by "x". The list of splicing factors was curated based on Agafonov et al. [264], Wahl et al. [445] and the HGNC database [398].

Similarly to the observation upon dTAG7, proteins of the U1 and U2 snRNPs were robustly detected and, with a few exceptions, depleted from the chromatin upon dBET6 treatment (Figure 38). Particularly, U2AF2, a subunit of the U2 auxiliary factor necessary

to recognize the intronic polypyrimidine tract downstream of the branch point, but also involved in PRP19 recruitment [446], and U2SURP, a poorly characterized interactor of CHERP and SPF45 (= RBM17) [447] were significantly depleted.

A significant reduction was also seen for EFTUD2 and SNRNP40 of the U5 snRNP. Additionally, SNRNP200 was just below the significance threshold ($p = 0.0571$). dBET6 treatment resulted in a consistent reduction of PRP19 subunits and associated factors at the chromatin. In particular, amounts of BUD31, CRNKL1, PPIE, RBM22 and XAB2 were significantly reduced. While XAB2 and PPIE, together with PRPF19, ISY1 and AQR that were not significantly depleted, form a physical complex involved in transcription-coupled DNA repair [448], the other three significantly depleted proteins are only loosely associated with PRP19 [449,450].
A significant decrease in CWC22 and CWC27 as well as of PPWD1 was found, paralleling the observation made upon dTAG7 treatment. Among other proteins, also CWC25 and the ATP-dependent helicases DHX5, DDX16 and DDX41. Overall, the $B_{activated}$ and C complexes appeared to be more severely affected by dBET6 than dTAG7 treatment.

Like in case of dTAG7 treatment, SR and hnRNP proteins, with few exceptions, were not dissociated from the chromatin upon dBET6 treatment. Intriguingly, hnRNPM was significantly depleted in both experiments.

With EIF4A3, RBM8A, MAGOH, MAGOHB and SAP18, five EJC proteins were significantly depleted upon dBET6 treatment, indicating a more pronounced effect than upon BRD4 degradation. Contrarily, THO complex proteins overall showed an insignificant increase, as was also observed upon dTAG7 treatment.
Lastly, we found QKI, BCLAF1, but also CCAR2, which was significantly affected also upon BRD4 degradation, to be significantly depleted upon pan-BET degradation. WTAP, a protein known to regulate splicing by recruiting RNA methyltransferases [451], which showed a mild, not significant increase upon dTAG7 treatment, was significantly increased upon dBET6 treatment.

In summary, the splicing factors SNRPD3, SNRPF, SF3B1, EFTUD2, SNRNP40, HNRNPM, RBM22, BUD31, CWC27, CWC22, PPWD1, MAGOHB and CCAR2 were depleted from the chromatin upon both treatments.

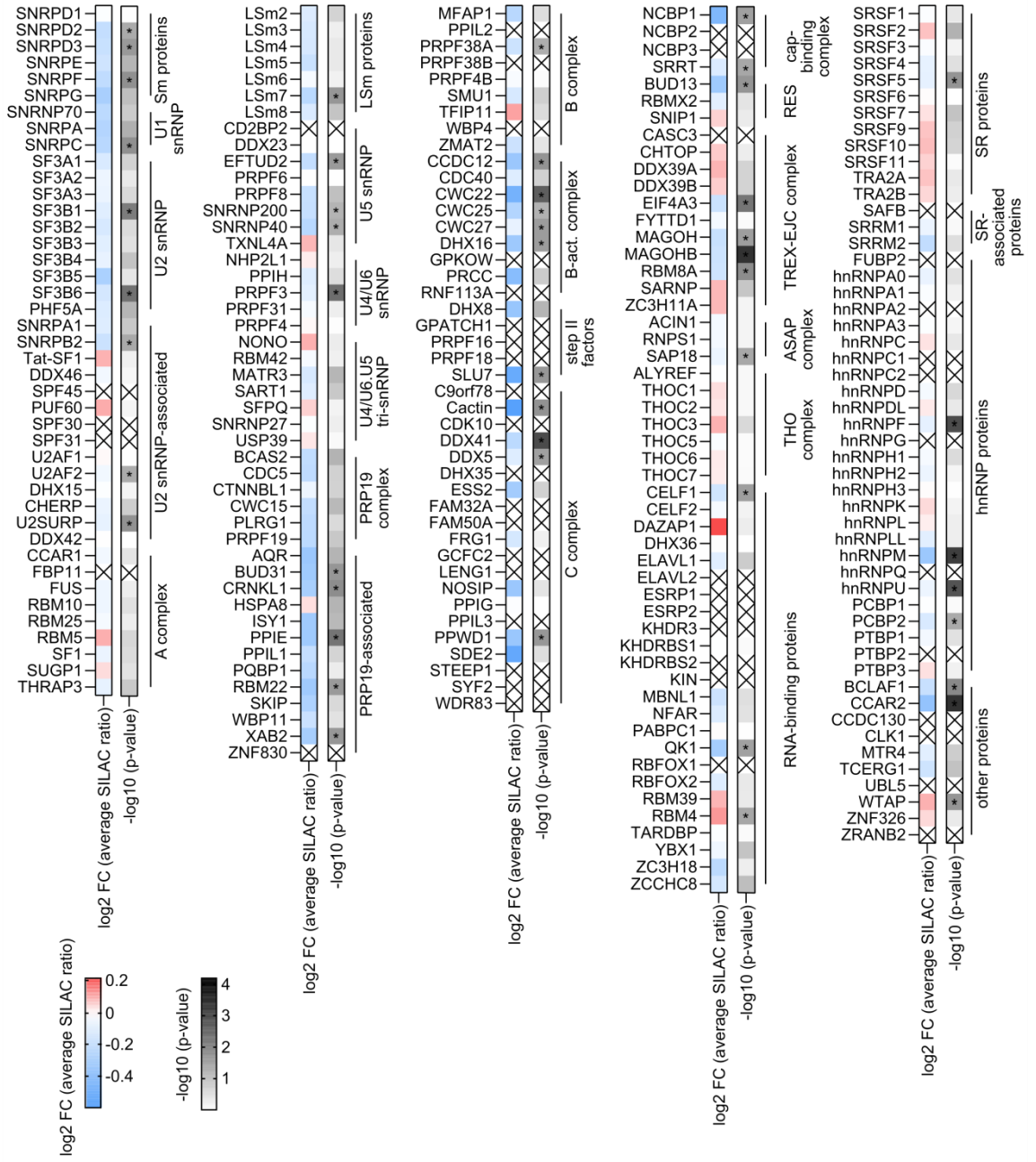


Figure 38: 120 min of dBET6 treatment reduces chromatin binding of some splicing factors. Chromatin of K562 dTAG-BRD4 cells treated with dBET6 in four biological replicates was isolated and analyzed by mass spectrometry. Statistical significance ($p \leq 0.05$) was calculated by a two-tailed one-sample t-test and is indicated by the gray scale and "*". An increase or decrease in the chromatin fraction upon treatment is indicated by shades of red or blue, respectively. Proteins not detecte or filtered out are marked by "x". The list of splicing factors was curated based on Agafonov et al. [264], Wahl et al. [445] and the HGNC database [398].

4.2.5 BET PROTEINS HAVE DIFFERENT BINDING PREFERENCES TOWARDS SPLICING FACTORS

To determine whether splicing factors and BET proteins co-localize within the cell, we reanalyzed our native IP-MS data with regard to splicing factors. Surprisingly, while numerous interactions were found, the overlap with factors that dissociated from chromatin upon degrader treatment was limited.

Overall, more snRNP factors interacted with BRD2 and BRD3 than with BRD4 (Figure 39), although a technical detection bias cannot be excluded. Interactions with the SF3 complex and U2AF and U2SURP were shared between all BET proteins, while interactions with most U4-, U5- and U6-associated proteins appeared to be specific for BRD2 and BRD3. Also, several proteins that are not core snRNP components but associate with the A and B complexes were consistently enriched in all three IP experiments, including RBM25, THRAP3 and MFAP1. Despite the pronounced effect of the degrader treatment on chromatin localization of proteins of the $B_{activated}$ complex, only CDC40, CWC27 and PRCC interacted with all BET proteins. In line with our chromatin-MS experiments, subunits and accessory proteins of the PRP19 complex were very consistently detected.

Lastly, we could detect numerous SR-rich and hnRNP proteins among the interactors of BRD2, BRD3 and BRD4.

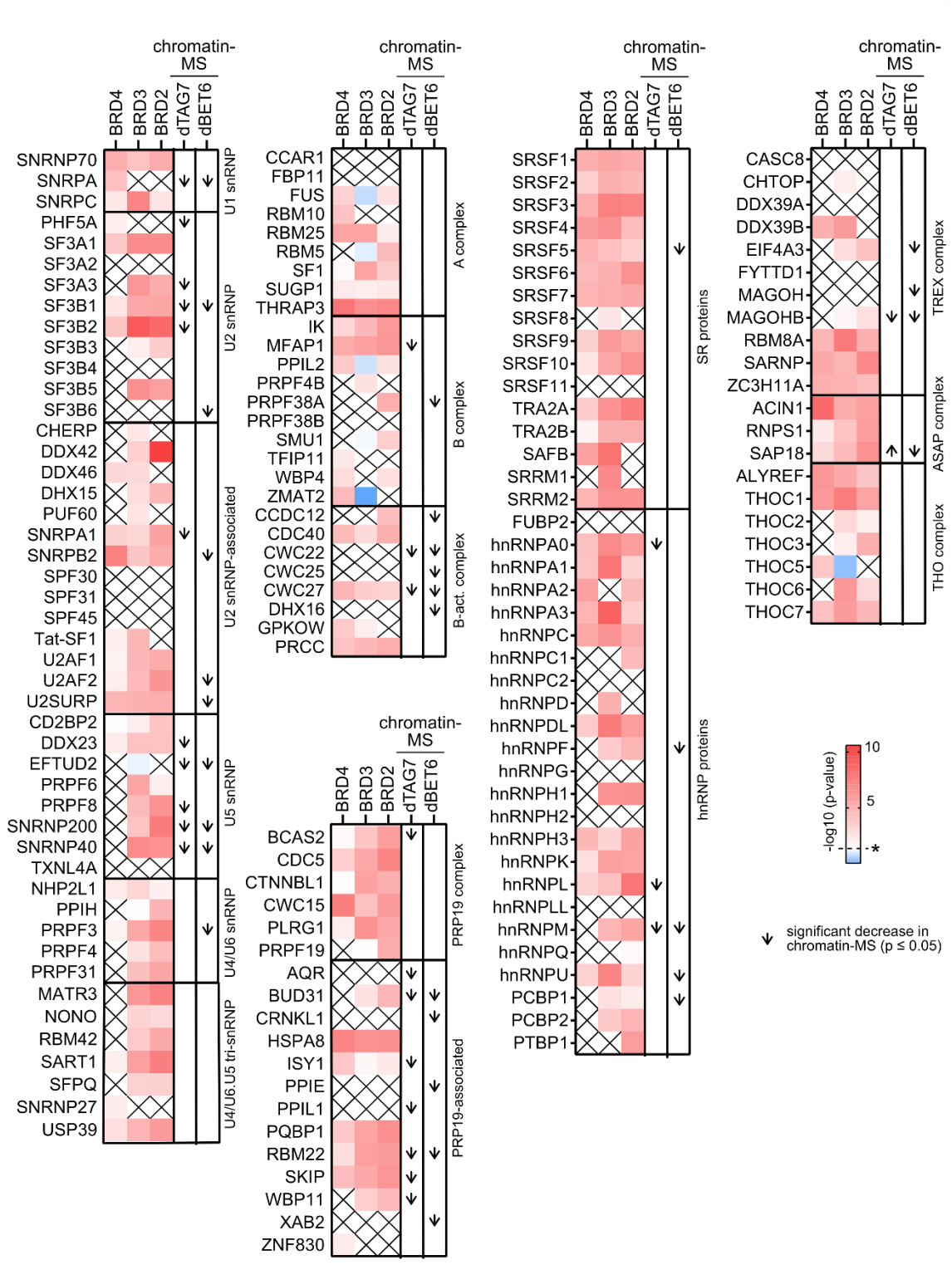


Figure 39: Splicing factors physically interact with BET proteins. Native IP-MS using antibodies targeted against BRD2, BRD3 and BRD4 was performed in five biological replicates. Significance of the IP enrichment (*: $p \leq 0.05$) was determined by two-tailed two-sample t-test and is indicated by color. Proteins not detected by IP-MS in a specific sample or filtered out are indicated by “x”. Arrows mark significant changes seen by chromatin-MS upon 120 min of dTAG7 or dBET6 treatment.

Surprised that only BRD2 and BRD3 but not BRD4 appeared to interact with the U5 snRNP in these proteome-wide experiments, we performed additional native IP experiments for PRPF8, an important scaffolding protein of the U5 snRNP [452], and the U2 snRNP protein SF3B1, followed by immunoblotting for BRD2, BRD3 and BRD4. Also, reverse IP experiments were performed. Interestingly, we found each of the BET proteins to interact with both tested splicing factors (Figure 40A, B). Despite the lower signal intensity in the input control, the BRD4 signal in the SF3B1 and PRPF8 lanes was comparable to the signal of BRD2 and BRD3. Also, as seen before, all BET proteins appeared to interact with each other. By and large, it appeared plausible that IP-MS did not detect interactions of BRD4 and the U5 snRNP for merely technical reasons.

Speculating that the interaction between BRD4 and some of the splicing factors could occur in the promoter-proximal region, we reanalyzed ChIP-seq data for PRPF8 that were generated in K562 dTAG-BRD4 cells under control conditions. Interestingly, an enrichment of PRPF8 was found downstream of the TSS (Supplemental Figure 7).

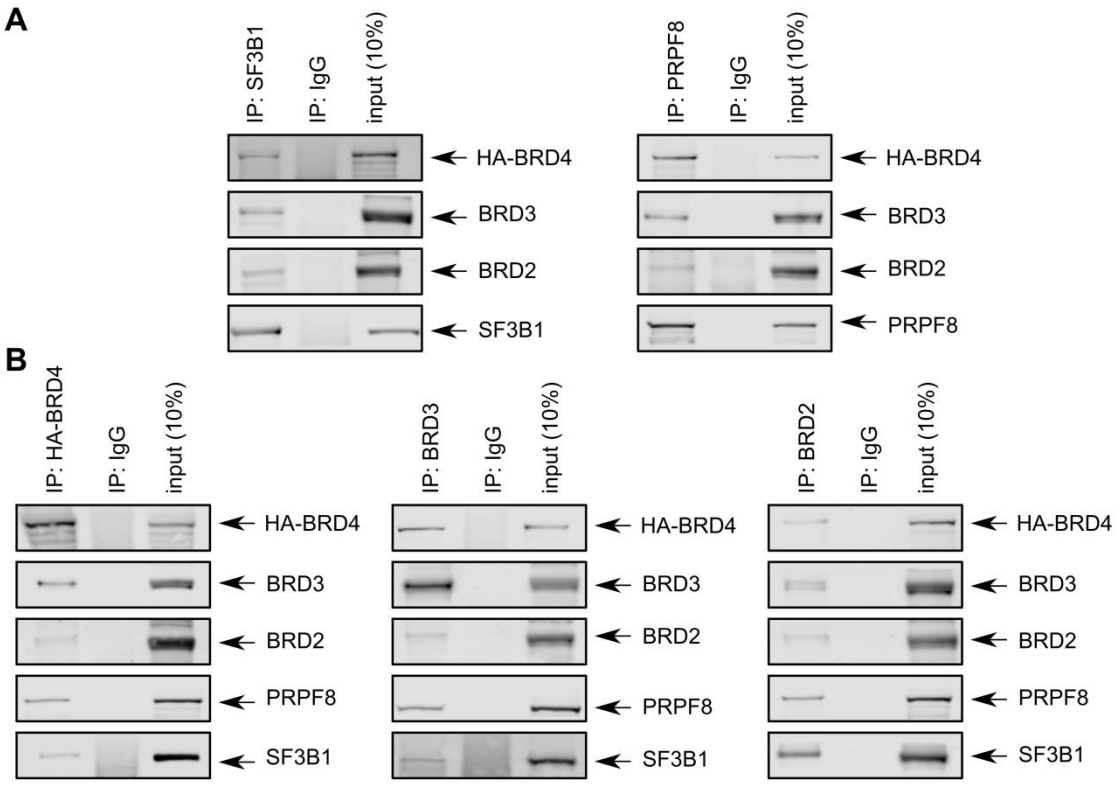


Figure 40: BRD2, BRD3 and BRD4 interact with PRPF8 and SF3B1. Native IP was performed using antibodies targeted against PRPF8 and SF3B1 (A) or BRD2, BRD3 and BRD4 (B), followed by SDS-PAGE.

4.3 BRD4 AND ENHANCER TRANSCRIPTION

4.3.1 POL II AND SPT5 ACCUMULATE AT ENHANCERS UPON TARGETED BRD4 DEGRADATION

Pol II is known to undergo pausing not only at gene loci but also at enhancers [453]. Conversely, enhancers, specifically those producing eRNA, have been implicated in regulating early elongation at genes [315], for instance, by changing the local distribution of pausing factors [454]. Also, both SPT5 and PAF1 have been shown to bind enhancer regions [151,153,159]. In light of this, we asked whether BRD4 could affect transcription in a similar manner at genes and enhancers.

Indeed, the trends seen in the promoter-proximal region were partially recapitulated at enhancer centers. Since SI-NET-seq read coverage at enhancers was relatively low, we used input-normalized RPB2 ChIP-Rx data as a substitute. Upon BRD4-selective depletion, we observed an increase in signal at the center of annotated active enhancers in K562 [408], as seen in meta-gene plots (Figure 41A). Interestingly, a slight decrease was seen distally to the enhancer center, suggesting an elongation defect. Also, SPT5 locally accumulated upon BRD4 depletion (Figure 41B). Specifically, the SPT5 occupancy showed local maxima flanking the enhancer center, which indicated an association of SPT5 with bidirectionally transcribing Pol II. For PAF1 and SPT6, a decrease in input-normalized signal was seen, although the signal intensity overall was low.

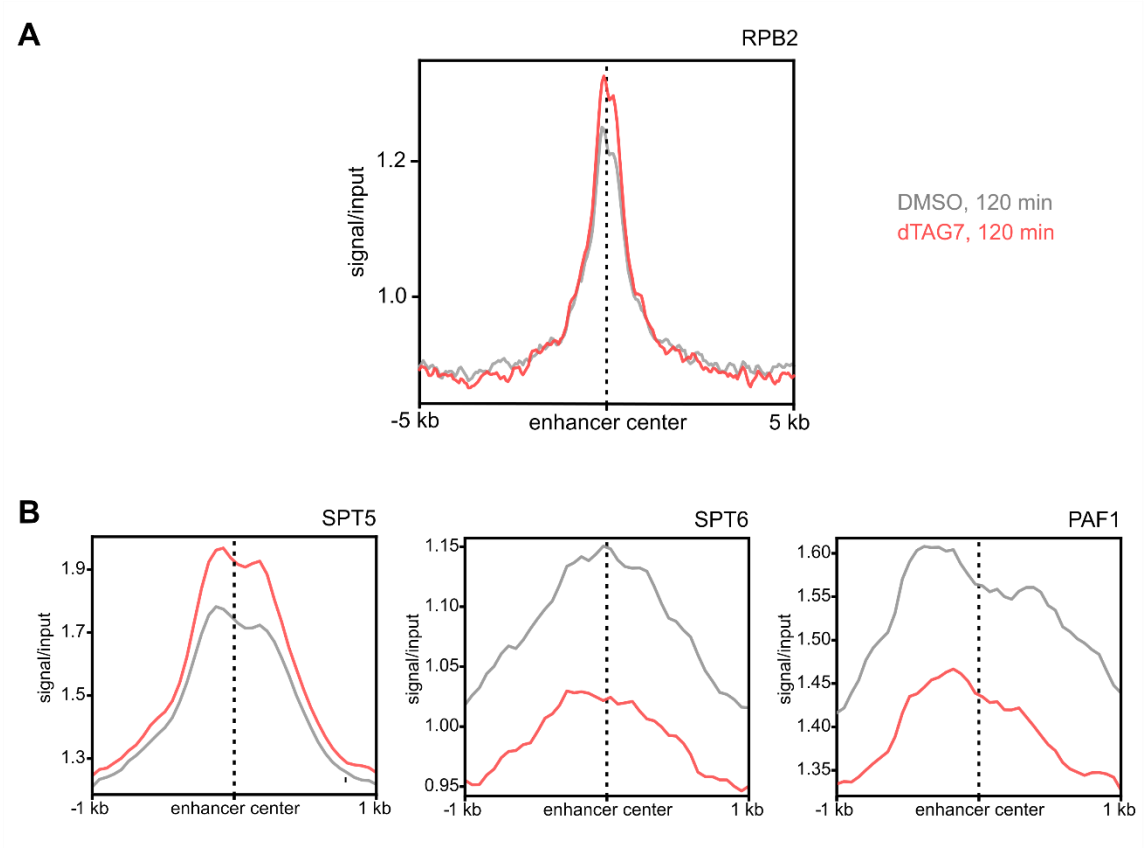


Figure 41: Similarly to the effect in the promoter-proximal region of genes, binding of Pol II and SPT5, but not SPT6 and PAF1, is locally increased at enhancer centers upon targeted BRD4 ablation. (A) Meta-gene representation of the input-normalized RPB2 occupancy aligned at the center of 3,228 enhancer regions in K562 upon dTAG7 treatment (red) and under control conditions (grey). **(B)** Input-normalized ChIP-Rx signal of SPT5, SPT6 and PAF1 at annotated enhancer centers.

4.3.2 BRD4 ABLATION HAS NO IMMEDIATE EFFECT ON CHROMATIN ACCESSIBILITY UNDER THE TESTED CONDITIONS

Transcription requires chromatin to be relatively open, as is reflected by the high degree of accessibility usually seen in promoter and enhancer regions. On the other hand, transcription itself evokes the local decompaction of the chromatin structure [455,456]. To test if BRD4 directly or indirectly influences chromatin accessibility, K562 dTAG-BRD4 cells were treated for 120 or 360 min with dTAG7 or DMSO, followed by an assay for transposase-accessible chromatin with sequencing (ATAC-seq).

Unexpectedly, all four samples irrespectively of the treatment or sampling time point were highly correlated (Pearson correlation coefficient r = 0.99). Likewise, visual inspection of the accessible regions did not reveal obvious differences. Accessibility correlated well with known regulatory sites, indicating reasonable data quality. Open regions at the

exemplary *c-MYB* locus showed a remarkable overlap with the putative TSS and binding sites of EP300 and its recruitment factor C/EBP-β and the hematopoietic transcription factor GATA1 (Figure 42A). A weaker but still clear association was found with binding sites of PU.1 and ETO2, antagonizers of GATA1 and the expression of erythroid genes [457,458]. Also, an accessible site 14 kb downstream of the *c-MYB* promoter previously characterized as a key enhancer of *c-MYB* transcription [459] was detected.

To quantitatively compare the samples, we performed a per-bin RPKM normalization of the ATAC-seq data. Notably, at both time points, a mild increase in accessibility was observed upstream of the TSS, putatively in the promoter regions, upon treatment (Figure 42B). A similar effect was seen at the center position of active enhancers (Figure 42C). Chromatin accessibility was slightly more strongly increased after 360 min of treatment at both types of regulatory elements, possibly suggesting indirect effects.

Lastly, we assessed whether DNA–DNA contacts were affected by 120 min of BRD4-selective degradation. To obtain a higher coverage at regulatory regions, Hi-ChIP was performed using an H3K27ac-specific antibody. Interestingly, a decrease in the interaction frequency was seen for interactions between promoters and intergenic enhancers as well as within each of these groups of elements (Figure 43). Interactions involving intragenic enhancers, gene body regions or 3’ gene ends showed a weaker reduction in the absence of BRD4.

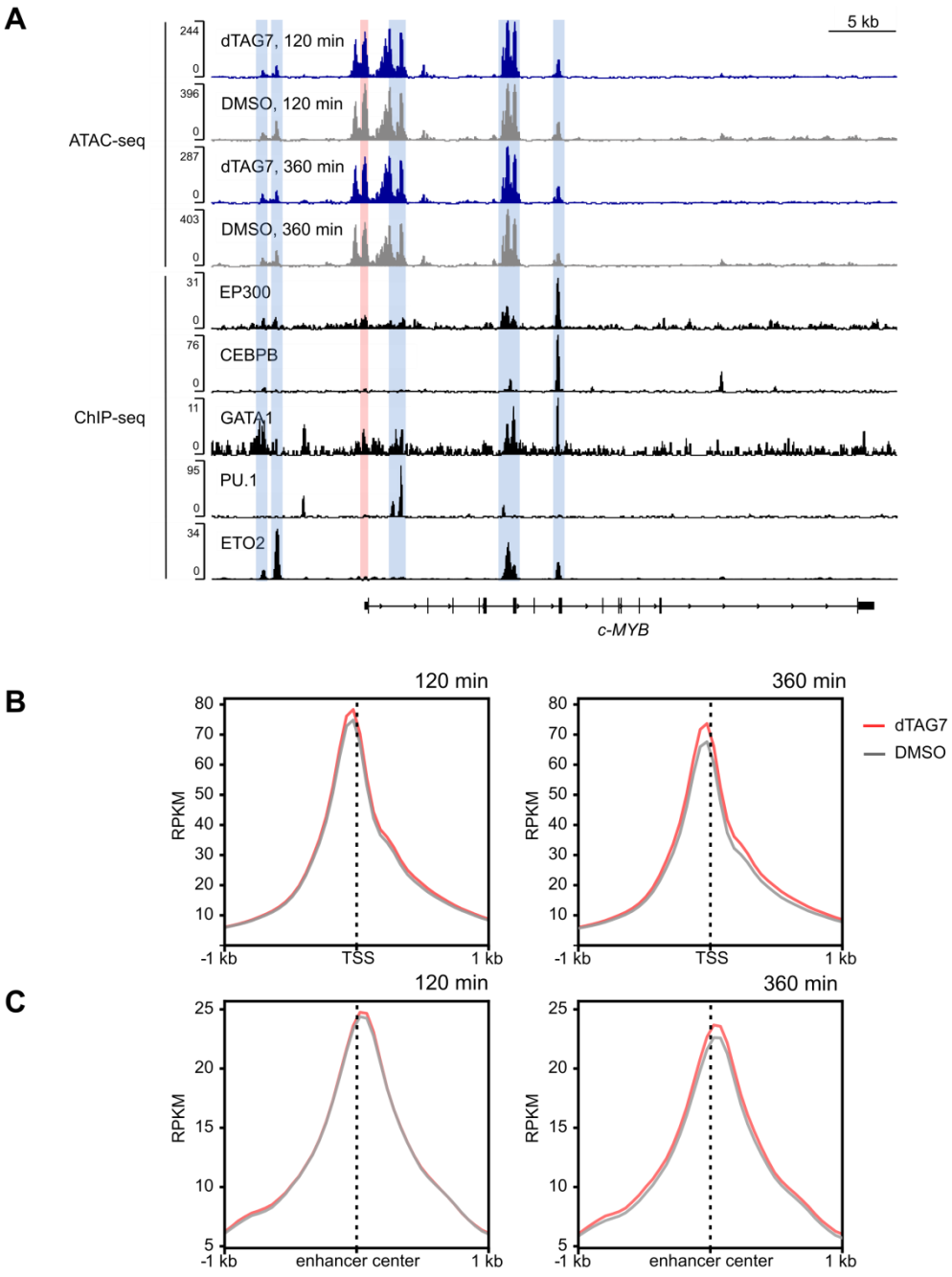


Figure 42: BRD4 ablation for 120 or 360 min has no apparent effect on chromatin accessibility. **(A)** Transposase-accessible sites at the exemplary *c-MYB* locus overlapped with binding sites of TFs involved in hematopoiesis. ATAC-seq was performed upon 120 min or 360 min of dTAG7 treatment in K562 dTAG-BRD4 cells. All ChIP-seq data shown were acquired from ENCODE. Putative regulatory regions and the promoter-proximal regiom are shaded in blue and red, respectively. **(B)** Meta-gene representation of the Tn5-accessible regions upon dTAG7 treatment (red) or under control conditions (grey). 13,400 gene regions actively transcribed in K562 were considered in this analysis. **(C)** Tn5-accessible regions detected upon dTAG7 treatment (red) or under control conditions (grey) at enhancer centers. Coordinates of 3,228 enhancer regions in K562 (indicated by dotted lines) were taken from Lidschreiber et al. [408].

