Wayne State University Dissertations

1-1-2013

An Investigation Into The Role Of Cfia 3' End Processing Complex In The Termination And Initiation/reinitiation Of Transcription

Nadra Al-Husini
Wayne State University,

Follow this and additional works at: http://digitalcommons.wayne.edu/oa_dissertations

Part of the Biology Commons

AN INVESTIGATION INTO THE ROLE OF CFIA 3' END PROCESSING COMPLEX IN THE TERMINATION AND INITIATION/REINITIATION OF TRANSCRIPTION

by

NADRA AL-HUSINI

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirement

for the degree of

DOCTOR OF PHILOSOPHY

2014

MAJOR: BIOLOGICAL SCIENCES

Approved by:

_____

Adviser   Date

_____

_____

_____

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

I.1. EUKARYOTIC TRANSCRIPTION

In eukaryotes, transcription is driven by three distinct RNA polymerases: RNA polymerase I (RNAPI), RNA polymerase II (RNAPII), and RNA polymerase III (RNAPIII) [1]. Although possessing a similar structural design, they transcribe different species of RNA. RNAPI and RNAPIII transcribe genes encoding ribosomal RNA (rRNA), transfer RNA (tRNA), and small nuclear RNA (U6 snRNA), whereas RNAPII is responsible for the transcription of the protein-coding genes, the majority of small nuclear RNAs (snRNAs), and small nucleolar RNAs (snoRNAs). In addition, RNAPII transcribes a wide range of RNA species with no apparent coding potential. These non-coding RNAs include small interfering RNAs (siRNAs), micro-RNAs (miRNAs), cryptic unstable transcripts (CUTs), stable unannotated transcripts (SUTs), Xrn1 stabilized transcripts (XUTs), meiotic unannotated transcripts (MUTs), and *Ssu72*-restricted transcripts (SRTs) ([2-8]. In addition, there are other RNAPII-transcribed non-coding RNA species that do not fall into any of the categories described above. The biological role of many of these non-coding RNAs is not yet clear [9].

I.2. RNA POLYMERASE II

RNAPII is a large, multisubunit enzyme that is highly conserved from yeast to human. It has a molecular weight of ~0.5 MDa. It is composed of 12 subunits. The catalytic core of the enzyme is formed by 10 subunits; Rpb1, Rpb2, Rpb3, Rpb5, Rpb6, Rpb8, Rpb9, Rpb10, Rpb11 and Rpb12. The core enzyme can efficiently elongate transcripts, but is deficient in promoter-associated initiation and the termination

of transcription. The dimer of Rpb4/Rpb7 associates with the core to make a form of enzyme that can accomplish all steps of transcription from initiation to termination. The Rpb4/Rpb7 dimer is crucial for initiation as well as termination steps of transcription [10-

Figure I.1. Order of recruitment of the general transcription factors (GTFs) during the formation of the pre initiation complex
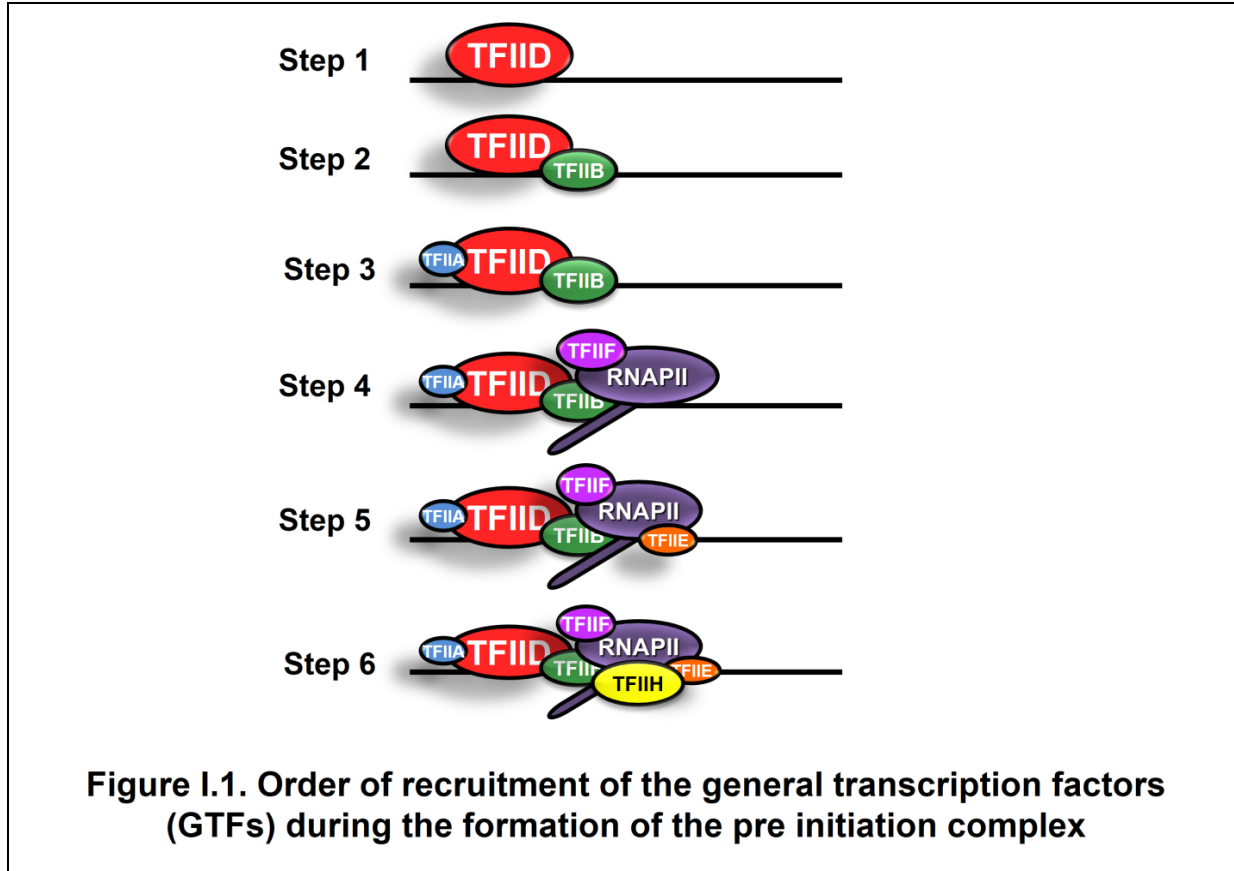
14]. The carboxy-terminal-domain (CTD) of the largest subunit of RNAPII (Rpb1) is critical for both transcription and cotranscriptional RNA processing. This CTD consists of tandem heptapeptide repeats of the sequence $Y_1S_2P_3T_4S_5P_6S_7$. The number of repeats varies from 26 in yeast to 52 in mammalian systems. During the transcription cycle, the CTD is subjected to several post-transcriptional modifications that work as recognition marks for binding of the factors required for the execution of different steps of transcription [15]. On its own, RNAPII can unwind DNA, synthesize RNA, proofread the nascent transcript, and rewind DNA. It cannot, however, recognize the regulatory

elements located at the 5' and 3' ends of the gene, or respond to regulatory cues [16,17]. The recognition of regulatory elements and response to environmental cues requires a group of accessory factors.

I.3. RNAPII TRANSCRIPTION CYCLE

The transcription cycle of RNAPII can be divided into four major steps; initiation, elongation, termination and reinitiation [18] (Figure I.2). These steps occur sequentially in a coordinated manner [12,18-20]. The transcription cycle starts with the binding of an

Figure I.2. Major steps in the RNAP II transcription cycle
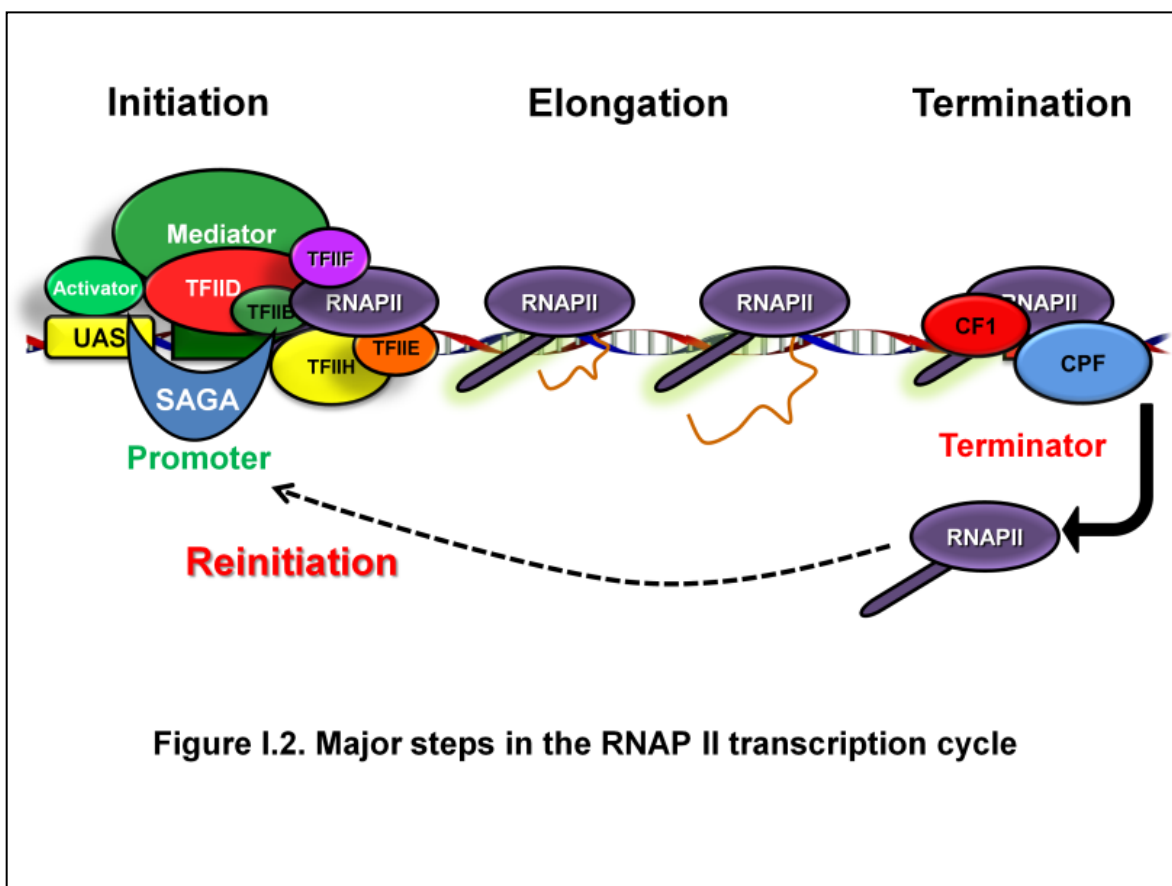
activator to the upstream activating sequence (UAS) or enhancer element. The activator facilitates the recruitment of RNAPII and the general transcription factors (GTFs); TFIID, TFIIB, TFIIA, TFIIF, TFIIE and TFIIH, to the promoter to form a ~2 MDa pre-initiation complex (PIC). The PIC assembles in the following order; TFIID-TFIIB-TFIIA-RNAPII

and TFIIF-TFIIE-TFIIH (Figure I.1) [21-24]. The GTFs get recruited to the specific DNA sequences flanking the transcription start sites (TSSs) in the promoter region. These DNA elements are collectively called 'Core Promoter Elements' (CPEs), and include the TATA-box, the initiator (INR)-element, TFIIB-recognition elements (BRE), downstream promoter elements (DPE), and a variety of other gene specific elements. The combination of these regulatory elements is gene specific, and it seems that each promoter has a distinct architecture. The first promoter element identified was the TATA-box (TATAA) located ~25 bp upstream of the TSS. The general transcription factor that is recruited first on the promoter during the transcription cycle is TFIID. It is a megadalton complex containing the TATA-binding protein (TBP) and the TBP associated factors (TAFs). The binding of TBP to the TATA-box bends DNA in the promoter region. TBP may also be recruited on some promoters as a component of the SAGA complex. TFIID then interacts with the TFIIA via TAF40-TFIIA interaction, and helps stabilize TFIID-DNA interaction [16,25]. In the absence of a TATA-box, TFIID initiates PIC assembly by binding to the initiator (INR)-element and the DPE. The interaction of TFIID with INR-element and DPE is through TAFs [25,26]. The next general transcription factor to arrive on the promoter is TFIIB, which is recruited through its interaction with the BRE^u and BRE^d elements that flank the TATA-box. TFIIB binding results in the formation of a stable ternary complex (DNA-TFIID-TFIIB complex). The ternary complex is sufficient to facilitate the recruitment of RNAPII. The interaction of the RNAPII with the ternary complex requires TFIIF [25]. The crystal structure has revealed that the N-terminal region of TFIIB has a loop structure, the "B-finger", reaching into the active site in the cleft of the polymerase. The B-finger plays a very

critical role in start site selection by RNAPII. TFIIE and TFIIH are recruited next. TFIIH is a 12 subunit complex with a molecular weight of ~500 KD [27]. The helicase activity of TFIIH is required for promoter clearance. RNAPII and the above mentioned GTFs are able to direct accurate initiation of transcription from the promoter and establish a basal level of transcription. They, however, fail to respond to gene-specific activators. 'Mediator' helps activator communicate with the general transcription machinery. Mediator is a ~1 MDa complex composed of more than 24 subunits. It works as a transducer of the regulatory information from activators/repressors to the PIC [28]. Considering its recently demonstrated role in basal transcription, Mediator is now considered one of the general transcription factors.

I.3.1. INITIATION

Initiation is an early step in the transcription cycle, and the one most targeted by the regulatory signals. The successful assembly of a functional PIC on the promoter region is a pre-requisite for the initiation of transcription. However, the mere formation of a PIC is not a guarantee for the successful initiation of transcription. For initiation to occur, the DNA bound by the active site of the polymerase has to unwind so that the single stranded region is exposed and can act as a template. This step results in the formation of a bubble at the TSS (-9 to +2), and an 'open complex' is formed. The unwinding of DNA during bubble formation is an intrinsic activity of the polymerase. The maintenance of the bubble needs TFIIF, which binds to the non-template strand and assists in establishing the open promoter conformation. TFIIB has also been proposed to play a crucial role in this step [25,29]. After bubble formation, RNAPII initiates transcription. The fate of the initiation event depends on the length of the nascent

transcript [30]. Transcription starts with the repetitive synthesis and abortion of short transcripts. The transcripts of less than 5 nucleotides are often unstable and are frequently aborted. When the transcript length reaches 6 nucleotides, it starts clashing with the B-finger loop of TFIIB that is inserted into the RNA exit channel of the polymerase [29]. During the transition from initiation to elongation, the B-finger of TFIIB is pushed out of the RNA exit channel and the bubble collapses. The bubble collapse is facilitated by the helicase activity of the Ssl2 (XPB in higher eukaryotes) subunit of TFIIH that acts as a wrench to unwind the DNA downstream of the bubble. A recent report implicated TFIIH in the formation of the bubble as well [23].

In addition to the helicase activity, TFIIH possesses another enzymatic activity, the kinase activity. The Kin28 (Cdk7 in mammals) subunit of TFIIH is the kinase specific for Ser-5 ($Y_1S_2P_3T_4S_5P_6S_7$) of the CTD. The Ser5- phosphorylated CTD acts as a loading dock for the capping enzyme [15]. The capping enzyme adds a methylguanosine cap to the 5' end of the nascent transcript as soon as it emerges from the RNAPII exit channel, which happens when the length of the transcript is 17-25 nucleotides. The 25 nucleotide long capped transcript marks the end of initiation, and the beginning of the elongation step [31,32]. Up to this point, RNAPII is still tethered to the promoter-bound initiation complex. In order to escape from the promoter, RNAPII must sever its ties with the promoter-bound factors. The promoter escape/clearance requires both of the enzymatic activities of TFIIH [25,33]. Upon promoter clearance, a number of GTFs remain at the promoter, forming a reinitiation 'scaffold' [20,23,34]. The reinitiation scaffold consists of TFIID, TFIIA, TFIIE, TFIIH and Mediator. The scaffold serves as a launching pad for re-entry of RNAPII, TFIIB and TFIIF during subsequent

rounds of transcription. This facilitates faster reinitiation and increases transcription efficiency by bypassing the need for the *de novo* PIC assembly [20].

I.3.2. ELONGATION

Soon after promoter clearance, RNAPII encounters a number of barriers, such as inhibitory factors, arrest sequences, and nucleosomes. In most of the eukaryotes,

Figure I.3. 3' end processing/termination by RNAP II
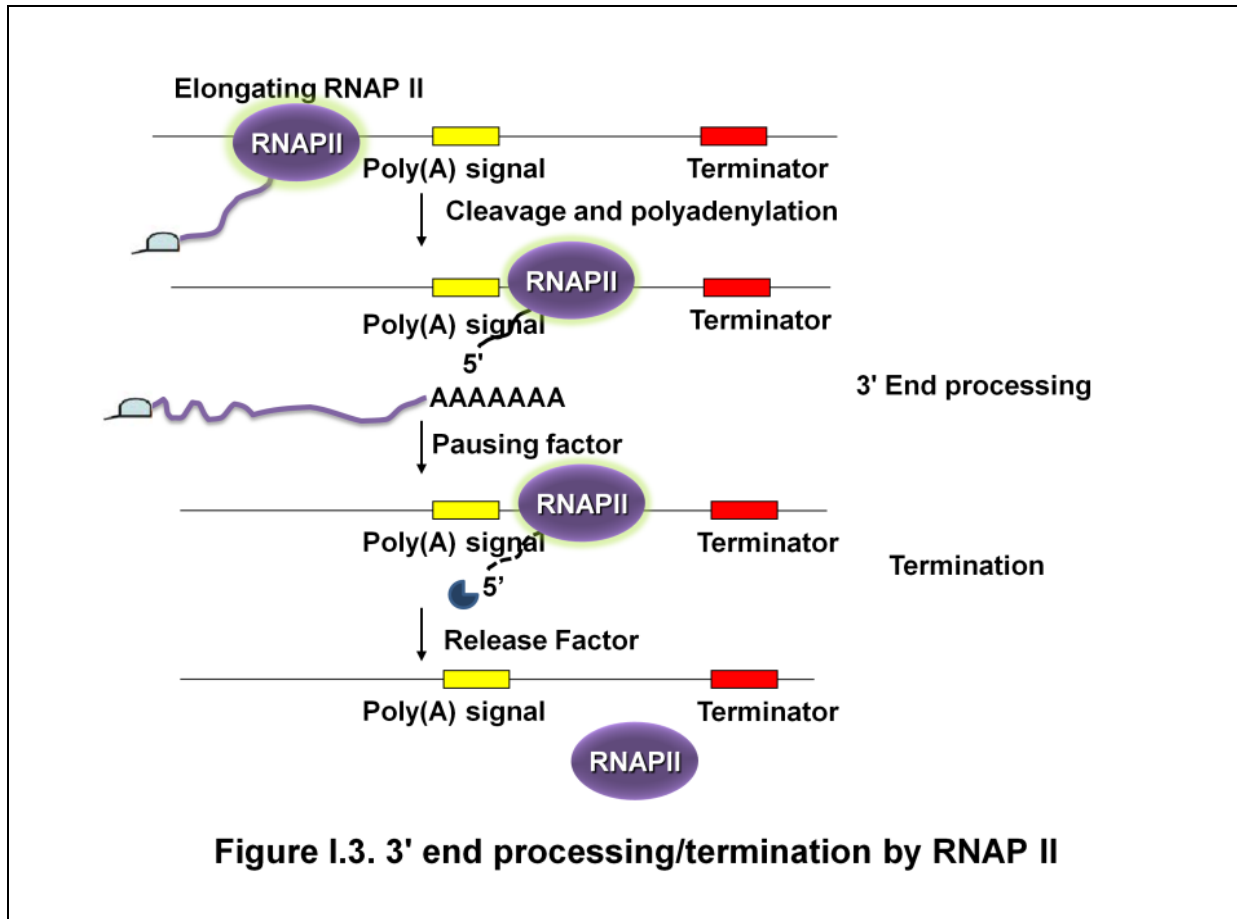
RNAPII pauses at ~50 bp from the TSS. This pausing could be a check point to allow time for the recruitment of the capping enzyme [35]. It is induced by two factors, DSIF (DRB-sensitive inducing factor) and NELF (Negative elongation factor). The pause is alleviated by the positive elongation factor P-TEFb (Ctk1/Bur1 in yeast) which mediates the phosphorylation of Ser2 of the CTD as well as DSIF and NELF thereby neutralizing

their inhibitory effect. Another common impediment for the elongating polymerase is offered by AT-rich sequences, where the polymerase backtracks and gets arrested due to the misalignment of the active site with the 3' end of the growing RNA chain. This elongation arrest is overcome by TFIIS, which stimulates the 3' to 5' exonuclease activity of the backtracked polymerase, thereby aligning the active site with the 3' end of the growing RNA chain [36,37]. The elongating RNAPII also needs to overcome the nucleosomal barrier during elongation. At least three different types of activities help the polymerase get around the nucleosomal obstruction. These are ATP-dependent chromatin remodelers, histone chaperones and histone modifying enzymes. The ATP-dependent chromatin remodelers like RSC and Chd1 modify chromatin using the energy of ATP hydrolysis, while histone modifying enzymes like Set, Set2 and HATs post-translationally modify the histone tails by adding a chemical group.