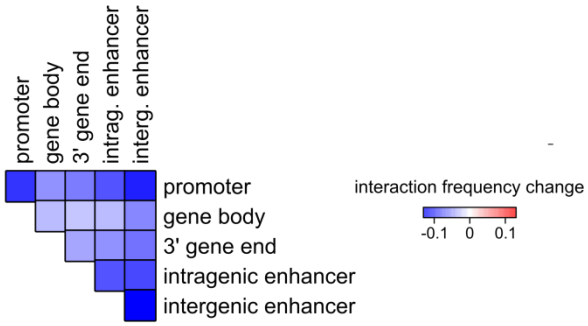


Figure 43: 120 min of dTAG7 treatment reduces DNA–DNA contacts primarily between promoters and enhancers. Changes in interaction frequency between gene regions and regulatory elements, as determined by Hi-ChIP using an H3K27ac-specific antibody upon 120 min of dTAG7 treatment or under control conditions. Gene regions were categorized using ChomHMM [417]. The analysis was performed by Annkatrin Bressin.

5 DISCUSSION

5.1 MECHANISTIC INSIGHTS INTO THE ROLE OF BRD4 IN EARLY ELONGATION

5.1.1 BRD4 AND BET DEPLETION GLOBALLY AFFECT ELONGATION BY PREVENTING EFFICIENT PAUSE RELEASE

Since the year 2000, BRD4 was mentioned in >2000 articles enlisted at NCBI Pubmed, predominantly in the context of human diseases. In this light, how limited our insights into the mechanistic action of BRD4 still are, is astonishing. Only part of this knowledge gap could be filled employing BET bromodomain-specific inhibitors, which induced incomplete phenotypes that—in case of more extended treatment times—reflected both direct and indirect molecular effects. Usage of PROTAC degraders critically advanced our understanding, as they cause rapid and near-complete protein removal. Here, by combining BET and BRD4-selective targeted degradation with different complementary high-resolution methods, we could uncover some of the functions these proteins fulfill in transcription.

Confirming earlier findings made in different cellular systems, BET ablation resulted in a global collapse of productive elongation. The selective depletion of BRD4 recapitulated this effect qualitatively, but not entirely quantitatively, such that more Pol II remained present in the gene body region. Given the similarity of both molecular phenotypes, loss of BRD4 likely also underlay the impact of dBET6 on transcription. Observing such a widespread elongation defect was surprising given that erythroleukemic K562 cells are often classified as relatively BET-insensitive [328].

Strikingly, a significant reduction of Pol II engaged in transcription occurred already within 40 to 50 min of dBET6 treatment. This was unexpected in view of published studies quantifying PROTAC degradation kinetics [420] and our immunoblots showing a reduction of cellular BRD4 levels by >40 % after 40 min of degrader treatment. However, ChIP-Rx upon 40 min of dTAG7 treatment revealed faster dissociation of BRD4 from the chromatin, possibly due to interference of the degrader and cereblon binding with recognition of acetylated histones by the bromodomains. Overall, the rapidness of the transcriptional effects indicates that BET depletion exerts an immediate effect on

transcription. Thus, BET proteins and specifically BRD4, rather than functioning as a relay, appear to be directly involved in transcription regulation.

A thorough analysis of the SI-NET-seq data also revealed changes in the promoter-proximal region. While the pausing index defined as the ratio of Pol II signal in the promoter-proximal to that in the gene body region is traditionally calculated to prove a defect in pause release, we, due to the use of spike-ins, were able to detect the absolute changes in Pol II occupancy upon treatment. Particularly, BRD4-selective degradation increased the Pol II signal in the promoter-proximal region. In contrast, levels of promoter-proximally paused Pol II were lower upon dBET6 treatment relative to the DMSO control. This increase in Pol II occupancy in the promoter-proximal region and the concomitant decrease at the gene body point to a defect in promoter-proximal pause release. In particular, Pol II in the absence of BRD4 appears capable of entering promoter-proximal pausing, but cannot transition into productive elongating efficiently, resulting in a "pile-up" of Pol II downstream of the TSS. Notably, although no absolute increase in paused Pol II was seen upon dBET6 treatment, the ratio of Pol II occupancy in the promoter-proximal and the gene body region nevertheless increased, indicating a pause release defect. Thus, the same molecular mechanism can be assumed to underlie the effects of both treatments.

Our results are in agreement with findings by Zheng et al. [137], who observed a promoter-proximal increase in Pol II occupancy upon AID-mediated depletion of the long BRD4 isoform, while degradation of either BRD2 or BRD3 resulted in the contrary effect.

5.1.2 BRD4 PROMOTES THE ASSEMBLY OF THE ELONGATION COMPLEX

We chose a quantitative, unbiased and proteome-wide approach to determine whether the chromatin association of elongation factors depends on BRD4. Although similar strategies have been described in the literature [460], we adapted the cell fractionation protocol of the NET-seq workflow to be used for chromatin analysis by mass spectrometry, and moreover combined it with SILAC labeling. Originally, this protocol has been specifically optimized to quantitatively (>95%) recover Pol II [59], what appeared as an advantage for the purification of transcription regulators. Furthermore, it does not rely on chemical crosslinking, which can increase background signal [461] or reduce protein detectability [418]. Lastly, using a nuclease instead of ultrasound in the final chromatin

solubilization step minimized protein damage but also protein losses, as no insoluble material was removed by centrifugation and thereby excluded from the MS analysis. Based on the results of chromatin-MS, the effects on selected proteins were confirmed by ChIP-Rx, which provided spatial resolution and offered better strategies for normalization.

A 120 min treatment with either dTAG7 or dBET6 treatment very consistently resulted in a significant decrease of core transcription elongation factors of the well-characterized activated elongation complex [87] from the chromatin. Upon each treatment, we found Pol II subunits to be most strongly depleted, which corresponds well with the observed elongation defect. To rule out that RPB1 was polyubiquitylated and degraded, like it has been observed upon acute depletion of SPT5 [151,152], we consulted our whole cell MS data collected upon dTAG7 treatment but could not detect a decrease in RPB1 or other subunits.

A similarly pronounced reduction like in Pol II was seen in DSIF subunits, which was not surprising given their stable association with Pol II across the transcribed region [57]. ChIP-Rx performed upon dTAG7 treatment confirmed the overall reduction but revealed local differences: SPT5 levels were increased in the promoter-proximal region but decreased along the gene body and near the pA site. However, occupancy changes partially mirrored the underlying binding pattern of Pol II: only the decrease in SPT5 binding in the termination zone was on average more pronounced than the reduction of Pol II, although a more pronounced reduction in the promoter-proximal and gene body region of individual genes cannot be excluded. In line with the idea that BRD4 aids SPT5 binding to the chromatin directly, immunoprecipitation revealed physical interactions between both proteins. An alternative possibility that was not tested in this study is that BRD4 affects SPT5 phosphorylation and thereby its activity [144].

A role of SPT5 in the BRD4-dependent elongation defect at first glance seems to be in contrast to the finding by Muhar et al. [325] that SPT5 occupancy near the TSS in untreated cells is a strong predictor of BET insensitivity at the gene level. However, we propose that the relative change in SPT5 upon treatment, rather than the absolute levels under control conditions, could be linked to the elongation defect. The transcriptional defect we observed upon dTAG7 treatment is not comparable to the effect of targeted SPT5 ablation [151,152] in that the latter perturbation prevented Pol II accumulation in the promoter-proximal region. This discrepancy argues that BRD4 does not recruit SPT5 prior to promoter-proximal pausing but might stabilize it during pause release and early

elongation. Nevertheless, the possibility remains that the decrease in SPT5 caused by BRD4 depletion—similarly to the effect seen upon targeted SPT5 depletion by Aoi et al. [152]—leads to a reduction in processivity, which would also manifest in a decrease in Pol II occupancy. Lastly, an involvement of SPT5 appears likely also in view of the correlation of BET sensitivity with gene length, as Pol II processivity in a narrow window 15 to 20 kb downstream of the TSS is particularly dependent on SPT5, resulting in the most pronounced effects of SPT5 depletion at long genes [150].

In support of a functional link between BRD4 and NELF, most of its subunits were identified in our two BRD4 IP-MS experiments, which is in accordance with findings by Lambert et al.[310]. Contrariwise, there was no evidence that BRD4 stabilizes NELF at the chromatin, as subunits of the NELF complex were not depleted by dBET6 treatment. Also, no shift in the pausing position, which has been proposed to be a characteristic upon targeted NELF degradation [120], was seen in the absence of BRD4, for instance, by RPB2 ChIP-Rx. Overall, our findings suggest that BRD4, rather than establishing or maintaining the paused state via regulating NELF, acts downstream of it.

Also PAF, another key elongation factor, is depleted from the chromatin in the absence of BRD4 or BET proteins. PAF core subunits consistently showed a fold-change slightly lower than that of Pol II subunits, while the dissociable, independently functional RTF1 [462] remained undetected. ChIP-Rx for PAF1 revealed a decrease in the promoter-proximal and gene body region and the termination zone upon BRD4-selective degradation, which was more pronounced than the decrease in RPB2, demonstrating that PAF was inefficiently recruited to Pol II. The fine structure of the ChIP-Rx profile at the 5' gene end suggested PAF1 association not with the paused but rather with the just-released Pol II, which is in accordance with the notion of a step-wise assembly of the elongation complex [88]. In agreement with Yu et al. [156] and Dawson et al. [328], we found several PAF subunits to interact with BRD4 by IP-MS and IP-WB. Based on recent cryo-EM structures, simultaneous binding of NELF and PAF to Pol II is sterically hindered, offering a mechanism that prevents re-association of NELF with the active elongation complex [87]. Detecting interactions of BRD4 with both complexes is surprising and not easily interpretable. One hypothesis is that BRD4 forms two distinct complexes with either PAF or NELF. Alternatively, BRD4 could be involved in the transition from pausing into elongation.

Functional studies on PAF present an inconsistent picture, which complicates the evaluation of our data. Our data appear to be in accordance with Yu et al. [156], who demonstrated a pause release-promoting role of PAF, and thus disagree with Žumer et al. [102], who proposed that destabilization of PAF results in a general decrease in Pol II velocity and therefore an increase in Pol II occupancy at the gene body, rather than a pause release defect.

Lastly, a strong reduction of SPT6 and its co-factor IWS1 was detected by quantitative chromatin-MS. Based on ChIP-Rx, PAF1 and SPT6 showed a similar binding pattern, as both accumulate slightly downstream of the promoter-proximal pause site. Additionally, also SPT6 was more strongly depleted from all tested gene regions than Pol II (not shown), suggesting an actual recruitment defect. Like all of the aforementioned elongation factors, SPT6 appears to interact with BRD4.
Also targeted SPT6 degradation has been shown to reduce the transcriptional output globally. Moreover, it leads to an increase in Pol II occupancy in the promoter-proximal region but to a decrease at the gene body [102,165], which greatly resembles the effect seen upon BET or BRD4 depletion. Accumulation of Pol II upstream of the +1 nucleosome, which has been seen upon SPT6 degradation [102], was not observed in our experiment as Pol II remains bound at the first, promoter-proximal pausing site.

Taken together, BRD4 depletion globally reduces the recruitment or stable binding of key elongation factors, each of which serves specific functions in early elongation. Thus, the effect of BRD4-selective depletion on transcription might be seen as a composite of the individual effects (Figure 44). Whether additional relevant factors are lost from the chromatin upon dTAG7 treatment remains to be tested.

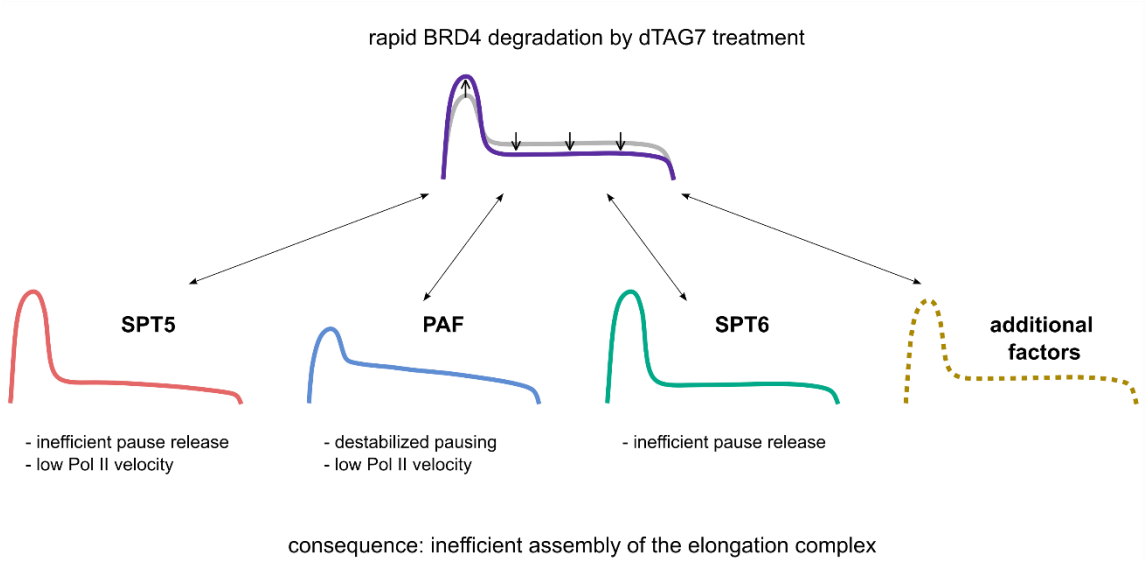


Figure 44: Loss of SPT5, SPT6 and the PAF complex from the chromatin contribute to the characteristic effect of targeted BRD4 depletion on transcription. (Model) Based on the current literature, rapid depletion of crucial elongation factors imposes distinct changes on Pol II in the promoter-proximal and the gene body region, which are not identical to but resemble the changes seen upon BRD4 degradation. Therefore, loss of SPT5, SPT6 and PAF from chromatin likely underlies the elongation defect seen upon BRD4 depletion.

BRD4 initially was claimed to recruit P-TEFb to the paused Pol II, but more recent literature has shown otherwise [135–137]. By native IP-MS, we identified interactions of BRD4 with CDK9, cyclin T1 and subunits of the super-elongation complex, but not HEXIM1. This finding is in accordance with the notion that the interaction of HEXIM1 and the long BRD4 isoform with CDK9 is mutually exclusive [133,334], whereas the short isoform of BRD4—which was not captured by the antibody used in our native IP experiment—does interact with HEXIM1 [130]. Also, in agreement with our observations, an association of BRD4 with SEC proteins and co-localization at some genes has been demonstrated by others [137,328]. Despite these interactions, no loss of CDK9, but rather a significant increase after 120 min and 360 min of dBET6 treatment was observed by chromatin-MS, arguing against a CDK9 recruitment defect.

Additionally, we found no evidence for lower CDK9 activity upon loss of BRD4. Under treatment conditions, levels of p-Ser2—similar to the levels of p-Tyr1 and p-Thr4, which depend not on CDK9 but likely on ABL1 and PLK3, respectively [463,464]—were increased at the 5’ gene end, indicating that CTD phosphorylation of promoter-proximally paused Pol II does not depend on BRD4. Notably, the increase in phosphorylation cannot be explained solely by the increase in total Pol II in this region. Moreover, it is in agreement with earlier studies employing BRD4-selective degradation [325,331]. Our findings also suggest that the

kinase activity of BRD4 is not required to maintain normal p-Ser2 levels on a short time scale. The kinase activity of BRD4 towards the RPB1 CTD has been demonstrated *in vitro* as well as *in vivo*, but only at low temporal resolution [335,465]. Whether the rapid ablation of BRD4 affects CTD phosphorylation immediately has not been tested previously.

5.1.3 Other processes might contribute to the loss of Pol II from the gene body

In principle, also changes in elongation rate or processivity could underlie the decrease in SI-NET-seq signal at the gene body [466]. A decrease in processivity in the stricter sense manifests as premature termination either in the promoter-proximal region [171,426] or at the gene body [467,468]. In fact, a high turnover of Pol II undergoing initiation or promoter-proximal pausing has been described *in vitro* and *in vivo* [168–170].

In view of the loss of critical elongation factors from the chromatin, a positive effect of the degrader treatment on elongation rate appears improbable.
Since no assay to approximate premature termination, *e.g.*, 4sU-seq [176], was performed in this study, we cannot exclude it as a possibility. However, indirect evidence coming from our chromatin-MS experiments argues against it. In particular, the binding of RNA 3' end processing factors near the TSS, which has been implicated in premature termination [438,469], was decreased upon either treatment. This was particularly the case for FIP1, a protein suggested to also function in alternative polyadenylation site selection [470]. Consequently, the decrease in these proteins is expected to result rather in a decrease in premature termination. We did not detect other factors involved, for instance, the decapping proteins DCP1a and DCP2 [471,472], but found a significant decrease in NCBP1 of the cap-binding complex. The most parsimonious explanation for the significant decrease in XRN2 seen upon 120 min dBET6 treatment appears to be the inefficient cleavage at the pA site, which prevents its engagement with nascent RNA (discussed below). Moreover, dBET6 treatment led to a significant reduction of the Integrator subunits INTS1 and INTS11, a paralog of CPSF73, after 120 min and of additional subunits after 360 min, some of which were also among the interactors of BRD4. However, while these observations are interesting, a decrease in premature termination is rather expected upon Integrator loss [425].

Moreover, we can exclude that the increase in SI-NET-seq reads mapping to the promoter-proximal region upon dTAG7 treatment was caused by an increase in premature termination rather than promoter-proximal pausing, *i.e*., that SI-NET-seq also detected RNA associated with prematurely terminating Pol II, since also ChIP-Rx revealed an increase in Pol II occupancy in that region.

5.1.4 BRD4-SELECTIVE AND BET DEGRADATION HAVE DISTINCT EFFECTS ON POL II IN THE PROMOTER-PROXIMAL REGION

Interestingly, treatment with dBET6 resulted in a more pronounced elongation defect than treatment with dTAG7, but not in an accumulation of Pol II in the promoter-proximal region. In fact, promoter-proximal Pol II levels were lower upon dBET6 treatment than under control conditions. Given the similar impact of both treatments on the chromatin proteome, inefficient assembly of the elongation machinery—possibly mainly driven by the BRD4 loss—presumably underlies the elongation defect in both cases but seems to be further modulated in the absence of BRD2 and BRD3. A recent study provided evidence for distinct functions of BRD2 and BRD3 in early elongation: targeted degradation of either BRD2 or BRD3 in colorectal adenocarcinoma DLD1 cells resulted in a decrease in Pol II in the promoter-proximal region but a mild increase at the gene body, indicating that BRD2 and BRD3 stabilize paused Pol II [137]. dBET6 treatment resulted in a weak increase in promoter-proximal signal in the said study, suggesting also cell line-specific differences. Notably, additivity of the putatively contrary effects of BRD2, BRD3 and BRD4 depletion can only explain the distinct effect of dBET6 treatment on the promoter-proximal but not the gene body region, where we observed a more substantial decrease in Pol II than upon dTAG7 treatment.

In line with the idea that not only BRD4 regulates transcription, we found BET proteins to interact with a similar set of elongation regulators, including SPT5, the PAF complex, CDK9 and Cyclin T1, whereas interactions with other subunits of the super-elongation complex in accordance with published literature were BRD4-specific [328,473]. How these interactions mechanistically could exert contrary effects on pause release remains to be investigated.

Also a reduction in transcription initiation might decrease Pol II occupancy in the promoter-proximal region. Due to the low mappability of sequence reads of <20 bp, SI-

NET-seq does not inform on defects during the first phase of transcription. However, in support of this hypothesis, we found BRD4 and BRD3, and to a lesser extent BRD2, to interact with TFIID. In contrast, Mediator almost exclusively interacted with BRD4, which is in accordance with an earlier report [310]. However, these interactions are not required for chromatin binding, because pan-BET or BRD4-selective degradation did not induce dissociation of initiation factors or Mediator from the chromatin, as we saw by chromatin-MS. This is in contrast to earlier studies that found Mediator modules to be lost from chromatin upon 120 min of JQ1 treatment in AML cells [318,319]. Moreover, deletion of Bdf1 in *S. cerevisiae* leads to a loss of TFIID at some promoter clusters [474]. Thus, transcription initiation appears to be less sensitive to dTAG7 treatment than elongation under the tested conditions.

Another possibility is that not general initiation factors but context-specific TFs are deregulated by BET ablation. Oncogenic driver TFs are disproportionally dependent on BET proteins in various leukemia cell lines. For example, pan-BET inhibition downregulates BCL2, CDK6, c-MYC and c-MYB transcriptionally in AML and multiple myeloma cell lines, thereby inducing cell cycle arrest or apoptosis [328,475,476]. Strikingly, transcription elongation at the *c-MYB* locus was most strongly affected by dBET6 treatment at each of the tested time points. Although *c-MYB* also showed the strongest decrease in Pol II occupancy upon dTAG7 treatment, the absolute change was less pronounced. The effect on protein levels was not determined, as c-MYB escaped detection by mass spectrometry. However, c-MYB protein has a half-life of only 30 min [477] and, therefore, could be realistically depleted during a 120 min treatment. In K562, c-MYB is among the ten most highly transcribed TFs (Fantom5) and peaks at enhancers and promoters associated with >12,000 genes in K562 [478], further underlining its regulatory importance in this cell line. While c-MYC has been implicated in regulating promoter-proximal pause release [57] and productive elongation [413,479], such roles have not been described for c-MYB. Thus, the reduction in c-MYB unlikely mediated the elongation defects observed upon dBET6 and dTAG7 treatment, but could have an effect on initiation.

Lastly, we cannot exclude recycling defects that, due to the more pronounced readthrough transcription, are expected to be more severe in case of dBET6 treatment. Such a mechanism has been demonstrated to cause reduced binding of TBP, TFIIB and CDK9 at the 5' gene end, when termination was disrupted by PAS mutation or PCF11 knockdown [480].

5.2 A BRD4-DEPENDENT CONTROL POINT OF RNA 3' END PROCESSING AND TERMINATION

Unexpectedly, both degraders also induced a global termination defect. Termination defects as a consequence of BET ablation have not been described in the literature so far—likely, because they might be minimal upon BET inhibition which leads only to an incomplete dissociation of BRD4 from the chromatin. Furthermore, readthrough transcripts were likely not polyadenylated and thus unstable and therefore required a method like SI-NET-seq for quantitative detection. A tendency towards readthrough transcription was seen after only 40 min of dBET6 treatment, indicating that it was a direct effect of BET depletion and not, for instance, a response to cellular stress.

Chromatin-MS performed upon 120 min of dBET6 or dTAG7 treatment revealed a significant depletion specifically of CstF and CPSF subunits, including the subunits directly mediating signal site recognition (CstF64, CPSF30, WDR33) and cleavage of the nascent transcript (CPSF73), as well as symplekin. Detecting symplekin was interesting as it is known to travel with Pol II and promote dephosphorylation of p-Ser5 of the RPB1 CTD by stimulating SSU72 [201]. Overall, chromatin-MS suggested an extensive defect in the assembly of the RNA 3' processing machinery rather than an effect on individual subunits. Consequently, failure to recognize the polyadenylation signal or additional signal sites and to catalyze the hydrolytic cleavage of the RNA phosphodiester backbone led to a decrease in RNA cleavage efficiency, as was detected by an RT-qPCR-based assay. Additionally, XRN2 was depleted from the chromatin upon dBET6 treatment. Thus, our findings agree with the current unified model after which activity of CPSF73 is a prerequisite for both creating an entry site for XRN2 and allosterically slowing Pol II [184,244]. Specifically, it is in accordance with the observation that rapid depletion of CPSF73 causes a more extensive, global downstream shift of the termination zone than XRN2 depletion [184]. Also, similar but less severe effects were also seen upon siRNA-mediated knockdown of CPSF73, but also of both CstF64 and CstF64τ [222,425].

Hypothesizing that BRD4 regulates the co-transcriptional association of RNA 3' processing factors with the elongation machinery [481] rather than with nascent RNA, we performed ChIP-Rx rather than an RNA-centric assay, *e.g.*, RNA crosslinking and immunoprecipitation (CLIP). Strikingly, ChIP-Rx for CPSF73, FIP1 and CstF64 indicated recruitment of these factors already in the promoter-proximal region, specifically

somewhat downstream of the promoter-proximal pausing site. A significant, dTAG7-induced decrease was seen across the promoter-proximal region, the gene body and the termination zone, confirming our mass spectrometry-based findings.

According to the unified model of termination, deceleration of Pol II facilitates transcription termination. Although SI-NET-seq—likely due to low sequencing coverage at the 3' gene end—did not reveal Pol II pausing or slowing in this gene region, total Pol II ChIP-Rx signal was somewhat higher downstream of the pA site than at the gene body. Moreover, BRD4 ablation resulted in a decrease in Pol II occupancy downstream of the pA site, indicating that Pol II failed to undergo the allosteric alteration required for slowing, possibly due to reduced CPSF binding. Thr4 phosphorylation of the RPB1 CTD has been suggested to be mechanistically linked to efficient termination as it peaks downstream of the pA site in a CPSF73-dependent manner [184,258]. Therefore, we profiled p-Thr4 by ChIP-Rx and indeed observed a decrease upon dTAG7 treatment that was more pronounced than the decrease in p-Tyr1 or p-Ser2.

Because also the dephosphorylation of SPT5 at its C-terminal repeat downstream of the pA site reduces the elongation rate and thereby facilitates transcription termination [238,248], we also focused on the responsible phosphatase complex and found a significant global decrease in the regulatory protein PPP1R10 upon dBET6 treatment, but no change of PP1. Thus, it appears possible that BRD4 affects not only chromatin association of SPT5 but also its activity.

Co-localization of CPSF and CstF proteins with BRD4 and their significant decrease upon dTAG7 treatment near the 5' gene end indicate a role for BRD4 in the initial recruitment of at least some CstF and CPSF proteins. Moreover, our data suggest that these proteins travel along the gene with Pol II, although we cannot exclude that a fraction of RNA 3' end processing factors is recruited directly to the pA site. Our genome-wide findings are in accordance with previous reports of an association of CPSF73, CPSF30, CstF77, CstF64 and PCF11 with 5' ends of selected genes [212,221,224,225]. Also, CPSF and CstF proteins have been shown to interact with GTFs specifically [220,482]. Whether RNA 3' end processing factors "linger" on the elongation complex to increase the local concentration and accelerate RNA 3' end maturation or also influence, for instance, elongation or splicing remains to be tested.

Ser2 phosphorylation has been suggested to be involved in recruiting RNA 3' end processing factors [225,483]. However, our ChIP-Rx data argues against this possibility because the recruitment of RNA 3' processing factors was decreased despite the increased phosphorylation levels in the promoter-proximal region that was observed. Additionally, CDK9 inhibition by 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB) treatment has been shown earlier to be inconsequential for the recruitment of CPSF and CstF proteins in this region [224]. However, we cannot exclude that the decrease in p-Ser2 at the gene body and in the termination zone contributed to the termination defect. Ser2 phosphorylation in these regions likely depends on CDK12 rather than CDK9 [142] and was shown to promote the assembly of the RNA 3' end processing machinery [60,225,252,255]. Notably, in contrast to Laitem et al. [60], the reduction in p-Ser2 and CstF64 near the pA site in our system did not lead to an upstream shift of the termination zone but to the opposite outcome.

Given the high correlation between the effects of dTAG7 and dBET6 treatments as seen by SI-NET-seq or chromatin-MS, the termination defect observed upon dBET6 treatment might be primarily due to the loss of BRD4, although a contribution of BRD2 and BRD3 cannot be excluded. In fact, while different IP-MS protocols revealed interactions of BRD4 and proteins of the RNA 3' end processing machinery, these interactions were also found for BRD2 and BRD3 and showed even a higher level of significance.

Based on our IP-MS data, recruitment of RNA 3' end processing factors could happen directly via BRD4 but possibly also via elongation factors. Especially PAF appears to be a likely candidate because of its treatment-induced dissociation at all regions across the transcribed region and its binding profile, which resembles that of CPSF and CstF subunits. In human cells, the PAF subunit CDC73 has been shown to be required to recruit CPSF and CstF proteins to the 5' end of an exemplary gene [250]. The same publication also showed RNA-independent interactions of almost all RNA 3' end processing factors especially with CDC73, which is in accordance with our IP-MS and IP-WB data. Extending these findings, we could show that these interactions do not depend on BRD4.

Also SPT5 has been implicated in recruiting RNA 3' end processing factors [251] and controlling termination, as its depletion causes readthrough transcription [150,151]. On the other hand, no loss in interactions between RPB1 and CPSF as well as CstF subunits is seen upon targeted SPT5 depletion [152]. Given that the decrease in SPT5 binding was

confined to the termination zone, it was likely not the main cause of the general recruitment defect of RNA 3' processing factors in our study.

Furthermore, SPT6, which based on chromatin-MS and ChIP-Rx was strongly depleted upon treatment and also appeared as a BRD4 interactor, was likely involved in the termination defect. Like PAF1, SPT6 upon dTAG7 treatment showed a reduction in binding in the promoter-proximal region, and with regard to its binding profile resembled 3' processing factors. However, we could not detect interactions with CPSF and CstF subunits, possibly because IP conditions were not optimized explicitly for this experiment. A contribution of SPT6 would be in accordance with a recent finding that its targeted depletion induces readthrough transcription and diminishes interactions of Pol II with the CstF module, symplekin and XRN2 [165].

It seems incongruous at the first glance how processes at the 5' and 3' gene end can be coordinated, particularly in view of the elongation defect. Part of the solution might be that the block of pause release is only incomplete, as can be seen from the SI-NET-seq experiments conducted after 120 min of treatment. Additionally, elongation and termination might show different sensitivity to BRD4 depletion: for instance, a mild reduction in BRD4 levels might still support efficient recruitment of elongation factors but could already profoundly affect RNA 3' processing factors.

An interesting open question is how elongation and RNA 3' end processing factors interact on the biochemical level. Based on immunoprecipitation experiments, binding of CPSF to the body of Pol II or to CstF, which contacts the Pol II CTD, are mutually exclusive [194]. One interpretation thereof is that CPSF travels with Pol II until it reaches the hexameric PAS, while CstF binds to the complex only upon transcription of the G/U-rich region. Subsequently, CPSF releases the Pol II body but binds to CstF [194]. Thus, how also CstF can travel with Pol II needs to be investigated. In light of the propensity of BRD4 and the RPB1 CTD to form heterotypic condensates, one could speculate that they also bind RNA 3' end processing factors via weak, multivalent interactions rather than by specific interaction interfaces [484].

In conclusion, BRD4 has a central role in a transcriptional control point located in the promoter-proximal region that ensures the proper assembly of the elongation complex. Specifically, pause release is only enabled for complexes competent to elongate but also terminate efficiently.

5.3 CURRENT EVIDENCE FOR FUNCTIONAL REDUNDANCY BETWEEN BRD2, BRD3 AND BRD4

While the identification of unique roles of the different BET species is a reasonable goal, characterization of the functional overlap might be as important. The effects of pan-BET and BRD4-selective degradation on transcription elongation and termination were similar but not identical, suggesting that BRD2 or BRD3, to some extent, could compensate for the loss of BRD4. Combining chromatin-MS with dTAG7 and dBET6 treatment time courses allowed us to monitor changes in chromatin composition with temporal resolution. Given a median temporal delay of >8 h between maximal RNA and protein amounts [485], changes were likely not confounded by transcriptional or post-transcriptional regulation. Interestingly, we identified groups of functionally related proteins which displayed similar trends that were indicative of cellular compensation processes. In particular, while proteins of cluster III (splicing factors and proteins involved in rRNA maturation) were significantly depleted upon 120 min of treatment with either degrader, only a prolonged treatment with dBET6 resulted in more severe effects. In case of an extended treatment with dTAG7, protein levels remained relatively stable, suggesting that BRD2 or BRD3 could partially compensate for the loss of BRD4. Also, this appears to be in agreement with the finding that all BET species interacted with a similar set of splicing factors.

Moreover, we found some indication for BET-independent compensation. Depletion of proteins of cluster V, which predominantly contained core transcription proteins, was more severe after 120 than 360 min of dBET6 treatment, possibly indicating a mechanism to restore transcription.
Additionally, we identified one cluster (cluster I) containing structural proteins like CTCF or the cohesin subunit STAG2 and proteins of the MCM complex involved in DNA replication licensing that increased significantly in amount upon treatment, hypothetically to maintain chromatin integrity.

Overall, these experiments demonstrated how a rapid treatment targeted at a transcription regulator by affecting protein localization could have profound effects also on the proteome level. Thus, not only the short half-lives of transcription factors and

chromatin-modifying enzymes [486] but also their fast recruitment provides a mechanism to react dynamically to cellular changes.

We asked whether rapid compensation processes for the loss of BRD4 could require a shift or increase in BRD2 or BRD3 occupancy. As an approximation, we quantified the intracellular distribution of BRD2 and BRD3. Compared to the control condition, no changes were seen within 120 min of treatment, indicating that compensation processes on this time scale likely involve the redistribution of already chromatin-associated BRD2 or BRD3. However, levels of both BRD2 and BRD3 in the chromatin fraction were increased after 24 h of dTAG7 treatment, which could be indicative of additional chromatin recruitment. How BRD2 and BRD3 of the nucleoplasmic pool are recruited to formerly BRD4-bound sites could be an interesting question for future research. One could speculate that, for instance, histone acetylation marks, which are similarly well recognized by all BET species [310], are involved. A similar process has been described in in *S. cerevisiae*, where the BET homologs Bdf1 and Bdf2 normally bind distinct genomic locations, *i.e.*, specific groups of promoters [487]. However, deletion of the Bdf1 causes Bdf2 to occupy the original Bdf1 binding sites, leading to partial functional compensation [474].

5.4 Molecular links between BET proteins and pre-mRNA splicing

Because of the newly discovered BRD4-dependent mechanism of RNA 3' end processing and termination control, it is tempting to speculate that BRD4 is also involved in other co-transcriptional processes and promotes, for instance, the recruitment of splicing factors.

In support of this, chromatin-MS performed after 120 min of BRD4 depletion revealed splicing factors to be among the most robustly depleted proteins. Specifically, the U1, U2 and U5 snRNPs as well as the PRP19 complex were found to be depleted from chromatin, which was remarkable, as the physical association of some spliceosomal snRNPs with Pol II is well established. For instance, the structure of human U1 snRNP bound to Pol II has been solved and, based on predictions, can also accommodate the U2 snRNP [488]. Also, functional studies demonstrated an intimate interaction of Pol II and the U2 snRNP required to regulate promoter-proximal pausing [489]. Thus, one might hypothesize that the elongating machinery in the absence of BRD4 lacks association not only with elongation factors but also with key splicing factors. Moreover, p-Ser5 has been

shown to bind to the U5 snRNP [167], suggesting that the recruitment of this complex occurs already early during the transcription cycle. In fact, ChIP-seq revealed a moderate enrichment of PRPF8 near the TSS, although its BRD4 dependency still needs to be tested. Lastly, also PRP19 is known to interact with the Pol II CTD [446] and regulate transcription elongation [268].

Surprisingly, certain proteins, *e.g.*, SR proteins and subunits of the exon junction and THO complexes, consistently were detected at the chromatin in somewhat higher amounts upon BRD4 or BET degradation. One potential explanation that could be easily tested by a CLIP-based approach is that readthrough transcripts likely contain a large number of cryptic splice sites which could recruit and sequester splicing factors at the chromatin.

Native IP-MS did not reveal a clear preference of any of the BET proteins for specific splicing complexes, although BRD4 appeared to interact less with spliceosomal snRNP factors than BRD2 or BRD3. In line with this, the effects of BET and BRD4-selective degradation on chromatin-bound splicing factor levels were well correlated. Based on the current data, one should assume BET proteins to have overlapping rather than specialized functions in splicing. Of note, many of the proteins depleted from chromatin upon dTAG7 or dBET6 treatment based on native IP-MS did not interact with any of the BET proteins, which might indicate that ablation of BET proteins affects chromatin-association of these proteins only indirectly. Conversely, most of the BET interactors were not depleted, suggesting rather passive co-localization.

It remains unclear to what extent the decrease in splicing factor binding could also contribute to the elongation and termination defects observed. In particular, short-term inhibition of SF3B1 has been shown to induce accumulation of elongating Pol II in the promoter-proximal region, indicative of a defect in pause release [489]. Moreover, through a process named telescripting, U1 snRNP acts to prevent the usage of actionable but usually inactive pA sites within the transcribed region [438]. Whether an imbalance between U1 snRNP and CPSF and CFIm factors in the absence of BRD4 could result in premature termination needs to be tested in the future.

Ultimately, whether BRD4 should be considered a splicing regulator depends on whether its depletion causes splicing defects. Spike-in-normalized RNA-seq data collected by Winter et al. [135] and re-analyzed by Uppal et al. [490] shows that dBET6 treatment affects

alternative splicing, specifically inclusion or skipping of exons, in MOLT4 cells. Roughly ten times more genes were affected after 360 than after 120 min. Preliminary data of K562 dTAG-BRD4 cells showed a similar effect upon 120 min of dTAG7 treatment, but no increase in severity with longer treatment time (not shown), suggesting compensation by BRD2 or BRD3. While we cannot rule out cell type-specific effects—based on cell line databases, MOLT4 cells have accumulated non-silent mutations in core subunits of the splicing machinery which might have a destabilizing effect—the somewhat weaker effect of the dTAG7 treatment likely argues against a unique role of BRD4 in splicing.

5.5 A ROLE OF BRD4 IN ENHANCER TRANSCRIPTION

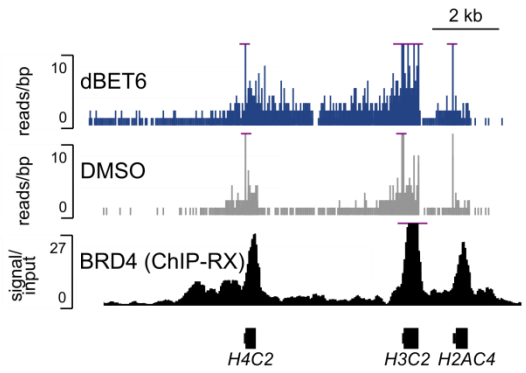
Upon 120 and 360 min of dTAG7 treatment, we observed a slight but consistent increase in accessibility at TSSs and active enhancers. This might be in accordance with a function of promoter-proximal pausing in maintaining an open chromatin state [73] but clearly requires further research. Our observation, however, is in contrast to the widespread decrease in accessibility seen upon targeted SPT5 depletion [151,153] and PAF1 knockdown [157]. A possible explanation is that PAF can affect the nucleosome structure not only by regulating transcription but also via its role in H2B ubiquitylation [491,492]. Also, changes in chromatin structure, which are likely mediated by chromatin remodelers, have been previously demonstrated to happen more slowly than transcriptional changes [493].

An appealing hypothesis is that BRD4 exerts similar effects on transcription at promoters and enhancers. Upon BRD4-selective degradation, total Pol II occupancy was increased at the enhancer center but weakly decreased at the flanking regions, resembling the increase in promoter-proximal pausing at protein-coding and lncRNA genes. Since the enhancer center is only an approximation of the actual start site of enhancer transcription, our analysis probably underestimated the effect. Pol II pausing recently was discovered to be a feature of enhancers that regulate target genes which show promoter-proximal pausing in *Drosophila* and mammalian cells [453]. Overall, this finding is in line with the notion of the genetic and epigenetic similarity between promoters and enhancers [494–496]. We also found SPT5 to accumulate at enhancer centers upon targeted BRD4 degradation, mirroring the effects seen at 5’ gene ends. PAF1 and SPT6 signal were detectable but noisy at enhancers and therefore should not be interpreted quantitatively. Altogether, this argues for a role of the elongation machinery and BRD4 in enhancer

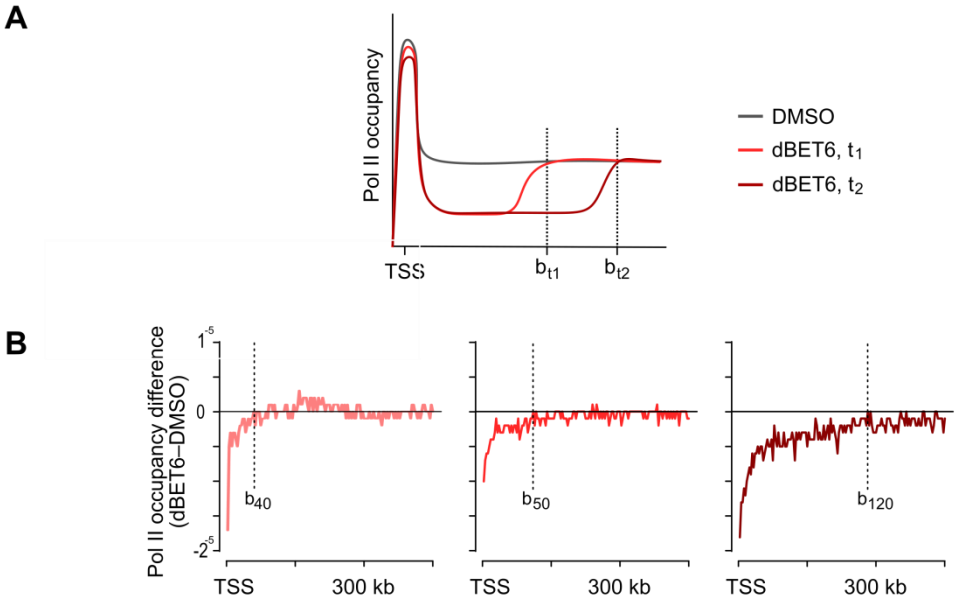
transcription and is in agreement with a recent study showing that knockdown of SPT5 as well as NELF diminishes paused Pol II from enhancers [453].
Also, promoter-proximal pausing was suggested to facilitate rapid and synchronous transcription induction upon a stimulus [74,75]. Given that enhancer transcription precedes transcription at the target gene in various cellular systems [497], one could speculate that Pol II pausing has a similar function also at enhancers.

It remains unclear if BRD4 regulates enhancers and promoters individually or mediates interactions between both elements. Our Hi-ChIP data suggested that rapid BRD4 ablation was sufficient to decrease the interaction frequency between promoters and intergenic as well as intragenic enhancers on a global scale, and thereby agreed with a recent study that combined a 4.5 h treatment with a BRD4-selective degrader with Hi-C in neural crest cells [313]. However, in clear contrast, no changes in contacts for selected promoter–enhancer pairs were seen upon 90 min of BET inhibitor or 24 h of BET degrader treatment in a second study [341]. Possibly, this discrepancy is due to methodological differences. For instance, our unbiased approach possibly could more sensitively detect changes in weak interactions, which might be absent from the set of annotated promoter–enhancer pairs used by Crump et al. [341]. Alternatively, following the reasoning of El Khattabi et al. [498] with regard to the function of Mediator, BRD4 might mainly act as a functional rather than an architectural bridge, *i.e.*, by creating a local high concentration of transcription regulators.

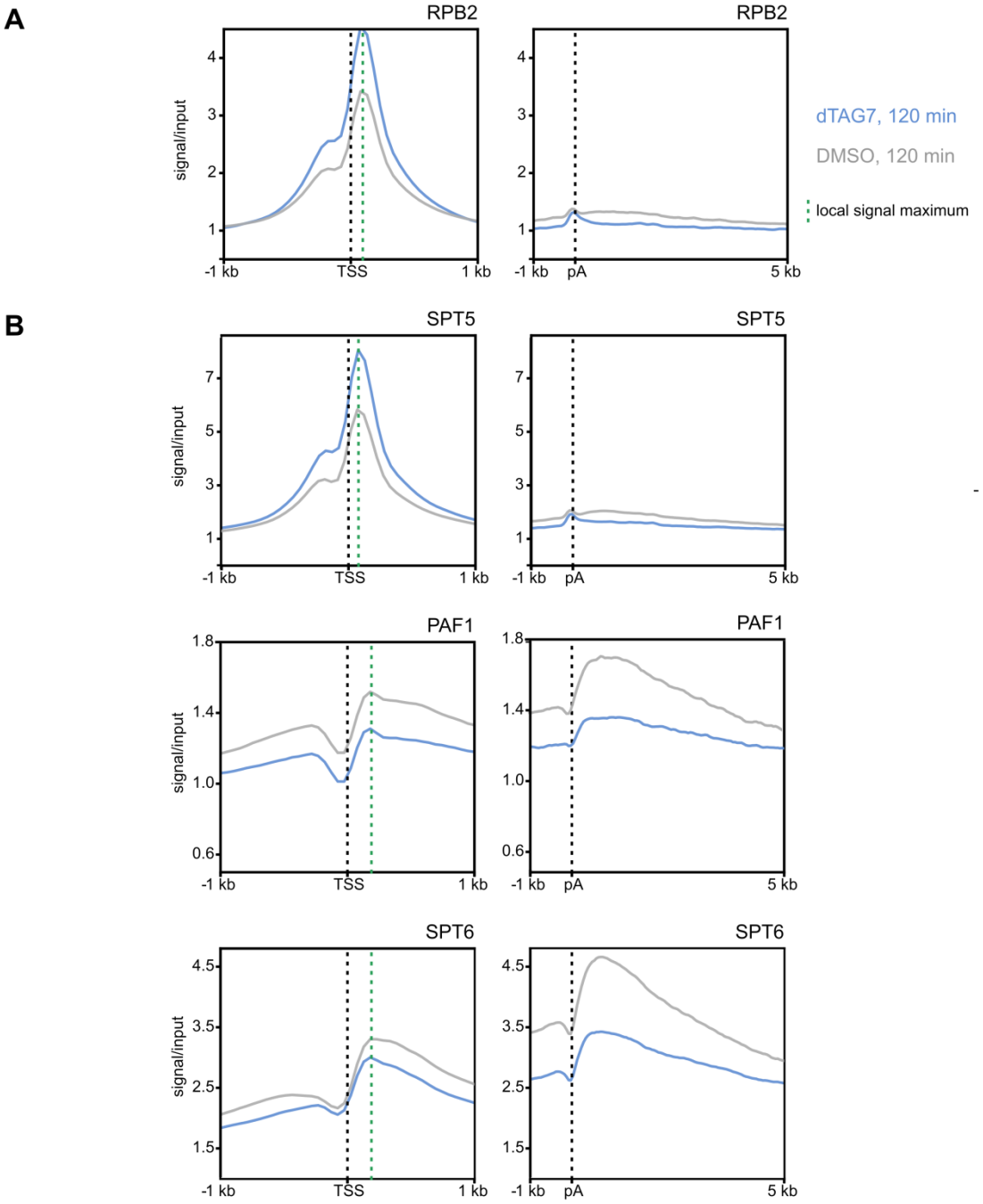
6 SUPPLEMENTAL DATA



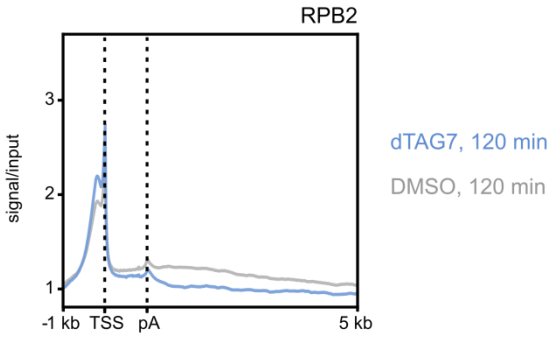
Supplemental Figure 1: No decrease in Pol II occupancy is seen at exemplary histone gene loci upon 120 min of dBET6 treatment. Spike-in-normalized SI-NET-seq signal in reads per bp and input-normalized BRD4 ChIP-Rx signal at three histone genes of the H1-1 cluster. Particularly high signal is cropped (violet lines). All genes shown are transcribed from right to left in this representation.



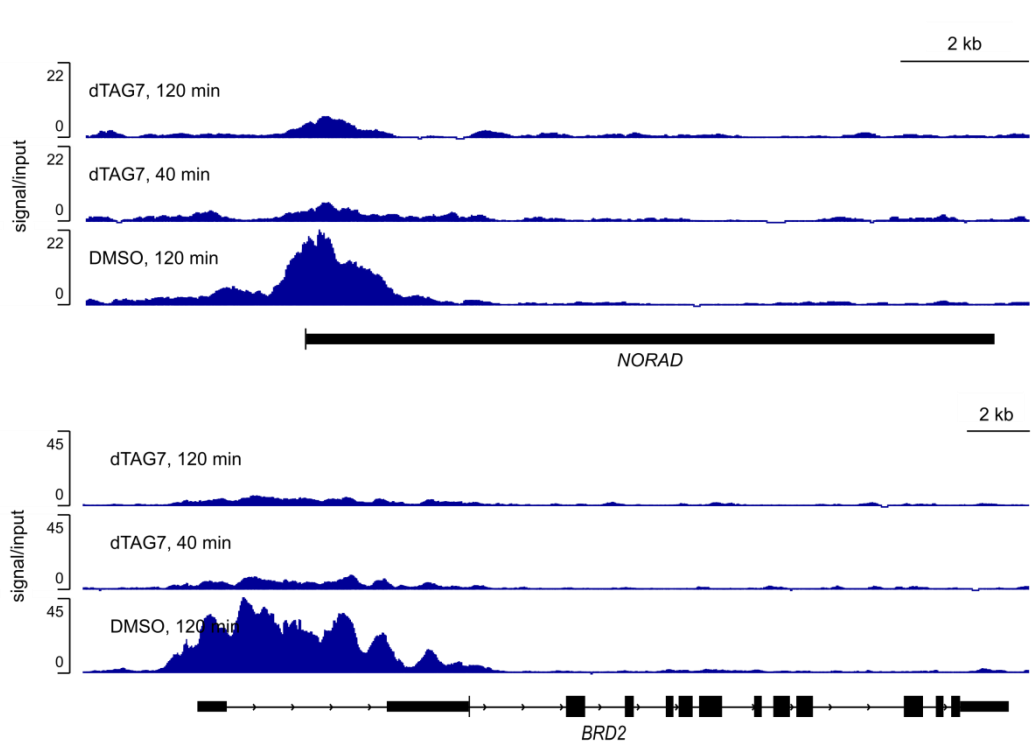
Supplemental Figure 2: Pol II released into productive elongation before the elongation defect manifests itself appears as a "wave". (A) Schematic definition of Pol II "wave". The intersection of the Pol II binding curves determined upon treatment and under control conditions indicates the position of the back edge of the Pol II wave (labelled b_{t1}, b_{t2}) at the time points t_1 and t_2 of dBET6 treatment. **(B)** Pol II occupancy difference upon 40, 50 and 120 min of dBET6 treatment. The analysis was performed by Annkatrin Bressin and with minor changes included in Arnold et al. [368].



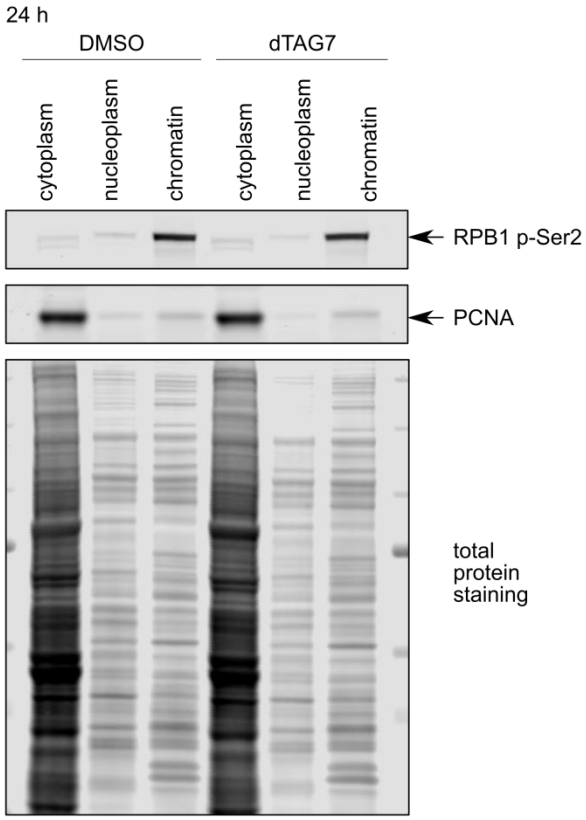
Supplemental Figure 3: Targeted BRD4 degradation affects binding of RPB2, SPT5, PAF1 and SPT6 at 5' and 3' gene ends. Meta-gene representation of the input-normalized ChIP-Rx signal, centered at the annotated TSS or pA site (indicated by a black line). A local signal maximum downstream of the TSS is indicated in green. 13,040 active genes were included in the analysis. Four, three and one biological replicates were analyzed for RPB2 and SPT5, PAF1 and SPT6, respectively.