I.3.3. TERMINATION

There are two distinct, although coupled, events occurring at the end of each RNAPII transcriptional cycle: (1) the 3' end processing of the nascent transcript, and (2) termination of transcription (Figure I.3). 3' end processing is the endonucleolytic cleavage of RNA at the 3' end followed by the addition of about 50-60 adenine nucleotides (180-200 in higher eukaryotes) to the cleaved end. The polyadenylated transcript is then released from the template. The elongating polymerase may still keep on transcribing the template. The termination of transcription is not accomplished until the polymerase also dissociates from the template. Thus, termination of transcription involves the release of transcript as well as the polymerase from the template [38]

Although the factors required for the 3' end processing of mRNA are relatively well characterized, a thorough understanding of the factors required for the termination of transcription has eluded us. The cleavage and polyadenylation of pre-mRNA is critical for cell growth and viability. It is a pre-requisite for transport of mRNA from the nucleus to the cytoplasm [39]. The polyadenylation confers stability to the transcript by preventing exonucleolytic cleavage of mRNA from the 3' end [40,41]. The

Figure I.4. 3' end processing/termination machinery in yeast. The complexes and their components are color coded.
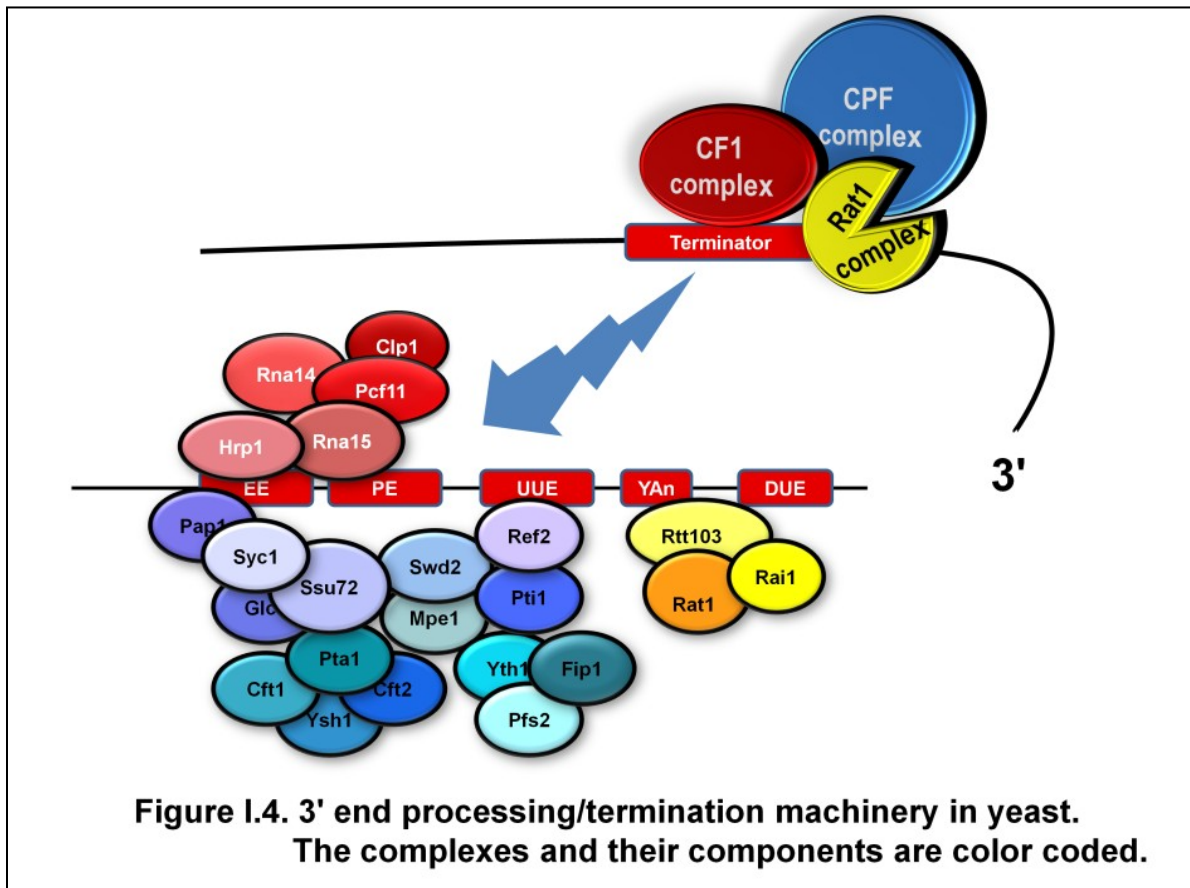
polyadenylated transcripts are also a better substrate for translation by ribosomes [41]. 3' end processing is often a pre-requisite for termination of transcription [42,43]. In fact, cleavage and polyadenylation are generally coupled to termination. The first clue in this regard came when it was shown that both processes are dependent upon the same DNA sequence elements at the 3' end of genes [44-49]. The interdependency of these

two events was further reinforced when it was found that a number of cleavage and polyadenylation factors are also required for termination [50-52].

Termination is a crucial step in the eukaryotic transcription cycle. Efficient termination is essential for the maintenance of the overall integrity of the transcriptome [53]. It ensures that a pool of free polymerase molecules is available for reinitiation during subsequent rounds of transcription [54]. Termination also prevents transcription interference among neighboring genes [55,56]. In yeast, this is of particular significance due to the compact nature of the genome [57]. Furthermore, termination prevents 3' end-initiated antisense transcription that may interfere with the normal sense transcription [54]. In addition, proper termination prevents the synthesis of extended, nonfunctional RNAs [58].

I.3.3.1 *Cis-acting* elements required for 3' end processing/termination

Despite its overwhelming significance, termination is the least understood step of the eukaryotic transcription cycle [54]. The termination of transcription requires cis-acting elements and trans-acting factors. Many of the factors required for termination are the same as required for 3' end processing of precursor mRNA. The 3' end processing is directed by the sequence elements embedded within the coding region and the 3' untranslated region (3'UTR) of the pre-mRNA [59] (Figure I.4). Disruption of these elements reduces the efficiency of 3' end processing of RNA [60]. The *cis* acting elements required for 3' end processing are quite similar in yeast and mammals (Figures I.4 and I.5). However, they exhibit major differences as well. In yeast, the analysis of the 3' end *cis* acting elements is complicated due to a high degree of variability and redundancy [59]. The poly(A) site, which is an important element required

for both 3' end processing and termination, is defined by four elements: the AU-rich efficiency element (EE), the A-rich positioning element (PE), the cleavage site [Y(A)n], and two U-rich elements flanking the cleavage site, namely, the upstream U-rich element (UUE) and the downstream U-rich element (DUE) [61,62]. In contrast, the

Figure I.5. 3' end processing/termination machinery in mammals. The complexes and their components are color coded.
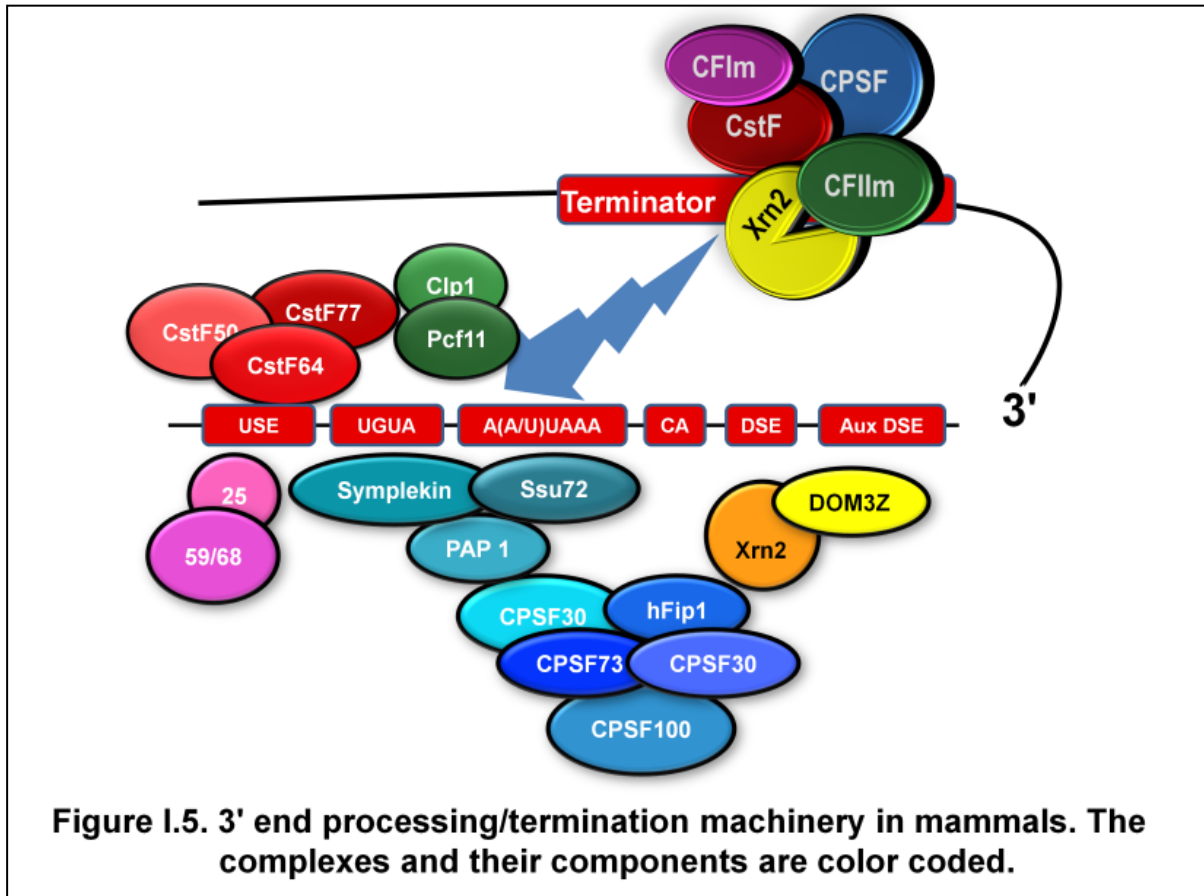
poly(A) site in mammals contains three primary and two auxiliary elements that determine and regulate the 3' end processing reaction. The three primary elements are the polyadenylation signal (PAS), AAUAAA, the cleavage site, CA, and the downstream element (DSE), mainly composed of G/U rich region. The two auxiliary elements are upstream U-rich element and the downstream G-rich element [61,62].

The *cis*-acting elements required for termination of transcription are not as well characterized as the ones required for the 3' end processing. However, some studies

have uncovered some sequences involved specifically in termination. First, a study done by the Sherman lab revealed a consensus element (TAG…TAGT…TTT) that is present in about 14 genes as a key termination signal [47,63]. Second, it has been demonstrated that a sequence downstream of the poly(A) site in the gastrin gene can cause transcription termination regardless of the presence of the poly(A) site [64,65]. Third, a sequence in the adenovirus genome called CCAAT-box was shown to be necessary for termination [66]. In general, 3' end processing and termination are thought to be the function of both a functional poly(A) signal and a downstream RNAPII pausing site [42,64,67,68]. In general, a poly(A) site is required for the termination of transcription of most RNAPII-transcribed genes. It has been reported that termination does not occur at a specific site or distance from the poly(A) site, but takes place at variable distances from the poly(A) site (about 100-150 bp downstream of the poly(A) site in yeast, and from a few bases to several kilo bases from the AAUAAA site in mammals) [42,63,64,69,70].

I.3.3.2*. Trans-acting* factors required for 3'end processing/termination

Although the reactions taking place at the 3` end of a gene are seemingly simple, a megadalton complex is required to execute these reactions (Figures I.4 and I.5). The CFI (Cleavage Factor I), CPF (Cleavage Polyadenylation Factor), and Rat I complexes are required for both cleavage-polyadenylation as well as termination of transcription In budding yeast [71]. These complexes together are made up of more than 23 proteins [61]. The CFI complex is composed of five subunits; Rna15, Rna14, Pcf11, Hrp1, and Clp1 [72]. The chromatographic analysis revealed that CFI complex can be further separated into the CFIA (Rna14, Rna15, Pcf11 and Clp1) and CFIB (Hrp1)

subcomplexes [59,73]. The CPF complex consists of seven subunits organized in two

subcomplexes; PFI (Polyadenylation Factor 1) and CFII (Cleavage Factor II). The PFI

Figure I.6. The Composition, stoichiometry, and topology of Cleavage
Factor I complex (CFI) , adapted from Gordon et al., 2011
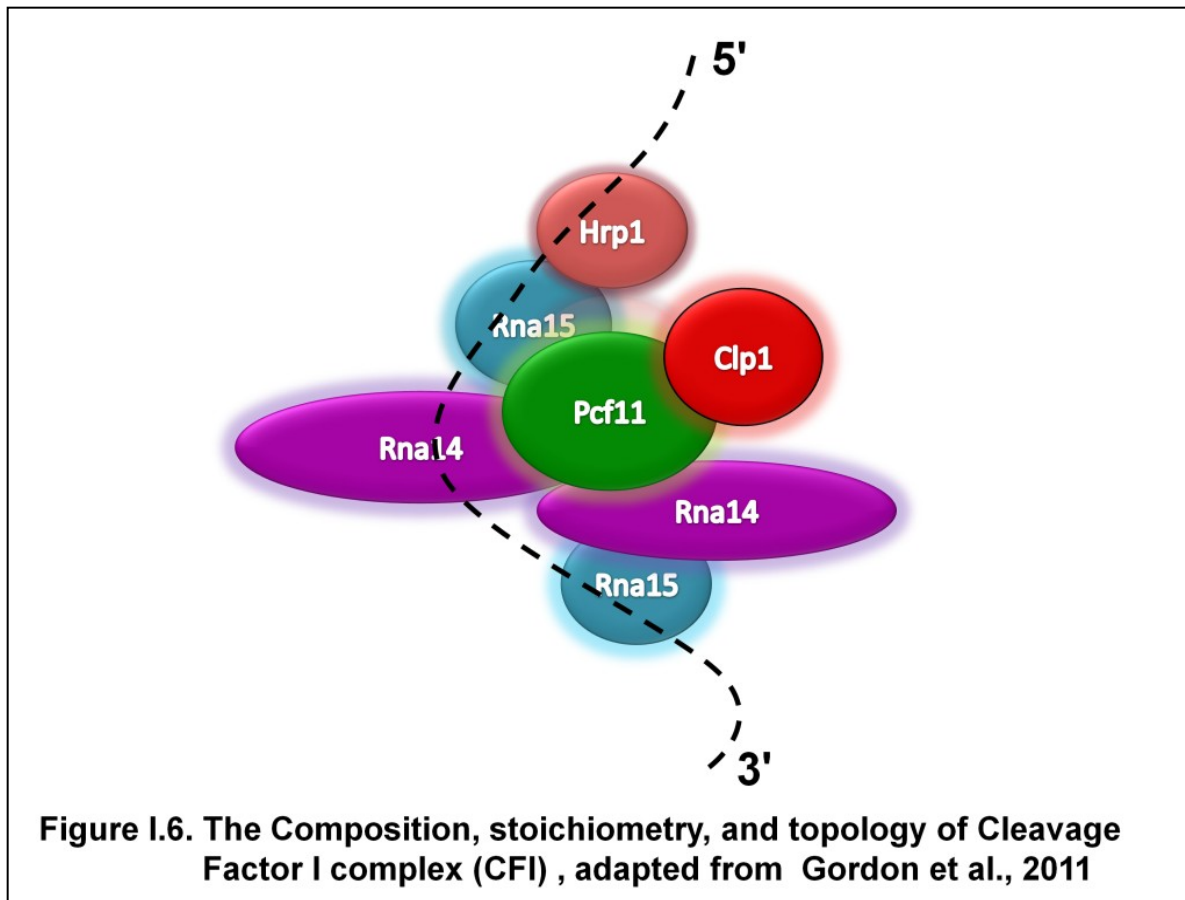
subcomplex contains three subunits, Fip1, Yth1 and Pfs2, while the CFII subcomplex

contains four subunits, Pta1, Yhh1, Ydh1, and Ysh1. In addition to these two

subcomplexes, CPF contains other subunits that are not part of either of these

subcomplexes. These are Pap1, Pti1, Ssu72, Glc7, Syc1, Swd2 and Mpe1 [59,61,74].

In vitro studies showed that the cleavage reaction requires CFIA, CFIB and CFII, while

polyadenylation requires CPF, CFIA, CFIB and Pap1. The third complex, the Rat1

complex, is a dedicated termination complex and contains three subunits, Rat1, Rtt103,

and Rai1 [75]. As described earlier, the Rat1 subunit of this complex possesses the 5' to

3' exoribonuclease activity.

The 3' end processing/termination machineries in yeast and mammals exhibit a lot of similarities (Figures I.4 and I.5) [61,62]. The mammalian machinery includes several complexes namely, Cleavage and Polyadenylation Specificity Factor (CPSF), Cleavage Stimulatory Factor (CstF), Cleavage Factor I (CF Im), and Cleavage Factor II (CF IIm) [61,62,71]. In addition to these complexes, poly(A) polymerase (PAP1), Poly(A) Binding Proteins (PABP), symplekin and the CTD of RNAPII are also needed for 3' end processing (Figure I.5.) [61,62]. Many of the yeast subunits have homologues in mammals. For example, subunits of the yeast CPF subunits are homologous to the CPSF subunits, and yeast CFI complex exhibits homology with mammalian CstF complex. In addition, Rat1 and Rai1 subunits of yeast Rat1 complex are homologous to mammalian Xrn2 and Dom3z subunits respectively [75]. Despite these homologies, there exist significant differences between the yeast and mammalian complexes. The yeast termination factors mentioned above are those required for the termination of transcription of a majority of mRNA transcripts. The termination of transcription of snRNAs, snoRNAs, CUTs and short mRNAs in yeast occurs through a distinctive pathway that requires a different set of termination factors. The core factors required in this pathway are Nrd1, Nab3, and Sen1. In addition to these factors, termination of at least some of the non-coding RNA species also requires mRNA 3' end processing/termination factors [75].