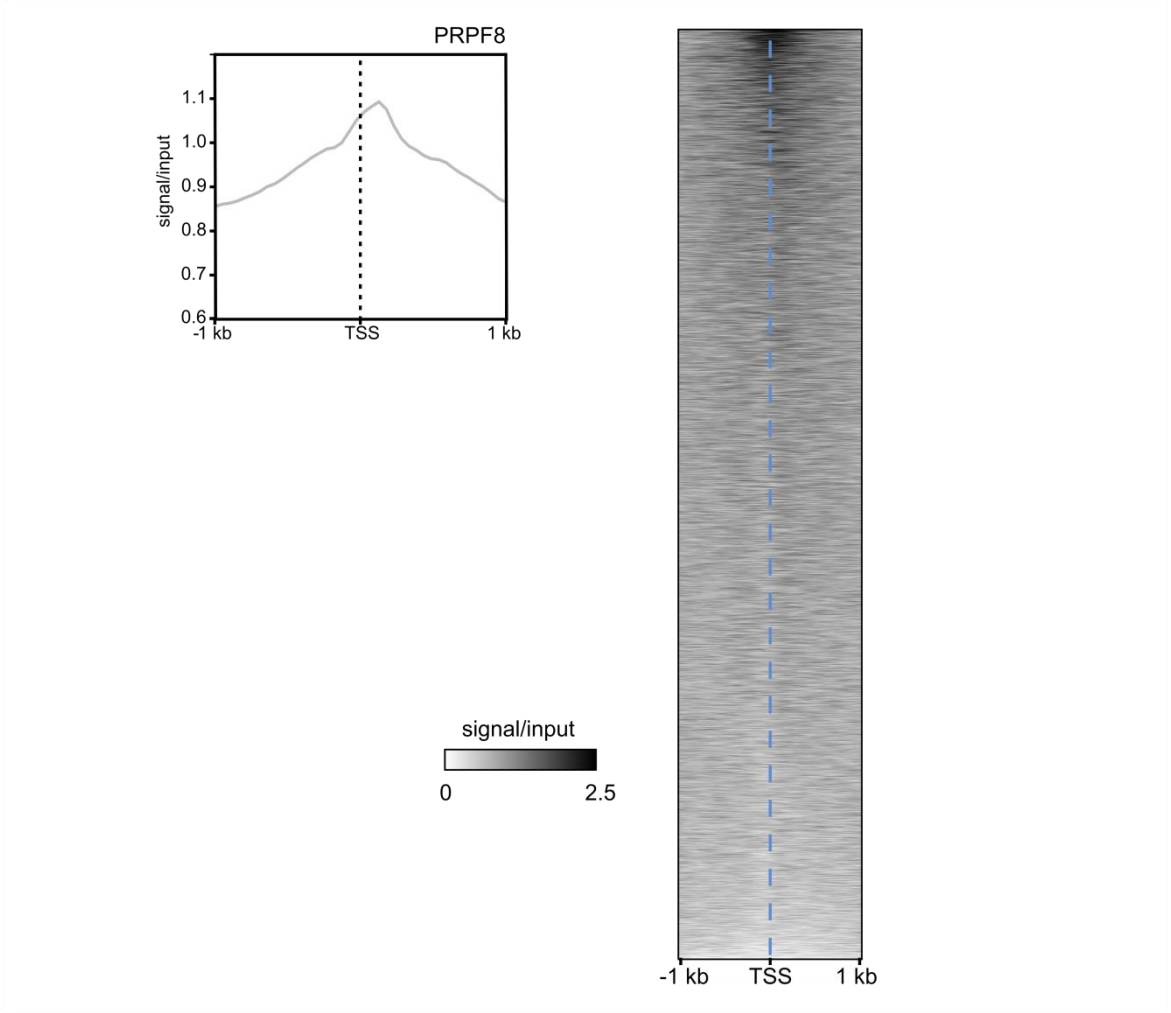
Supplemental Figure 4: ChIP-Rx performed upon BRD4 ablation reveals an increase in RPB2 binding in the promoter-proximal and a decreased in the gene body region. Meta-gene representation of the input-normalized signal centered at the TSS and the pA site without scaling. Data of one representative biological replicate (out of two) are shown.



Supplemental Figure 5: Upon 40 and 120 min of TAG7 treatment, BRD4 is lost from the exemplary *NORAD* and *BRD2* loci. Input-normalized ChIP-Rx signal of one representative biological replicate (out of two) that was collected in K562 dTAG-BRD4 cells is shown. Both genes are transcribed from left to right in this representation. The data were visualized using the IGV genome browser [399].



Supplemental Figure 6: A 24 h treatment with dTAG7 does not affect RPB1 p-Ser2 levels or PCNA. p-Ser2 marks the CTD of actively elongating Pol II, while PCNA is an essential protein in DNA replication. A total protein staining was done as a loading control.



Supplemental Figure 7: PRPF8 binds to the chromatin near the TSS. Meta-gene and heat map representation of the input-normalized occupancy centered at the TSS in untreated K562 dTAG-BRD4 cells. 13,040 active genes were included in the analysis. One biological replicate was performed.

Supplemental Table 1: Localization changes of selected proteins involved in transcription initiation, elongation or RNA processing, as determined by quantitative chromatin-MS. Proteins were sorted by complex or functional category. Significant changes, as determined by two-tailed t-test, are indicated in grey.

protein name	complex, category	dBET6, 120 min		dTAG7, 120 min	
		log2 FC (average SILAC ratio)	-log10 (p-value)	log2 FC (average SILAC ratio)	-log10 (p-value)
MED1	Mediator	0.087145	0.184053		
MED12		0.022833	0.026261		
MED17		0.061013	0.239744		
MED25		-0.00863	0.011643		
MED26		0.006908	0.018036		
GTF2H1	GTFs	0.050762	0.223025		
GTF2H2		0.115065	1.34585		
GTF2H3		0.251645	1.13574		
ERCC3		0.049621	0.153694		
GTF2I		-0.01539	0.026611	0.02222	0.314977
TBP;TBPL2		-0.0316	0.261214		

protein name	complex, category	dBET6, 120 min		dTAG7, 120 min	
		log2 FC (average SILAC ratio)	-log10 (p-value)	log2 FC (average SILAC ratio)	-log10 (p-value)
TAF1;TAF1L		-0.07527	0.152012		
TAF4		-0.03532	0.120502		
TAF5		-0.05417	1.18588		
TAF5L		0.050298	0.188035		
TAF6		-0.0846	0.076367	-0.00215	0.005957
TAF9;TAF9B		-0.09093	0.238654		
TAF10		0.185022	0.843991		
TAF15		0.055629	0.730816	0.052176	0.850284
GTF2E1		0.064396	0.121617		
GTF2H1		0.050762	0.223025		
GTF2H2		0.115065	1.34585		
GTF2H3		0.251645	1.13574		
ERCC3		0.049621	0.153694		
CDK7		0.023663	0.760242		
GTF2I		-0.01539	0.026611		
PPP1CA	PP1	0.036568	1.13682	0.030055	0.622658
PPP1CB		0.059359	0.587857	-0.0129	0.058671
PPP1CC		0.001195	0.008706	0.061397	1.00943
PPP1R10		-0.518800	1.96339	-0.22167	1.12054
PPP2CA	PP2	0.052885	0.239465	-0.01739	0.180537
PPP2R1A		0.063557	0.341519	-0.01323	0.077737
PPP2R1B		-0.17362	0.173365		
PPP2R2A; PPP2R2D		0.126223	0.497192		
PPP2R5C		0.173655	0.431264		
PPP2R5D		0.189535	0.538279		
PPP2R5E		-0.09039	0.836556		
NCBP1	CBC	-0.451	1.55945	-0.15159	1.6645
INTS1	Integrator	-0.22818	1.47955		
INTS2		0.412656	0.30483		
INTS3				0.087594	0.320706
INTS4		-0.01843	0.051316		
INTS6		-0.15952	0.85505	0.021218	0.112709
INTS7		0.012621	0.035107		
INTS8		-0.09075	0.724523		
INTS9		-0.02504	0.074978		
INTS10		-0.06894	0.385601		
INTS12		-0.05659	0.252088	-0.03063	0.119705
RPRD1A	PRPD	-0.16563	1.33129	-0.0623	1.04717
RPRD1B		-0.28028	1.50534	-0.03825	0.138828
RPRD2		-0.35508	1.27744	-0.11083	0.962558

7 REFERENCES

1. O'Brien J, Hayder H, Zayed Y, Peng C. Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front Endocrinol (Lausanne)*. 2018;9:402. https://www.frontiersin.org/article/10.3389/fendo.2018.00402.

2. Svejstrup JQ. The RNA polymerase II transcription cycle: Cycling through chromatin. *Biochim Biophys Acta - Gene Struct Expr*. 2004;1677(1-3):64-73. doi:10.1016/j.bbaexp.2003.10.012

3. Knezetic JA, Luse DS. The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell*. 1986;45(1):95-104. doi:https://doi.org/10.1016/0092-8674(86)90541-6

4. Lorch Y, Kornberg RD. Chromatin-remodeling for transcription. *Q Rev Biophys*. 2017;50:1-15. doi:10.1017/S003358351700004X

5. Nguyen VQ, Ranjan A, Liu S, et al. Spatiotemporal coordination of transcription preinitiation complex assembly in live cells. *Mol Cell*. 2021;81(17):3560-3575.e6. doi:https://doi.org/10.1016/j.molcel.2021.07.022

6. Zaret KS, Carroll JS. Pioneer transcription factors: Establishing competence for gene expression. *Genes Dev*. 2011;25(21):2227-2241. doi:10.1101/gad.176826.111

7. He HH, Meyer CA, Shin H, et al. Nucleosome dynamics define transcriptional enhancers. *Nat Genet*. 2010;42(4):343-347. doi:10.1038/ng.545

8. Li E, Liu H, Huang L, et al. Long-range interactions between proximal and distal regulatory regions in maize. *Nat Commun*. 2019;10(1):2633. doi:10.1038/s41467-019-10603-4

9. Zhang T, Zhang Z, Dong Q, Xiong J, Zhu B. Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biol*. 2020;21(1):1-7. doi:10.1186/s13059-020-01957-w

10. Core LJ, Martins AL, Danko CG, Waters CT, Siepel A, Lis JT. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat Genet*. 2014;46(12):1311-1320. doi:10.1038/ng.3142

11. Carninci P, Sandelin A, Lenhard B, et al. Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet*. 2006;38(6):626-635. doi:10.1038/ng1789

12. Rach EA, Winter DR, Benjamin AM, et al. Transcription initiation patterns indicate divergent strategies for gene regulation at the chromatin level. *PLoS Genet*. 2011;7(1). doi:10.1371/journal.pgen.1001274

13. Lenhard B, Sandelin A, Carninci P. Metazoan promoters: Emerging characteristics and insights into transcriptional regulation. *Nat Rev Genet*. 2012;13(4):233-245. doi:10.1038/nrg3163

14. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014;507(7493):455-461. doi:10.1038/nature12787

15. Hahn S, Buratowski S, Sharp PA, Guarente L. Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. *Proc Natl Acad Sci U S A*. 1989;86(15):5718-5722. doi:10.1073/pnas.86.15.5718

16. Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res*. 2006;16(1):1-10. doi:10.1101/gr.4222606

17. Schug J, Schuller WP, Kappen C, Salbaum JM, Bucan M, Stoeckert CJ. Promoter features related to tissue specificity as measured by Shannon entropy. *Genome Biol*. 2005;6(4). doi:10.1186/gb-2005-6-4-r33

18. Javahery R, Khachi A, Lo K, Zenzie-Gregory B, Smale ST. DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Mol Cell Biol*. 1994;14(1):116-127. doi:10.1128/mcb.14.1.116-127.1994

19. Smale ST, Baltimore D. The “initiator” as a transcription control element. *Cell*. 1989;57(1):103-113. doi:https://doi.org/10.1016/0092-8674(89)90176-1

20. Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene*. 2007;389(1):52-65. doi:10.1016/j.gene.2006.09.029

21. Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH. New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev* . 1998;12(1):34-44. doi:10.1101/gad.12.1.34

22. Kutach AK, Kadonaga JT. The Downstream Promoter Element DPE Appears To Be as Widely Used as the TATA Box in Drosophila Core Promoters . *Mol Cell Biol*. 2000;20(13):4754-4764. doi:10.1128/mcb.20.13.4754-4764.2000

23. Sainsbury S, Bernecky C, Cramer P. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol*. 2015;16(3):129-143. doi:10.1038/nrm3952

24. Buratowski S, Hahn S, Guarente L, Sharp PA. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell*. 1989;56(4):549-561. doi:https://doi.org/10.1016/0092-8674(89)90578-3

25. Bleichenbacher M, Tan S, Richmond TJ. Novel Interactions Between the Components of Human and Yeast TFIIA/TBP/DNA Complexes. *J Mol Biol*. 2003;332(4):783-793. doi:https://doi.org/10.1016/S0022-2836(03)00887-8

26. Bushnell DA, Westover KD, Davis RE, Kornberg RD. Structural Basis of Transcription: An RNA Polymerase II-TFIIB Cocrystal at 4.5 Angstroms. *Science*. 2004;303(5660):983-988. doi:10.1126/science.1090838

27. Kostrewa D, Zeller ME, Armache K-J, et al. RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature*. 2009;462(7271):323-330. doi:10.1038/nature08548

28. Zhang Z, English BP, Grimm JB, et al. Rapid dynamics of general transcription factor TFIIB binding during preinitiation complex assembly revealed by single-molecule analysis. *Genes Dev*. 2016;30(18):2106-2118. doi:10.1101/gad.285395.116

29. Čabart P, Újvári A, Pal M, Luse DS. Transcription factor TFIIF is not required for

initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIB in early elongation complexes. *Proc Natl Acad Sci*. 2011;108(38):15786-15791. doi:10.1073/pnas.1104591108

30. Chen H-T, Warfield L, Hahn S. The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat Struct Mol Biol*. 2007;14(8):696-703. doi:10.1038/nsmb1272

31. Ohkuma Y, Roeder RG. Regulation of TFIIH ATPase and kinase activities by TFIIE during active initiation complex formation. *Nature*. 1994;368(6467):160-163. doi:10.1038/368160a0

32. Goodrich JA, Tjian R. Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell*. 1994;77(1):145-156. doi:https://doi.org/10.1016/0092-8674(94)90242-9

33. Petrenko N, Jin Y, Dong L, Wong KH, Struhl K. Requirements for RNA polymerase II preinitiation complex formation in vivo. *Elife*. 2019;8:e43654. doi:10.7554/eLife.43654

34. Allen BL, Taatjes DJ. E i V The Mediator complex : a central integrator of transcription. *Nat Publ Gr*. 2015;16(3):155-166. doi:10.1038/nrm3951

35. Baek HJ, Kang YK, Roeder RG. Human mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. *J Biol Chem*. 2006;281(22):15172-15181. doi:10.1074/jbc.M601983200

36. Johnson KM, Wang J, Smallwood A, Arayata C, Carey M. TFIID and human mediator coactivator complexes assemble cooperatively on promoter DNA. *Genes Dev*. 2002;16(14):1852-1863. doi:10.1101/gad.995702

37. Esnault C, Ghavi-Helm Y, Brun S, et al. Mediator-Dependent Recruitment of TFIIH Modules in Preinitiation Complex. *Mol Cell*. 2008;31(3):337-346. doi:10.1016/j.molcel.2008.06.021

38. Takahashi H, Parmely TJ, Sato S, et al. Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell*. 2011;146(1):92-104. doi:10.1016/j.cell.2011.06.005

39. Kremer SB, Kim S, Jeon JO, et al. Role of Mediator in regulating pol ii elongation and nucleosome displacement in Saccharomyces cerevisiae. *Genetics*. 2012;191(1):95-106. doi:10.1534/genetics.111.135806

40. Eick D, Geyer M. The RNA Polymerase II Carboxy-Terminal Domain (CTD) Code. *Chem. Rev*. 2013.113(11):8456-8490. doi: 10.1021/cr400071f

41. Glover-Cutter K, Larochelle S, Erickson B, et al. TFIIH-Associated Cdk7 Kinase Functions in Phosphorylation of C-Terminal Domain Ser7 Residues, Promoter-Proximal Pausing, and Termination by RNA Polymerase II. *Mol Cell Biol*. 2009;29(20):5455-5464. doi:10.1128/mcb.00637-09

42. Wong KH, Jin Y, Struhl K. TFIIH Phosphorylation of the Pol II CTD Stimulates Mediator Dissociation from the Preinitiation Complex and Promoter Escape. *Mol Cell*. 2014;54(4):601-612. doi:https://doi.org/10.1016/j.molcel.2014.03.024

43. McCracken S, Fong N, Yankulov K, et al. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*. 1997;385(6614):357-360. doi:10.1038/385357a0

44. Nilson KA, Guo J, Turek ME, et al. THZ1 Reveals Roles for Cdk7 in Co-transcriptional Capping and Pausing. *Mol Cell*. 2015;59(4):576-587. doi:10.1016/j.molcel.2015.06.032

45. Boeing S, Rigault C, Heidemann M, Eick D, Meisterernst M. RNA polymerase II C-terminal heptarepeat domain Ser-7 phosphorylation is established in a mediator-dependent fashion. *J Biol Chem*. 2010;285(1):188-196. doi:10.1074/jbc.M109.046565

46. Rimel JK, Poss ZC, Erickson B, et al. Selective inhibition of CDK7 reveals high-confidence targets and new models for TFIIH function in transcription. *Genes Dev* . 2020;34(21-22):1452-1473. http://genesdev.cshlp.org/content/34/21-22/1452.abstract.

47. Core LJ, Waterfall JJ, Lis JT. Nascent RNA Sequencing Reveals Widespread Pausing and Divergent Initiation at Human Promoters. *Science*. 2008;322(5909):1845-1848. doi:10.1126/science.1162228

48. Seila AC, Calabrese JM, Levine SS, et al. Divergent transcription from active promoters. *Science*. 2008;322(5909):1849-1851. doi:10.1126/science.1162253

49. Scruggs BS, Gilchrist DA, Nechaev S, et al. Bidirectional Transcription Arises from Two Distinct Hubs of Transcription Factor Binding and Active Chromatin. *Mol Cell*. 2015;58(6):1101-1112. doi:10.1016/j.molcel.2015.04.006

50. Flynn RA, Almada AE, Zamudio JR, Sharp PA. Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. *Proc Natl Acad Sci*. 2011;108(26):10460-10465. doi:10.1073/pnas.1106630108

51. Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature*. 2013;499(7458):360-363. doi:10.1038/nature12349

52. Ntini E, Järvelin AI, Bornholdt J, et al. Polyadenylation site–induced decay of upstream transcripts enforces promoter directionality. *Nat Struct Mol Biol*. 2013;20(8):923-928. doi:10.1038/nsmb.2640

53. Trinklein ND, Force Aldred S, Hartman SJ, Schroeder DI, Otillar RP, Myers RM. An abundance of bidirectional promoters in the human genome. *Genome Res*. 2004;14(1):62-66. doi:10.1101/gr.1982804

54. Albrecht AS, Ørom UA. Bidirectional expression of long ncRNA/protein-coding gene pairs in cancer. *Brief Funct Genomics*. 2016;15(3):167-173. doi:10.1093/bfgp/elv048

55. Austenaa LMI, Barozzi I, Simonatto M, et al. Transcription of Mammalian cis-Regulatory Elements Is Restrained by Actively Enforced Early Termination. *Mol Cell*. 2015;60(3):460-474. doi:10.1016/j.molcel.2015.09.018

56. Ptashne M, Gann A. Transcriptional activation by recruitment. *Nature*. 1997;386(6625):569-577. doi:10.1038/386569a0

57. Rahl PB, Lin CY, Seila AC, et al. C-Myc regulates transcriptional pause release. *Cell*. 2010;141(3):432-445. doi:10.1016/j.cell.2010.03.030

58. Jonkers I, Kwak H, Lis JT. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife*. 2014;2014(3):1-25. doi:10.7554/eLife.02407

59. Mayer A, Di Iulio J, Maleri S, et al. Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell*. 2015;161(3):541-544. doi:10.1016/j.cell.2015.03.010

60. Laitem C, Zaborowska J, Isa NF, Kufs J, Dienstbier M, Murphy S. CDK9 inhibitors define elongation checkpoints at both ends of RNA polymerase II-transcribed genes. *Nat Struct Mol Biol*. 2015;22(5):396-403. doi:10.1038/nsmb.3000

61. Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet*. 2012;13(10):720-731. doi:10.1038/nrg3293

62. Margaritis T, Holstege FCP. Poised RNA Polymerase II Gives Pause for Thought. *Cell*. 2008;133(4):581-584. doi:10.1016/j.cell.2008.04.027

63. Mayer A, Landry HM, Churchman LS. Pause & go: from the discovery of RNA polymerase pausing to its functional implications. *Curr Opin Cell Biol*. 2017;46:72-80. doi:10.1016/j.ceb.2017.03.002

64. Henriques T, Gilchrist DA, Nechaev S, et al. Stable pausing by rna polymerase II provides an opportunity to target and integrate regulatory signals. *Mol Cell*. 2013;52(4):517-528. doi:10.1016/j.molcel.2013.10.001

65. Saldi T, Cortazar MA, Sheridan RM, Bentley DL. Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. *J Mol Biol*. 2016;428(12):2623-2635. doi:10.1016/j.jmb.2016.04.017

66. Muniz L, Nicolas E, Trouche D. RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. *EMBO J*. 2021;40(15):1-21. doi:10.15252/embj.2020105740

67. Fong N, Kim H, Zhou Y, et al. Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev*. 2014;28(23):2663-2676. doi:10.1101/gad.252106.114

68. Hazelbaker DZ, Marquardt S, Wlotzka W, Buratowski S. Kinetic Competition between RNA Polymerase II and Sen1-Dependent Transcription Termination. *Mol Cell*. 2013;49(1):55-66. doi:10.1016/j.molcel.2012.10.014

69. Fong N, Brannan K, Erickson B, et al. Effects of Transcription Elongation Rate and Xrn2 Exonuclease Activity on RNA Polymerase II Termination Suggest Widespread Kinetic Competition. *Mol Cell*. 2015;60(2):256-267. doi:10.1016/j.molcel.2015.09.026

70. Core L, Adelman K. Promoter-proximal pausing of RNA polymerase II: A nexus of gene regulation. *Genes Dev*. 2019;33(15-16):960-982. doi:10.1101/gad.325142.119

71. Sabari BR, Sabari BR, Agnese AD, et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science*. 2018;3958(6400):1-17.

72. Gilchrist DA, Nechaev S, Lee C, et al. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev* . 2008;22(14):1921-1933. doi:10.1101/gad.1643208

73. Gilchrist DA, Santos G Dos, Fargo DC, et al. Pausing of RNA Polymerase II Disrupts DNA-Specified Nucleosome Organization to Enable Precise Gene Regulation. *Cell*. 2010;143(4):540-551. doi:10.1016/j.cell.2010.10.004

74. Boettiger AN, Levine M. Synchronous and stochastic patterns of gene activation in the drosophila embryo. *Science*. 2009;325(5939):471-473. doi:10.1126/science.1173976

75. Williams LH, Fromm G, Gokey NG, et al. Pausing of RNA Polymerase II Regulates Mammalian Developmental Potential through Control of Signaling Networks. *Mol Cell*. 2015;58(2):311-322. doi:10.1016/j.molcel.2015.02.003

76. Gariglio P, Bellard M, Chambon P. Clustering of RNA polymerase B molecules in the 5′ moiety of the adult β-globin gene of hen erythrocytes. *Nucleic Acids Res*. 1981;9(11):2589-2598. doi:10.1093/nar/9.11.2589

77. Gilmour DS, Lis JT. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. *Mol Cell Biol*. 1986;6(11):3984-3989. doi:10.1128/mcb.6.11.3984-3989.1986

78. Rasmussen EB, Lis JT. In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. *Proc Natl Acad Sci*. 1993;90(17):7923-7927. doi:10.1073/pnas.90.17.7923

79. 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 Dev* . 1992;6(11):2201-2213. doi:10.1101/gad.6.11.2201

80. Strobl LJ, Eick D. Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-myc in vivo. *EMBO J*. 1992;11(9):3307-3314. doi:https://doi.org/10.1002/j.1460-2075.1992.tb05409.x

81. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. *Cell*. 2007;130(1):77-88. doi:10.1016/j.cell.2007.05.042

82. Muse GW, Gilchrist DA, Nechaev S, et al. RNA polymerase is poised for activation across the genome. *Nat Genet*. 2007;39(12):1507-1511. doi:10.1038/ng.2007.21

83. Zeitlinger J, Stark A, Kellis M, et al. RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. *Nat Genet*. 2007;39(12):1512-1516. doi:10.1038/ng.2007.26

84. Kwak H, Lis JT. Control of Transcriptional Elongation. *Annu Rev Genet*. 2013;47(1):483-508. doi:10.1146/annurev-genet-110711-155440

85. Dangkulwanich M, Ishibashi T, Liu S, et al. Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. *Elife*. 2013;2:e00971. doi:10.7554/eLife.00971

86. Noe Gonzalez M, Blears D, Svejstrup JQ. Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nat Rev Mol Cell Biol*. 2021;22(1):3-21. doi:10.1038/s41580-020-00308-8

87. Vos SM, Farnung L, Boehning M, et al. Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature*. 2018;560(7720):607-612. doi:10.1038/s41586-018-0440-4

88. Vos SM, Farnung L, Linden A, Urlaub H, Cramer P. Structure of complete Pol II–DSIF–PAF–SPT6 transcription complex reveals RTF1 allosteric activation. *Nat Struct Mol Biol*. 2020;27(7):668-677. doi:10.1038/s41594-020-0437-1

89. Hendrix DA, Hong J-W, Zeitlinger J, Rokhsar DS, Levine MS. Promoter elements

associated with RNA Pol II stalling in the Drosophila embryo. *Proc Natl Acad Sci*. 2008;105(22):7762-7767. doi:10.1073/pnas.0802406105

90. Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, Adelman K. Global Analysis of Short RNAs. *Science*. 2010;327(5963):335-338. doi:10.1126/science.1181421

91. Watts JA, Burdick J, Daigneault J, et al. cis Elements that Mediate RNA Polymerase II Pausing Regulate Human Gene Expression. *Am J Hum Genet*. 2019:1-12. doi:10.1016/j.ajhg.2019.08.003

92. Gajos M, Jasnovidova O, van Bömmel A, Freier S, Vingron M, Mayer A. Conserved DNA sequence features underlie pervasive RNA polymerase pausing. *Nucleic Acids Res*. 2021;8(7):4402-4420. doi:10.1093/nar/gkab208

93. Zamft B, Bintu L, Ishibashi T, Bustamante, Carlos. Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases. *Proc Natl Acad Sci*. 2012;109(23):8948-8953. doi:10.1073/pnas.1205063109

94. Kulish D, Struhl K. TFIIS Enhances Transcriptional Elongation through an Artificial Arrest Site In Vivo. *Mol Cell Biol*. 2001;21(13):4162-4168. doi:10.1128/mcb.21.13.4162-4168.2001

95. Cheung ACM, Cramer P. Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature*. 2011;471(7337):249-253. doi:10.1038/nature09785

96. Danko CG, Hah N, Luo X, et al. Signaling Pathways Differentially Affect RNA Polymerase II Initiation, Pausing, and Elongation Rate in Cells. *Mol Cell*. 2013;50(2):212-222. doi:10.1016/j.molcel.2013.02.015

97. Fuchs G, Voichek Y, Benjamin S, Shlomit G, Amit I, Moshe O. 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biol*. 2014;15(5):R69.