I.3.3.3. CLEVAGE FACTOR I COMPLEX (CFI)

The focus of this study is the CFI complex. It is comprised of five subunits organized into two subcomplexes, CFIA and CFIB [72,73]. The CFI complex can be assembled *in vitro* from purified, recombinant components, which suggests that the

 assembly may not be dependent on other cellular factors. The CFIA subcomplex is composed of two subunits each of Rna14 and Rna15, and one subunit each of Pcf11 and Clp1 [76,77], while CFIB subcomplex is constituted of a single protein, Hrp1 (Figure

Figure I.7. Structure and domains of the Cleavage Polyadenylation 1 factor (Clp1), adapted from Noble et al., 2007
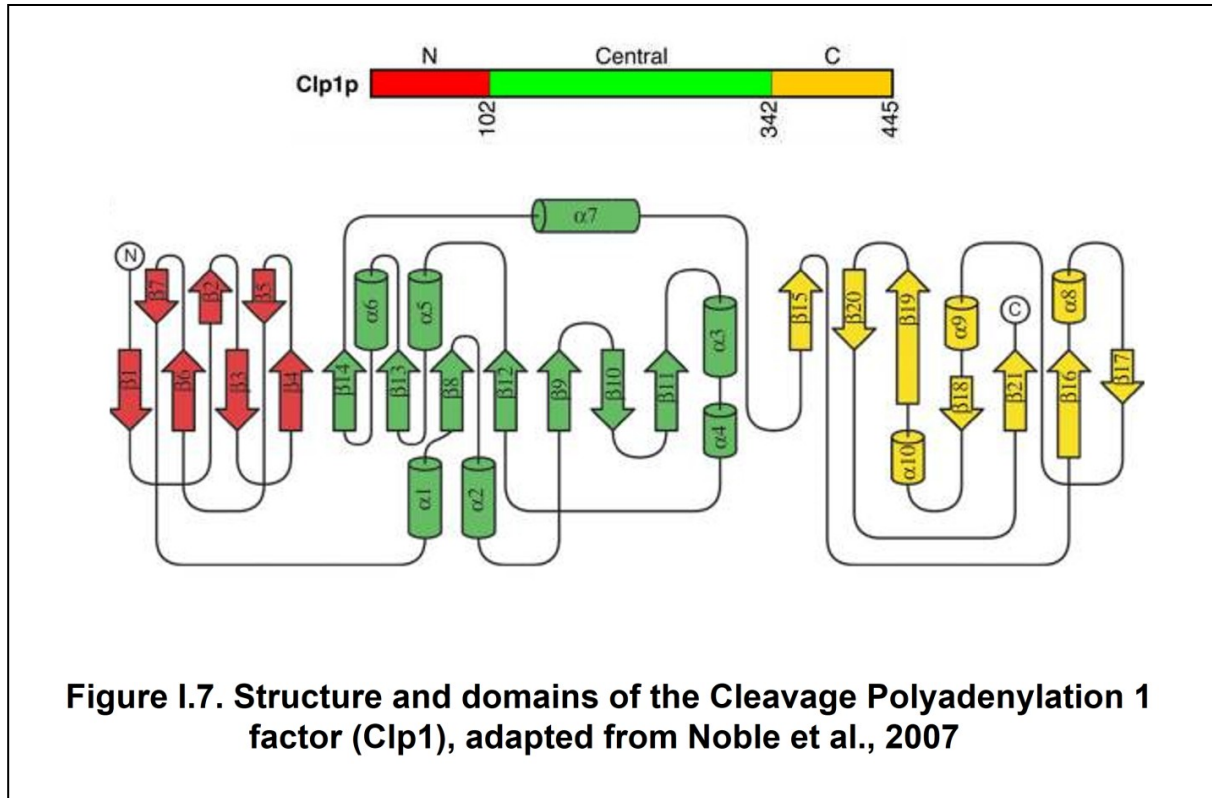
I.6). None of the five subunits of the CFI complex possess endoribonuclease activity. The CFI complex has been implicated both in 3' end processing of nascent mRNA as well as termination of transcription. The precise role of the CFI complex in termination and RNA processing is not clear yet. It has been suggested that the function of the CFI complex is to facilitate the recruitment of the CPF complex, which then executes both 3' end processing and the termination step of transcription [78]. The CFI complex has also been implicated in the export of mRNA from the nucleus to the cytoplasm [79-81]. The Hrp1 and Rna15 subunits of CFI complex have RNA recognition motifs (RRM), which help them to bind the nascent mRNA and position the processing/termination machinery

to the 3' end of transcribing RNA. The Pcf11 subunit has a CTD interaction domain (CID) that interacts preferentially with the Ser2-phosphorylated form of the CTD. Pcf11 is one of the best studied subunits of the CFI complex. It has been implicated both in poly(A)-dependent and poly(A)-independent termination pathways [48,82-84]. It is the only subunit that makes direct contact with three other CFI subunits, namely Rna14, Rna15, and Clp1. It has also been found to interact with the Pta1, Ssu72, Ysh1, Cft1, and Cft2 subunits of CPF complex, although a direct binding to these factors has not been proved [85]. The ability of Pcf11 to interact with the CTD suggests that it may be the factor that facilitates the CTD-dependent recruitment of the CFI and CPF complexes to the 3' end of RNA.

In addition to Pcf11, two other components of the CFI complex, Rna14 and Rna15, have also been the focus of intense investigation. Rna15 recognizes the A-rich EE in the nascent transcript through its RRM. Dimerization of Rna14 with Rna15 helps guide Rna15 to the EE because Rna14 works as a bridge between Rna15 and Hrp1. The mammalian counterparts of Rna14 and Rna15, called CstF64 and CstF77, also form a dimer [86].

Of all the CFI subunits, Clp1 is the least explored. Clp1 was first reported as a subunit of CFI complex in 1997 by Walter Keller's laboratory [87]. However, it was not the focus of intense investigation until recently. This could be attributed, in part, to the non-availability of reliable conditional mutants for this subunit [76]. Clp1 has been conserved during evolution from yeast to humans. Yeast Clp1 shares 23% identity with its human counterpart [61,88]. Clp1 possess a large central domain through which it interacts with a number of 3' end processing/termination factors. The small N-terminal

and C-terminal domains are crucial for cell viability [89]. The central domain of Clp1 contains an evolutionarily conserved ATP binding motif called 'Walker A motif' (Figure I.7) [76]. While the human Clp1 was shown to have an ATPase activity, biochemical assays demonstrated that yeast Clp1 lacks such activity at least *in vitro* [90]. The precise function of the ATP-binding motif of yeast Clp1 and its effect on cell viability is controversial [89,91]. Clp1 interacts with a number of factors belonging to both CFI and CPF complexes through its N-terminal and central domains. It interacts with Pta1, Cft1, Pfs2, and Pcf11 via its central domain and with Ysh1 and Ssu72 via the N-terminus domain. It has been shown to play a central role in the assembly of the CFIA complex [89,92]. The crucial role of Clp1 in the assembly of the CFIA complex, the central position of Clp1p in the processing/termination machinery, and its multiple interactions with other 3' end processing/termination factors makes it a suitable factor for mediating the interaction of the CFI and CPF complexes [89,91,92]. *In vitro* studies have revealed a role for Clp1 in both the cleavage and polyadenylation steps of 3' end processing [72,89]. In addition, *in vivo* analysis showed that Clp1 plays a role in termination of transcription of snoRNAs and CUTs [92]. However, the precise function for Clp1 in the termination of protein coding genes has not been fully addressed yet *in vivo*. Apart from the five proteins making the CFI complex, another protein, poly(A) polymerase 1 (Pap1), has been found to be functionally related to this complex [38]. Additionally, studies from our lab found Pap1 in a complex with CFI and TFIIB [93]. Pap1 is an authentic component of the CPF complex. It is an essential 3' end processing enzyme that is required along with CFI, PFI and PFII for the addition of poly(A) tail to the 3'-OH group generated after the cleavage of the nascent transcript [38,94]. In the absence of CFI,

PFI and PFII, Pap1 can add a poly(A) tail nonspecifically to any RNA which indicates that CFI and PFII confer substrate specificity to Pap1p.

I.3.3.4. Pathways for the termination of transcription by RNAPII
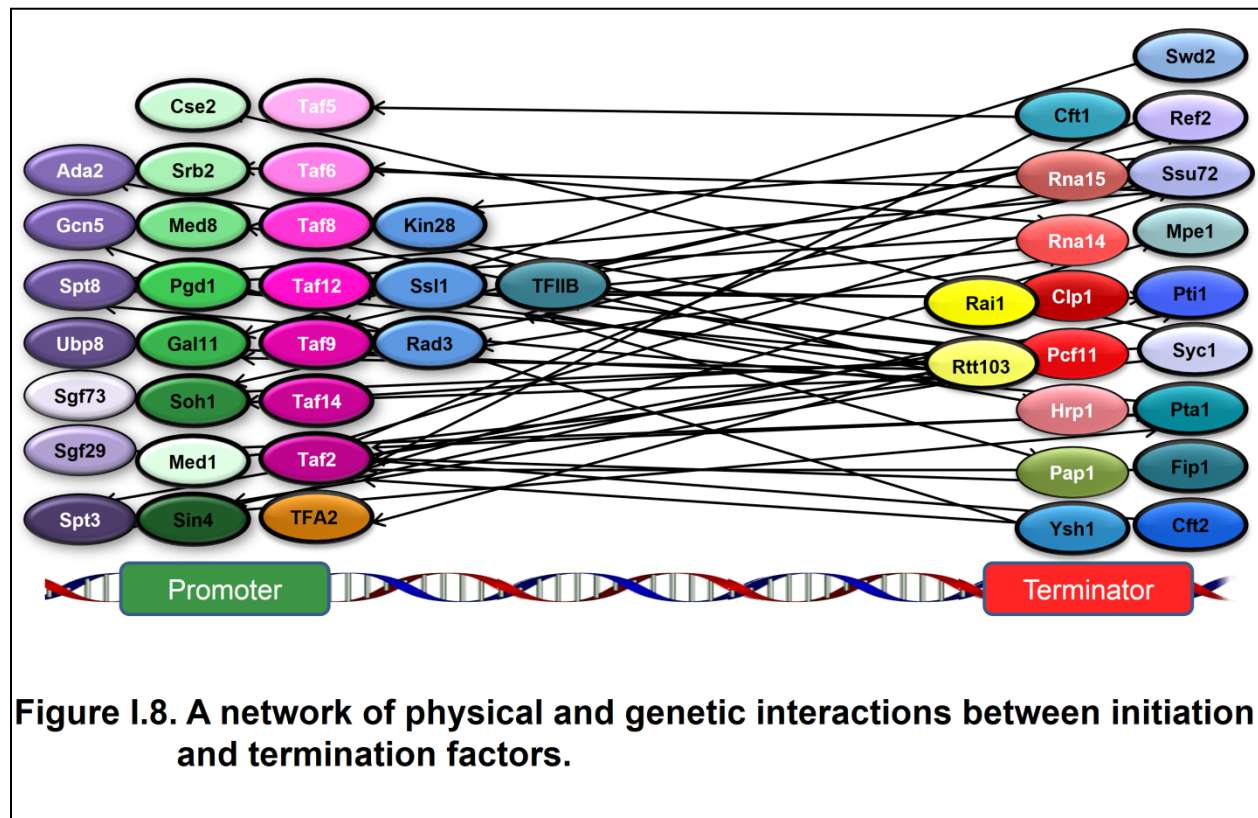
Figure I.8. A network of physical and genetic interactions between initiation and termination factors.

Termination of transcription occurs through at least two different pathways; a poly(A)-dependent pathway, and a poly(A)-independent pathway. There are two models explaining the termination of transcription by poly(A)-dependent pathway; the 'allosteric model' and the 'torpedo model' [42,44,95,96]. Both models connect termination directly or indirectly to poly(A)-dependent 3' end processing. The allosteric model states that termination is the consequence of a conformational change in the elongation complex triggered by transcription through the poly(A) site. The recruitment of negative elongation factors or the release of anti-termination factors may contribute to the conformational changes in the elongation complex. The net result is the destabilization

of the elongation complex, thereby facilitating termination [54,97]. The torpedo model, on the other hand, suggests that the termination of transcription by RNAPII occurs in a manner similar to rho-dependent termination in bacteria. The endonucleolytic cleavage of the nascent mRNA at the poly(A) site results in polyadenylation followed by the release of processed mRNA. The downstream cleavage product ,RNA with an

Figure I.9. Initiation and termination factors crosslink to the distal ends of a gene.
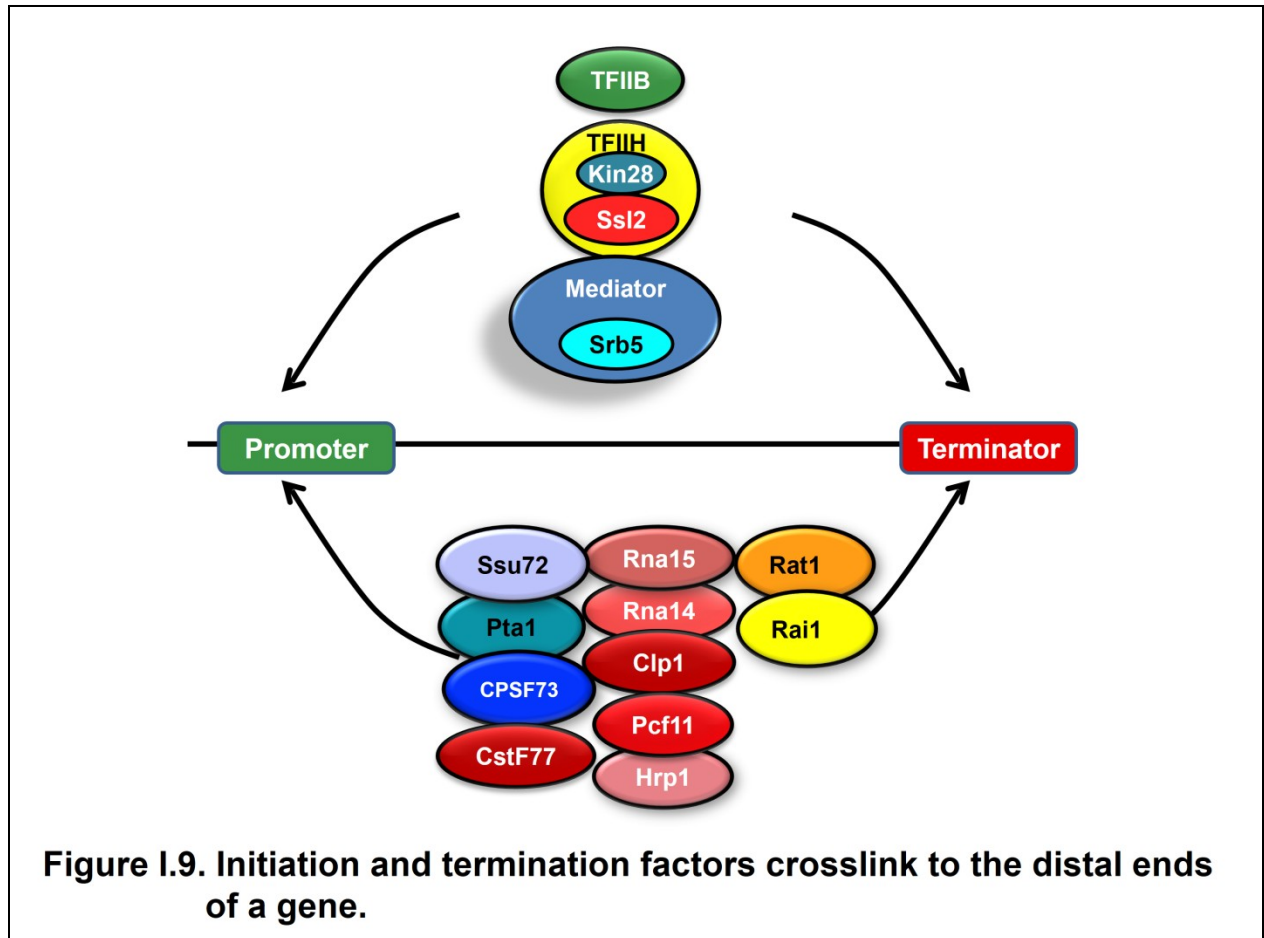
uncapped 5' end, however, is still attached to the elongating polymerase. The torpedo model suggests the degradation of this downstream cleavage product is critical for termination of transcription. Rat1 (Xrn2 in mammals) is an exonuclease with 5' to 3' exoribonuclease activity. Rat1 binds to the uncapped 5'-monophosphate end of the elongating RNA and starts cleaving the transcript from 5' to 3' end until it catches up

with the advancing polymerase. The polymerase then dissociates from the template in a manner that is still not completely understood [98,99]. There is experimental evidence in support of both models. The actual method of termination may be a 'hybrid' of the two proposed models [82,100]. The poly(A)-independent termination pathway, which has been shown to operate in budding yeast, is dependent on Sen1. The Sen1-dependent pathway has been implicated in termination of transcription of small stable non-coding RNAs (sn/snoRNAs), and cryptic unstable transcripts (CUTs) [53,54]. It is also involved in the termination of short protein coding genes [101,102]. This termination pathway requires the function of exosome-TRAMP complex, which utilizes its exonuclease activity to either trim the snRNA/snoRNA transcript to the mature stable RNA or to terminate and degrade the CUTs after endoribonucleolytic cleavage [53,64,75].

The transcription cycle of RNA polymerase II can be divided into a number of discreet steps. These are the assembly of the preinitiation complex, initiation, promoter clearance, elongation and termination. A well orchestrated coordinated execution of these steps is essential for completion of the transcription cycle. The steps of transcription are interlinked, and affect each other. It was known for a long time that initiation affects elongation, and elongation is important for termination. It was, however, not appreciated that termination affects reinitiation of the next round of transcription. Recent evidence suggests that a network of intricate interactions exists between the initiation and termination steps of transcription (Figuer I.8).

The experimental evidence that suggest a termination-reinitiation link can be divided into four broad categories. First, genetic screens revealed multiple interactions between the initiation and termination factors [103-118]. The most well studied of these

interactions is that of TFIIB with Ssu72. Second, a number of studies using independent experimental approaches have found a network of physical interactions between initiation and termination factors [74,93,107,112,114,115,117-131]. Many of these physical interactions were also observed in the genetic screens. Third, crosslinking studies have found a number of initiation and termination factors occupying both the ends of genes [67,93,120,132-136]. (Figure I.9). Fourth, Chromosome Conformation Capture (CCC) analysis of multiple genes has revealed that the promoter and the terminator regions of genes are juxtaposed during transcription. Such promoter-terminator interactions result in the formation of a looped gene conformation [7,93,120,132,137-150]. These findings strongly suggest a crosstalk of the termination and reinitiation steps during the transcription cycle. Termination of a round of transcription is therefore not the end of the transcription cycle, but rather the beginning of the next round of transcription.

I.3.4. REINITIATION

Following the initial round of transcription, the chromatin template is utilized by multiple reinitiation events to make multiple copies of the transcript from the same gene. The studies suggest that the first and the subsequent rounds of transcription are not identical, as many steps required for *de novo* initiation are bypassed during reinitiation. Reinitiation is crucial for the persistence of an activated state through multiple transcription cycles, and is therefore a significant determinant of the level of RNA in the cell. Despite its critical role, reinitiation is still a poorly characterized aspect of the transcription cycle. During the first round of transcription, the general transcription factors and RNAPII assemble on the promoter, forming a preinitiation complex (PIC).

Following the initiation of transcription, most of these factors are left behind on the promoter, forming a 'scaffold' that serves as the entry point for RNAPII during reinitiation [34]. Since multiple steps involving the recruitment of GTFs are bypassed during the second and subsequent rounds of transcription, reinitiation is always faster than initiation. Reinitiation efficiency can be augmented further if RNAPII is somehow directly transferred from the terminator to the promoter without being released from the template [151]. This hyperprocessive reinitiation could potentially boost the transcription rate by many fold.

The reinitiation strategies described above are supported by several studies that showed that the number of RNA polymerase molecules in eukaryotic cells does not considerably exceed the number of active transcriptional units at a given time [152-156]. These findings argue against the presence of a large pool of free polymerase molecules available for de novo initiation events. Recent studies strongly suggest that the recycling of RNA polymerase is facilitated by the interactions of the initiation and termination machineries [132,137]. These interactions result in the formation of dynamic structures termed 'gene loops' that may facilitate RNAPII recycling through multiple transcription events [157].