98. Ring BZ, Yarnell WS, Roberts JW. Function of E. coli RNA Polymerase σ Factor-σ70 in Promoter-Proximal Pausing. *Cell*. 1996;86(3):485-493. doi:https://doi.org/10.1016/S0092-8674(00)80121-X

99. Zhilina E, Esyunina D, Brodolin K, Kulbachinskiy A. Structural transitions in the transcription elongation complexes of bacterial RNA polymerase during σ-dependent pausing. *Nucleic Acids Res*. 2012;40(7):3078-3091. doi:10.1093/nar/gkr1158

100. Yamaguchi Y, Takagi T, Wada T, et al. NELF, a Multisubunit Complex Containing RD, Cooperates with DSIF to Repress RNA Polymerase II Elongation. *Cell*. 1999;97(1):41-51. doi:https://doi.org/10.1016/S0092-8674(00)80713-8

101. Yamaguchi Y, Inukai N, Takashi N, 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. *Mol Cell Biol*. 2002;22(9):2918-2927. doi:10.1128/MCB.22.9.2918-2927.2002

102. Žumer K, Maier KC, Farnung L, et al. Two distinct mechanisms of RNA polymerase II elongation stimulation in vivo. *Mol Cell*. 2021;81(15):3096-3109.e8. doi:10.1016/j.molcel.2021.05.028

103. Weber CM, Ramachandran S, Henikoff S. Nucleosomes are context-specific, H2A.Z-Modulated barriers to RNA polymerase. *Mol Cell*. 2014;53(5):819-830. doi:10.1016/j.molcel.2014.02.014

104. Kwak H, Nicholas JF, Core LJ, Lis JT. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing *Science*. 2013;339(6122):950. doi:10.1126/science.1229386

105. Hall MA, Shundrovsky A, Bai L, Fulbright RM, Lis JT, Wang MD. High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat Struct Mol Biol*. 2009;16(2):124-129. doi:10.1038/nsmb.1526

106. Farnung L, Vos SM, Cramer P. Structure of transcribing RNA polymerase II-nucleosome complex. *Nat Commun*. 2018;9(1):1-6. doi:10.1038/s41467-018-07870-y

107. Ehara H, Kujirai T, Fujino Y, Shirouzu M, Kurumizaka H, Sekine S ichi. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science*. 2019.363(6428):744-747. doi:10.1126/science.aav8912

108. Mason PB, Struhl K. The FACT Complex Travels with Elongating RNA Polymerase II and Is Important for the Fidelity of Transcriptional Initiation In Vivo. *Mol Cell Biol*. 2003;23(22):8323-8333. doi:10.1128/MCB.23.22.8323-8333.2003

109. Jamai A, Puglisi A, Strubin M. Histone Chaperone Spt16 Promotes Redeposition of the Original H3-H4 Histones Evicted by Elongating RNA Polymerase. *Mol Cell*. 2009;35(3):377-383. doi:https://doi.org/10.1016/j.molcel.2009.07.001

110. Wang T, Liu Y, Edwards G, Krzizike D, Scherman H, Luger K. The histone chaperone FACT modulates nucleosome structure by tethering its components. *Life Sci Alliance*. 2018;1(4):e201800107. doi:10.26508/lsa.201800107

111. Skene PJ, Hernandez AE, Groudine M, Henikoff S. The nucleosomal barrier to promoter escape by RNA polymerase ii is overcome by the chromatin remodeler Chd1. *Elife*. 2014;2014(3):1-19. doi:10.7554/eLife.02042

112. Teves SS, Weber CM, Henikoff S. Transcribing through the nucleosome. *Trends Biochem Sci*. 2014;39(12):577-586. doi:https://doi.org/10.1016/j.tibs.2014.10.004

113. Kujirai T, Kurumizaka H. Transcription through the nucleosome. *Curr Opin Struct Biol*. 2020;61:42-49. doi:https://doi.org/10.1016/j.sbi.2019.10.007

114. Bortvin A, Winston F. Evidence That Spt6p Controls Chromatin Structure by a Direct Interaction with Histones. *Science*. 1996;272(5267):1473-1476. doi:10.1126/science.272.5267.1473

115. Bernecky C, Plitzko JM, Cramer P. Structure of a transcribing RNA polymerase II-DSIF complex reveals a multidentate DNA-RNA clamp. *Nat Struct Mol Biol*. 2017;24(10):809-815. doi:10.1038/nsmb.3465

116. Vos SM, Farnung L, Urlaub H, Cramer P. Structure of paused transcription complex Pol II–DSIF–NELF. *Nature*. 2018;560(7720):601-606. doi:10.1038/s41586-018-0442-2

117. Crickard JB, Fu J, Reese JC. Biochemical analysis of yeast suppressor of Ty 4/5 (Spt4/5) reveals the importance of nucleic acid interactions in the prevention of RNA polymerase II arrest. *J Biol Chem*. 2016;291(19):9853-9870. doi:10.1074/jbc.M116.716001

118. Wada T, Takagi T, Yamaguchi Y, et al. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* . 1998;12(3):343-356. http://genesdev.cshlp.org/content/12/3/343.abstract.

119. Core LJ, Waterfall JJ, Gilchrist DA, et al. Defining the Status of RNA Polymerase at Promoters. *Cell Rep*. 2012;2(4):1025-1035. doi:https://doi.org/10.1016/j.celrep.2012.08.034

120. Aoi Y, Smith ER, Shah AP, et al. NELF Regulates a Promoter-Proximal Step Distinct from RNA Pol II Pause-Release. *Mol Cell*. 2020;78(2):261-274.e5. doi:10.1016/j.molcel.2020.02.014

121. Wei P, Garber ME, Fang S-M, 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):451-462. doi:https://doi.org/10.1016/S0092-8674(00)80939-3

122. Marshall NF, Price DH. Purification of P-TEFb, a Transcription Factor Required for the Transition into Productive Elongation (∗). *J Biol Chem*. 1995;270(21):12335-12338. doi:https://doi.org/10.1074/jbc.270.21.12335

123. Luo Z, Lin C, Shilatifard A. The super elongation complex (SEC) family in transcriptional control. *Nat Rev Mol Cell Biol*. 2012;13(9):543-547. doi:10.1038/nrm3417

124. Yik JHN, Chen R, Pezda AC, Zhou Q. Compensatory Contributions of HEXIM1 and HEXIM2 in Maintaining the Balance of Active and Inactive Positive Transcription Elongation Factor b Complexes for Control of Transcription. *J Biol Chem*. 2005;280(16):16368-16376. doi:https://doi.org/10.1074/jbc.M500912200

125. Yik JHN, Chen R, Nishimura R, Jennings JL, Link AJ, Zhou Q. Inhibition of P-TEFb (CDK9/Cyclin T) Kinase and RNA Polymerase II Transcription by the Coordinated Actions of HEXIM1 and 7SK snRNA. *Mol Cell*. 2003;12(4):971-982. doi:https://doi.org/10.1016/S1097-2765(03)00388-5

126. Zhou K, Zhuang S, Liu F, et al. Disrupting the Cdk9/Cyclin T1 heterodimer of 7SK snRNP for the Brd4 and AFF1/4 guided reconstitution of active P-TEFb. *Nucleic Acids Res*. 2022;50(2):750-762. doi:10.1093/nar/gkab1228

127. D'Orso I, Frankel AD. RNA-mediated displacement of an inhibitory snRNP complex activates transcription elongation. *Nat Struct Mol Biol*. 2010;17(7):815-821. doi:10.1038/nsmb.1827

128. Faust TB, Li Y, Bacon CW, et al. The HIV-1 Tat protein recruits a ubiquitin ligase to reorganize the 7SK snRNP for transcriptional activation. Jones KA, ed. *Elife*. 2018;7:e31879. doi:10.7554/eLife.31879

129. Cho S, Schroeder S, Kaehlcke K, et al. Acetylation of cyclin T1 regulates the equilibrium between active and inactive P-TEFb in cells. *EMBO J*. 2009;28(10):1407-1417. doi:10.1038/emboj.2009.99

130. Schröder S, Cho S, Zeng L, et al. Two-pronged binding with bromodomain-containing protein 4 liberates positive transcription elongation factor b from inactive ribonucleoprotein complexes. *J Biol Chem*. 2012;287(2):1090-1099. doi:10.1074/jbc.M111.282855

131. Eberhardy SR, Farnham PJ. Myc Recruits P-TEFb to Mediate the Final Step in the Transcriptional Activation of the cad Promoter. *J Biol Chem*. 2002;277(42):40156-40162. doi:https://doi.org/10.1074/jbc.M207441200

132. Jang MK, Kazuki M, Meisheng Z, Jeong H-S, Brady JN, Ozato K. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and

stimulates RNA polymerase II-dependent transcription. *Mol Cell*. 2005;19(4):523-534. doi:10.1016/j.molcel.2005.06.027

133. Yang Z, Yik JHN, Chen R, et al. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell*. 2005;19(4):535-545. doi:10.1016/j.molcel.2005.06.029

134. Peterlin BM, Price DH. Controlling the Elongation Phase of Transcription with P-TEFb. *Mol Cell*. 2006;23(3):297-305. doi:10.1016/j.molcel.2006.06.014

135. Winter GE, Mayer A, Buckley DL, et al. BET Bromodomain Proteins Function as Master Transcription Elongation Factors Independent of CDK9 Recruitment. *Mol Cell*. 2017;67(1):5-18.e19. doi:10.1016/j.molcel.2017.06.004

136. Kanno T, Kanno Y, LeRoy G, et al. BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. *Nat Struct Mol Biol*. 2014;21(12):1047-1057. doi:10.1038/nsmb.2912

137. Zheng B, Aoi Y, Shah AP, et al. Acute perturbation strategies in interrogating RNA polymerase II elongation factor function in gene expression. *Genes Dev*. 2021;35(3-4):273-285. doi:10.1101/gad.346106.120

138. Kim JB, Sharp PA. Positive Transcription Elongation Factor b Phosphorylates hSPT5 and RNA Polymerase II Carboxyl-terminal Domain Independently of Cyclin-dependent Kinase-activating Kinase. *J Biol Chem*. 2001;276(15):12317-12323. doi:https://doi.org/10.1074/jbc.M010908200

139. 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. *Mol Cell*. 2006;21(2):227-237. doi:10.1016/j.molcel.2005.11.024

140. Koh F, Dan I, Yehong H, Ran T, Takeshi K, Matija PB. Dynamics of Human Immunodeficiency Virus Transcription: P-TEFb Phosphorylates RD and Dissociates Negative Effectors from the Transactivation Response Element. *Mol Cell Biol*. 2004;24(2):787-795. doi:10.1128/MCB.24.2.787-795.2004

141. Martinez-Rucobo FW, Sainsbury S, Cheung ACM, Cramer P. Architecture of the RNA polymerase–Spt4/5 complex and basis of universal transcription processivity. *EMBO J*. 2011;30(7):1302-1310. doi:https://doi.org/10.1038/emboj.2011.64

142. Bartkowiak B, Liu P, Phatnani HP, et al. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev*. 2010;24(20):2303-2316. doi:10.1101/gad.1968210

143. Greenleaf AL. Human CDK12 and CDK13, multi-tasking CTD kinases for the new millenium. *Transcription*. 2019;10(2):91-110. doi:10.1080/21541264.2018.1535211

144. Lu X, Zhu X, Li Y, et al. Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. *Nucleic Acids Res*. 2016;44(14):6853-6867. doi:10.1093/nar/gkw571

145. Olson CM, Jiang B, Erb MA, et al. Pharmacological perturbation of CDK9 using selective CDK9 inhibition or degradation. *Nat Chem Biol*. 2018;14(2):163-170. doi:10.1038/nchembio.2538

146. Narita T, Yamaguchi Y, Yano K, et al. Human Transcription Elongation Factor NELF: Identification of Novel Subunits and Reconstitution of the Functionally Active Complex. *Mol Cell Biol*. 2003;23(6):1863-1873.

doi:10.1128/MCB.23.6.1863-1873.2003

147. Wier AD, Mayekar MK, Héroux A, Arndt KM, VanDemark AP. Structural basis for Spt5-mediated recruitment of the Paf1 complex to chromatin. *Proc Natl Acad Sci U S A*. 2013;110(43):17290-17295. doi:10.1073/pnas.1314754110

148. Shetty A, Kallgren SP, Demel C, et al. Spt5 Plays Vital Roles in the Control of Sense and Antisense Transcription Elongation. *Mol Cell*. 2017;66(1):77-88.e5. doi:10.1016/j.molcel.2017.02.023

149. Lindstrom DL, Squazzo SL, Muster N, et al. Dual Roles for Spt5 in Pre-mRNA Processing and Transcription Elongation Revealed by Identification of Spt5-Associated Proteins. *Mol Cell Biol*. 2003;23(4):1368-1378. doi:10.1128/mcb.23.4.1368-1378.2003

150. Fitz J, Neumann T, Pavri R. Regulation of RNA polymerase II processivity by Spt5 is restricted to a narrow window during elongation . *EMBO J*. 2018;37(8):1-20. doi:10.15252/embj.201797965

151. Hu S, Peng L, Xu C, Wang Z, Song A, Chen FX. SPT5 stabilizes RNA polymerase II, orchestrates transcription cycles, and maintains the enhancer landscape. *Mol Cell*. 2021;81(21):4425-4439.e6. doi:10.1016/j.molcel.2021.08.029

152. Aoi Y, Takahashi Y, Shah AP, et al. SPT5 stabilization of promoter-proximal RNA polymerase II. *Mol Cell*. 2021;82(21):4413-4424.e5. doi:10.1016/j.molcel.2021.08.006

153. Fitz J, Neumann T, Steininger M, et al. Spt5-mediated enhancer transcription directly couples enhancer activation with physical promoter interaction. *Nat Genet*. 2020;22(5):1-11. doi:10.1038/s41588-020-0605-6

154. Rozenblatt-Rosen O, Hughes CM, Nannepaga SJ, et al. The Parafibromin Tumor Suppressor Protein Is Part of a Human Paf1 Complex. *Mol Cell Biol*. 2005;25(2):612-620. doi:10.1128/mcb.25.2.612-620.2005

155. Kim M, Ahn SH, Krogan NJ, Greenblatt JF, Buratowski S. Transitions in RNA polymerase II elongation complexes at the 3′ ends of genes. *EMBO J*. 2004;23(2):354-364. doi:10.1038/sj.emboj.7600053

156. Yu M, Yang W, Ni T, et al. RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. *Science*. 2015;350(6266):1383-1386. doi:10.1126/science.aad2338

157. Hou L, Wang Y, Liu Y, et al. Paf1C regulates RNA polymerase II progression by modulating elongation rate. *Proc Natl Acad Sci U S A*. 2019;116(29):14583-14592. doi:10.1073/pnas.1904324116

158. Chen FX, Woodfin AR, Gardini A, et al. PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. *Cell*. 2015;162(5):1003-1015. doi:10.1016/j.cell.2015.07.042

159. Chen FX, Collings CK, Cao K, et al. PAF1 regulation of promoter-proximal pause release via enhancer activation. *Science*. 2017;357(6357):1294-1298. doi:10.1126/science.aan3269

160. Ding L, Paszkowski-Rogacz M, Mircetic J, Chakraborty D, Buchholz F. The Paf1 complex positively regulates enhancer activity in mouse embryonic stem cells. *Life Sci Alliance*. 2020;4(3):1-14. doi:10.26508/LSA.202000792

161. Sdano MA, Fulcher JM, Palani S, et al. A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. *Elife*. 2017;6:e28723. doi:10.7554/eLife.28723

162. Hartzog GA, Fu J. The Spt4-Spt5 complex: A multi-faceted regulator of transcription elongation. *Biochim Biophys Acta - Gene Regul Mech*. 2013;1829(1):105-115. doi:10.1016/j.bbagrm.2012.08.007

163. Jeronimo C, Poitras C, Robert F. Histone Recycling by FACT and Spt6 during Transcription Prevents the Scrambling of Histone Modifications. *Cell Rep*. 2019;28(5):1206-1218.e8. doi:10.1016/j.celrep.2019.06.097

164. Endoh M, Zhu W, Hasegawa J, et al. Human Spt6 Stimulates Transcription Elongation by RNA Polymerase II In Vitro. *Mol Cell Biol*. 2004;24(8):3324-3336. doi:10.1128/mcb.24.8.3324-3336.2004

165. Narain A, Bhandare P, Adhikari B, et al. Targeted protein degradation reveals a direct role of SPT6 in RNAPII elongation and termination. *Mol Cell*. 2021;81(15):3110-3127.e14. doi:10.1016/j.molcel.2021.06.016

166. Doris SM, Chuang J, Viktorovskaya O, et al. Spt6 Is Required for the Fidelity of Promoter Selection. *Mol Cell*. 2018;72(4):687-699.e6. doi:10.1016/j.molcel.2018.09.005

167. Nojima T, Rebelo K, Gomes T, Grosso AR, Proudfoot NJ, Carmo-Fonseca M. RNA Polymerase II Phosphorylated on CTD Serine 5 Interacts with the Spliceosome during Co-transcriptional Splicing. *Mol Cell*. 2018;72(2):369-379.e4. doi:10.1016/j.molcel.2018.09.004

168. 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. *Mol Cell*. 2017;67(3):411-422.e4. doi:10.1016/j.molcel.2017.06.027

169. Darzacq X, Shav-Tal Y, De Turris V, et al. In vivo dynamics of RNA polymerase II transcription. *Nat Struct Mol Biol*. 2007;14(9):796-806. doi:10.1038/nsmb1280

170. Erickson B, Sheridan RM, Cortazar M, Bentley DL. Dynamic turnover of paused Pol II complexes at human promoters. *Genes Dev*. 2018;32(17-18):1-11. doi:10.1101/gad.316810.118.

171. Chiu AC, Suzuki HI, Wu X, Mahat DB, Kriz AJ, Sharp PA. Transcriptional Pause Sites Delineate Stable Nucleosome-Associated Premature Polyadenylation Suppressed by U1 snRNP. *Mol Cell*. 2018;69(4):648-663.e7. doi:10.1016/j.molcel.2018.01.006

172. Lemay JF, Bachand F. Fail-safe transcription termination: Because one is never enough. *RNA Biol*. 2015;12(9):927-932. doi:10.1080/15476286.2015.1073433

173. Shearwin KE, Callen BP, Egan JB. Transcriptional interference - A crash course. *Trends Genet*. 2005;21(6):339-345. doi:10.1016/j.tig.2005.04.009

174. Mazo A, Hodgson JW, Petruk S, Sedkov Y, Brock HW. Transcriptional interference: An unexpected layer of complexity in gene regulation. *J Cell Sci*. 2007;120(16):2755-2761. doi:10.1242/jcs.007633

175. Creamer TJ, Darby MM, Jamonnak N, et al. Transcriptome-Wide Binding Sites for Components of the Saccharomyces cerevisiae Non-Poly(A) Termination Pathway: Nrd1, Nab3, and Sen1. *PLOS Genet*. 2011;7(10):e1002329. https://doi.org/10.1371/journal.pgen.1002329.

176. Schwalb B, Michel M, Zacher B, et al. TT-seq maps the human transient transcriptome. *Science*. 2016;352(6290):1225-1228. doi:10.1126/science.aad9841

177. Stewart M. Polyadenylation and nuclear export of mRNAs. *J Biol Chem*. 2019;294(9):2977-2987. doi:10.1074/jbc.REV118.005594

178. Connelly S, Manley JL. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev*. 1988;2(4):440-452. doi:10.1101/gad.2.4.440

179. Hirose Y, Manley JL. RNA polymerase II is an essential mRNA polyadenylation factor. *Nature*. 1998;395(6697):93-96. doi:10.1038/25786

180. Shi Y, Di Giammartino DC, Taylor D, et al. Molecular Architecture of the Human Pre-mRNA 3' Processing Complex. *Mol Cell*. 2009;33(3):365-376. doi:10.1016/j.molcel.2008.12.028

181. Osheim YN, Proudfoot NJ, Beyer AL. EM Visualization of Transcription by RNA Polymerase II. *Mol Cell*. 1999;3(3):379-387. doi:10.1016/s1097-2765(00)80465-7

182. Zhang H, Rigo F, Martinson HG. Poly(A) Signal-Dependent Transcription Termination Occurs through a Conformational Change Mechanism that Does Not Require Cleavage at the Poly(A) Site. *Mol Cell*. 2015;59(3):437-448. doi:10.1016/j.molcel.2015.06.008

183. Eaton JD, West S. An end in sight? Xrn2 and transcriptional termination by RNA polymerase II. *Transcription*. 2018;9(5):21541264.2018.1498708. doi:10.1080/21541264.2018.1498708

184. Eaton JD, Francis L, Davidson L, West S. A unified allosteric/torpedo mechanism for transcriptional termination on human protein-coding genes. *Genes Dev*. 2020;34(1-2):132-145. doi:10.1101/gad.332833.119

185. Whitelaw E, Proudfoot N. Alpha-thalassaemia caused by a poly(A) site mutation reveals that transcriptional termination is linked to 3′ end processing in the human alpha 2 globin gene. *EMBO J*. 1986;5(11):2915-2922. doi:https://doi.org/10.1002/j.1460-2075.1986.tb04587.x

186. Zarudnaya MI, Kolomiets IM, Potyahaylo AL, Hovorun DM. Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. *Nucleic Acids Res*. 2003;31(5):1375-1386. doi:10.1093/nar/gkg241

187. Zhu Y, Wang X, Forouzmand E, et al. Molecular Mechanisms for CFIm-Mediated Regulation of mRNA Alternative Polyadenylation. *Mol Cell*. 2018;69(1):62-74.e4. doi:https://doi.org/10.1016/j.molcel.2017.11.031

188. Hu J, Lutz CS, Wilusz J, Tian B. Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. *RNA*. 2005;11(10):1485-1493. doi:10.1261/rna.2107305

189. Zhang Y, Sun Y, Shi Y, Walz T, Tong L. Structural Insights into the Human Pre-mRNA 3'-End Processing Machinery. *Mol Cell*. 2019;77(4):800-806.e6. doi:10.1016/j.molcel.2019.11.005

190. Laishram RS. Poly(A) polymerase (PAP) diversity in gene expression - Star-PAP vs canonical PAP. *FEBS Lett*. 2014;588(14):2185-2197. doi:10.1016/j.febslet.2014.05.029

191. Clerici M, Faini M, Aebersold R, Jinek M. Structural insights into the assembly and polya signal recognition mechanism of the human CPSF complex. *Elife*. 2017;6:1-20. doi:10.7554/eLife.33111

192. Sun Y, Zhang Y, Hamilton K, et al. Molecular basis for the recognition of the human AAUAAA polyadenylation signal. *Proc Natl Acad Sci*. 2018;115(7):E1419-E1428. doi:10.1073/pnas.1718723115

193. Helmling S, Zhelkovsky A, Moore CL. Fip1 Regulates the Activity of Poly(A) Polymerase through Multiple Interactions. *Mol Cell Biol*. 2001;21(6):2026-2037. doi:10.1128/MCB.21.6.2026-2037.2001

194. Nag A, Narsinh K, Martinson HG. The poly(A)-dependent transcriptional pause is mediated by CPSF acting on the body of the polymerase. *Nat Struct Mol Biol*. 2007;14(7):662-669. doi:10.1038/nsmb1253

195. Mandel CR, Kaneko S, Zhang H, et al. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature*. 2006;444(7121):953-956. doi:10.1038/nature05363

196. Bienroth S, Keller W, Wahle E. Assembly of a processive messenger RNA polyadenylation complex. *EMBO J*. 1993;12(2):585-594. doi:https://doi.org/10.1002/j.1460-2075.1993.tb05690.x

197. Eckmann CR, Rammelt C, Wahle E. Control of poly(A) tail length. *WIREs RNA*. 2011;2(3):348-361. doi:https://doi.org/10.1002/wrna.56

198. Takagaki Y, Manley JL. Complex Protein Interactions within the Human Polyadenylation Machinery Identify a Novel Component. *Mol Cell Biol*. 2000;20(5):1515-1525. doi:10.1128/MCB.20.5.1515-1525.2000

199. Schönemann L, Kühn U, Martin G, et al. Reconstitution of CPSF active in polyadenylation: Recognition of the polyadenylation signal by WDR33. *Genes Dev*. 2014;28(21):2381-2393. doi:10.1101/gad.250985.114

200. Nagaike T, Logan C, Hotta I, Rozenblatt-Rosen O, Meyerson M, Manley JL. Transcriptional Activators Enhance Polyadenylation of mRNA Precursors. *Mol Cell*. 2011;41(4):409-418. doi:10.1016/j.molcel.2011.01.022

201. Xiang K, Nagaike T, Xiang S, et al. Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex. *Nature*. 2010;467(7316):729-733. doi:10.1038/nature09391

202. Sun Y, Zhang Y, Aik WS, et al. Structure of an active human histone pre-mRNA 3′-end processing machinery. *Science*. 2020;367(6478):700-703. doi:10.1126/science.aaz7758

203. Murthy KGK, Manley JL. The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev*. 1995;9(21):2672-2683. doi:10.1101/gad.9.21.2672

204. Gilmartin GM, Nevins JR. An ordered pathway of assembly of components required for polyadenylation site recognition and processing. *Genes Dev*. 1989;3(12 B):2180-2190. doi:10.1101/gad.3.12b.2180

205. Fong N, Bentley DL. Capping, splicing, and 3′ processing are independently stimulated by RNA polymerase II: Different functions for different segments of the CTD. *Genes Dev*. 2001;15(14):1783-1795. doi:10.1101/gad.889101

206. Kaneko S, Son J, Shen SS, Reinberg D, Bonasio R. PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. *Nat Struct Mol Biol*. 2013;20(11):1258-1264. doi:10.1038/nsmb.2700

207. Brown KM, Gilmartin GM. A Mechanism for the Regulation of Pre-mRNA 3′ Processing by Human Cleavage Factor Im. *Mol Cell*. 2003;12(6):1467-1476. doi:https://doi.org/10.1016/S1097-2765(03)00453-2

208. Yang Q, Gilmartin G, Doublié S. Structural basis of UGUA recognition by the Nudix protein CFIm25 and implications for a regulatory role in mRNA 3′ processing. *Proc Natl Acad Sci*. 2010;107(22):10062-10067. doi:10.1073/pnas.1000848107

209. Gruber AR, Martin G, Keller W, Zavolan M. Cleavage factor Im is a key regulator of 3′ UTR length. *RNA Biol*. 2012;9(12):1405-1412. doi:10.4161/rna.22570

210. Barillà D, Lee BA, Proudfoot NJ. Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae . *Proc Natl Acad Sci*. 2001;98(2):445-450. doi:10.1073/pnas.98.2.445

211. West S, Proudfoot NJ. Human Pcf11 enhances degradation of RNA polymerase II-associated nascent RNA and transcriptional termination. *Nucleic Acids Res*. 2008;36(3):905-914. doi:10.1093/nar/gkm1112

212. Kamieniarz-Gdula K, Gdula MR, Panser K, et al. Selective Roles of Vertebrate PCF11 in Premature and Full-Length Transcript Termination. *Mol Cell*. 2019;74(1):158-172.e9.. doi:10.1016/j.molcel.2019.01.027

213. Dominski Z, Yang X, Marzluff WF. The Polyadenylation Factor CPSF-73 Is Involved in Histone-Pre-mRNA Processing. *Cell*. 2005;123(1):37-48. doi:https://doi.org/10.1016/j.cell.2005.08.002

214. Yang X-C, Sabath I, Dębski J, et al. A Complex Containing the CPSF73 Endonuclease and Other Polyadenylation Factors Associates with U7 snRNP and Is Recruited to Histone Pre-mRNA for 3′-End Processing. *Mol Cell Biol*. 2013;33(1):28-37. doi:10.1128/MCB.00653-12

215. Sullivan KD, Steiniger M, Marzluff WF. A Core Complex of CPSF73, CPSF100, and Symplekin May Form Two Different Cleavage Factors for Processing of Poly(A) and Histone mRNAs. *Mol Cell*. 2009;34(3):322-332. doi:10.1016/j.molcel.2009.04.024

216. Kolev NG, Steitz JA. Symplekin and multiple other polyadenylation factors participate in 3′-end maturation of histone mRNAs. *Genes Dev* . 2005;19(21):2583-2592. doi:10.1101/gad.1371105

217. Bentley DL. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr Opin Cell Biol*. 2005;17(3):251-256. doi:https://doi.org/10.1016/j.ceb.2005.04.006

218. Calvo O, Manley JL. Strange bedfellows: polyadenylation factors at the promoter. *Genes Dev* . 2003;17(11):1321-1327. doi:10.1101/gad.1093603

219. He X, Khan AU, Cheng H, Pappas DL, Hampsey M, Moore CL. Functional interactions between the transcription and mRNA 3′ end processing machineries mediated by Ssu72 and Sub1. *Genes Dev* . 2003;17(8):1030-1042. doi:10.1101/gad.1075203

220. Dantonel JC, Murthy KGK, Manjey JL, Tora L. Transcription factor TFIID recruits factor CPSF for formation of 3 end of mRNA. *Nature*. 1997;389(6649):399-402.

doi:10.1038/38763

221. Swinburne IA, Meyer CA, Liu XS, Silver PA, Brodsky AS. Genomic localization of RNA binding proteins reveals links between pre-mRNA processing and transcription. *Genome Res*. 2006;16(7):912-921. doi:10.1101/gr.5211806

222. Nojima T, Gomes T, Grosso ARF, et al. Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell*. 2015;161(3):526-540. doi:10.1016/j.cell.2015.03.027

223. Venkataraman K, Brown KM, Gilmartin GM. Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev* . 2005;19(11):1315-1327. doi:10.1101/gad.1298605

224. 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. *Nat Struct Mol Biol*. 2008;15(1):71-78. doi:10.1038/nsmb1352

225. Davidson L, Muniz L, West S. 3' end formation of pre-mRNA and phosphorylation of Ser2 on the RNA polymerase II CTD are reciprocally coupled in human cells. *Genes Dev*. 2014;28(4):342-356. doi:10.1101/gad.231274.113

226. O'Sullivan JM, Tan-Wong SM, Morillon A, et al. Gene loops juxtapose promoters and terminators in yeast. *Nat Genet*. 2004;36(9):1014-1018. doi:10.1038/ng1411

227. Ansari A, Hampsey M. A role for the CPF 3′-end processing machinery in RNAP II-dependent gene looping. *Genes Dev* . 2005;19(24):2969-2978. doi:10.1101/gad.1362305

228. Singh BN, Hampsey M. A Transcription-Independent Role for TFIIB in Gene Looping. *Mol Cell*. 2007;27(5):806-816. doi:https://doi.org/10.1016/j.molcel.2007.07.013

229. Tan-Wong SM, French JD, Proudfoot NJ, Brown MA. Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Proc Natl Acad Sci*. 2008;105(13):5160-5165. doi:10.1073/pnas.0801048105

230. Perkins KJ, Lusic M, Mitar I, Giacca M, Proudfoot NJ. Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. *Mol Cell*. 2008;29(1):56-68. doi:10.1016/j.molcel.2007.11.030

231. West S, Gromak N, Proudfoot NJ. Human 5′ → 3′ exonuclease Xm2 promotes transcription termination at co-transcriptional cleavage sites. *Nature*. 2004;432(7016):522-525. doi:10.1038/nature03035

232. Kim M, Krogan NJ, Vasiljeva L, et al. The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*. 2004;432(7016):517-522. doi:10.1038/nature03041

233. Pearson EL, Moore CL. Dismantling Promoter-driven RNA Polymerase II Transcription Complexes in Vitro by the Termination Factor Rat1. *J Biol Chem*. 2013;288(27):19750-19759. doi:https://doi.org/10.1074/jbc.M112.434985

234. Dengl S, Cramer P. Torpedo Nuclease Rat1 Is Insufficient to Terminate RNA Polymerase II in Vitro*. *J Biol Chem*. 2009;284(32):21270-21279. doi:https://doi.org/10.1074/jbc.M109.013847

235. Skourti-Stathaki K, Proudfoot NJ, Gromak N. Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-

Dependent Termination. *Mol Cell*. 2011;42(6):794-805. doi:https://doi.org/10.1016/j.molcel.2011.04.026

236. Eaton JD, Davidson L, Bauer DLV, Natsume T, Kanemaki MT, West S. Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes Dev*. 2018;32(2):127-139. doi:10.1101/gad.308528.117

237. Logan J, Falck-Pedersen E, Darnell JE, Shenk T. A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proc Natl Acad Sci*. 1987;84(23):8306-8310. doi:10.1073/pnas.84.23.8306

238. Cortazar MA, Sheridan RM, Erickson B, et al. Control of RNA Pol II Speed by PNUTS-PP1 and Spt5 Dephosphorylation Facilitates Termination by a '" Sitting Duck Torpedo " Mechanism. *Mol Cell*. 2019;76(6):896-808.e4. doi:10.1016/j.molcel.2019.09.031

239. Gregersen LH, Mitter R, Ugalde AP, Agami R, Stewart A, Svejstrup. SCAF4 and SCAF8, mRNA Anti-Terminator Proteins. *Mol. Cell.* 2019;177(7):1797-1813.e18. doi:10.1016/j.cell.2019.04.038

240. Zhang Z, Fu J, Gilmour DS. CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. *Genes Dev* . 2005;19(13):1572-1580. doi:10.1101/gad.1296305

241. Kaneko S, Rozenblatt-Rosen O, Meyerson M, Manley JL. The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3′ processing and transcription termination. *Genes Dev* . 2007;21(14):1779-1789. http://genesdev.cshlp.org/content/21/14/1779.abstract.

242. Richard P, Manley JL. Transcription termination by nuclear RNA polymerases. *Genes Dev*. 2009;23(11):1247-1269. doi:10.1101/gad.1792809

243. Luo W, Johnson AW, Bentley DL. The role of Rat1 in coupling mRNA 3′-end processing to transcription termination: Implications for a unified allosteric-torpedo model. *Genes Dev*. 2006;20(8):954-965. doi:10.1101/gad.1409106

244. Eaton JD, West S. Termination of Transcription by RNA Polymerase II: BOOM! *Trends Genet*. 2020;36(9):664-675. doi:10.1016/j.tig.2020.05.008

245. Spies N, Nielsen CB, Padgett RA, Burge CB. Biased Chromatin Signatures around Polyadenylation Sites and Exons. *Mol Cell*. 2009;36(2):245-254. doi:https://doi.org/10.1016/j.molcel.2009.10.008

246. Anamika K, Gyenis À, Poidevin L, Poch O, Tora L. RNA Polymerase II Pausing Downstream of Core Histone Genes Is Different from Genes Producing Polyadenylated Transcripts. *PLoS One*. 2012;7(6):e38769. https://doi.org/10.1371/journal.pone.0038769.

247. Enriquez-Harris P, Levitt N, Briggs D, Proudfoot NJ. A pause site for RNA polymerase II is associated with termination of transcription. *EMBO J*. 1991;10(7):1833-1842. doi:10.1002/j.1460-2075.1991.tb07709.x

248. Parua PK, Booth GT, Sansó M, et al. A Cdk9-PP1 switch regulates the elongation-termination transition of RNA polymerase II. *Nature*. 2018;558(7710):460-464. doi:10.1038/s41586-018-0214-z

249. Nordick K, Hoffman MG, Betz JL, Jaehning JA. Direct interactions between the Paf1 complex and a cleavage and polyadenylation factor are revealed by

dissociation of Paf1 from RNA polymerase II. *Eukaryot Cell*. 2008;7(7):1158-1167. doi:10.1128/EC.00434-07

250. Rozenblatt-Rosen O, Nagaike T, Francis JM, et al. The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3′ mRNA processing factors. *Proc Natl Acad Sci*. 2009;106(3):755-760. doi:10.1073/pnas.0812023106

251. Mayer A, Schreieck A, Lidschreiber M, Leike K, Martin DE, Cramer P. The Spt5 C-Terminal Region Recruits Yeast 3' RNA Cleavage Factor I. *Mol Cell Biol*. 2012;32(7):1321-1331. doi:10.1128/MCB.06310-11

252. Ahn SH, Kim M, Buratowski S. Phosphorylation of Serine 2 within the RNA Polymerase II C-Terminal Domain Couples Transcription and 3′ End Processing. *Mol Cell*. 2004;13(1):67-76. doi:10.1016/S1097-2765(03)00492-1

253. Lunde BM, Reichow SL, Kim M, et al. Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat Struct Mol Biol*. 2010;17(10):1195-1201. doi:10.1038/nsmb.1893

254. Meinhart A, Cramer P. Recognition of RNA polymerase II carboxy-terminal domain by 3′-RNA-processing factors. *Nature*. 2004;430(6996):223-226. doi:10.1038/nature02679

255. Gu B, Eick D, Bensaude O. CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic Acids Res*. 2013;41(3):1591-1603. doi:10.1093/nar/gks1327

256. Shah N, Maqbool MA, Yahia Y, et al. Tyrosine-1 of RNA Polymerase II CTD Controls Global Termination of Gene Transcription in Mammals. *Mol Cell*. 2018;69(1):113-125.e6. doi:10.1016/j.molcel.2017.12.009

257. Heidemann M, Hintermair C, Vo?? K, Eick D, Voß K, Eick D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim Biophys Acta - Gene Regul Mech*. 2013;1829(1):55-62. doi:10.1016/j.bbagrm.2012.08.013

258. Schlackow M, Nojima T, Gomes T, Dhir A, Carmo-Fonseca M, Proudfoot NJ. Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. *Mol Cell*. 2017;65(1):25-38. doi:10.1016/j.molcel.2016.11.029

259. Hsin J-P, Sheth A, Manley JL. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science*. 2011;334(6056):683-686. doi:10.1126/science.1206034

260. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. *Nature*. 2010;463(7280):457-463. doi:10.1038/nature08909

261. Mercer TR, Clark MB, Andersen SB, et al. Genome-wide discovery of human splicing branchpoints. *Genome Res* . 2015;25(2):290-303. doi:10.1101/gr.182899.114

262. Mount SM. A catalogue of splice junction sequences. *Nucleic Acids Res*. 1982;10(2):459-472. doi:10.1093/nar/10.2.459

263. Wang Z, Burge CB. Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *Rna*. 2008;14(5):802-813. doi:10.1261/rna.876308

264. Agafonov DE, Deckert J, Wolf E, et al. Semiquantitative Proteomic Analysis of the Human Spliceosome via a Novel Two-Dimensional Gel Electrophoresis Method.

Mol Cell Biol. 2011;31(13):2667-2682. doi:10.1128/mcb.05266-11

265. Cvitkovic I, Jurica MS. Spliceosome Database: a tool for tracking components of the spliceosome. *Nucleic Acids Res*. 2013;41(D1):D132-D141. doi:10.1093/nar/gks999

266. Will CL, Lührmann R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol*. 2011;3(7):1-2. doi:10.1101/cshperspect.a003707

267. Chan SP, Cheng SC. The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *J Biol Chem*. 2005;280(35):31190-31199. doi:10.1074/jbc.M505060200

268. Chanarat S, Seizl M, Sträßer K. The prp19 complex is a novel transcription elongation factor required for TREX occupancy at transcribed genes. *Genes Dev*. 2011;25(11):1147-1158. doi:10.1101/gad.623411

269. Wuarin J, Schibler U. Physical Isolation of Nascent RNA Chains Transcribed by RNA Polymerase II: Evidence for Cotranscriptional Splicing. *Mol Cell Biol*. 1994:7219-7225.

270. Herzel L, Ottoz DSM, Alpert T, Neugebauer KM. Splicing and transcription touch base: Co-transcriptional spliceosome assembly and function. *Nat Rev Mol Cell Biol*. 2017;18(10):637-650. doi:10.1038/nrm.2017.63

271. Tellier M, Maudlin I, Murphy S. Transcription and splicing: A two-way street. *Wiley Interdiscip Rev RNA*. 2020;11(5):e1593. doi:10.1002/wrna.1593

272. Moehle EA, Braberg H, Krogan NJ, Guthrie C. Adventures in time and space. *RNA Biol*. 2014;11(4):313-319. doi:10.4161/rna.28646

273. Neugebauer KM. Nascent RNA and the Coordination of Splicing with Transcription. *Cold Spring Harb Perspect Biol* . 2019;11(8). doi:10.1101/cshperspect.a032227

274. Strikoudis A, Lazaris C, Trimarchi T, et al. Regulation of transcriptional elongation in pluripotency and cell differentiation by the PHD-finger protein Phf5a. *Nat Cell Biol*. 2016;18(11):1127-1138. doi:10.1038/ncb3424

275. Fong YW, Zhou Q. Stimulatory effect of splicing factors on transcriptional elongation. *Nature*. 2001;414(6866):929-933. doi:10.1038/414929a

276. Kyburz A, Friedlein A, Langen H, Keller W. Direct Interactions between Subunits of CPSF and the U2 snRNP Contribute to the Coupling of Pre-mRNA 3′ End Processing and Splicing. *Mol Cell*. 2006;23(2):195-205. doi:10.1016/j.molcel.2006.05.037

277. McCracken S, Lambermon M, Blencowe BJ. SRm160 Splicing Coactivator Promotes Transcript 3′-End Cleavage. *Mol Cell Biol*. 2002;22(1):148-160. doi:10.1128/mcb.22.1.148-160.2002

278. Hirose Y, Tacke R, Manley JL. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev*. 1999;13(10):1234-1239. http://genesdev.cshlp.org/content/13/10/1234.abstract.

279. Harlen KM, Trotta KL, Smith EE, Mosaheb MM, Fuchs SM, Churchman LS. Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Rep*. 2016;15(10):2147-2158. doi:10.1016/j.celrep.2016.05.010

280. Yu Y, Reed R. FUS functions in coupling transcription to splicing by mediating an interaction between RNAP II and U1 snRNP. *Proc Natl Acad Sci U S A*. 2015;112(28):8608-8613. doi:10.1073/pnas.1506282112

281. Dujardin G, Lafaille C, de la Mata M, et al. How Slow RNA Polymerase II Elongation Favors Alternative Exon Skipping. *Mol Cell*. 2014;54(4):683-690. doi:10.1016/j.molcel.2014.03.044

282. De La Mata M, Alonso CR, Kadener S, et al. A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell*. 2003;12(2):525-532. doi:10.1016/j.molcel.2003.08.001

283. Alexander RD, Innocente SA, Barrass JD, Beggs JD. Splicing-Dependent RNA polymerase pausing in yeast. *Mol Cell*. 2010;40(4):582-593. doi:10.1016/j.molcel.2010.11.005

284. Drexler HL, Choquet K, Churchman LS. Splicing Kinetics and Coordination Revealed by Direct Nascent RNA Sequencing through Nanopores. *Mol Cell*. 2020;77(5):985-998.e8. doi:https://doi.org/10.1016/j.molcel.2019.11.017

285. Dey A, Chitsaz F, Abbasi A, Misteli T, Ozato K. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc Natl Acad Sci U S A*. 2003;100(15):8758-8763. doi:10.1073/pnas.1433065100

286. Houzelstein D, Bullock SL, Lynch DE, et al. Growth and Early Postimplantation Defects in Mice Deficient for the Bromodomain-Containing Protein Brd4. 2002;22(11):3794-3802. doi:10.1128/MCB.22.11.3794

287. Jones MH, Numata M, Shimane M. Identification and Characterization of BRDT: A Testis-Specific Gene Related to the Bromodomain Genes RING3 and Drosophila fsh. *Genomics*. 1997;45(3):529-534. doi:https://doi.org/10.1006/geno.1997.5000

288. Federation AJ, Nandakumar V, Searle BC, et al. Highly Parallel Quantification and Compartment Localization of Transcription Factors and Nuclear Report Highly Parallel Quantification and Compartment Localization of Transcription Factors and Nuclear Proteins. *Cell Reports*. 2020;30(8):2463-2471.e5. doi:10.1016/j.celrep.2020.01.096

289. Wu SY, Chiang CM. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J Biol Chem*. 2007;282(18):13141-13145. doi:10.1074/jbc.R700001200

290. Dhalluin C, Carlson JE, Zeng L, et al. Structure and ligand of a histone acetyltransferase bromodomain. *Nature*. 1999;399(6735):491-496. doi:10.1038/20974

291. Bisgrove DA, Mahmoudi T, Henklein P, Verdin E. Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. *Proc Natl Acad Sci U S A*. 2007;104(34):13690-13695. doi:10.1073/pnas.0705053104

292. Gaucher J, Boussouar F, Montellier E, et al. Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J*. 2012;31(19):3809-3820. doi:10.1038/emboj.2012.233

293. Floyd SR, Pacold ME, Huang Q, et al. The bromodomain protein Brd4 insulates chromatin from DNA damage signalling. *Nature*. 2013;498(7453):246-250. doi:10.1038/nature12147

294. Wu SY, Lee CF, Lai HT, et al. Opposing Functions of BRD4 Isoforms in Breast

Cancer. *Mol Cell*. 2020;78(6):1114-1132.e10. doi:10.1016/j.molcel.2020.04.034

295. Nakamura Y, Umehara T, Nakano K, et al. Crystal structure of the human BRD2 bromodomain: Insights into dimerization and recognition of acetylated histone H4. *J Biol Chem*. 2007;282(6):4193-4201. doi:10.1074/jbc.M605971200

296. Filippakopoulos P, Picaud S, Mangos M, et al. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell*. 2012;149(1):214-231. doi:10.1016/j.cell.2012.02.013

297. Vollmuth F, Blankenfeldt W, Geyer M. Structures of the dual bromodomains of the P-TEFb-activating protein Brd4 at atomic resolution. *J Biol Chem*. 2009;284(52):36547-36556. doi:10.1074/jbc.M109.033712

298. Huang H, Zhang J, Shen W, et al. Solution structure of the second bromodomain of Brd2 and its specific interaction with acetylated histone tails. *BMC Struct Biol*. 2007;7:57. doi:10.1186/1472-6807-7-57

299. Liu Y, Wang X, Zhang J, et al. Structural Basis and Binding Properties of the Second Bromodomain of Brd4 with Acetylated Histone Tails. *Biochemistry*. 2008;47(24):6403-6417. doi:10.1021/bi8001659

300. Jung M, Philpott M, Müller S, et al. Affinity map of bromodomain protein 4 (BRD4) interactions with the histone H4 tail and the small molecule inhibitor JQ1. *J Biol Chem*. 2014. doi:10.1074/jbc.M113.523019

301. Lin Y, Umehara T, Inoue M, et al. Solution structure of the extraterminal domain of the bromodomain-containing protein BRD4. *Protein Sci*. 2008;17(12):2174-2179. doi:10.1110/ps.037580.108

302. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

303. Kanno T, Kanno Y, Siegel RM, Jang MK, Lenardo MJ, Ozato K. Selective Recognition of Acetylated Histones by Bromodomain Proteins Visualized in Living Cells. *Mol Cell*. 2004;13(1):33-43. doi:https://doi.org/10.1016/S1097-2765(03)00482-9

304. Huang B, Yang X-D, Zhou M-M, Ozato K, Chen L-F. Brd4 Coactivates Transcriptional Activation of NF-κB via Specific Binding to Acetylated RelA. *Mol Cell Biol*. 2009;29(5):1375-1387. doi:10.1128/mcb.01365-08

305. Roe JS, Mercan F, Rivera K, Pappin DJ, Vakoc CR. BET Bromodomain Inhibition Suppresses the Function of Hematopoietic Transcription Factors in Acute Myeloid Leukemia. *Mol Cell*. 2015;58(6):1028-1039. doi:10.1016/j.molcel.2015.04.011

306. Shi J, Wang Y, Zeng L, et al. Disrupting the Interaction of BRD4 with Diacetylated Twist Suppresses Tumorigenesis in Basal-like Breast Cancer. *Cancer Cell*. 2014;25(2):210-225. doi:10.1016/j.ccr.2014.01.028

307. Lamonica JM, Deng W, Kadauke S, et al. Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proc Natl Acad Sci*. 2011;108(22):E159-E168. doi:10.1073/pnas.1102140108

308. LeRoy G, Rickards B, Flint SJ. The Double Bromodomain Proteins Brd2 and Brd3 Couple Histone Acetylation to Transcription. *Mol Cell*. 2008;30(1):51-60. doi:10.1016/j.molcel.2008.01.018

309. Morinière J, Rousseaux S, Steuerwald U, et al. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature*. 2009;461(7264):664-668. doi:10.1038/nature08397

310. Lambert J-P, Picaud S, Fujisawa T, et al. Interactome Rewiring Following Pharmacological Targeting of BET Bromodomains. *Mol Cell*. 2019;7(73):621-638.e17. doi:10.1016/J.MOLCEL.2018.11.006

311. Rahman S, Sowa ME, Ottinger M, et al. The Brd4 Extraterminal Domain Confers Transcription Activation Independent of pTEFb by Recruiting Multiple Proteins, Including NSD3. *Mol Cell Biol*. 2011;31(13):2641-2652. doi:10.1128/MCB.01341-10

312. Wai DCC, Szyszka TN, Campbell AE, et al. The BRD3 ET domain recognizes a short peptide motif through a mechanism that is conserved across chromatin remodelers and transcriptional regulators. *J Biol Chem*. 2018;293(19):7160-7175. doi:10.1074/jbc.RA117.000678

313. Linares-Saldana R, Kim W, Bolar NA, et al. BRD4 orchestrates genome folding to promote neural crest differentiation. *Nat Genet*. 2021;53(October):1480-1492. doi:10.1038/s41588-021-00934-8

314. Natoli G, Andrau J-C. Noncoding Transcription at Enhancers: General Principles and Functional Models. *Annu Rev Genet*. 2012;46(1):1-19. doi:10.1146/annurev-genet-110711-155459

315. Meng H, Bartholomew B. Emerging roles of transcriptional enhancers in chromatin looping and promoter-proximal pausing of RNA polymerase II. *J Biol Chem*. 2018;293(36):13786-13794. doi:10.1074/jbc.R117.813485

316. Hsu E, Zemke NR, Berk AJ. Promoter-specific changes in initiation, elongation, and homeostasis of histone H3 acetylation during CBP/p300 inhibition. *Elife*. 2021;10:1-21. doi:10.7554/eLife.63512

317. Donner AJ, Ebmeier CC, Taatjes DJ, Espinosa JM. CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nat Struct Mol Biol*. 2010;17(2):194-201. doi:10.1038/nsmb.1752

318. Bhagwat AS, Roe JS, Mok BYL, Hohmann AF, Shi J, Vakoc CR. BET Bromodomain Inhibition Releases the Mediator Complex from Select cis-Regulatory Elements. *Cell Rep*. 2016;15(3):519-530. doi:10.1016/j.celrep.2016.03.054

319. Lovén J, Hoke HA, Lin CY, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*. 2013;153(2):320-334. doi:10.1016/j.cell.2013.03.036

320. Di Micco R, Fontanals-Cirera B, Low V, et al. Control of embryonic stem cell identity by brd4-dependent transcriptional elongation of super-enhancer-associated pluripotency genes. *Cell Rep*. 2014;9(1):234-247. doi:10.1016/j.celrep.2014.08.055

321. Shu S, Lin CY, He HH, et al. Response and resistance to BET bromodomain inhibitors in triple-negative breast cancer. *Nature*. 2016;529(7586):413-417. doi:10.1038/nature16508

322. Sooraj D, Sun C, Doan A, et al. MED12 and BRD4 cooperate to sustain cancer growth upon loss of mediator kinase. *Mol Cell*. 2022;82(1):123-139.e7. doi:https://doi.org/10.1016/j.molcel.2021.11.015

323. Pelish HE, Liau BB, Nitulescu II, et al. Mediator kinase inhibition further activates super-enhancer-associated genes in AML. *Nature*. 2015;526(7572):273-276. doi:10.1038/nature14904

324. Lee JE, Park YK, Park S, et al. Brd4 binds to active enhancers to control cell identity gene induction in adipogenesis and myogenesis. *Nat Commun*. 2017;8(1):1-12. doi:10.1038/s41467-017-02403-5

325. Muhar M, Anja E, Neumann T, et al. SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science*. 2018;360(6390):800-805. doi:10.1126/science.aao2793

326. Nguyen VT, Kiss T, Michels AA, Bensaude O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*. 2001;414(6861):322-325. doi:10.1038/35104581

327. Michels AA, Fraldi A, Li Q, et al. Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J*. 2004;23(13):2608-2619. doi:https://doi.org/10.1038/sj.emboj.7600275

328. Dawson MA, Prinjha RK, Dittmann A, et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature*. 2011;478(7370):529-533. doi:10.1038/nature10509

329. Anand P, Brown JD, Lin CY, et al. BET bromodomains mediate transcriptional pause release in heart failure. *Cell*. 2013;154(3):569-582. doi:10.1016/j.cell.2013.07.013

330. Bartman CR, Hamagami N, Keller CA, et al. Transcriptional Burst Initiation and Polymerase Pause Release Are Key Control Points of Transcriptional Regulation. *Mol Cell*. 2019;73(3):519-532.e4. doi:10.1016/j.molcel.2018.11.004

331. Nabet B, Roberts JM, Buckley DL, et al. The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol*. 2018;14(5):431-441. doi:10.1038/s41589-018-0021-8

332. Krueger BJ, Varzavand K, Cooper JJ, Price DH. The Mechanism of Release of P-TEFb and HEXIM1 from the 7SK snRNP by Viral and Cellular Activators Includes a Conformational Change in 7SK. *PLoS One*. 2010;5(8):e12335. https://doi.org/10.1371/journal.pone.0012335.