I.4. GENE LOOPING

Several findings have revealed cross-talk between the promoter and the terminator regions of several eukaryotic genes during transcription (Figure I.10) [137,140,147,149,150]. The interaction of the promoter and terminator regions of a gene resulting in the formation of a looped conformation is referred to as gene looping [132,137,139]. First reported in yeast, it was later discovered in higher eukaryotes as

well [138,140-142,145,147-150]. Gene loops have been observed in mammalian systems for a number of genes, including the HIV proviral gene, *BRCA1*, *CD68, COX2* and *MMP13* genes [138,140-142]. The promoter and 3′ terminal regions of the Progesterone Receptor (PR) coding gene were also found to be in close proximity, and

Figure I.10. Gene looping model
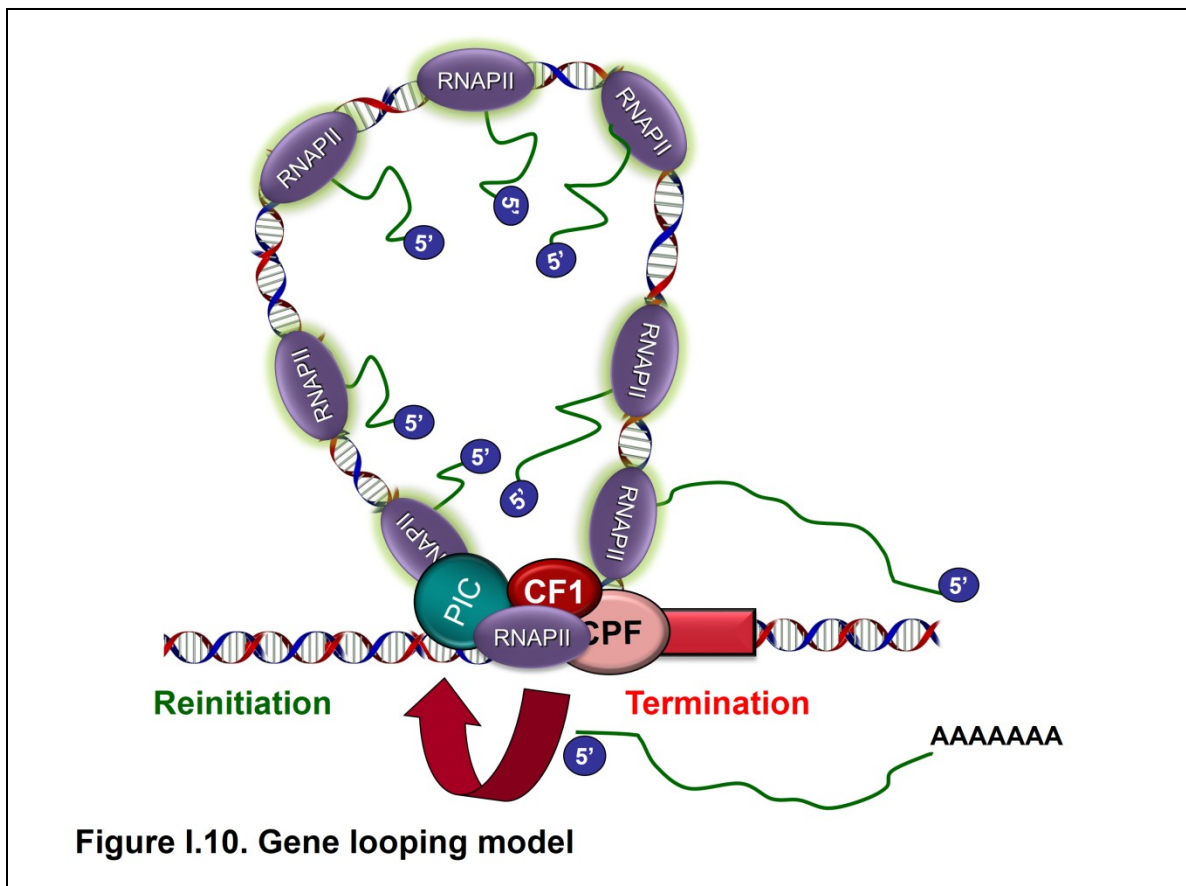
RNAi machinery was implicated in formation of this gene loop [145]. How RNAi machinery facilitates looping of PR gene needs further investigation. Type II collagen coding gene (Col2a1), mainly expressed in chondrocytes, has also been shown to form a loop between the 3' UTR and the promoter [147]. Since an enhancer element is located in the 3' UTR of Col2a1, it is not clear whether this gene loop is due to the promoter-terminator interaction or the enhancer-promoter interaction. In addition, the optimal expression of the RARβ2 gene requires the formation of a gene loop [148]. The

gene looping here required the nucleotide excision repair (NER) factors. Apart from yeast and mammalian systems, gene looping has also been found in *Drosophila melanogaster*, wherein the *polo* gene was found to form a TFIIB-dependent gene loop in a manner reminiscent of that in yeast [149]. Last but not least, gene looping was recently demonstrated in plants for the floral repressor gene *FLC*. [150]. Gene looping may be a general, possibly ubiquitous, transcriptional regulatory mechanism in eukaryotic systems.

Although the exact biological role of gene looping is not yet clear, it has been implicated in a variety of cellular functions including activation of transcription, repression of transcription, transcription memory, termination of transcription, directionality of promoter-initiated transcription and intron-mediated enhancement of transcription. One potential role of gene looping that is conserved across a spectrum of eukaryotic systems is the enhancement of transcription. The transcriptional activation of genes have been found to coincide with the formation of a looped architecture in yeast, HIV provirus, mammalian systems, flies and plants [120,132,137,139,140,142-144,147,150]. In budding yeast, it was demonstrated that the enhanced transcription was compromised in a looping defective strain [93]. In other organisms, however, it was not clear if transcriptional activation was the cause or the effect of gene looping. Further studies are needed to elucidate the role of gene looping in transcriptional stimulation. In yeast, it has been shown that gene looping accompanies activator-dependent transcription [120,143,144]. In the absence of gene looping, the kinetics of activator-dependent transcription exhibited a lag phase [143,144]. It has been proposed that the gene loop confers a sort of transcriptional memory to the gene [143,144,158]. Gene

looping has also been shown to play a crucial role in the Intron-Mediated Enhancement (IME) of transcription [159].   Introns fail to enhance transcription of a gene in the absence of gene looping, thereby suggesting that it is not the intron-facilitated splicing but intron-dependent gene looping that was responsible for activation of the gene. Gene looping does not always lead to activation of the gene. The looping of the mammalian *BRCA1* gene coincided with the repression of the gene in a transcription-dependent manner [141]. In *Drosophila*, gene looping was found to play a role in the termination of transcription and in coordinating the expression of tandem genes [149]. The transcription of *Drosophila polo-snap* tandem genes is negatively correlated. When *polo* is transcribed it forms a loop and this loop aids in keeping the downstream gene, *snap*, in a repressed state by masking its promoter by proteins involved in the *polo* loop formation. Consequently, the loss of polo gene looping exposes snap promoter elements, which leads to *snap* transcription.

Last but not least, gene looping was recently implicated in conferring directionality to promoter-associated transcription. It was a long standing question why the promoter-bound polymerase tends to move in only one direction resulting in sense transcription. A recent study suggests that it is gene looping that prevents polymerase from transcribing in anti-sense direction [7]. Thus, gene looping seems to play different regulatory roles in different cellular contexts.

Accordingly, our hypothesis is that gene looping facilitates rapid reinitiation events which increase the overall efficiency of transcription. Gene looping might execute this task by coupling termination to reinitiation. This coupling might result in a direct recycling of RNAPII from the terminator to the promoter; bypassing the rate

limiting step of de-novo RNAPII recruitment to the promoter. Transcriptional activation via coupling termination to reinitiation has been demonstrated for RNAPIII, RNAPI, mitochondrial RNA polymerase and archael polymerase [151,160-163]. There is some evidence indicating that a similar connection between termination and initiation exists during the RNAPII transcription cycle [164]. How termination was linked to initiation was, however, not clear in this study. The focus of this investigation is whether the termination-coupled reinitiation of transcription is dependent on gene looping. Gene looping-mediated enhancement of transcription may be through termination-reinitiation coupling. Recent evidence suggests that gene looping may also enhance transcription of a gene by conferring directionality to otherwise intrinsically bidirectional eukaryotic promoters [7]. Thus, gene looping mediated transcriptional activation may be the consequence of its effect on both termination-coupled reinitiation as well as promoter directionality.

Given its occurrence in several diverse systems such as yeast, plants, flies and mammals, and its involvement in different transcriptional regulatory contexts, elucidating the mechanism and the role of gene looping will add to our understanding of how a cell can program a rapid response to environmental and developmental signals.

I.5. DIVERGENT TRANSCRIPTION

The methodological breakthroughs in genomewide transcriptome profiling analyses such as high resolution strand-specific tiling arrays, strand-specific global run on, and RNA-Seq combined with advanced high-throughput sequencing approaches, revealed the bidirectional nature of the promoters of eukaryotic RNAPII transcribed genes in species as diverse as yeast and mammals [2,3,70,165-169]. Promoter

bidirectionality results in transcription happening divergently from the promoter region of RNAPII-transcribed genes. The high prevalence of divergent transcription supports the proposal of them having a role in transcriptional regulation.

In yeast, single gene as well as genomewide studies revealed the presence and prevalence of bidirectional promoters [3,168]. Promoter bidirectionality has been found to be the major source of the non-coding RNA (ncRNA) species called CUTs (cryptic unstable transcripts) [170,171]. CUTs are a principal category of RNAPII transcripts that are highly unstable due to RNA surveillance pathways in the cell, which explains their escape from detection in genomic assays. Immediately after synthesis, CUTs are targeted for degradation by the Nrd1-TRAMP-exosome complexes [4,170,172]. Accordingly, their detection was possible only when these surveillance mechanisms are compromised. Genome-wide studies utilizing different approaches have revealed that the CUTs are generally 200-600 nucleotides long and have heterogeneous 3` ends. They generally exist in low copy numbers, but are widespread within a cell. They originate predominantly from the nucleosome free regions (NFRs) that are the hallmark of most eukaryotic promoters. Although divergent transcription can happen as a consequence of the inherent nucleosome free nature of promoter regions which result in a chromatin architecture permissive for transcription, some evidence suggests that transcription contributes to formation of NFRs [2].

ChIP-Seq studies revealed enrichment of the chromatin marks for transcription initiation like H3K4me2, H3K4me3 and H3 and H4 acetylation, at the bidirectional promoters. The H3K79me3, H3K36me3 and H2B ubiquitination elongation marks, however, are found only downstream of the promoter in the sense direction

[70,165,173-175]. The lack of elongation marks in the antisense direction suggests that, although initiation occurs in both directions, elongation occurs exclusively in the sense direction.

I.5.1. TRANSCRIPTIONAL TERMINATION OF CUTs

RNAPII transcribes many different classes of ncRNAs including TSSa-RNAs, siRNA, miRNA, SUTs, XUTs, CUTs, snoRNA, snRNA, SRTs, and MUTs [2-7]. Transcription termination of SUTs and XUTs is believed to involve the poly(A)-site (PAS) dependent pathway [5,176,177]. In contrast, termination of CUTs, snRNA, snoRNA occurs through the Nrd1-dependent pathway [2,4,92,178,179]. Although the CPF and CFI complexes are integral components of the PAS-dependent termination pathway, recent evidence suggests their involvement in the Nrd1-dependent pathway as well [179]. The mechanism by which these three complexes act collectively to terminate the transcription of non-coding RNAs is not clear yet. Depletion of Clp1, a component of the CFI complex, resulted in defective transcription of all analyzed RNAPII transcribed genes including some CUTs ([92]. Whether the CFI and CPF complexes are generally required for the termination of CUTs needs further investigation [83,92,116,136,178,180-183].

There are some observations that support a potential role for CFI and CPF in terminating the promoter-associated anti-sense transcripts (CUTs). First, a number of studies revealed the presence of components of the CFI and CPF complexes and their mammalian counterparts, CPSF and CstF, in the vicinity of the promoter regions [67,93,132-136,184]. Second, two recent studies independently reported the presence of functional PASs near the promoter region [184,185]. Third, a number of studies

demonstrated a role for the components of the CFI complex in termination of some CUTs. Fourth, there are some indications that components the CFI and CPF complexes are required for transcriptional directionality. Taken together, these observations strongly suggest that the CFI and CPF complexes are involved in the termination of the promoter-initiated anti-sense transcripts, thereby restricting their elongation and providing directionality to the bidirectional eukaryotic promoters.

CHAPTER II

ROLE OF CFIB AND POLY(A) POLYMERASE (PAP1) IN TRANSCRIPTION AND GENE LOOPING

II.1. ABSTRACT

During transcriptional activation of a gene, the promoter and terminator regions of the gene physically interact with each other to form a looped structure. We have earlier demonstrated that this topology is formed by the interaction of the initiation and the termination machineries occupying the distal ends of a gene. We further found the general transcription factor TFIIB, the CFI subunits and the poly(A) polymerase (Pap1) exist in one megacomplex called 'holo-TFIIB complex' [93]. Earlier we examined the role of three CFIA subunits namely, Rna15, Rna14, and Pcf11 in promoter localization and gene looping [93,120]. Here, we extend this investigation to assess the role of the remaining CFI subunit Hrp1, which is generally referred to as the CFIB complex, as well as poly(A) polymerase (Pap1) in gene loop formation and other aspects of transcription. We found both Hrp1 and Pap1 crosslinked to the promoter as well as the terminator regions of a transcriptionally active gene. Utilizing the temperature sensitive mutants of Hrp1 and Pap1, we carried out ChIP (Chromatin Immunoprecipitation) and Chromosome Conformation Capture (CCC) analysis to determine the role of these factors in gene loop formation. Our results suggest that Pap1, just like CFIA subunits, contacts the 5' end of a gene. Since the presence of Pap1 at the distal ends coincided with gene looping, and looping was abolished in a mutant of Pap1, we conclude that Pap1 is required for gene looping. Hrp1 also crosslinks to the promoter end of a gene, but its role in gene loop formation needs further investigation.

II.2. INTRODUCTION

The role of the CFI complex and poly(A) polymerase in 3` end processing and termination has been the focus of intense investigation [38,93]. The role of four of the five subunits forming CFI complex, Rna14, Rna15, Pcf11 and Hrp1 in RNA processing and the termination of transcription is well established [48,72,186]. The function of Pap1 in the polyadenylation of mRNA has also been unequivocally demonstrated [48,187]. The CFI subunits Rna14, Rna15 and Pcf11 were found occupying the distal ends of a gene during transcriptionally activated state of a gene [93,120]. Accordingly, Rna14, Rna15, Pcf11 were found essential for gene looping [93,120]. The function of the remaining CFI subunits, Hrp1 and Clpl, as well as Pap1 in termination and gene looping, however, remained unclear. The role of Hrp1 and Clp1 as well as Pap1 in gene looping and aspects of transcription will be the subject of this and the following chapters. This chapter focuses on the role of Hrp1 and Pap1, whereas the role of Clp1 will be the subject of the next chapter (Chapter III).

Hrp1 is a heterogeneous ribonucleoprotein that binds the positioning element (PE) in the poly(A)-site through two internal RNA recognition motifs (RRMs) [188,189]. Hrp1 is a 73 KDa protein that was initially discovered as a suppressor of a temperature-sensitive mutant of NPL3, which is an RNA export protein [81,190]. This suggested a possible role of Hrp1 in mRNA nuclear export. The role of Hrp1 in RNA transport gained further support when it was found shuttling between the nucleus and the cytoplasm [186]. Chromatographic analysis found Hrp1 in the fraction containing RNA processing factor CFIB [73,190]. Later on, the CFIB RNA processing function was assigned to Hrp1 [186]. Hrp1 does not have a homolog in higher eukaryotes despite its essential role in

both cleavage and polyadenylation reactions in budding yeast [61,62,186]. In the context of its function as a 3' end processing factor, Hrp1 was found to bind to the PE and facilitate the correct positioning of the CFI complex on the nascent mRNA transcript. Both *in vivo* and *in vitro* studies have confirmed its function in cleavage as well as polyadenylation reactions [72,189]. The involvement of Hrp1 in gene loop formation however remains to be elucidated.

Pap1p is an essential and highly conserved protein required for the addition of the poly(A) tail to the 3′ end of mRNA [191-193]. In principle, Pap1 has a nonspecific polyadenylation activity, and can polyadenylate any RNA substrate *in vitro*. Its mRNA specificity in vivo is conferred by the CFI and CPF 3' end processing complexes. The identity of Pap1 as the poly(A) polymerase enzyme was determined twenty years after finding the polyadenylation activity in yeast cell extracts[191,193]. Pap1p was found dispensable for the mRNA cleavage and the termination steps, but essential for the polyadenylation reaction [48,187]. Although, Pap1 is a component of the CPF complex, it has also been found associated with the CFI complex [38,93]. The role of Pap1 in gene looping has not been tested.

Using temperature sensitive mutants, we found that Pap1 is required for gene looping of *INO1* or *MET16* while Hrp1 is not. However, both the factors crosslinked to the 5' end of the two genes. These results strongly suggest a role for pap1 in gene looping in budding yeast.

II.3. RESULTS

II.3.1. Role of Hrp1 and Pap1 in transcription of *MET16* and *INO1*

To have an insight into the physiological function of Hrp1 and Pap1 in transcription and associated gene architecture, we used the temperature-sensitive

Figure II.1. Role of Hrp1 and Pap1 in transcription of *MET16* and *INO1*
(A) and E) Schematic representation of *MET16* and *INO1* showing the positions of the primer pairs used in RT-PCR analysis. (B) and (F) RT-PCR analysis of *MET16* and *INO1* in *hrp1-5* and *pap1-1* as well as the isogenic wild type strains during induced and non-induced transcription at the 25°C and 37°C temperatures. (C) and (G) Quantification of the *hrp1-5* data shown in B and F respectively. (D) and (H) Quantification of the *pap1-1* data shown in B and F respectively. The transcript level of 18S was used as control. Black bars represent the transcription repression condition whereas gray bars represent the transcription activation conditions. *Met*, methionine; *Ino*, inositol. Error bars indicate one standard deviation.

mutants of Hrp1 and pap1 called *hrp1-5* and *pap1-1* respectively. *hrp1-5* carries a point mutation in the first RRM of Hrp1. The length of the poly(A) tail in this mutant is severely affected. The length of poly(A) tails was decreased in the *hrp1-5* mutant strain [186]. *pap1-1* mutant was kindly provided by Dr. Claire Moore and has been described in[187]. Both the mutant strains grow well at 25°C (permissive temperature), but stop growing following the transfer of cells to 37°C (non-permissive temperature).

To determine the effect of these mutations on transcription and gene looping we used *MET16* and *INO1* genes. *MET16* and *INO1* are involved in the biosynthesis of

methionine and inositol respectively [194,195]. We have previously used these two genes successfully in our analysis for a number of reasons. First, transcription of both these genes can be easily regulated by simply changing the growth medium. The transcription of *MET16* and *INO1* is activated in the absence of methionine and inositol respectively. Second, their transcription is well understood, and their transcription activators are well characterized. Third, their size and restriction site locations makes them a suitable candidate for performing CCC analysis.