333. Lee S, Liu H, Hill R, et al. JMJD6 cleaves MePCE to release positive transcription elongation factor b (P-TEFb) in higher eukaryotes. *Elife*. 2020;9:e53930. doi:10.7554/eLife.53930

334. Liu W, Ma Q, Wong K, et al. Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell*. 2013;155(7):1581-1595. doi:10.1016/j.cell.2013.10.056

335. Devaiah BN, Lewis BA, Cherman N, et al. BRD4 is an atypical kinase that phosphorylates Serine2 of the RNA Polymerase II carboxy-terminal domain. *Proc Natl Acad Sci*. 2012;109(18):6927-6932. doi:10.1073/pnas.1120422109

336. Devaiah BN, Mu J, Akman B, et al. MYC protein stability is negatively regulated by BRD4. *Proc Natl Acad Sci U S A*. 2020;117(24):13457-13467. doi:10.1073/pnas.1919507117

337. Koch F, Fenouil R, Gut M, et al. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol*.

2011;18(8):956-963. doi:10.1038/nsmb.2085

338. Kulaeva OI, Nizovtseva E V., Polikanov YS, Ulianov S V., Studitsky VM. Distant Activation of Transcription: Mechanisms of Enhancer Action. *Mol Cell Biol*. 2012;32(24):4892-4897. doi:10.1128/MCB.01127-12

339. Zippo A, Serafini R, Rocchigiani M, Pennacchini S, Krepelova A, Oliviero S. Histone Crosstalk between H3S10ph and H4K16ac Generates a Histone Code that Mediates Transcription Elongation. *Cell*. 2009;138(6):1122-1136. doi:10.1016/j.cell.2009.07.031

340. Lee K, Hsiung CCS, Huang P, Raj A, Blobel GA. Dynamic enhancer–gene body contacts during transcription elongation. *Genes Dev*. 2015;29(19):1992-1997. doi:10.1101/gad.255265.114

341. Crump NT, Ballabio E, Godfrey L, et al. BET inhibition disrupts transcription but retains enhancer-promoter contact. *Nat Commun*. 2021;12(1):223. doi:10.1038/s41467-020-20400-z

342. Rahnamoun H, Lee J, Sun Z, et al. RNAs interact with BRD4 to promote enhanced chromatin engagement and transcription activation. *Nat Struct Mol Biol*. 2018;25(8):687-697. doi:10.1038/s41594-018-0102-0

343. Wu SY, Lee AY, Lai HT, Zhang H, Chiang CM. Phospho switch triggers brd4 chromatin binding and activator recruitment for gene-specific targeting. *Mol Cell*. 2013;49(5):843-857. doi:10.1016/j.molcel.2012.12.006

344. Nagarajan S, Hossan T, Alawi M, et al. Bromodomain Protein BRD4 Is Required for Estrogen Receptor-Dependent Enhancer Activation and Gene Transcription. *Cell Rep*. 2014;8(2):460-469. doi:10.1016/j.celrep.2014.06.016

345. Brown JD, Feldman ZB, Doherty SP, et al. BET bromodomain proteins regulate enhancer function during adipogenesisQiong,. *Proc Natl Acad Sci*. 2018;115(9):2144-2149. doi:10.1073/pnas.1711155115

346. Hnisz D, Abraham BJ, Lee TI, et al. Super-enhancers in the control of cell identity and disease. *Cell*. 2013;155(4):934-947. doi:10.1016/j.cell.2013.09.053

347. Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013;153(2):307-319. doi:10.1016/j.cell.2013.03.035

348. Olley G, Ansari M, Bengani H, et al. BRD4 interacts with NIPBL and BRD4 is mutated in a Cornelia de Lange-like syndrome. *Nat Genet*. 2018;50(3):329-332. doi:10.1038/s41588-018-0042-y

349. Sartorelli V, Lauberth SM. Enhancer RNAs are an important regulatory layer of the epigenome. *Nat Struct Mol Biol*. 2020;27(6):521-528. doi:10.1038/s41594-020-0446-0

350. Preker P, Almvig K, Christensen MS, et al. PROMoter uPstream Transcripts share characteristics with mRNAs and are produced upstream of all three major types of mammalian promoters. *Nucleic Acids Res*. 2011;39(16):7179-7193. doi:10.1093/nar/gkr370

351. Pombo A, Jackson DA, Hollinshead M, Wang Z, Roeder RG, Cook PR. Regional specialization in human nuclei: Visualization of discrete sites of transcription by RNA polymerase III. *EMBO J*. 1999;18(8):2241-2253. doi:10.1093/emboj/18.8.2241

352. Moss T, Stefanovsky VY. At the Center of Eukaryotic Life. *Cell*. 2002;109(5):545-548. doi:https://doi.org/10.1016/S0092-8674(02)00761-4

353. Turowski TW, Tollervey D. Transcription by RNA polymerase III: insights into mechanism and regulation. *Biochem Soc Trans*. 2016;44(5):1367-1375. doi:10.1042/BST20160062

354. Kim TH, Barrera LO, Zheng M, et al. A high-resolution map of active promoters in the human genome. *Nature*. 2005;436(7052):876-880. doi:10.1038/nature03877

355. Gilchrist DA, Dos Santos G, Fargo DC, et al. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell*. 2010;143(4):540-551. doi:10.1016/j.cell.2010.10.004

356. Radonjic M, Andrau JC, Lijnzaad P, et al. Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon S. cerevisiae stationary phase exit. *Mol Cell*. 2005;18(2):171-183. doi:10.1016/j.molcel.2005.03.010

357. Tietjen JR, Zhang DW, Rodríguez-Molina JB, et al. Chemical-genomic dissection of the CTD code. *Nat Struct Mol Biol*. 2010;17(9):1154-1161. doi:10.1038/nsmb.1900

358. Weiner A, Lara-Astiaso D, Krupalnik V, et al. Co-ChIP enables genome-wide mapping of histone mark co-occurrence at single-molecule resolution. *Nat Biotechnol*. 2016;34(9):953-961. doi:10.1038/nbt.3652

359. Orlando DA, Chen MW, Brown VE, et al. Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep*. 2014;9(3):1163-1170. doi:10.1016/j.celrep.2014.10.018

360. Rhee HS, Pugh BF. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature*. 2012;483(7389):295-301. doi:10.1038/nature10799

361. He Q, Johnston J, Zeitlinger J. ChIP-nexus enables improved detection of in vivo transcription factor binding footprints. *Nat Biotechnol*. 2015;33(4):395-401. doi:10.1038/nbt.3121

362. Kwak H, Fuda NJ, Core LJ, Lis JT. Precise Maps of RNA Polymerase Reveal How Promoters Direct Initiation and Pausing. *Science*. 2013;339(6122):950-953. doi:10.1126/science.1229386

363. Kovelman R, Roeder RG. Sarkosyl defines three intermediate steps in transcription initiation by RNA polymerase III: application to stimulation of transcription by E1A. *Genes Dev* . 1990;4(4):646-658. doi:10.1101/gad.4.4.646

364. Zhang J, Cavallaro M, Hebenstreit D. Timing RNA polymerase pausing with TV-PRO-seq. *Cell Reports Methods*. 2021;1(6):100083. doi:https://doi.org/10.1016/j.crmeth.2021.100083

365. Tzerpos P, Daniel B, Nagy L. Global Run-on Sequencing (GRO-Seq). In: Borggrefe T, Giaimo BD, eds. *Enhancers and Promoters*. Humana, New York, NY.; 2021:25-39.

366. Bernecky C, Herzog F, Baumeister W, Plitzko JM, Cramer P. Structure of transcribing mammalian RNA polymerase II. *Nature*. 2016;529(7587):551-554. doi:10.1038/nature16482

367. Lindell T, Weinberg F, Morris PW, Roeder R, Rutter WJ. Specific Inhibition of

Nuclear RNA Polymerase II by α-Amanitin. *Science*. 1970;170(3956):447-449. doi:10.1126/science.170.3956.447

368. Arnold M, Bressin A, Jasnovidova O, Meierhofer D, Mayer A. A BRD4-mediated elongation control point primes transcribing RNA polymerase II for 3′-processing and termination. *Mol Cell*. 2021;81(17):3589-3603.e13. doi:10.1016/j.molcel.2021.06.026

369. Jasnovidova O, Arnold M, Mayer A. Illuminating Enhancer Transcription at Nucleotide Resolution with Native Elongating Transcript Sequencing (NET-Seq). In: Borggrefe T, Giaimo BD, eds. *Enhancers and Promoters*. Humana, New York, NY.; 2021.

370. Reimer KA, Mimoso CA, Adelman K, Neugebauer KM. Co-transcriptional splicing regulates 3′ end cleavage during mammalian erythropoiesis. *Mol Cell*. 2021;81(5):998-1012.e7. doi:10.1016/j.molcel.2020.12.018

371. Herzel L, Straube K, Neugebauer KM. Long-read sequencing of nascent RNA reveals coupling among RNA processing events. *Genome Res* . 2018;28(7):1008-1019. doi:10.1101/gr.232025.117

372. Hirabayashi S, Bhagat S, Matsuki Y, et al. NET-CAGE characterizes the dynamics and topology of human transcribed cis-regulatory elements. *Nat Genet*. 2019;51(9):1369-1379. doi:10.1038/s41588-019-0485-9

373. Wang J, Zhao Y, Zhou X, Hiebert SW, Liu Q, Shyr Y. Nascent RNA sequencing analysis provides insights into enhancer-mediated gene regulation. *BMC Genomics*. 2018;19(1):633. doi:10.1186/s12864-018-5016-z

374. Gressel S, Schwalb B, Decker TM, et al. CDK9-dependent RNA polymerase II pausing controls transcription initiation. *Elife*. 2017;6:1-24. doi:10.7554/eLife.29736

375. Mayer A, Churchman LS. Genome-wide profiling of RNA polymerase transcription at nucleotide resolution in human cells with native elongating transcript sequencing. *Nat Protoc*. 2016;11(4):813-833. doi:10.1038/nprot.2016.047

376. Winter GE, Buckley DL, Paulk J, et al. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science*. 2015;348(6241):1376-1381. doi:10.1126/science.aab1433

377. Krönke J, Udeshi ND, Narla A, et al. Lenalidomide Causes Selective Degradation of IKZF1 and IKZF3 in Multiple Myeloma Cells. *Science*. 2014;343(6168):301-306. doi:10.1126/science.1244851

378. Buckley DL, Van Molle I, Gareiss PC, et al. Targeting the von Hippel–Lindau E3 Ubiquitin Ligase Using Small Molecules To Disrupt the VHL/HIF-1α Interaction. *J Am Chem Soc*. 2012;134(10):4465-4468. doi:10.1021/ja209924v

379. Yesbolatova A, Natsume T, Hayashi K, Kanemaki MT. Generation of conditional auxin-inducible degron (AID) cells and tight control of degron-fused proteins using the degradation inhibitor auxinole. *Methods*. 2019;164-165:73-80. doi:https://doi.org/10.1016/j.ymeth.2019.04.010

380. Wada T, Asahi T, Sawamura N. Nuclear cereblon modulates transcriptional activity of Ikaros and regulates its downstream target, enkephalin, in human neuroblastoma cells. *Biochem Biophys Res Commun*. 2016;477(3):388-394. doi:10.1016/j.bbrc.2016.06.091

381. Duan DR, Humphrey JS, Chen DY, et al. Characterization of the VHL tumor

suppressor gene product: localization, complex formation, and the effect of natural inactivating mutations. *Proc Natl Acad Sci*. 1995;92(14):6459-6463. doi:10.1073/pnas.92.14.6459

382. Samarasinghe KTG, Crews CM. Targeted protein degradation: A promise for undruggable proteins. *Cell Chem Biol*. 2021;28(7):934-951. doi:https://doi.org/10.1016/j.chembiol.2021.04.011

383. Liu J, Chen H, Kaniskan HÜ, et al. TF-PROTACs Enable Targeted Degradation of Transcription Factors. *J Am Chem Soc*. 2021;143(23):8902-8910. doi:10.1021/jacs.1c03852

384. Yesbolatova A, Saito Y, Kitamoto N, et al. The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. *Nat Commun*. 2020;11(1). doi:10.1038/s41467-020-19532-z

385. Buckley DL, Raina K, Darricarrere N, et al. HaloPROTACS: Use of Small Molecule PROTACs to Induce Degradation of HaloTag Fusion Proteins. *ACS Chem Biol*. 2015;10(8):1831-1837. doi:10.1021/acschembio.5b00442

386. Clackson T, Yang W, Rozamus LW, et al. Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. *Proc Natl Acad Sci U S A*. 1998;95(18):10437-10442. doi:10.1073/pnas.95.18.10437

387. Poser I, Sarov M, Hutchins JR a, et al. BAC TrangeneOmics: a high-troughput method for exploration of protein function in mammals. *Nat Methods*. 2008;5(5):409-415. doi:10.1038/nmeth.1199.BAC

388. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842. doi:10.1093/bioinformatics/btq033

389. Smedley D, Haider S, Ballester B, et al. BioMart – biological queries made easy. *BMC Genomics*. 2009;10(1):22. doi:10.1186/1471-2164-10-22

390. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357-359. doi:10.1038/nmeth.1923

391. Ghandi M, Huang FW, Jané-Valbuena J, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature*. 2019;569(7757):503-508. doi:10.1038/s41586-019-1186-3

392. Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011;7(1):539. doi:https://doi.org/10.1038/msb.2011.75

393. Concordet J-P, Haeussler M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res*. 2018;46(W1):W242-W245. doi:10.1093/nar/gky354

394. Martin M. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. *EMBnet.journal*. 2011;17(1).

395. Ramírez F, Ryan DP, Grüning B, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res*. 2016;44(W1):W160-W165. doi:10.1093/nar/gkw257

396. Afgan E, Baker D, van den Beek M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res*. 2016;44(W1):W3-W10. doi:10.1093/nar/gkw343

397. Stansfield JC, Cresswell KG, Vladimirov VI, Dozmorov MG. HiCcompare: an R-package for joint normalization and comparison of HI-C datasets. *BMC Bioinformatics*. 2018;19(1):279. doi:10.1186/s12859-018-2288-x

398. Tweedie S, Braschi B, Gray K, et al. Genenames.org: the HGNC and VGNC resources in 2021. *Nucleic Acids Res*. 2021;49(D1):D939-D946. doi:10.1093/nar/gkaa980

399. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-26. doi:10.1038/nbt.1754

400. Zhang Y, Liu T, Meyer CA, et al. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137. doi:10.1186/gb-2008-9-9-r137

401. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008;26(12):1367-1372. doi:10.1038/nbt.1511

402. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13(1):134. doi:10.1186/1471-2105-13-134

403. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res*. 2019;47(D1):D419-D426. doi:10.1093/nar/gky1038

404. Tyanova S, Temu T, Sinitcyn P, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods*. 2016;13(9):731-740. doi:10.1038/nmeth.3901

405. Wang R, Nambiar R, Zheng D, Tian B. PolyA_DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Res*. 2018;46(D1):D315-D319. doi:10.1093/nar/gkx1000

406. Untergasser A, Cutcutache I, Koressaar T, et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res*. 2012;40(15):e115-e115. doi:10.1093/nar/gks596

407. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607-D613. doi:10.1093/nar/gky1131

408. Lidschreiber K, Jung LA, von der Emde H, et al. Transcriptionally active enhancers in human cancer cells. *Mol Syst Biol*. 2021;17(1):e9873. doi:https://doi.org/10.15252/msb.20209873

409. Sakuma T, Nakade S, Sakane Y, Suzuki K-IT, Yamamoto T. MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat Protoc*. 2016;11(1):118-133. doi:10.1038/nprot.2015.140

410. Zorita E, Cuscó P, Filion GJ. Starcode: sequence clustering based on all-pairs search. *Bioinformatics*. 2015;31(12):1913-1919. doi:10.1093/bioinformatics/btv053

411. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21. doi:10.1093/bioinformatics/bts635

412. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8

413. Baluapuri A, Hofstetter J, Dudvarski Stankovic N, et al. MYC Recruits SPT5 to RNA Polymerase II to Promote Processive Transcription Elongation. *Mol Cell*. 2019;74(4):674-687.e11. doi:10.1016/j.molcel.2019.02.031

414. Corces MR, Trevino AE, Hamilton EG, et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods*. 2017;14(10):959-962. doi:10.1038/nmeth.4396

415. Mumbach MR, Rubin AJ, Flynn RA, et al. HiChIP: Efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods*. 2016;13(11):919-922. doi:10.1038/nmeth.3999

416. Servant N, Varoquaux N, Lajoie BR, et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 2015;16(1):259. doi:10.1186/s13059-015-0831-x

417. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods*. 2012;9(3):215-216. doi:10.1038/nmeth.1906

418. Mohammed H, Taylor C, Brown GD, Papachristou EK, Carroll JS, D'Santos CS. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat Protoc*. 2016;11(2):316-326. doi:10.1038/nprot.2016.020

419. Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat Protoc*. 2007;1(6):2650-2660. doi:10.1038/nprot.2006.427

420. Riching KM, Mahan S, Corona CR, et al. Quantitative Live-Cell Kinetic Degradation and Mechanistic Profiling of PROTAC Mode of Action. *ACS Chem Biol*. 2018;13(9):2758-2770. doi:10.1021/acschembio.8b00692

421. Schick S, Grosche S, Kohl KE, et al. Acute BAF perturbation causes immediate changes in chromatin accessibility. *Nat Genet*. 2021;53(3):269-278. doi:10.1038/s41588-021-00777-3

422. Porrua O, Libri D. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol*. 2015;16(3):190-202. doi:10.1038/nrm3943

423. Kustatscher G, Hégarat N, Wills KLH, et al. Proteomics of a fuzzy organelle: Interphase chromatin. *EMBO J*. 2014;33(6):648-664. doi:10.1002/embj.201387614

424. Thul PJ, Åkesson L, Wiking M, et al. A subcellular map of the human proteome. *Science*. 2017;356(6340):eaal3321. doi:10.1126/science.aal3321

425. Lykke-Andersen S, Žumer K, Molska EŠ, et al. Integrator is a genome-wide attenuator of non-productive transcription. *Mol Cell*. 2021;81(3):514-529.e6. doi:10.1016/j.molcel.2020.12.014

426. Elrod ND, Henriques T, Huang KL, et al. The Integrator Complex Attenuates Promoter-Proximal Transcription at Protein-Coding Genes. *Mol Cell*. 2019;76(5):738-752.e7. doi:10.1016/j.molcel.2019.10.034

427. Vervoort SJ, Welsh SA, Devlin JR, et al. The PP2A-Integrator-CDK9 axis fine-tunes transcription and can be targeted therapeutically in cancer. *Cell*. 2021;184(12):3143-3162.e32. doi:10.1016/j.cell.2021.04.022

428. Huang K-L, Jee D, Stein CB, et al. Integrator Recruits Protein Phosphatase 2A to Prevent Pause Release and Facilitate Transcription Termination. *Mol Cell*.

2020;80(2):345-358.e9. doi:https://doi.org/10.1016/j.molcel.2020.08.016

429. Ni Z, Olsen JB, Guo X, et al. Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B. *Transcription*. 2011;2(5):201-206. doi:10.4161/trns.2.5.17803

430. Ali I, Ruiz DG, Ni Z, et al. Crosstalk between RNA Pol II C-Terminal Domain Acetylation and Phosphorylation via RPRD Proteins. *Mol Cell*. 2019;74(6):1164-1174.e4. doi:10.1016/j.molcel.2019.04.008

431. Yu G, Wang L-G, He Q-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*. 2015;31(14):2382-2383. doi:10.1093/bioinformatics/btv145

432. O'Reilly FJ, Rappsilber J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nat Struct Mol Biol*. 2018;25(11):1. doi:10.1038/s41594-018-0147-0

433. Fields S, Song OK. A novel genetic system to detect protein-protein interactions. *Nature*. 1989;340(6230):245-246. doi:10.1038/340245a0

434. Mohammed H, D'Santos C, Serandour AA, et al. Endogenous Purification Reveals GREB1 as a Key Estrogen Receptor Regulatory Factor. *Cell Rep*. 2013;3(2):342-349. doi:10.1016/j.celrep.2013.01.010

435. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol Cell Proteomics*. 2014;13(9):2513-2526. doi:10.1074/mcp.M113.031591

436. Greenleaf AL. Positive patches and negative noodles: linking RNA processing to transcription? *Trends Biochem Sci*. 1993;18(4):117-119. doi:10.1016/0968-0004(93)90016-G

437. Descostes N, Heidemann M, Spinelli L, et al. Tyrosine phosphorylation of RNA Polymerase II CTD is associated with antisense promoter transcription and active enhancers in mammalian cells. *Elife*. 2014;2014(3):1-19. doi:10.7554/eLife.02105

438. So BR, Di C, Cai Z, et al. A Complex of U1 snRNP with Cleavage and Polyadenylation Factors Controls Telescripting, Regulating mRNA Transcription in Human Cells. *Mol Cell*. 2019;76(4):1-10. doi:10.1016/j.molcel.2019.08.007

439. Cretu C, Schmitzová J, Ponce-Salvatierra A, et al. Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *Mol Cell*. 2016;64(2):307-319. doi:10.1016/j.molcel.2016.08.036

440. Chanarat S, Sträßer K. Splicing and beyond: The many faces of the Prp19 complex. *Biochim Biophys Acta - Mol Cell Res*. 2013;1833(10):2126-2134. doi:https://doi.org/10.1016/j.bbamcr.2013.05.023

441. Du P, Wang L, Sliz P, Gregory RI. A Biogenesis Step Upstream of Microprocessor Controls miR-17∼92 Expression. *Cell*. 2015;162(4):885-899. doi:https://doi.org/10.1016/j.cell.2015.07.008

442. Schütze T, Ulrich AKC, Apelt L, et al. Multiple protein–protein interactions converging on the Prp38 protein during activation of the human spliceosome. *RNA* . 2016;22(2):265-277. doi:10.1261/rna.054296.115

443. Busetto V, Barbosa I, Basquin J, et al. Structural and functional insights into

CWC27/CWC22 heterodimer linking the exon junction complex to spliceosomes. *Nucleic Acids Res*. 2021;48(10):5670-5683. doi:10.1093/NAR/GKAA267

444. Close P, East P, Dirac-Svejstrup AB, et al. DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation. *Nature*. 2012;484(7394):386-389. doi:10.1038/nature10925

445. Wahl MC, Will CL, Lührmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell*. 2009;136(4):701-718. doi:10.1016/j.cell.2009.02.009

446. David CJ, Boyne AR, Millhouse SR, Manley JL. The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev*. 2011;25(9):972-982. doi:10.1101/gad.2038011

447. De Maio A, Yalamanchili HK, Adamski CJ, et al. RBM17 Interacts with U2SURP and CHERP to Regulate Expression and Splicing of RNA-Processing Proteins. *Cell Rep*. 2018;25(3):726-736.e7. doi:https://doi.org/10.1016/j.celrep.2018.09.041

448. Kuraoka I, Ito S, Wada T, et al. Isolation of XAB2 Complex Involved in Pre-mRNA Splicing, Transcription, and Transcription-coupled Repair. *J Biol Chem*. 2008;283(2):940-950. doi:10.1074/jbc.M706647200

449. Bertram K, Agafonov DE, Liu WT, et al. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature*. 2017;542(7641):318-323. doi:10.1038/nature21079

450. Chung S, Zhou Z, Huddleston KA, et al. Crooked neck is a component of the human spliceosome and implicated in the splicing process. *Biochim Biophys Acta - Gene Struct Expr*. 2002;1576(3):287-297. doi:https://doi.org/10.1016/S0167-4781(02)00368-8

451. Liu J, Yue Y, Han D, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol*. 2014;10(2):93-95. doi:10.1038/nchembio.1432

452. Bertram K, Agafonov DE, Dybkov O, et al. Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. *Cell*. 2017;170(4):701-713.e11. doi:10.1016/j.cell.2017.07.011

453. Henriques T, Scruggs BS, Inouye MO, et al. Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev*. 2018;32(1):26-41. doi:10.1101/gad.309351.117

454. Schaukowitch K, Joo JY, Liu X, Watts JK, Martinez C, Kim TK. Enhancer RNA facilitates NELF release from immediate early genes. *Mol Cell*. 2014;56(1):29-42. doi:10.1016/j.molcel.2014.08.023

455. Li B, Carey M, Workman JL. The Role of Chromatin during Transcription. *Cell*. 2007;128(4):707-719. doi:10.1016/j.cell.2007.01.015

456. Clark DJ, Felsenfeld G. A nucleosome core is transferred out of the path of a transcribing polymerase. *Cell*. 1992;71(1):11-22. doi:10.1016/0092-8674(92)90262-B

457. Arinobu Y, Mizuno S, Chong Y, et al. Reciprocal Activation of GATA-1 and PU.1 Marks Initial Specification of Hematopoietic Stem Cells into Myeloerythroid and Myelolymphoid Lineages. *Cell Stem Cell*. 2007;1(4):416-427. doi:https://doi.org/10.1016/j.stem.2007.07.004

458. Wheat JC, Sella Y, Willcockson M, et al. Single-molecule imaging of transcription dynamics in somatic stem cells. *Nature*. 2020;583(7816):431-436. doi:10.1038/s41586-020-2432-4

459. Sanda T, Lawton LN, Barrasa MI, et al. Core Transcriptional Regulatory Circuit Controlled by the TAL1 Complex in Human T Cell Acute Lymphoblastic Leukemia. *Cancer Cell*. 2012;22(2):209-221. doi:https://doi.org/10.1016/j.ccr.2012.06.007

460. Wierer M, Mann M. Proteomics to study DNA-bound and chromatin-associated gene regulatory complexes. *Hum Mol Genet*. 2016;25(R2):R106-R114. doi:10.1093/hmg/ddw208

461. Sutherland BW, Toews J, Kast J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom*. 2008;43(6):699-715. doi:10.1002/jms.1415

462. Cao Q-F, Yamamoto J, Isobe T, et al. Characterization of the Human Transcription Elongation Factor Rtf1: Evidence for Nonoverlapping Functions of Rtf1 and the Paf1 Complex. *Mol Cell Biol*. 2015;35(20):3459-3470. doi:10.1128/mcb.00601-15

463. Hintermair C, Heidemann M, Koch F, et al. Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *EMBO J*. 2012;31(12):2784-2797. doi:https://doi.org/10.1038/emboj.2012.123

464. Baskaran R, Dahmus ME, Wang JY. Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proc Natl Acad Sci*. 1993;90(23):11167-11171. doi:10.1073/pnas.90.23.11167

465. Weissman JD, Singh AK, Devaiah BN, Schuck P, LaRue RC, Singer DS. The intrinsic kinase activity of BRD4 spans its BD2-B-BID domains. *J Biol Chem*. 2021;297(5):101326. doi:10.1016/j.jbc.2021.101326

466. Ehrensberger AH, Kelly GP, Svejstrup JQ. Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. *Cell*. 2013;154(4):713-715. doi:10.1016/j.cell.2013.07.032

467. Kamieniarz-Gdula K, Proudfoot NJ. Transcriptional Control by Premature Termination: A Forgotten Mechanism. *Trends Genet*. 2019;35(8):553-564. doi:10.1016/j.tig.2019.05.005

468. Oh J-M, Di C, Venters CC, et al. U1 snRNP telescripting regulates a size–function-stratified human genome. *Nat Struct Mol Biol*. 2017;24(11):993-999. doi:10.1038/nsmb.3473

469. Martin G, Gruber AR, Keller W, Zavolan M. Genome-wide Analysis of Pre-mRNA 3' End Processing Reveals a Decisive Role of Human Cleavage Factor I in the Regulation of 3' UTR Length. *Cell Rep*. 2012;1(6):753-763. doi:10.1016/j.celrep.2012.05.003

470. Li W, You B, Hoque M, et al. Systematic Profiling of Poly(A)+ Transcripts Modulated by Core 3' End Processing and Splicing Factors Reveals Regulatory Rules of Alternative Cleavage and Polyadenylation. *PLoS Genet*. 2015;11(4):1-28. doi:10.1371/journal.pgen.1005166

471. Brannan K, Kim H, Erickson B, et al. MRNA Decapping Factors and the Exonuclease Xrn2 Function in Widespread Premature Termination of RNA Polymerase II Transcription. *Mol Cell*. 2012;46(3):311-324.

doi:10.1016/j.molcel.2012.03.006

472. Davidson L, Kerr A, West S. Co-transcriptional degradation of aberrant pre-mRNA by Xrn2. *EMBO J*. 2012;31(11):2566-2578. doi:https://doi.org/10.1038/emboj.2012.101

473. Cheung KL, Zhang F, Jaganathan A, et al. Distinct Roles of Brd2 and Brd4 in Potentiating the Transcriptional Program for Th17 Cell Differentiation. *Mol Cell*. 2017;65(6):1068-1080.e5. doi:10.1016/j.molcel.2016.12.022

474. Durant M, Pugh BF. NuA4-Directed Chromatin Transactions throughout the Saccharomyces cerevisiae Genome . *Mol Cell Biol*. 2007;27(15):5327-5335. doi:10.1128/mcb.00468-07

475. Delmore JE, Issa GC, Lemieux ME, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell*. 2011;146(6):904-917. doi:10.1016/j.cell.2011.08.017

476. Zuber J, Shi J, Wang E, et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature*. 2011;478(7370):524-528. doi:10.1038/nature10334

477. Corradini F, Cesi V, Bartella V, et al. Enhanced proliferative potential of hematopoietic cells expressing degradation-resistant c-Myb mutants. *J Biol Chem*. 2005;280(34):30254-30262. doi:10.1074/jbc.M504703200

478. Lemma RB, Ledsaak M, Fuglerud BM, Sandve GK, Eskeland R, Gabrielsen OS. Chromatin occupancy and target genes of the haematopoietic master transcription factor MYB. *Sci Rep*. 2021;11(1):9008. doi:10.1038/s41598-021-88516-w

479. Jaenicke LA, von Eyss B, Carstensen A, et al. Ubiquitin-Dependent Turnover of MYC Antagonizes MYC/PAF1C Complex Accumulation to Drive Transcriptional Elongation. *Mol Cell*. 2016;61(1):54-67. doi:10.1016/j.molcel.2015.11.007

480. Mapendano CK, Lykke-Andersen S, Kjems J, Bertrand E, Jensen TH. Crosstalk between mRNA 3′ End Processing and Transcription Initiation. *Mol Cell*. 2010;40(3):410-422. doi:10.1016/j.molcel.2010.10.012

481. Komarnitsky P, Cho EJ, Buratowski S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*. 2000;14(19):2452-2460. doi:10.1101/gad.824700

482. Wang Y, Fairley JA, Roberts SGE. Phosphorylation of TFIIB Links Transcription Initiation and Termination. *Curr Biol*. 2010;20(6):548-553. doi:10.1016/j.cub.2010.01.052

483. Sansó M, Levin RS, Lipp JJ, et al. P-TEFb regulation of transcription termination factor Xrn2 revealed by a chemical genetic screen for Cdk9 substrates. *Genes Dev*. 2016;30(1):117-131. doi:10.1101/gad.269589.115

484. Guo YE, Manteiga JC, Henninger JE, et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature*. 2019. doi:10.1038/s41586-019-1464-0

485. Mahdessian D, Cesnik AJ, Gnann C, et al. Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature*. 2021;590(7847):649-654. doi:10.1038/s41586-021-03232-9

486. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene

expression control. *Nature*. 2011;473(7347):337-342. doi:10.1038/nature10098

487. Matangkasombut O, Buratowski RM, Swilling NW, Buratowski S. Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. *Genes Dev* . 2000;14(8):951-962. doi:10.1101/gad.14.8.951

488. Zhang S, Aibara S, Vos SM, Agafonov DE, Lührmann R, Cramer P. Structure of a transcribing RNA polymerase II-U1 snRNP complex. *Science*. 2021;371(6526):305-309. doi:10.1126/science.abf1870

489. Caizzi L, Monteiro-Martins S, Schwalb B, et al. Efficient RNA polymerase II pause release requires U2 snRNP function. *Mol Cell*. 2021;81(9):1920-1934.e9. doi:10.1016/j.molcel.2021.02.016

490. Uppal S, Gegonne A, Chen Q, et al. The Bromodomain Protein 4 Contributes to the Regulation of Alternative Splicing. *Cell Rep*. 2019;29(8):2450-2460.e5. doi:10.1016/j.celrep.2019.10.066

491. Wu L, Li L, Zhou B, Qin Z, Dou Y. H2B Ubiquitylation Promotes RNA Pol II Processivity via PAF1 and pTEFb. *Mol Cell*. 2014;54(6):920-931. doi:https://doi.org/10.1016/j.molcel.2014.04.013

492. Fuchs G, Hollander D, Voichek Y, Ast G, Oren M. Cotranscriptional histone H2B monoubiquitylation is tightly coupled with RNA polymerase II elongation rate. *Genome Res*. 2014;24(10):1572-1583. doi:10.1101/gr.176487.114

493. Vaid R, Wen J, Mannervik M. Release of promoter-proximal paused Pol II in response to histone deacetylase inhibition. *Nucleic Acids Res*. 2020;48(9):4877-4890. doi:10.1093/nar/gkaa234

494. Andersson R. Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model. *BioEssays*. 2015;37(3):314-323. doi:10.1002/bies.201400162

495. Andersson R, Sandelin A, Danko CG. A unified architecture of transcriptional regulatory elements. *Trends Genet*. 2015;31(8):426-433. doi:10.1016/j.tig.2015.05.007

496. Kim TK, Shiekhattar R. Architectural and Functional Commonalities between Enhancers and Promoters. *Cell*. 2015;162(5):948-959. doi:10.1016/j.cell.2015.08.008

497. Arner E, Daub CO, Vitting-Seerup K, et al. Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. *Science (80-)*. 2015;347(6225):1010-1014. doi:10.1126/science.1259418

498. El Khattabi L, Zhao H, Kalchschmidt J, et al. A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. *Cell*. 2019:1-14. doi:10.1016/j.cell.2019.07.011

499. Kim TK, Hemberg M, Gray JM, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010;465(7295):182-187. doi:10.1038/nature09033

500. Gregersen LH, Mitter R, Svejstrup JQ. Using TTchem-seq for profiling nascent transcription and measuring transcript elongation. *Nat Protoc*. 2020;15(2):604-627. doi:10.1038/s41596-019-0262-3

501. Kafri R, Springer M, Pilpel Y. Genetic Redundancy: New Tricks for Old Genes. *Cell*. 2009;136(3):389-392. doi:10.1016/j.cell.2009.01.027

502. Nowak MA, Boerlijst MC, Cooke J, Smith JM. Evolution of genetic redundancy. *Nature*. 1997;388(6638):167-171. doi:10.1038/40618

8 PUBLICATIONS

Arnold M*, Bressin A*, Jasnovidova O, Meierhofer D, Mayer A. A BRD4-mediated elongation control point primes transcribing RNA polymerase II for 3′-processing and termination. *Mol Cell.* 2021.81(17):3589-3603.e13.

Jasnovidova O*, **Arnold M***, Mayer A. Illuminating Enhancer Transcription at Nucleotide Resolution with Native Elongating Transcript Sequencing (NET-Seq). In: Borggrefe T, Giaimo BD, eds. *Enhancers and Promoters. Methods Mol Biol.* Humana, New York, NY.; 2021.

(*: equal contribution)

Eischer N., **Arnold M**, Mayer A. Emerging roles of BET proteins in transcription and co-transcriptional RNA processing. *Wiley Interdiscip Rev RNA*. 2022. (accepted for publication).

9 CONTRIBUTIONS

Annkatrin Bressin processed and all sequencing data, except ATAC-seq and performed analyses. Also, Figure 21, Figure **23** and Figure **43** and Supplemental Figure 2 were prepared by her.
Mario Rubio performed additional analyses on the RPB1 phosho-CTD ChIP-Rx data.
Andreas Mayer developed SI-NET-seq.
Olga Jasnovidova optimized the DNA barcode design and helped optimizing SI-NET-seq.
Thomas Kratz performed the IP-MS experiments for BRD2, BRD3 and BRD4 under my supervision.
All mass spectrometric analyses and processing of the raw data were performed by the mass spectrometry core facility of the Max Planck Institute for Molecular Genetics.
All sequencing was done at the sequencing core facility of Max Planck Institute for Molecular Genetics.

10 APPENDIX

10.1 ABBREVIATIONS

AML	acute myeloid leukemia
ATD	average termination distance
BD	bromodomain
BET	bromodomain and extra-terminal domain
blastiR	blasticidin resistance gene
BRD4	bromodomain-containing protein 4
cDNA	complementary DNA
CFIm	mammalian cleavage factor I
circDNA	circularized DNA
CPSF	cleavage and polyadenylation specificity factor
CstF	cleavage stimulation factor
dNTP	deoxy-nucleoside triphosphate
DRB	5,6-Dichlorobenzimidazole 1-β-D-ribofuranoside
dTAG	degradation tag
ET	extra-terminal
eRNA	enhancer RNA
FC	fold-change
FDR	false discovery rate
GO	gene ontology
GTF	general transcription factor
HA	hemagglutinin
IDR	instrinsically disordered region
LFQ	label-free quantification
mRNA	messenger RNA
lincRNA	long intergenic non-conding RNA
lncRNA	long non-coding RNA
MMEJ	microhomology-mediated end joining
MS	mass spectrometry
NDR	nucleosome-depleted region
pA	polyadenylation site
PAF	RNA polymerase II-associated factor
PAS	polyadenylation signal

PCR	polymerase chain reaction
PIC	pre-initiation complex
Pol II	RNA polymerase II
puroR	puromycin resistance gene
PROTAC	proteolysis-targeting chimera
RPKM	reads per kilobase million
rRNA	ribosomal RNA
RTI	readthrough index
RT-qPCR	reverse transcription-based quantitative polymerase chain reaction
SE	super-enhancer
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
TF	transcription factor
TSS	transcription start site
UMI	unique molecular identifier

10.2 LIST OF FIGURES

10.3 DATA SETS GENERATED IN THIS PROJECT

Table 24: Sequencing-based data.

data type	specification	internal ID	GEO ID (if available)
SI-NET-seq	dTAG7, 120 min, HEK293T	MRA01	
SI-NET-seq	DMSO, 120 min, HEK293T	MRA02	
SI-NET-seq	dTAG7, 120 min, HEK293T	MRA03	
SI-NET-seq	DMSO, 120 min, HEK293T	MRA04	
SI-NET-seq	DMSO, 120 min, K562 wt	MRA05	
SI-NET-seq	dBET6, 120 min, K562 wt	MRA06	
SI-NET-seq	DMSO, 120 min, K562 wt	MRA07	
SI-NET-seq	dBET6, 120 min, K562 wt	MRA08	
SI-NET-seq	sea urchin	MRA09	
Hi-ChIP	dTAG7, 120 min, K562 dTAG-BRD4, H3K27ac	MRA10	
Hi-ChIP	dTAG7, 120 min, K562 dTAG-BRD4, H3K27ac	MRA11	
Hi-ChIP	dTAG7, 120 min, K562 dTAG-BRD4, H3K27ac	MRA12	
Hi-ChIP	DMSO, 120 min, K562 dTAG-BRD4, H3K27ac	MRA13	
Hi-ChIP	DMSO, 120 min, K562 dTAG-BRD4, H3K27ac	MRA14	
Hi-ChIP	DMSO, 120 min, K562 dTAG-BRD4, H3K27ac	MRA15	
SI-NET-seq	dTAG7, 120 min, K562 dTAG-BRD4	MRA16	GSM4816670
SI-NET-seq	dBET6, 120 min, K562 dTAG-BRD4	MRA17	GSM4816668
SI-NET-seq	DMSO, 120 min, K562 dTAG-BRD4	MRA18	GSM4816665
SI-NET-seq	dTAG7, 120 min, K562 dTAG-BRD4	MRA19	
SI-NET-seq	dBET6, 120 min, K562 dTAG-BRD4	MRA20	GSM4816667
SI-NET-seq	DMSO, 120 min, K562 dTAG-BRD4	MRA21	GSM4816666
SI-NET-seq	dTAG7, 120 min, K562 dTAG-BRD4	MRA22	GSM4816669

SI-NET-seq	DMSO, 120 min, K562 dTAG-BRD4	MRA23	
ChIP-Rx	dBET6, 50 min, K562 dTAG-BRD4	MRA24	GSM4816678
ChIP-Rx	DMSO, 50 min, K562 dTAG-BRD4	MRA25	GSM4816672
ChIP-Rx	dBET6, 50 min, K562 dTAG-BRD4	MRA26	
ChIP-Rx	DMSO, 50 min, K562 dTAG-BRD4	MRA27	
ChIP-Rx	SPT5, dTAG7, 120 min, K562 dTAG-BRD4	MRA34	
ChIP-Rx	SPT5, DMSO, 120 min, K562 dTAG-BRD4	MRA35	
ChIP-Rx	FIP1, dTAG7, 120 min, K562 dTAG-BRD4	MRA36	GSM4816697
ChIP-Rx	FIP1, DMSO, 120 min, K562 dTAG-BRD4	MRA37	GSM4816695
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA38	GSM4816710
SI-NET-seq	dBET6, 50 min, K562 dTAG-BRD4	MRA39	
SI-NET-seq	DMSO, 50 min, K562 dTAG-BRD4	MRA40	GSM5233640
SI-NET-seq	dBET6, 50 min, K562 dTAG-BRD4	MRA41	GSM4816677
SI-NET-seq	DMSO, 50 min, K562 dTAG-BRD4	MRA42	GSM4816671
ChIP-Rx	PAF1, dTAG7, 120 min, K562 dTAG-BRD4	MRA43	
ChIP-Rx	PAF1, DMSO, 120 min, K562 dTAG-BRD4	MRA44	
ChIP-Rx	p-Ser2, dTAG7, 120 min, K562 dTAG-BRD4	MRA45	GSM5233640
ChIP-Rx	p-Ser2, DMSO, 120 min, K562 dTAG-BRD4	MRA46	GSM5233639
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA47	GSM5233641
ChIP-Rx	CstF77, dTAG7, 120 min, K562 dTAG-BRD4	MRA48	
ChIP-Rx	CstF77, DMSO, 120 min, K562 dTAG-BRD4	MRA49	
ChIP-Rx	RPB2, dTAG7, 120 min, K562 dTAG-BRD4	MRA50	GSM4816685
ChIP-Rx	RPB2, DMSO, 120 min, K562 dTAG-BRD4	MRA51	GSM4816683
ChIP-Rx	CstF64, dTAG7, 120 min, K562 dTAG-BRD4	MRA52	GSM4816705
ChIP-Rx	CstF64, DMSO, 120 min, K562 dTAG-BRD4	MRA53	GSM4816703
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA54	GSM4816712
ChIP-Rx	CPSF73, dTAG7, 120 min, K562 dTAG-BRD4	MRA55	GSM4816701
ChIP-Rx	CPSF73, DMSO, 120 min, K562 dTAG-BRD4	MRA56	GSM4816699
ChIP-Rx	SPT5, dTAG7, 120 min, K562 dTAG-BRD4	MRA57	
ChIP-Rx	SPT5, DMSO, 120 min, K562 dTAG-BRD4	MRA58	
SI-NET-seq	dBET6, 40 min, K562 dTAG-BRD4	MRA59	GSM4816676
SI-NET-seq	dBET6, 60 min, K562 dTAG-BRD4	MRA60	
SI-NET-seq	DMSO, 60 min, K562 dTAG-BRD4	MRA61	GSM4816673
SI-NET-seq	dBET6, 40 min, K562 dTAG-BRD4	MRA62	GSM4816675
SI-NET-seq	dBET6, 60 min, K562 dTAG-BRD4	MRA63	
SI-NET-seq	DMSO, 60 min, K562 dTAG-BRD4	MRA64	GSM4816674
ChIP-Rx	RPB2, dTAG7, 120 min, K562 dTAG-BRD4	MRA65	GSM4816686
ChIP-Rx	RPB2, DMSO, 120 min, K562 dTAG-BRD4	MRA66	GSM4816684
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA67	
ChIP-Rx	FIP1, dTAG7, 120 min, K562 dTAG-BRD4	MRA68	GSM4816698
ChIP-Rx	FIP1, DMSO, 120 min, K562 dTAG-BRD4	MRA69	GSM4816696
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA70	GSM4816711
ChIP-Rx	CstF64, dTAG7, 120 min, K562 dTAG-BRD4	MRA71	GSM4816706
ChIP-Rx	CstF64, DMSO, 120 min, K562 dTAG-BRD4	MRA72	GSM4816704
ChIP-Rx	PAF1, dTAG7, 120 min, K562 dTAG-BRD4	MRA73	GSM4816693
ChIP-Rx	PAF1, DMSO, 120 min, K562 dTAG-BRD4	MRA74	GSM4816691
ChIP-Rx	CPSF30, dTAG7, 120 min, K562 dTAG-BRD4	MRA75	
ChIP-Rx	CPSF30, DMSO, 120 min, K562 dTAG-BRD4	MRA76	
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA77	GSM4816709
ChIP-Rx	CPSF30, dTAG7, 120 min, K562 dTAG-BRD4	MRA78	
ChIP-Rx	CPSF30, DMSO, 120 min, K562 dTAG-BRD4	MRA79	
ChIP-Rx	CstF77, dTAG7, 120 min, K562 dTAG-BRD4	MRA80	
ChIP-Rx	CstF77, DMSO, 120 min, K562 dTAG-BRD4	MRA81	
ChIP-Rx	SPT5, dTAG7, 120 min, K562 dTAG-BRD4	MRA82	GSM4816689
ChIP-Rx	SPT5, DMSO, 120 min, K562 dTAG-BRD4	MRA83	GSM4816687
ChIP-Rx	input dTAG7, 120 min, K562 dTAG-BRD4	MRA84	GSM4816707
ChIP-Rx	CstF77, dTAG7, 120 min, K562 dTAG-BRD4	MRA85	
ChIP-Rx	CstF77, DMSO, 120 min, K562 dTAG-BRD4	MRA86	
ChIP-Rx	CPSF73, dTAG7, 120 min, K562 dTAG-BRD4	MRA87	GSM4816702
ChIP-Rx	CPSF73, DMSO, 120 min, K562 dTAG-BRD4	MRA88	GSM4816700
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA89	GSM4816713
ChIP-Rx	CstF77, dTAG7, 120 min, K562 dTAG-BRD4	MRA90	
ChIP-Rx	CstF77, DMSO, 120 min, K562 dTAG-BRD4	MRA91	

ChIP-Rx	SPT5, dTAG7, 120 min, K562 dTAG-BRD4	MRA92	GSM4816690
ChIP-Rx	SPT5, DMSO, 120 min, K562 dTAG-BRD4	MRA93	GSM4816688
ChIP-Rx	input dTAG7, 120 min, K562 dTAG-BRD4	MRA94	GSM4816708
ChIP-Rx	PAF1, dTAG7, 120 min, K562 dTAG-BRD4	MRA95	GSM4816694
ChIP-Rx	PAF1, DMSO, 120 min, K562 dTAG-BRD4	MRA96	GSM4816692
ATAC-seq	dTAG7, 120 min, K562 dTAG-BRD4	MRA97	
ATAC-seq	DMSO, 120 min, K562 dTAG-BRD4	MRA98	
ATAC-seq	dTAG7, 360 min, K562 dTAG-BRD4	MRA99	
ATAC-seq	DMSO, 360 min, K562 dTAG-BRD4	MRA100	
RNA-seq	dTAG7, 120 min, K562 dTAG-BRD4	MRA101	
RNA-seq	dTAG7, 120 min, K562 dTAG-BRD4	MRA102	
RNA-seq	DMSO, 120 min, K562 dTAG-BRD4	MRA103	
RNA-seq	DMSO, 120 min, K562 dTAG-BRD4	MRA104	
RNA-seq	dTAG7, 360 min, K562 dTAG-BRD4	MRA105	
RNA-seq	dTAG7, 360 min, K562 dTAG-BRD4	MRA106	
RNA-seq	DMSO, 360 min, K562 dTAG-BRD4	MRA107	
RNA-seq	DMSO, 360 min, K562 dTAG-BRD4	MRA108	
ChIP-Rx	BRD4, dTAG7, 120 min, K562 dTAG-BRD4	MRA109	
ChIP-Rx	BRD4, dTAG7, 40 min, K562 dTAG-BRD4	MRA110	
ChIP-Rx	BRD4, DMSO, 120 min, K562 dTAG-BRD4	MRA111	
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA112	
ChIP-Rx	BRD4, dTAG7, 120 min, K562 dTAG-BRD4	MRA113	
ChIP-Rx	BRD4, dTAG7, 40 min, K562 dTAG-BRD4	MRA114	
ChIP-Rx	BRD4, DMSO, 120 min, K562 dTAG-BRD4	MRA115	
ChIP-Rx	input dTAG7, 120 min, K562 dTAG-BRD4	MRA116	
ChIP-seq	PRPF8, K562 dTAG-BRD4	MRA118	
ChIP-seq	PRPF8, K562 dTAG-BRD4	MRA119	
HiS-NET-seq	dTAG7, 40 min, K562 dTAG-BRD4	MRA125	
HiS-NET-seq	DMSO, 40 min, K562 dTAG-BRD4	MRA126	
HiS-NET-seq	dTAG7, 40 min, K562 dTAG-BRD4	MRA127	
HiS-NET-seq	DMSO, 40 min, K562 dTAG-BRD4	MRA128	
ChIP-Rx	p-Tyr1, dTAG7, 120 min, K562 dTAG-BRD4	MRA129	
ChIP-Rx	p-Tyr1, DMSO, 120 min, K562 dTAG-BRD4	MRA130	
ChIP-Rx	p-Thr4, dTAG7, 120 min, K562 dTAG-BRD4	MRA131	
ChIP-Rx	p-Thr4, DMSO, 120 min, K562 dTAG-BRD4	MRA132	
ChIP-Rx	p-Tyr1, dTAG7, 120 min, K562 dTAG-BRD4	MRA133	
ChIP-Rx	p-Tyr1, DMSO, 120 min, K562 dTAG-BRD4	MRA134	
ChIP-Rx	p-Thr4, dTAG7, 120 min, K562 dTAG-BRD4	MRA135	
ChIP-Rx	p-Thr4, DMSO, 120 min, K562 dTAG-BRD4	MRA136	
ChIP-Rx	SPT6, dTAG7, 120 min, K562 dTAG-BRD4	MRA137	
ChIP-Rx	SPT6, DMSO, 120 min, K562 dTAG-BRD4	MRA138	
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA139	
ChIP-Rx	input, dTAG7, 120 min, K562 dTAG-BRD4	MRA140	
ChIP-seq	PRPF8, K562 dTAG-BRD4	MRA142	
ChIP-seq	PRPF8, K562 dTAG-BRD4	MRA143	
HiS-NET-seq	dTAG7, 40 min, K562 dTAG-BRD4	MRA144	
HiS-NET-seq	DMSO, 40 min, K562 dTAG-BRD4	MRA145	
HiS-NET-seq	dTAG7, 40 min, K562 dTAG-BRD4	MRA146	
HiS-NET-seq	DMSO, 40 min, K562 dTAG-BRD4	MRA147	
Hi-ChIP	dTAG7, 120 min, K562 dTAG-BRD4, H3K27ac	MRA148	
Hi-ChIP	dTAG7, 120 min, K562 dTAG-BRD4, H3K27ac	MRA149	
Hi-ChIP	dTAG7, 120 min, K562 dTAG-BRD4, H3K27ac	MRA150	
Hi-ChIP	DMSO, 120 min, K562 dTAG-BRD4, H3K27ac	MRA151	
Hi-ChIP	DMSO, 120 min, K562 dTAG-BRD4, H3K27ac	MRA152	
Hi-ChIP	DMSO, 120 min, K562 dTAG-BRD4, H3K27ac	MRA153	

Table 25: Mass spectrometry-based data.

data type	specification
chromatin-MS	dTAG7, 120 min, K562 dTAG-BRD4; control: DMSO
chromatin-MS	dTAG7, 360 min, K562 dTAG-BRD4; control: DMSO
chromatin-MS	dBET6, 120 min, K562 dTAG-BRD4; control: DMSO
chromatin-MS	dBET6, 360 min, K562 dTAG-BRD4; control: DMSO
native IP-MS	K562 dTAG-BRD4, IP for BRD4; control: IgG
native IP-MS	K562 dTAG-BRD4, IP for BRD3; control: IgG
native IP-MS	K562 dTAG-BRD4, IP for BRD2; control: IgG
crosslink IP-MS	K562 dTAG-BRD4, IP for HA tag fused to BRD4; control: K562 wt, IP for (not expressed) HA-tag
native IP-MS	HeLa "8880", IP for GFP fused to CstF64; control: HeLa S3 wt, IP for (not expressed) GFP

10.4 Ongoing projects and collaborations

10.4.1 Spatial and temporal control of non-coding transcription

(a collaboration with Annkatrin Bressin, Olga Jasnovidova and Elisabeth Altendorfer)

While transcription is a long-known feature of enhancers [10,499], neither its function for enhancer activity and target gene activation nor its dynamic regulation is fully understood. Additionally, the catalog of, for instance, weak or intragenic enhancers is incomplete. Using SI-NET-seq as a starting point, a high-sensitivity version called high sensitivity (HiS-) NET-seq was developed, which involves metabolic labeling of the newly synthesized RNA with 4-thiouridine (4sU) [500] prior to chromatin isolation. This additional enrichment step, together with an improved library preparation protocol [92], results in more complex libraries and consequently a higher read coverage at a higher fraction of Pol II-transcribed units. In this project, HiS-NET-seq is used to capture quantitative changes in active transcription at enhancers and their target genes and the temporal order of their activation, for instance upon BRD4-selective perturbation. Additionally, it resolves Pol II pausing sites better than previous methods, thereby helping the understanding of regulatory pauses.

Within the scope of this project, I prepared HiS-NET-seq, BRD4 ChIP-Rx and Hi-ChIP libraries (partially discussed in 4.1.13, 4.3.2).

10.4.2 UNIQUE AND REDUNDANT FUNCTIONS OF BET PROTEINS IN TRANSCRIPTION AND CO-TRANSCRIPTIONAL PROCESSING

(a collaboration with Nicole Eischer and Annkatrin Bressin)

Comparative analyses of the functions of BRD2, BRD3 and BRD4 are relatively rare, mainly because most of the available small molecule inhibitors and degraders cannot discern between the BET species. Also, whether functions ascribed to the individual BET proteins based on knockout or knockdown studies are truly unique functions or partially redundant is largely unclear. Based on the fundamental role of BET proteins in transcription regulation, it is tempting to speculate that redundancy increases robustness of the early steps of gene expression. In fact, although redundancy is more often found among other functional groups of proteins, there are examples of redundant transcription factors [501]. On the other hand, redundancy is considered evolutionarily unstable [502], which appears in conflict with the observation that distinct BET paralogs are expressed in numerous species.

Aiming at elucidating the shared and unique BET protein functions in leukemia cells, orthologues protein degradation strategies are combined in this project to selectively and sequentially ablate the function of individual and groups of BET proteins. HiS-NET-seq and ChIP-Rx are used to monitor changes in transcription and in localization of those BET species that are not targeted for degradation and might cause redundancy.

For this project, I performed a chromatin-MS treatment time course (discussed in 4.2.1) and RNA-seq.