To assess the effect of the mutant alleles, *hrp1-5* and *pap1-1*, on the steady state transcription level of *MET16* and *INO1* under their different transcriptional states, we carried out RT-PCR analysis of *MET16* and *INO1* under activation conditiones in the

Figure II.2. Increasing the deactivation time, the time at which the cells were grown at 37°C, for the *hrp1-5* mutant does not intensify its effect on transcription. RT-PCR analysis of *INO1* in *rna14-1*, *hrp1-5* and the isogenic wild type strains during induced transcription at different time points (1, 2, and 3 hours) at 25°C and 37°C temperatures.
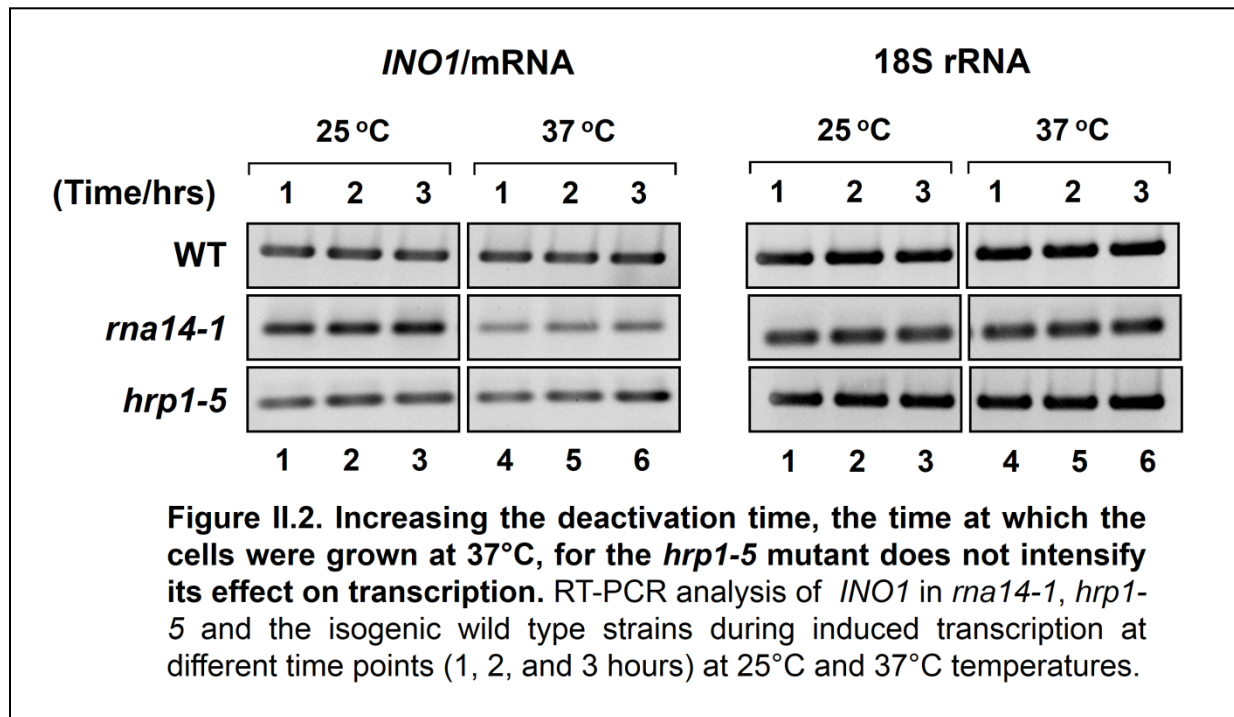
wild type and the mutant strains. RT-PCR analysis revealed only a very minor decrease in RNA levels of either *MET16* or *INO1* in *hrp1-5* cells following the temperature shift to 37°C (Figure II.1, B and F; lane 2; and panels C and G, gray bars). In contrast, *pap1-1*

cells showed a three to five fold decrease in the mRNA levels of the two genes upon shifting of cells to the non-permissive temperature (Figure II.1, B and F; lane 2; and *panels D and H, gray bars*). No such decrease in the transcript level of genes was observed in the isogenic wild type strain at 37°C (Figure II.1, B and F; lane 2). These results suggest that Pap1 is required for the optimal transcription of *MET16* and *INO1* while Hrp1 is not.

Given the previously documented defect in the length of the poly(A) tail in *hrp1-5* mutant [186], our RT-PCR results were surprising. To determine whether a longer deactivation time is needed to see the expected effect on the steady state transcript level of these genes, we performed a deactivation time course for *hrp1-5*. RT-PCR analysis of *INO1* was performed in *hrp1-5* and the isogenic wild type cells induced at 25°C or shifted to 37°C for 1, 2, and 3 hours (Figure II.2, Lanes 4, 5 and 6). A similar deactivation time course was performed for the *rna14-1* strain as a positive control. As expected, the *rna14-1* mutant displayed a dramatic decrease in the transcript level of *INO1* upon shifting of cells to the elevated temperature for 3 hours. No such decrease in the RNA level was observed in *hrp1-5* even after 3 hours of deactivation.

II.3.2. Effect of *pap1-1* and *hrp1-5* mutants on gene looping

Using the CCC approach, we have previously demonstrated the interaction of the promoter regions of *MET16* and *INO1* with their terminator sites in a transcription dependent manner [93,120]. Using this approach, we demonstrated the requirement of the Rna15, Rna14 and Pcf11 components of CFI complex in gene looping. In this approach, a PCR product obtained using divergent primers P1 and T1 is taken as a measure of gene looping (Figure II.3, A and E). A PCR product obtained using the F1

and R1 primers, a region lacking a restriction site, is taken as a control (Figure II.3, A

and E).

To examine the roles of Hrp1 and Pap1 in gene looping, CCC analysis was

carried out for *MET16* and *INO1* under induced conditions in *hrp1-5* and *pap1-1* strains

in the cells grown at the permissive and the non-permissive temperatures. A distinct P1-

T1 PCR signal was obtained for both *MET16* and *INO1* when *pap1-1* cells were grown

Figure II.3. Effect of *pap1-1* and *hrp1-5* mutants on gene looping
(A and E) schematic representation of *MET16* and *INO1,* respectively, indicating the positions of AluI restriction sites (*vertical lines*) and PCR primers (*arrows*) used in CCC analysis. (B and F) CCC analysis of *MET16* and *INO1* for detecting gene looping in W303-1a (wild type) and Hrp1 and poly(A) polymerase mutant strains (*hrp1-5*) and (*pap1-1*) following 120 min of induction followed by incubation at either permissive (25°C) or non-permissive (37°C) temperatures. *P1T1 PCR* reflects the looping signal, whereas control PCR represents the loading control. (C and G) quantification of the CCC results for *hrp1-5* shown in B *and* F, representing *P1T1 PCR* signal/*Control PCR* signal. (D,H) quantification of the CCC results for *pap1-1* shown in B *and* F, representing *P1T1 PCR* signal/Control PCR signal. *Error bars* indicate one standard deviation. *Met*, methionine; *Ino*, inositol. Error bars indicate one standard deviation.

at 25°C (Figure II.3, B and F, lane 1; and panels D and H, gray bars). Upon shifting the

cells to 37°C, P1-T1 looping signals decreased by about 5-fold for both

MET16 and *INO1* in the *pap1-1* strain (Figure II.3., B and F, lane 2; and panels D and

H, gray bars). No such diminution in looping signal was detected in the isogenic wild

type cells following the temperature shift to 37°C (Figure II.3., B and F, lane 2). When

the CCC experiment was repeated in *hrp1-5* mutant, no decrease in looping signal was

Figure II.4. Hrp1 and poly(A) polymerase occupy the distal ends of a gene in looped configuration. *A* and *D*, schematic depiction of *MET16* and *INO1* indicating the positions of ChIP primer pairs A, B, C, and D. *B* and *E*, quantification of the ChIP analysis data showing cross-linking of Hrp1 to different regions of *MET16* and *INO1* during repressed (*black bars*) and activated (*gray bars*) transcription. *C* and *F*, quantification of the ChIP analysis data showing cross-linking of Pap1 to different regions of *MET16* and *INO1* during repressed (*black bars*) and activated (*gray bars*) transcription. *Met*, methionine; *Ino*, inositol. *Error bars* indicate one standard deviation.

observed either for *MET16* and *INO1* following temperature shift to 37°C (Figure II.3., B

and F, lane 2; and panels C and G, gray bars). Thus, of the two factors investigated

here, we found only Pap1 essential for gene looping. Although we did not observe any

decrease in looping signal of genes in *hrp1-5* strain at elevated temperature, we cannot

make a conclusion regarding the role of Hrp1 in loop formation. To confirm the role of

Hrp1 in gene looping, we need to check other temperature-sensitive mutant alleles of

Hrp1.

II.3.3. Hrp1 and poly(A) polymerase occupy the distal ends of a gene in looped configuration

Next, we wanted to see whether Hrp1 and Pap1 are present in the vicinity of the promoter regions of *MET16* and *INO1*, as is the case with other CFI subunits. ChIP analysis was therefore performed to determine the presence of these two factors on different regions of *MET16* and *INO1* during the repressed and activated transcriptional states of genes. Our results show that both Hrp1 and Pap1 crosslinked to the terminator as well as the promoter regions of *MET16* during activated transcription (Figure II.4., panels B and C, regions A and D). Identical results were obtained with *INO1* (Figure II.4. d, panels E and F, regions A and D). Collectively, our previous and current chromatin immunoprecipitation studies suggest that in addition to their authentic position at the 3' end of genes, all the CFI subunits are also localized to the promoter region. This suggests a direct or indirect role for these subunits in the gene loop formation.

II.4. DISCUSSION

The results presented here demonstrate that both Hrp1 and Pap1 subunits of the previously reported holo-TFIIB complex are localized to the 5′ ends of genes only when they are in a looped conformation. Furthermore, we showed that the Pap1 plays a role in gene loop formation. Our results, however, were not conclusive regarding the role of Hrp1 in facilitating gene looping. We did not observe any decrease in the looping signal in the mutant strain for Hrp1. This does not completely rule out a role of Hrp1 in loop formation. It is quite possible that the point mutation in *hrp1-5* does not interfere with its role in gene looping or it even enhanced this activity. Other conditional mutants for Hrp1

might be needed to demonstrate conclusively the role of Hrp1 in promoter-terminator interaction. Our results also emphasize the connection between gene looping and transcription. Whenever a gene is found in looped conformation, an increase in transcription of the gene is observed. Whether gene looping is the cause or the effect of enhanced transcription remains to be elucidated.

CHAPTER III

A ROLE FOR CFIA 3` END PROCESSING COMPLEX IN PROMOTER ASSOCIATED TRANSCRIPTION

This chapter has been published:

Al Husini N, Kudla P, Ansari A (2013) A role for CF1A 3' end processing complex in promoter-associated transcription. PLoS Genet 9: e1003722.

III.1. ABSTRACT

The Cleavage Factor 1A (CFIA) complex, which is required for the termination of transcription in budding yeast, occupies the 3′ end of transcriptionally active genes. We recently demonstrated that CFIA subunits also crosslink to the 5′ end of genes during transcription. The presence of CFIA complex at the promoter suggested its possible involvement in the initiation/reinitiation of transcription. To check this possibility, we performed transcription run-on assay, RNAPII-density ChIP and strand-specific RT-PCR analysis in a mutant of CFIA subunit Clp1. As expected, RNAPII read through the termination signal in the temperature-sensitive mutant of *clp1* at elevated temperature. The transcription readthrough phenotype was accompanied by a decrease in the density of RNAPII in the vicinity of the promoter region. With the exception of TFIIB and TFIIF, the recruitment of the general transcription factors onto the promoter, however, remained unaffected in the *clp1* mutant. These results suggest that the CFIA complex affects the recruitment of RNAPII onto the promoter for reinitiation of transcription. Simultaneously, an increase in synthesis of promoter-initiated divergent antisense transcript was observed in the *clp1* mutant, thereby implicating CFIA complex in providing directionality to the promoter-bound polymerase. Chromosome Conformation Capture (CCC) analysis revealed a physical interaction of the promoter and terminator

regions of a gene in the presence of a functional CFIA complex. Gene looping was completely abolished in the *clp1* mutant. On the basis of these results, we propose that the CFIA-dependent recruitment of RNAPII onto the promoter for reinitiation and the regulation of directionality of promoter-associated transcription are accomplished through gene looping.

III.2. INTRODUCTION

The process of transcription can be divided into three principal steps; initiation, elongation and termination [33]. The accomplishment of each of these steps during the RNAPII-mediated transcription cycle requires a number of accessory factors. The initiation of transcription requires gene specific transcription factors as well as general transcription factors (GTFs): TFIID, TFIIB, TFIIA, TFIIF, TFIIE, TFIIH and Mediator complex, that assemble on the promoter to form the preinitiation complex [19,20,196,197]. The termination of transcription, which is intimately linked to the cleavage and polyadenylation of precursor mRNA, exhibits a similar requirement for a group of termination factors organized into two macromolecular complexes called Cleavage-Polyadenylation-Factor (CPF) complex and Cleavage Factor-1 (CFI) complex in yeast [54,61,62,75,198]. The initiation and termination factors have been remarkably conserved during evolution. The generally accepted view is that the initiation factors operate exclusively at the 5′ end of a gene and are committed to starting the transcription cycle, while termination factors have a dedicated role in ending the transcription cycle at the 3′ end of a gene. A number of recently published reports, however, challenge this dogma. It is evident that at least some initiation factors are also

necessary for termination, and the termination factors likewise may have a role in the

initiation or reinitiation step of the transcription cycle [75,199-202].

An increasing amount of biochemical, genetic and functional evidence suggests

the existence of a network of complex interactions between initiation and termination

Figure III 1. The amounts of the general transcription factor TFIIB and CF1A subunits Pcf11, Rna14, Rna15 are not altered in the *clp1ts* mutant at the non-permissive temperature. *clp1ts* mutant cells were grown at 25°C and then shifted to 37°C for 2 hours. Following cell lysis, Myc-tagged TFIIB or CF1A subunits were immunoprecipitated using anti-Myc antibodies from the cells grown at 25°C and 37°C. Equal amounts of protein were used in the immunoprecipitation. Immunoprecipitated proteins were fractionated on SDS-polyacrylamide gels and Western blot analysis was performed. (A) Western blot analysis for the cells grown at 25°C. (B) Western blot analysis for the cells grown at 37°C. Clp1 could not be detected in the cell lysate following the temperature shift to 37°C, while only a marginal change in the amounts of TFIIB and other CFIA subunits was observed in the mutant cells

factors. The general transcription factor TFIIB, for example, exhibits multiple genetic

and physical interactions with the factors operating at the 3′ end of genes

[93,103,120,203]. These studies suggested a plausible role for TFIIB in the termination

process. Accordingly, it was recently demonstrated that TFIIB is indeed actively

engaged in termination of transcription in mammals and flies [149,203]. Yeast Mediator

subunit Srb5, which has a well-established function in the initiation of transcription,

likewise, crosslinks to the 3′ end of genes and participates in the termination process [204]. TFIID is another promoter-bound protein that contacts the factors operating at the 3′ end of genes. Biochemical analysis of mammalian TFIID has revealed its reciprocal interaction with the CPSF 3′ end processing complex [130]. The TFIID-CPSF interaction is evolutionarily conserved. A recent proteomic analysis of yeast TFIID complex identified multiple interactions of TFIID subunit TAF150 with the components of the CPF 3′ end processing complex, which is the yeast homologue of CPSF complex [105,119,124].

Like initiation factors, an array of termination factors also crosstalk with the 5′ end of genes. The foremost among them is Ssu72, which was discovered as a protein of unknown function that genetically interacts with TFIIB [103]. Later on, yeast proteomic analysis identified Ssu72 as a component of the CPF 3′ end processing complex [115,205,206]. Ssu72 crosslinks to the 5′ end of genes, and interacts with several promoter-bound factors [104,107,110,116,119,132,136,207]. Pta1, which is a subunit of CPF complex, and Rat1 are other terminator-bound factors that physically interact with the 5′ end of genes and the associated initiation factors [132,133]. In addition to CPF complex, both 3' end processing and termination of transcription also require CFI complex. At least three subunits of this complex (Rna14, Rna15 and Pcf11) associate with both ends of a transcriptionally engaged gene [93,135]. CFIA subunits exhibit genetic and physical interaction with several promoter-bound factors that include both the general transcription factors and gene specific factors [74,93,104,108,111,114,120,126]. Furthermore, CFIA subunits are also required for juxtaposition of the promoter and terminator regions to form a looped gene structure

[93]. The well-orchestrated interaction of the distal ends of a gene strongly suggests that the termination and initiation steps of transcription may operate in a cooperative manner.

Figure III.2. Transcript levels of *INO1* and *CHA1* are adversely affected in the *clp1ts* mutant. (A, D) Schematic representations of *INO1* and *CHA1* showing the positions of the primer pairs used in RT-PCR analysis. (B) and (E) RT-PCR analysis of *INO1* and *CHA1* in *clp1ts* and isogenic wild type strains during induced and non-induced transcription at 25°C and 37°C. (C) and (F) Quantification of the data shown in B and E respectively. The transcript level of 18S was used as the control. The results shown are an average of four PCRs from two cDNA preparations from two different RNA extractions. Error bars indicate one unit of standard deviation. AS=Ammonium sulfate, ST=Serine/Threonine, WT=Wild type

The presence of termination factors on the promoter region could influence the events taking place at the 5′ end of genes. One possible role of the termination factors at the 5′ end could be to regulate initiation or reinitiation of transcription. It was recently demonstrated that proper termination of transcription is required for efficient execution of the transcription cycle in mammalian cells [164]. In that study, a termination defect adversely affected the recruitment of the general transcription factors onto the promoter of the same gene leading to a decrease in initiation of transcription. In a related study, a

decrease in the density of RNAPII at the promoter region was observed in the termination-defective Ssu72-C15S mutant [208]. One possible interpretation of these results is that proper termination is important for the recruitment of polymerase at the promoter for reinitiation. It is conceivable that the physical proximity of the promoter and

Figure III.3. The recruitment of the CF1A subunits onto the *INO1* and *CHA1* genes is adversely affected in the *clp1^ts* mutant at elevated temperature. (A, C) Schematic depictions of *INO1* and *CHA1* indicating the position of ChIP primer pairs. (B, D) ChIP analysis showing crosslinking of the CF1A subunits Rna14, Pcf11 and Rna15 to the 5' and 3' end of *INO1* and *CHA1* in the *clp1* mutant at 25°C and 37°C

terminator regions, which results in a looped gene conformation, facilitates a direct transfer of the released polymerase from the 3′ end to the juxtaposed promoter [132]. This would help bypass the rate-limiting step of recruitment of polymerase on the promoter, leading to enhanced transcription of the gene. A transfer of polymerase molecules from the terminator to the promoter has, indeed, been shown for RNAPIII-transcribed genes [151]. We propose that a similar termination-reinitiation coupling is

taking place during RNAPII-mediated transcription as well. Another possible function of termination factors at the 5' end of genes could be in providing directionality to the promoter-bound RNAPII to transcribe the sense strand. Genome wide analysis of human and yeast systems revealed the unexpected finding that RNAPII tends to transcribe both in the sense as well as anti-sense direction from the promoter region [2,3,70,168]. The promoter initiated anti-sense transcription, however, is aborted, thereby favoring productive elongation of the sense transcript. What confers directionality to the promoter-bound polymerase remains unclear. A recent study carried out in budding yeast demonstrated that the termination factors inhibit transcription of the promoter-initiated anti-sense transcripts, thereby providing directionality to the promoter-bound polymerase [7].

Here we demonstrate the role of CFIA complex in the promoter-associated transcription in budding yeast. In a mutant of Clp1 subunit of the CFIA complex, recruitment of the whole CFIA complex at the 3′ end of genes was compromised, leading to a termination defect. The termination defect coincided with a decrease in the recruitment of RNAPII on the promoter indicating a possible initiation defect. Since there was no significant decrease in the recruitment of the general transcription factors onto the 5′ end of a gene in the *clp1* mutant, these results strongly suggest a novel role for the CFIA complex in reinitiation of transcription. We further found a role for CFIA complex in the inhibition of promoter-initiated anti-sense transcription. Thus, CFIA complex may have an additional function in providing directionality to bivalent yeast promoters. The CFIA-dependent promoter-based events coincide with the gene

assuming a looped conformation, thereby suggesting a possible role of gene looping in reinitiation of transcription in the sense direction.

III.3. RESULTS

CFIA is a hexameric complex comprised of two subunits each of Rna14 and

Figure III.4. The recruitment of the CF1A subunits onto the *INO1* and *CHA1* genes remains unaffected in the wild type cells at elevated temperature. (A, C) Schematic depictions of *INO1* and *CHA1* indicating the position of ChIP primer pairs. **(B, D)** ChIP analysis showing crosslinking of the CF1A subunits Rna14, Pcf11 and Rna15 to the 5' and 3' end of *INO1* and *CHA1* in the *clp1* mutant at 25°C and 37°C. The input signals represent DNA prior to immunoprecipitation.

Rna15, and one subunit each of Pcf11 and Clp1 [77]. The Rna14, Rna15 and Pcf11 subunits have been studied extensively due to the availability of conditional mutant alleles. In contrast, little is known about the physiological role of Clp1. Recent studies, however, have implicated Clp1 both in the 3′ end processing of precursor mRNA and in the termination of transcription [89,91,92]. Structural analysis using mutants revealed

that Clp1 makes a direct physical contact with the Pcf11 subunit of CFIA complex as well as with the Ssu72 and Ysh1 subunits of CPF complex [77,89,91].

III.3.1. Clp1 is required for the recruitment of a termination-competent CFIA complex on transcriptionally active genes

To further analyze the role of Clp1 in transcription, we used a temperature-sensitive mutant of the factor called *clp1-769-5* [209]. Western blot analysis revealed that the Clp1 protein almost completely disappeared from the mutant cells following the temperature shift to $37^{o}C$, but there was only a marginal change in the signal for other CFIA subunits at the elevated temperature (Figure III.1). We examined the transcription of *INO1* and *CHA1* in the mutant *clp1* strain in cells grown at the permissive ($25^{o}C$) and non-permissive ($37^{o}C$) temperatures. The selection of *INO1* and *CHA1* was simply because their transcriptional regulation is well understood and can be induced under laboratory conditions. Furthermore, *CHA1* is relatively isolated in the yeast genome and therefore is a good candidate to study upstream and downstream transcription by transcription run-on (TRO) assay. RT-PCR was carried out using primers A and B as shown in Figure III.2, A and D in the mutant and wild type strains at $25^{o}C$ and $37^{o}C$. RT-PCR analysis revealed that the transcript level of both *INO1* and *CHA1* decreased by about 4-8 fold upon shifting the mutant cells to $37^{o}C$ (Figure III.2, B and E, lane 4; Figure III.2, C and F). No such decrease in transcript level was observed upon shifting the wild type cells to elevated temperature (Figure III.2, B and E, lane 4; Figure III.2, C and F). Thus, Clp1 is essential for optimal transcription of both *INO1* and *CHA1* in yeast. Since there was no appreciable decrease in the amount of CFIA subunits Rna14, Rna15 and Pcf11 in the mutant cells at the elevated temperature (Figure III.2), we next

checked if CFIA complex is recruited at the 3′ end of genes in the mutant cells. ChIP

analysis revealed that the recruitment of Rna14, Pcf11 and Rna15 at the 3′ end of *INO1*

and *CHA1* exhibited a decline following the temperature shift to 37°C (Figure III.3, B and

Figure III.5. TRO analysis showing transcription readthrough and alteration in promoter-associated transcription in the *clp1ts* mutant. (A) Schematic depiction of the *CHA1* gene indicating the positions of the probes (A-I) used in the TRO assay. (B) TRO analysis of *CHA1* in the wild type strain at 25°C and 37°C under induced transcription of the gene. (C) TRO analysis of *CHA1* in the temperature-sensitive mutant of Clp1 at the permissive (25°C) and non-permissive (37°C) temperatures. (D) and (E) Quantification of the data shown in B and C respectively. Error bars indicate one unit of standard deviation. The results shown are an average of at least four independent replicates. ST=Serine/Threonine, WT=Wild type, TRO=Transcription run-on

D, lanes 4, 12 and 20). No such decrease in the recruitment of CFIA subunits was

observed in the wild type cells at elevated temperature (Figure III.4, B and D, lanes 4,

12 and 20). The overall conclusion of these results is that the normal expression of

INO1 and *CHA1* is dependent on Clp1, and that the recruitment of a functional CFIA

complex at the 3′ end of these two genes occurs in a Clp1-dependent manner.

To understand the role of Clp1 in the transcription cycle, we performed

transcription run-on (TRO) analysis of *CHA1* in the wild type and temperature-sensitive

clp1-769-5 strains during different transcriptional states of the gene. The transcription of *CHA1* is regulated by the nitrogen source in the growth medium. The gene is maintained in a transcriptionally repressed state in a medium containing ammonium sulfate as the nitrogen source, and is stimulated upon shifting cells to a medium containing serine and threonine [210]. The position of transcriptionally active RNAPII was monitored at the positions A to I as shown in Figure III.5, A. The TRO analysis found transcriptionally active RNAPII being almost uniformly distributed between the promoter and the terminator regions of *CHA1* in the wild type strain during induced transcription (Figure III.5, B, lanes 3-7 and 13-17; Figure III.5, D). In the *clp1-769-5* mutant, however, the polymerase read through the termination signal into the downstream region at elevated temperature (Figure III.5, C, lanes 38 and 39; Figure III.5, E). No such transcription readthrough was observed in the mutant strain at the permissive temperature   (Figure III.5.C, lanes 28 and 29; Figure III.5.E) or in the wild type cells at 37°C (Figure III.5, B, lanes 18 and 19; Figure III.5, D). Strand-specific RT-PCR analysis corroborated the presence of sense transcripts downstream of the termination signal of *CHA1* in the *clp1* mutant at elevated temperature (Figure III.6, B, region Z). No such readthrough transcripts were observed in the isogenic wild type strain under identical conditions (Figure III.6, C, region Z). Strand-specific RT-PCR analysis was carried out using primers shown in Figure III.6, A and described in the figure legend. A logical conclusion of these results is that Clp1 is a termination factor in budding yeast.

III.3.2. A functional CFIA complex is required for reinitiation of transcription

Recently, we demonstrated crosslinking of Rna14, Rna15 and Pcf11 subunits of CFIA complex to the distal ends of genes in a transcription-dependent manner [93]. Here we show that the Clp1 subunit also localizes to both the 5′ and 3′ ends of transcriptionally

Figure III.6. Strand specific RT-PCR showing transcription readthrough and promoter-initiated anti-sense transcription in the *clp1ts* mutant. (A) Schematic representation of the *CHA1* locus indicating the positions of primers. (B) Strand-specific RT-PCR analysis of *CHA1* in the *clp1* mutant at the permissive (25°C, black bars) and non-permissive (37°C, grey bars) temperatures. (C) Strand-specific RT-PCR analysis of *CHA1* in the isogenic wild type strain at 25°C (black bars) and 37°C (grey bars). Error bars indicate one unit of standard deviation. TSS=Transcription start site, WT=Wild type, subscript s=sense and as=antisense

active *INO1* and *CHA1*(Figure III.7, B and E, lanes 1 and 4; Figure III.7, C and F). The CFIA complex, being a cleavage-polyadenylation factor, is expected to bind to the 3′ end of genes. It was, however, intriguing to find the entire CFIA complex occupying the 5′ end of genes as well. A clue regarding the role of the CFIA complex at the 5′ end of genes came when we observed that the transcription readthrough phenotype of the mutant strain at the elevated temperature was accompanied by a decrease in the TRO signal in the promoter-proximal coding region (Figure III.5., C, lane 33). This result

strongly suggested a role for Clp1 in the initiation/reinitiation of transcription. To determine if the observed decrease in TRO signal near the 5′ end of *CHA1* in the mutant was due to a failure to recruit RNAPII onto the promoter or due to a post-recruitment defect, we performed RNAPII density ChIP during the transcriptionally activated state of *INO1* and *CHA1* in *clp1-769-5* strain at permissive and non-permissive temperatures. RNAPII ChIP was performed using primer pairs A, B, C, D, E, and F as indicated in Figure III.8, A and D. Our results show that there was indeed a decrease in the density of RNAPII at the promoter region of both *INO1* and *CHA1* at elevated temperature (Figure III.8, B, lanes 1, 2 and Figure III.8, C, regions A, B; Figure III.8, E, lane 1 and Figure III.8, F, region A). There was no such decrease in the polymerase density at the promoter region of genes in the wild type cells at $37^{o}C$ (Figure III.9, B and E lanes 1 and 2; Figure III.9, C and F). The RNAPII-ChIP experiment revealed a nearly 2-fold decrease in the polymerase signal at the 5' end of *CHA1* in the mutant at $37^{o}C$ (Figure III.8, F, region A). In contrast, the TRO assay showed an at least 5-fold decrease in the polymerase signal near the promoter region of *CHA1* under identical conditions (Figure III.5.E, region C). This discrepancy could be attributed to the presence of transcriptionally inactive paused polymerase near the 5' end of *CHA1* that can be detected by ChIP assay, but not by the TRO assay. The overall conclusion of both the TRO and RNAPII-density ChIP results is that there is clearly a decrease in the amount of polymerase at the 5' end of a gene in the *clp1* mutant at elevated temperature. A plausible interpretation of these results is that a functional CFIA complex facilitates the recruitment of RNAPII onto the promoter during transcription. Next we asked if CFIA-dependent recruitment of RNAPII on the promoter

occurs during the initiation or reinitiation of transcription. The recruitment of TFIID, TFIIB, TFIIA, TFIIF, RNAPII, TFIIE and TFIIH occurs in this order in a sequencial manner during preinitiation complex (PIC) assembly [211,212]. The recruitment of RNAPII occurs subsequent to the formation of a TFIID-TFIIB-TFIIA complex on the promoter. This is followed by the binding of TFIIE and TFIIH to form the PIC. Following initiation of transcription, RNAPII along with TFIIF is released from the complex for

Figure III.7. Clp1 is recruited to the promoter and the terminator regions of transcriptionally active *INO1* and *CHA1*. (A) and (D) Schematic depictions of *INO1* and *CHA1* showing the positions of ChIP primer pairs. (B) and (E) ChIP analysis showing cross-linking of Clp1 to different regions of *INO1* and *CHA1* following 120 minutes of induction. (C) and (F) Quantification of the data shown in B and E respectively. Error bars indicate one unit of standard deviation. IP=Immunoprecipitation. Input= The input signals represent DNA prior to immunoprecipitation.

elongation [34]. Simultaneously, TFIIB is also released from the complex, while the rest of the general transcription factors are left behind on the promoter forming a 'scaffold' that is used as a loading dock for the re-entry of RNAPII for reinitiation of transcription during subsequent transcription cycles. The composition of protein factors on the promoter, therefore, can distinguish an 'initiation complex' from the 'reinitiation scaffold'

[202]. The initiation complex will contain all general transcription factors along with RNAPII, while the reinitiation scaffold will have general transcription factors with the exception of TFIIB and TFIIF and no RNAPII. Thus, to determine if CFIA-dependent recruitment of RNAPII was occurring during the initiation or reinitiation of transcription, we examined the promoter occupancy of *INO1* and *CHA1* for TFIID, TFIIB, TFIIF, TFIIE and TFIIH in *clp1-769-5* strain at the permissive and non-permissive temperatures by ChIP assay using primer pairs indicated in Figures III.10, A and C. Our results demonstrate that TFIID, TFIIB, TFIIF, TFIIE and TFIIH occupied the promoter region of both genes in the mutant at 25°C as well as 37°C (Figure III.10, B and D, region A black bar). Similar results were observed in the isogenic wild type strain (Figure III.13). TFIIB also occupied the terminator region of both genes at 25°C (Figure III, 10 B and D, region D grey bar for TFIIB-ChIP panel). The presence of TFIIB at the 3' end of genes is linked to CFIA-dependent gene looping [18]. A decrease in TFIIB signal near the 3' end of both *INO1* and *CHA1* was observed in the *clp1* mutant at 37°C (Figure III.10, B and D, region D grey bar for TFIIB-ChIP panel). This is in accord with the observed decrease in the TFIIB occupancy of the terminator region of transcriptionally active genes in the mutants of CFIA subunits [18]. A 25% decrease in the crosslinking of TFIIB and TFIIF to the promoter region of both *INO1* and *CHA1* was also observed in the mutant following the temperature shift to 37°C (Figure III.10, B and D, region A

grey bar). This is in agreement with the reported release of TFIIB and TFIIF from the promoter following initiation of transcription [34]. There was no appreciable change in the promoter occupancy of the rest of the general transcription factors following a shift

to elevated temperature, despite a decrease in the promoter-bound RNAPII signal. These results suggest that it is the reinitiation of transcription that is adversely affected in the *clp1-769-5* cells at elevated temperature. The overall conclusion of these results is that a functional CFIA complex is required for the recruitment of polymerase to the promoter for reinitiation of transcription. The possibility of CFIA complex being required for the recruitment of TFIIB and TFIIF for reinitiation cannot be ruled out.

III.3.3. CFIA complex limits divergent anti-sense transcription at the promoter

During the transcription cycle, RNAPII in the promoter-bound initiation complex transcribes in the sense direction, producing mRNA. Genome wide analysis of

Figure III.8. RNAP II density in the promoter region remains unaffected in wild type cells at the elevated temperature. (A, D) Schematic depictions of *INO1* and *CHA1* showing the positions of ChIP primer pairs. (B, E) ChIP analysis showing polymerase density in different regions of *INO1* and *CHA1* in the wild type cells at the permissive (25°C, black bars) and non-permissive (37°C, grey bars) temperatures. (C and F) Quantification of data shown in B and E respectively. The input signals represent DNA prior to immunoprecipitation. The results shown are an average of at least eight independent PCRs from four separate immunoprecipitations from two independently grown cultures. Error bars indicate one unit of standard deviation. IP=immunoprecipitate

transcribing polymerases has identified RNAPII molecules in the region just upstream of the transcription start site in most eukaryotic genes [2,3,70,168]. These upstream polymerases are involved in divergent anti-sense transcription, producing non-coding

RNA (ncRNA). These promoter-initiated, anti-sense ncRNAs are capped, non-adenylated, heterogeneous in size and often belong to a class of RNA called CUTs

Figure III.9. RNAP II density in the promoter region remains unaffected in wild type cells at the elevated temperature. (A, D) Schematic depictions of *INO1* and *CHA1* showing the positions of ChIP primer pairs. (B, E) ChIP analysis showing polymerase density in different regions of *INO1* and *CHA1* in the wild type cells at the permissive (25°C, black bars) and non-permissive (37°C, grey bars) temperatures. (C and F) Quantification of data shown in B and E respectively. The input signals represent DNA prior to immunoprecipitation. The results shown are an average of at least eight independent PCRs from four separate immunoprecipitations from two independently grown cultures. The input signals represent DNA prior to immunoprecipitation. Error bars indicate one unit of standard deviation. IP=immunoprecipitate

(cryptic unstable transcripts) that are rapidly degraded by the RNA surveillance mechanism of the cell [171,213]. Having already implicated CFIA complex in the sense-transcription of mRNA, we next asked if CFIA complex has a role in the regulation of divergent, anti-sense transcription of ncRNA. To address the issue, we performed strand-specific RT-PCR for *CHA1* in wild type and *clp1-769-5* mutant as described in [169]. In wild type cells, we could not detect promoter-initiated anti-sense transcripts under any condition (Figure III.3, C, region W). In the clp1-769-5 mutant also, no appreciable divergent anti-sense RNA could be detected at 25°C (Figure III.6., B, region W, black bar). At the elevated temperature, however, a 5-fold increase in the signal for

promoter-associated anti-sense transcripts was observed in the mutant strain (Figure

III.6, B, region W, grey bar). These results were corroborated by TRO assay, which
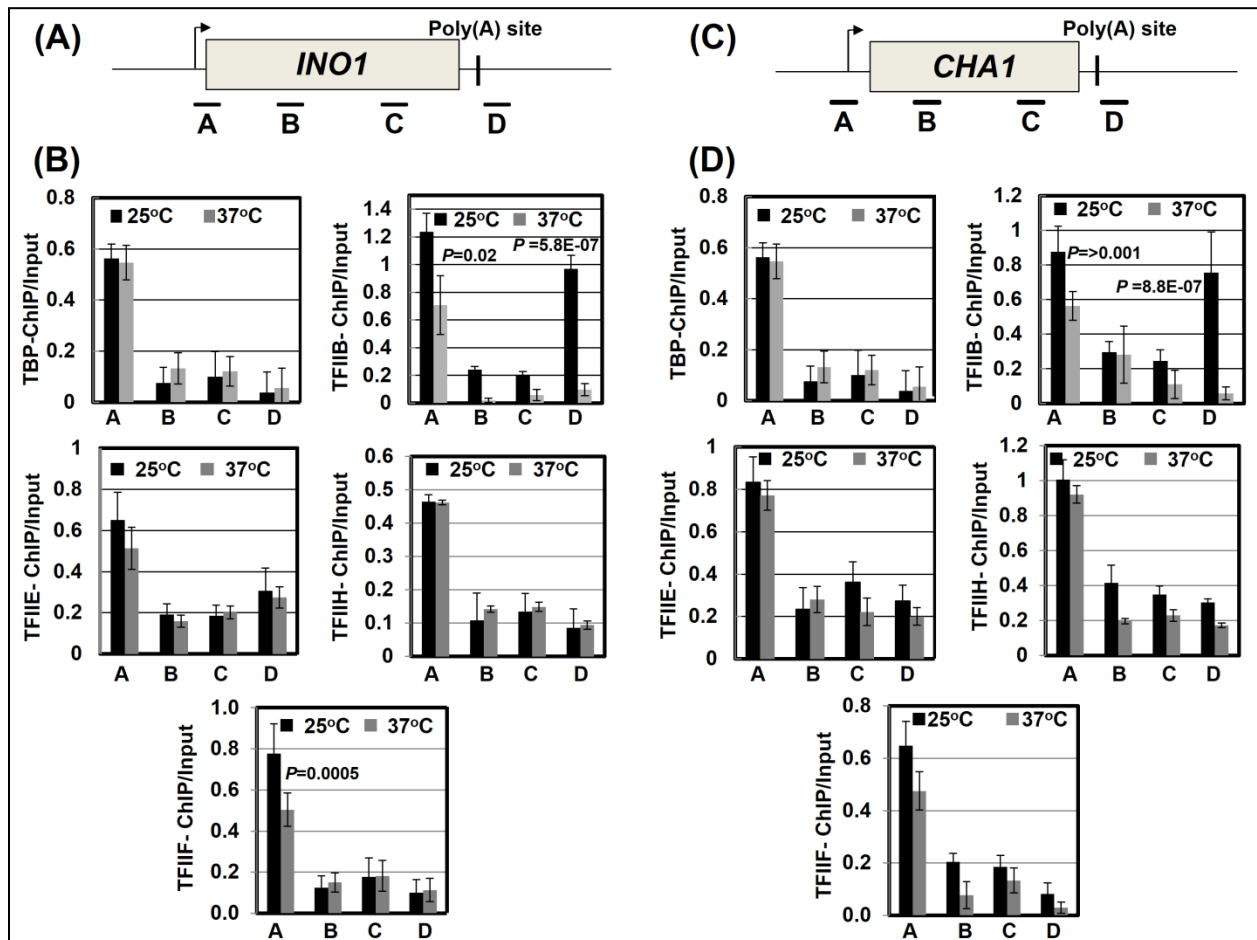
Figure III.10. The recruitment of the general transcription factors at the promoter of *INO1* and *CHA1* remains unaffected in the wild type cells at the elevated temperature. (A, C) Schematic depictions of *INO1* and *CHA1* indicating the position of ChIP primer pairs. (B, D) ChIP analysis showing crosslinking of the general transcription factors TFIID, TFIIB, TFIIF, TFIIE and TFIIH to different regions of *INO1* and *CHA1* in the wild type cells at 25ºC and 37ºC. The input signals represent DNA prior to immunoprecipitation. Error bars indicate one unit of standard deviation. IP=immunoprecipitate

detected the presence of transcriptionally engaged polymerase in the region upstream

of *CHA1* in the mutant strain at 37°C (Figure III.5, C, lane 31; Figure III.5, E region A).

The increase in the level of divergent anti-sense transcripts initiating from the 5' end of

the gene in the mutant could be attributed either to the stabilization of the transcripts or

to the synthesis of promoter-initiated anti-sense transcripts in the mutant. Since the
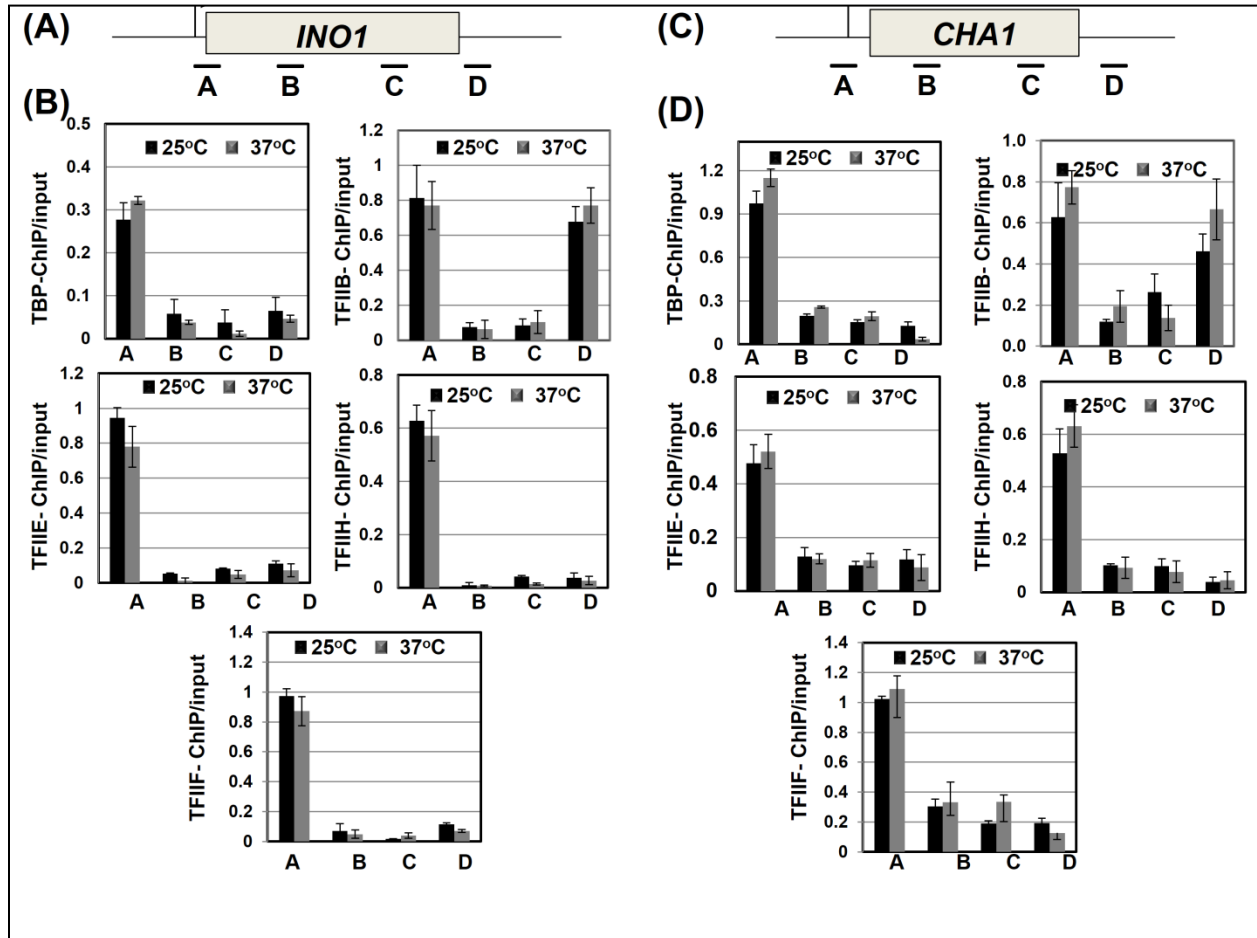
Figure III.11. The recruitment of the general transcription factors at the promoter of *INO1* and *CHA1* remains unaffected in the wild type cells at the elevated temperature. (A, C) Schematic depictions of *INO1* and *CHA1* indicating the position of ChIP primer pairs. (B, D) ChIP analysis showing crosslinking of the general transcription factors TFIID, TFIIB, TFIIF, TFIIE and TFIIH to different regions of *INO1* and *CHA1* in the wild type cells at 25°C and 37°C. The input signals represent DNA prior to immunoprecipitation. Error bars indicate one unit of standard deviation. IP=immunoprecipitate

TRO assay detected the presence of transcriptionally active RNAPII just upstream of

the promoter of *CHA1* in the *clp1* mutant at elevated temperature, it is reasonable to

conclude that the observed anti-sense transcripts were not the consequence of

stabilization of RNA, but the result of divergent anti-sense transcription initiating from

the 5' end of the gene. These results raise the possibility of a role for the CFIA complex

in limiting the transcription of promoter-associated anti-sense ncRNA, thereby favoring transcription of mRNA in the sense direction. We therefore propose that the CFIA complex may have an additional role in providing directionality to otherwise bidirectional yeast promoters. Our results are in agreement with a recent report that showed an increase in promoter-initiated divergent anti-sense transcription in termination-defective mutants [7].

Thus, in the absence of a functional CFIA complex in the *clp1-769-5* mutant, the promoter-associated downstream transcription of mRNA in the sense direction as well as the divergent upstream transcription of anti-sense RNA, exhibited an aberrant pattern.

III.3.4. A role for CFIA-dependent gene looping in promoter-associated transcription

A logical interpretation of the results described above is that the CFIA complex is not merely contacting the 5′ end of transcriptionally active genes, but is also influencing early events in the transcription cycle. Next we asked how the CFIA complex is recruited to the 5′ end of a gene. The binding of CFIA complex to the 5′ end could be independent of its recruitment at the 3′ end of a gene. Alternatively, gene looping, which is the transcription-dependent interaction of the promoter and the terminator regions of a gene, may facilitate positioning of the terminator-bound CFIA complex at the 5′ end of a gene [132]. We have earlier demonstrated the role of CFIA subunits Rna14, Rna15 and Pcf11 in gene looping [18]. To corroborate the role of CFIA complex in gene loop formation, we performed 3C analysis of *INO1* and *CHA1* in the *clp1-769-5* mutant at the permissive and non-permissive temperatures. Gene looping was monitored by the P1-

T1 primer pair shown in Figure III, 12, A and D, by the method described in [214]. A distinct P1-T1 PCR signal was obtained for both *INO1* and *CHA1* when the mutant cells were grown at 25°C (Figure III.12., B and E, lane 1; Figure III.12., C and F, black bar). The P1-T1 looping signal decreased by about 4-6 fold following transfer of cells to 37°C (Figure III.12., B and E, lane 2; Figure III.12., C and F, grey bar). These results confirmed that a functional CFIA complex is indispensable for gene loop formation in budding yeast.

III.4. DISCUSSION

The CFIA complex, which is known to localize and operate at the 3′ end of

Figure III.12. Gene looping of *INO1* and *CHA1* is compromised in *clp1ts* mutant. (A) and (D) Schematic depictions of *INO1* and *CHA1* indicating the position of restriction sites (vertical lines) and PCR primers (arrows) used in 3C analysis. (B) and (E) 3C analysis of *INO1* and *CHA1* to detect gene looping in the *clp1ts* mutant and the isogenic wild type strains at either permissive (25°C) or nonpermissive (37°C) temperatures. P1-T1 PCR reflects the looping signal while F1-R1 PCR represents the loading control indicating that an equal amount of template DNA was present in each of the 3C reactions. (C) and (F) Quantification of the 3C results shown in B and E. The results shown are an average of at least eight independent PCRs from four separate 3C replicates from two independently grown cultures. Error bars indicate one unit of standard deviation. WT=Wild type

RNAPII-transcribed genes in yeast, also contacts the 5′ end of genes. The promoter occupancy of the CFIA complex coincides with the gene assuming a looped

conformation. We recently purified a holo-TFIIB complex that contained all the CFI subunits and the general transcription factor TFIIB [93]. We showed that the holo-TFIIB complex mediates gene loop formation by simultaneously contacting the distal ends of a gene. Accordingly, gene looping was not observed in mutants of the Rna14, Rna15 and Pcf11 subunits of CFI complex. Here we show that gene looping is abolished in the *clp1* mutant as well. Whether the presence of CFIA at the 5′ end is the cause or the effect of gene looping is still unclear, but it is quite evident that the CFIA subunits at the 5′ end of a gene affect early events during the transcription cycle. The CFIA-dependent gene loop juxtaposes the terminator region of a gene with its cognate promoter. This

Figure III.13. A model showing the role of CF1-dependent gene looping in promoting reinitiation of sense transcription and in limiting promoter-initiated divergent anti-sense transcription.
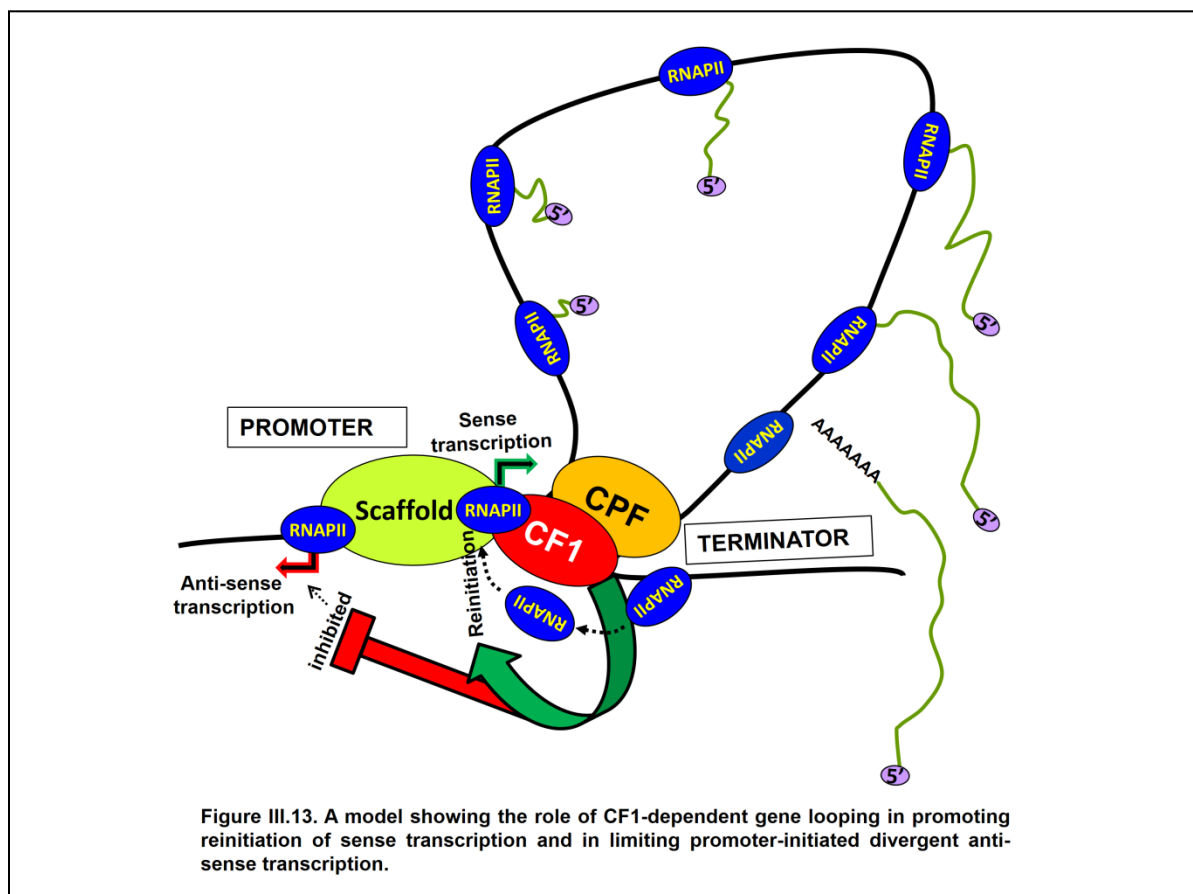
arrangement may facilitate binding of the RNAPII released from the terminator at the end of a transcription cycle to the promoter for starting the next round of transcription.

Accordingly, we observed a 2-fold decrease in the RNAPII density at the promoter in the absence of a functional CFIA complex. Since the promoter occupancy of the general transcription factors, with the exception of TFIIB and TFIIF, remained unaltered in the *clp1* mutant, we propose that the CFIA complex, by virtue of its role in gene looping, affects reinitiation rather than initiation of transcription. The possibility of CFIA subunits playing a role in the initiation, however, still cannot be ruled out. A similar study carried out in a mammalian system found termination factors affecting initiation rather than reinitiation of transcription [164]. The mechanism of termination-dependent initiation, however, was not clear in that study. Here we propose that the CFIA-dependent gene looping may account for the termination-reinitiation link.

Since a majority of eukaryotic promoters are intrinsically bidirectional, there should be some mechanism in the cell to favor transcription of mRNA in the sense direction, over the anti-sense transcription of ncRNA [2]. We found that the CFIA complex, while facilitating reinitiation in the sense direction, has an additional function in restricting transcription of the promoter-associated anti-sense RNA. The divergent, anti-sense transcription of ncRNA is widely believed to be terminated by the Nrd1-dependent pathway in yeast [4]. The CFIA complex, in general, is associated with the termination of mRNA synthesis by the poly(A)-dependent pathway [54,198]. Our results suggest that CFIA complex may be involved in the termination of anti-sense ncRNA synthesis as well. These results are in agreement with a recent report that demonstrated crosslinking of mammalian termination factors Xrn2 and TTF2 to the 5′ end of genes and their involvement in limiting promoter-initiated anti-sense transcription [177]. The regulation of transcriptional directionality by Ssu72, which is a subunit of the

CPF 3' end processing complex in yeast, further corroborates our results [7]. The limiting of promoter-initiated anti-sense transcription may direct the polymerase to move in the sense direction, thereby producing mRNA. Thus, CFIA complex may be involved in providing directionality to bivalent promoters.

Based on these results we propose a model of transcription by RNAPII (Figure III.13). The transcription-dependent promoter-terminator interaction places CFIA complex in the vicinity of the promoter. The promoter-bound CFIA affects transcription at two levels. First, CFIA-dependent termination releases RNAPII molecules from the 3′ end of gene near the promoter, thereby facilitating the recruitment of RNAPII to the promoter for reinitiation. Secondly, it provides directionality to the bidirectional promoter, thereby promoting the synthesis of mRNA over anti-sense ncRNA. Whether the CFIA complex limits promoter-initiated anti-sense transcripts by virtue of its termination activity needs further investigation. The net result is an upregulation of mRNA synthesis in the presence of a functional CFIA complex. Although a role for gene looping in facilitating transfer of polymerase from the terminator to the promoter for reinitiation has previously been hypothesized, this is the first instance where gene looping has actually been shown to help reinitiation of transcription.

CHAPTER IV

GENOMEWIDE ANALYSIS OF CLP1 FUNCTION IN TRANSCRIPTION IN BUDDING YEAST

IV.1. ABSTRACT

In budding yeast, both the 3' end processing of mRNA and the termination of transcription by RNAPII requires the function of the CFIA complex. The CFIA complex consists of four subunits, Rna14, Rna15, Pcf11, and Clp1. Unlike the first three subunits, the precise role of Clp1 subunit has not been thoroughly investigated. We have earlier demonstrated a role for this factor in transcriptional termination of the *CHA1* gene and found a decrease in the recruitment of the CFIA complex to the 3' end of this gene when its function is compromised. We further showed a function for this factor in promoter-associated transcription, wherein the function of Clp1 is required for the prevention and/or termination of the upstream antisense transcription initiated from the promoter of *CHA1*. To assess the generality of the observed functions of Clp1 in transcription, we tested the effect of Clp1 on transcription by RNAPII on a genomewide scale using the Genome Run-On (GRO-Seq) approach. Our results show a decrease in overall transcription when the function of this protein is compromised. Our results demonstrate a genomewide role for Clp1 in the termination of transcription, and further suggest that Clp1 is required for the RNAPII pausing that precedes the termination of transcription. Interestingly, we observed a dramatic increase in 3' initiated antisense transcription in the absence of a functional Clp1 protein. Additionally, the density of transcriptionally active polymerase at the 5' end of genes also exhibited a significant decrease in the Clp1 mutant at the elevated temperature. These results affirm the role of Clp1 in promoter-associated transcription on a genomewide scale.

IV.2. INTRODUCTION

The CFIA 3' end processing complex of budding yeast is composed of four subunits: Rna14, Rna15, Pcf11 and Clp1 [54,72]. Besides cleavage and polyadenylation, the complex has also been implicated in the termination of transcription [48]. The termination function of the CFIA complex was demonstrated using a nuclear run-on assay that revealed the readthrough of RNAPII beyond the poly(A) site in the mutants of all four CFIA subunits [48,215]. The transcription readthrough phenotype, however, has been demonstrated for just a few selected yeast genes. ChIP analysis identified all four subunits of the CFIA complex localized at the 3' end of selected genes in accordance with their role in 3' end processing and termination of transcription [93,215]. Genomewide analysis found Pcf11 crosslinked to the 3' end of a majority of transcriptionally active yeast genes [216]. This observation suggests that the CFIA complex may play a general role in the termination of transcription in budding yeast. More direct evidence in support of a role for the CFIA complex in the termination of transcription on a genomewide scale, however, is needed to firmly establish it as a general termination factor.

A vast majority of RNAPII-transcribed genes in yeast and higher eukaryotes exhibit anti-sense transcription initiating from the 5' as well as the 3' end of genes (Figure IV.1) [217,218]. In yeast, the 5' end initiated anti-sense transcripts are rapidly degraded by the RNA surveillance machinery of the cell and are therefore referred to as 'cryptic unstable transcripts' (CUTs) [7]. In contrast, the 3' end initiated anti-sense transcripts are stable and belong to the category of 'stable unannotated transcripts' (SUTs) [217]. At least 50% of SUTs in budding yeast are 3' end initiated anti-sense

transcripts. A number of genes in yeast and mammalian cells exhibit 3' end initiated antisense transcription under repressed conditions. The *GAL10* gene of budding yeast, for example, exhibits robust sense transcription in the presence of galactose [219,220].

Figure IV.1. A model of a gene showing the points of initiation of sense transcript, 5' initiated anti-sense transcript, and the 3' initiated anti-sense transcript.
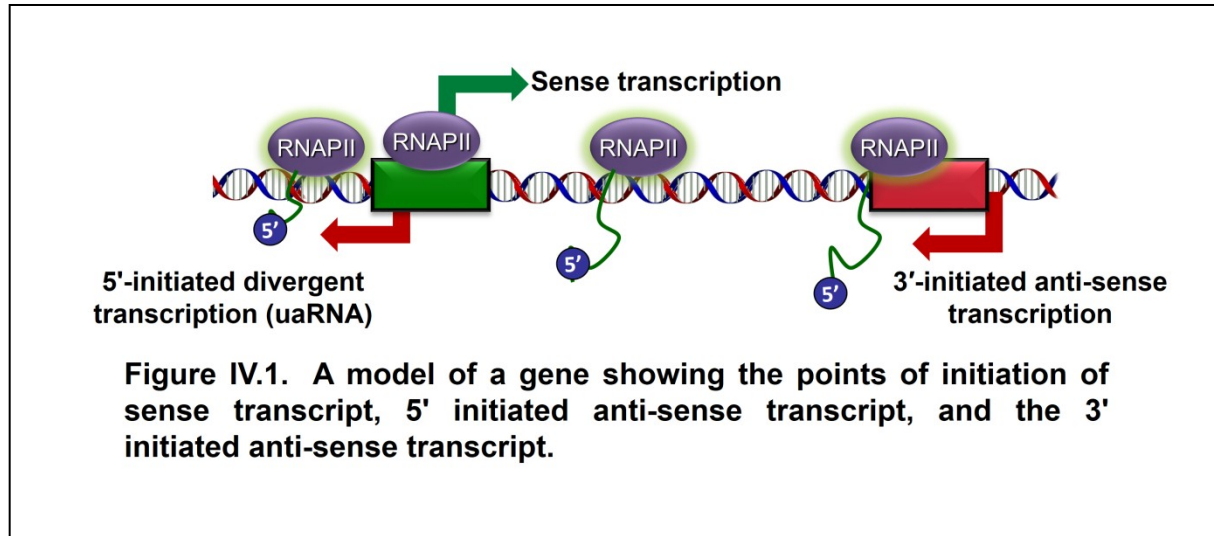
Upon shifting of the cells to glucose containing medium, however, sense transcription is almost completely inhibited, and the 3' end initiated anti-sense transcription

Figure IV.2. Growth curves for *clp1ts* and the isogenic wildtype strains at 25°C (A) and 37°C (B).
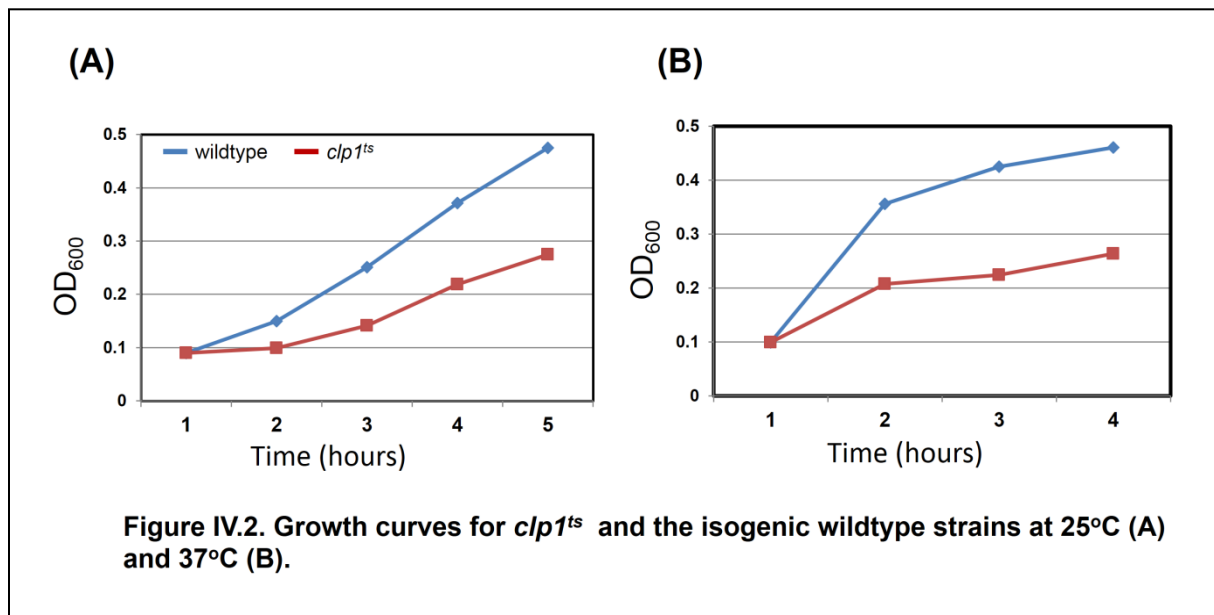
predominates. In mammalian cells, the assembly of the PIC at the promoter region of genes is adversely affected in the termination defective mutant cell lines [164]. Instead,

the assembly of the PIC was observed at the 3' end of genes. Such PICs formed at the 3' end of genes are capable of driving 3' end initiated anti-sense transcription. The initiation of sense transcription under these conditions is severely compromised. It has been proposed that when the 3' end initiated anti-sense transcript reaches the promoter end of the gene it may adversely affect initiation/reinitiation either through transcriptional interference or through histone modification [221-223]. Whether 3' end initiated anti-sense transcription is a general feature of termination defective mutants, however, is not clear. It is possible that proper termination is necessary to keep 3' end initiated anti-sense transcription under control. When the termination is compromised, anti-sense transcription from the 3' end is activated, which in turn adversely affects the initiation of transcription from the promoter.

Since the CFIA complex is involved in the cleavage-polyadenylation of mRNA and termination of transcription, it is expected to be present near the 3' end of genes. It was, however, intriguing to find all four subunits of the CFIA complex occupying the 5' end of genes as well [93]. Furthermore, the Ssu72 subunit of the CPF complex, which is also linked to 3' end processing and termination of transcription in budding yeast, has been localized to the promoter end of genes [132,224]. The presence of termination factors towards the 5' end of genes is an evolutionarily conserved feature as a number of 3' end processing/termination factors of higher eukaryotes also have been found crosslinked to the distal ends of genes [184]. The promoter occupancy of termination factors remained an enigma until recently when it was demonstrated that these factors provide directionality to transcription by limiting divergent antisense transcription initiating from the promoter region [7,184,215]. The majority of promoters of RNAPII-

transcribed genes in yeast and mammals are bidirectional. The transcription from these

promoters initiates in both the downstream sense direction as well as the upstream

Figure IV.3. Distribution of transcriptionally active RNA polymerase at the 3' end of protein coding genes of *clp1ts* cells grown at 25°C and 37°C. The GRO-Seq reads were aligned to the coding strand (blue for the 25°C, red for the 37°C) and the non-coding strand (green for the 25°C, fuchsia for 37°C). Shown are the reads aligned to +400/-400 bp window centered around the 3' end of genes. Genes with the 3' intergenic region of less than 700 bp were removed from the analysis. The *x axis (*Coverage) represents the number of reads that covers a nucleotide. The *Y axis* (Genome), represents the genomic coordinates and. showing a fixed window of -400/+ 400 centered around the 3' end of genes. The *p values* for the data comparing the number of reads obtained from the 25°C and 37°C libraries in both the sense and the antisense were less than <0.0001

antisense direction. The downstream sense transcription of the coding region produces

mRNA, while the upstream antisense transcription produces a ncRNA (called CUT in

yeast) that is rapidly degraded. Transcription in the upstream antisense direction is

terminated within a few hundred bases from the transcription start site by the

termination factors residing in the promoter region, while productive transcription in the

sense direction is allowed to proceed until the polymerase reaches the 3' end of the

gene. Thus, promoter linked termination factors provide directionality to the inherently

bidirectional eukaryotic promoters by limiting upstream antisense transcription [7,184]. A

recent report suggested that the CFIA complex may be involved in providing

directionality to bivalent yeast promoters [215]. Whether the CFIA complex has a general role in conferring promoter directionality during RNAPII-mediated transcription in budding yeast, however, needs further investigation.

Using a temperature-sensitive mutant of Clp1, we recently demonstrated the role of the CFIA complex in termination of transcription as well as in the promoter

Figure IV.4. Distribution of transcriptionally active RNA polymerase at the 3' end of protein coding genes of *clp1^{ts}* and WT cells grown at 37ᵒC. The GRO-Seq reads were aligned to the coding strand (blue for the wildtype, red for the mutant) and the non-coding strand (fuchsia for the wildtype and green for the mutant). Shown are the reads aligned to +400/-400 bp window centered around the 3' end of genes. Genes with the 3' intergenic region of less than 700 bp were removed from the analysis. The *x axis* (Coverage) represents the number of reads that covers a nucleotide. The *Y axis* (Genome), represents the genomic coordinates and shows a fixed window of -400/+ 400 centered around the 3' end of genes. The *p values* for the data comparing the number of reads obtained from the wildtype and the *clp1^{ts}* at 37ᵒC libraries in both the sense and the antisense were less than <0.0001

directionality of the *CHA1* gene in budding yeast [215]. Here we extend this study to show that the CFIA complex has a general role in the termination of transcription in budding yeast. Employing the GRO-Seq approach, we further showed that the CFIA complex suppresses anti-sense transcription initiating from the 3' end of genes. It is also involved in providing directionality to the promoter of a subset of yeast genes. This study will serve as a paradigm to probe the role of other 3' end processing factors of

yeast in the termination of transcription as well as in promoter directionality on a genomewide scale.

IV.3. RESULTS AND DISCUSSION

Figure IV.5. Representative genes showing 5' initiated anti-sense transcripts. (A) A representative gene showing 5' initiated anti-sense transcripts, the termination defect (readthrough phenotype and decreased sense transcript level. (B) A representative gene showing the 5' initiated anti-sense transcripts, and a decrease in sense transcription.
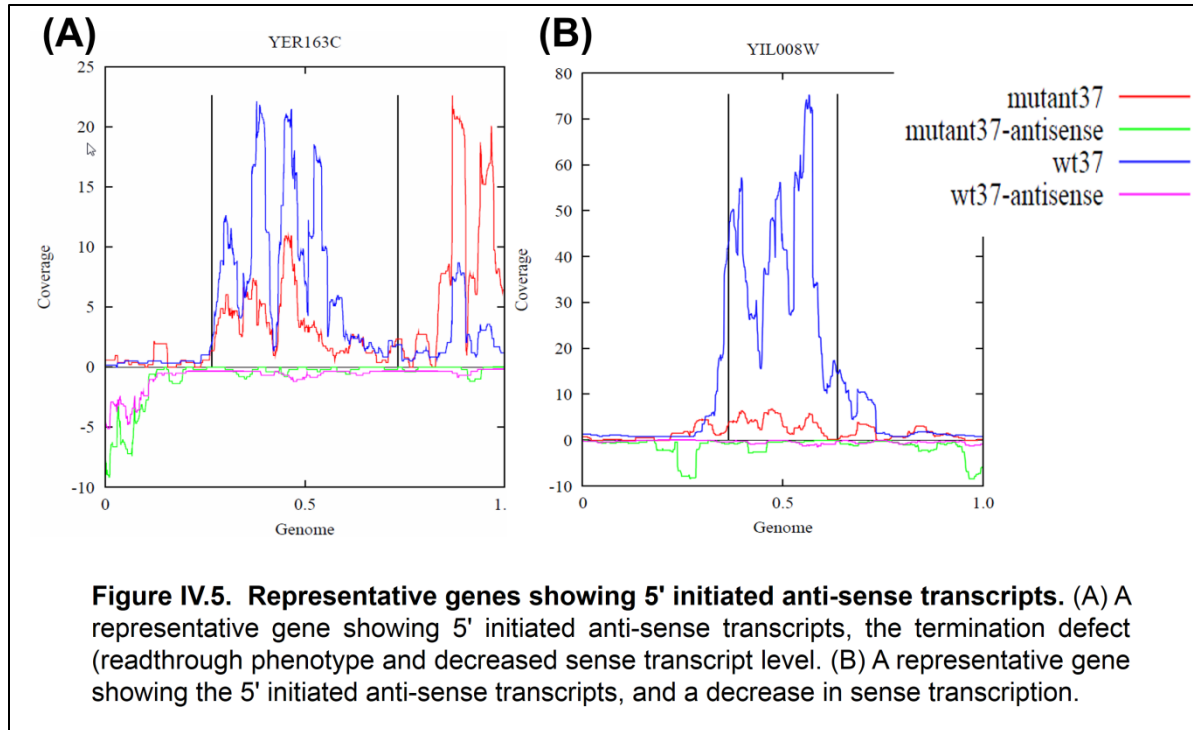
Of the three steps of the RNAPII transcription cycle (initiation, elongation and termination), initiation is the most well understood step. The general initiation factors are well characterized, and it is possible to perform initiation of transcription from a defined promoter using highly purified factors under *in vitro* conditions. In contrast, termination is the least understood step of the transcription cycle. The accessory factors required for termination are not thoroughly characterized. Consequently, termination of transcription under *in vitro* conditions using purified yeast factors has not been achieved so far. Like initiation, termination also requires *cis* acting DNA elements. The most common DNA element required for termination is the poly(A) site, which in higher eukaryotes is characterized by a conserved hexameric sequence, AAUAAA. Besides the poly(A) site,

other *cis* acting elements are required for efficient termination, both in yeast and in higher eukaryotes. Following transcription of the poly(A) site, RNAPII pauses (Figure I.3). A combination of biochemical and genetic approaches has identified two

Table IV.1. Number of transcribed genes in the clp1^ts mutant and the isogenic wildtype strains

Strain	Temperature	Replicate #	No. of transcribed genes
WT	25°C	1	5209
WT	37°C	1	5261
clp1^ts	25°C	1	4584
clp1^ts	37°C	1	3556
clp1^ts	25°C	2	4911
clp1^ts	37°C	2	3858

Table IV.2. Correlation coefficient of GRO-Seq analysis of *clp1^ts* replicates.

25°C-sense	25°C-antisense	37°C-sense	37°C-antisense
0.95	0.95	0.90	0.83

macromolecular complexes called CFI and CPF being essential for both 3' end processing and termination of transcription [43,52,54]. Phosphorylation of RNAPII CTD plays a crucial role in the recruitment of the 3' end processing/termination factors. The CFI complex is recruited to the termination site due to the interaction of Pcf11 subunit of the CFI complex with the serine-2 phosphorylated CTD [84]. The interaction of the Rna15 and Hrp1 subunits of the CFI complex stabilizes the association of the CFI complex with the elongating RNA. The Clp1 subunit of the CFI complex facilitates recruitment of the CPF complex to the site of termination [92]. This is followed by

cleavage and polyadenylation of the mRNA. The processed mRNA is released, and the Rat1 containing complex facilitates dissociation of polymerase from the template [133]. It has been shown that a mutation in any subunit of the CFI complex and some subunits of the CPF complex adversely affects termination [48]. When termination is defective, RNAPII does not pause and dissociate from the template beyond the poly(A) site, but reads through the poly (A) termination signal. This transcription readthrough phenotype is characteristic of defective termination, and has been widely used to determine the involvement of a factor in the termination of transcription.

IV.3.1. Clp1 is required for the termination of transcription on a genomewide scale

A number of experimental approaches can be used to detect the termination defect on a genome-wide scale [225]. The most popular among them are the Northern blot, RNA-Seq, RNAPII-ChIP-Seq, NET-Seq, and GRO-Seq. All these approaches may give similar results, but there are subtle differences in the precise information they reveal. Traditionally, and before the recent revolution in sequencing methodologies, two approaches were widely used to assess the termination defects; RNAPII-ChIP and the nuclear run-on (NRO) assay. The first is based on crosslinking, followed by shearing of the crosslinked chromatin, and purifying the RNAPII-bound DNA using antibodies specific for polymerase subunits. RNAPII-ChIP-Seq is a large scale version of this protocol, in which the tedious job of detecting the signal for individual genes is replaced by subjecting the coimmunoprecipitated DNA fragments to high-throughput sequencing [226]. This method, although effective in determining the location of RNAPII molecules in the genome, cannot distinguish transcriptionally active polymerase from the inactive molecules. In addition, it does not provide any information regarding the strand being

transcribed by the bound polymerase. The second traditional approach used for assessing the termination defect is the NRO assay [48]. This assay measures the density of transcriptionally active polymerase over a specific region of the genome. Briefly, this assay is done under conditions that prevent new initiation events while

Figure IV.6. Distribution of transcriptionally active RNA polymerase at the 5' end of protein coding genes of *clp1ts* cells grown at 25°C and 37°C.

The GRO-Seq reads were aligned to the coding strand (red for the 25°C, blue for the 37°C).Shown are the reads aligned to +400/-400 bp window centered around the 5' end of genes. Genes with the 5' intergenic region of less than 700 bp were removed from the analysis. The X and Y axis are as described previously.

allowing the engaged polymerase molecules to transcribe RNA in the presence of a radiolabeled nucleotide. The labeled nascent RNA is then allowed to hybridize to DNA probes on a nylon membrane. The probes are DNA fragments from the region under investigation. The intensity of the signal is detected using autoradiography. The genomewide version of this assay, called 'Global Run-On-Seq (GRO-Seq), overcomes the limitations in the NRO assay in genome coverage and the strand specificity [70]. In GRO-Seq, the newly synthesized transcripts incorporate BrUTP, which allow for the

affinity purification of these nascent transcripts on an anti-BrUTP column. The resultant nascent RNAs are subjected to high throughput sequencing. The outcome of GRO-Seq is the snapshot of the position as well as the density of the actively engaged polymerase in a strand specific manner. Another possible way of detecting the termination defects is the Native Elongating Transcript Sequencing (NET-Seq) approach [227]. In this technique, polymerase-bound nascent transcripts are coimmunoprecipitated and subjected to deep sequencing. This method allows the measurement of engaged polymerase. The major drawback of this method, however, is its inability to determine if the engaged polymerase is transcriptionally engaged or it is simply a paused or backtracked polymerase. The fourth method for detecting transcription defects is RNA-Seq [228]. In this method the whole RNA pool isolated from the cell is subjected to deep sequencing. Although this method can assess strand specificity and can be a very good indicator of the transcriptional state of the cell, it may not be able to detect transcripts with short half lives. To assess the role of the CFIA subunit Clp1 in transcription on a genomewide scale, we used the GRO-Seq approach, as it gives a snapshot of the position of transcriptionally engaged polymerase in the nucleus in a strand-specific manner. Comparing the GRO-Seq maps of the temperature-sensitive mutant of Clp1and the isogenic wild type strain at the permissive (25°C) and non-permissive (37°C) temperatures of the mutant can reveal if Clp1 is a universal termination factor like the general transcription factors, or its role in termination is restricted to a subset of genes. Accordingly, we performed GRO-Seq analysis in the Clp1 mutant and the isogenic wild type strains in the cells grown at 25°C and 37°C. The experiment with the Clp1 mutant was performed in duplicate. The

correlation between the replicates ranged from 83%-95% (Table IV.2). The number of transcriptionally active genes was determined by using an experimentally determined threshold of 25 reads per kilobase. We found that ~5200 of the 6693 annotated ORFs (~78 %) in the Saccharomyces Genome Database (SGD) were expressed as tags above the background (Table IV.1).

To investigate the role of Clp1 in the termination of transcription, we mapped

Figure IV.7. Distribution of transcriptionally active RNA polymerase at the 5' end of protein coding genes of *clp1^ts* and WT cells grown at 25ºC. The GRO-Seq reads were aligned to the coding strand (blue for the wildtype, red for the mutant). Shown are the reads aligned to +400/-400 bp window centered around the 5' end of genes. Genes with the 5' intergenic region of less than 700 bp were removed from the analysis. The X and Y axis are as described previously.

GRO-Seq reads to the annotated 3' end of genes. For this analysis, we removed all those genes whose next neighboring gene from the 3' end was equal to or less than 700 bp away. This was done because of the compact nature of the yeast genome, and often the terminator region of a gene overlaps with the promoter or terminator elements of the neighboring gene. We therefore restricted our analysis to 634 genes whose 3' end was at least 700 bp away from the neighboring ORF. We first compared the number of

genes transcribed in the Clp1 mutant at the permissive and non-permissive temperatures. Our results show that the number of genes transcribed in the mutant at $25^{o}C$ was reduced by about 25% upon shifting the cells to $37^{o}C$ (Table IV.1). This is in agreement with the observed decrease in growth rate and transcription in the mutant at the non-permissive temperature (Figure IV.2). Next we compared the transcription readthrough phenotype in the mutant at $25^{o}C$ and $37^{o}C$. We found that there was no decrease in the number of reads beyond the 3' end of genes upon shifting the cells to $37^{o}C$ (Figure IV.3). Furthermore, the number of reads beyond the 3' of the genes was more or less the same at the permissive and non-permissive temperature (Figure IV.3). Thus, the mutant was showing a readthrough phenotype even at the permissive temperature. It is not uncommon for temperature-sensitive mutants to exhibit the defective phenotype even at the permissive temperature. We therefore compared the readthrough phenotype in the mutant and the wild type strain at $37^{o}C$. The wild type strain exhibited a distinct peak towards the 3' end of genes (Figure IV.4). The peak was followed by a sharp drop off beyond the presumed polyadenylation signal. In contrast, there was neither a peak in the corresponding region at the 3' end, nor was there a decrease in the number of reads beyond the polyadenylation signal in the mutant at $37^{o}C$ (Figure IV.4). A logical interpretation of these data is that the polymerase reads through the termination signal in the mutant, but not in the isogenic wild type strain, at $37^{o}C$.

We draw two important conclusions from these results. First, Clp1 is a general termination factor. Of the 634 genes analyzed here, a majority of genes exhibit the readthrough phenotype in the Clp1 mutant at $37^{o}C$. No such readthrough was observed

in the wild type strain at 37°C. Second, Clp1 is required for the pausing of RNAPII near the 3' end of genes. It is generally believed that RNAPII pauses while transcribing the polyadenylation signal, and that this pausing is essential for the subsequent cleavage-polyadenylation reaction that precedes termination [54]. Here we show pausing of the polymerase towards the 3' end of genes on a genomewide scale in the wild type cells. No such pausing was observed in the mutant at 37°C. On the basis of these results, we propose that Clp1 is one of the factors that contribute to the pausing of polymerase before the termination of transcription in budding yeast. Since there is no report of Clp1 acting alone or being a part of any complex except for CFI, we extrapolate these results to propose that the CFI complex is an essential termination factor in budding yeast, and that it contributes to the pausing of the polymerase near the 3' end of genes in budding yeast.

IV.3.2 Clp1 regulates anti-sense transcription

Transcriptome analysis in yeast and higher eukaryotes has revealed that a vast majority of transcripts do not code for any protein [9]. These non-coding transcripts (ncRNAs) are the result of pervasive transcription. The anti-sense transcripts are a type of ncRNA that are produced from the anti-sense strand of annotated genes [217]. The anti-sense transcripts initiate from both the 5' end as well as the 3' end of genes (Figure IV.1). Those initiating from the 5' end are called pstream antisense RNA (uaRNA) [2]. A vast majority of RNAPII promoters in yeast as well as higher eukaryotes are bidirectional [2,3]. The transcription initiates from such promoters both in the sense as well as the anti-sense direction (Figure IV.1). The anti-sense transcripts are terminated when they are just a few hundred nucleotides long, while the sense transcription of the

coding region is allowed to proceed until the polymerase reaches the 3' end of the gene [168] [70,184]. Recent studies have revealed that the 3' end processing/termination machinery of the cell confers promoter directionality by terminating the transcription of uaRNA when it is just a few hundred nucleotides long [7,184,215]. The resultant uaRNA is extremely unstable, and is immediately degraded by the RNA surveillance machinery of the cell [168]. It has been demonstrated that gene looping plays a crucial role in the termination of uaRNA synthesis in budding yeast [7,215]. In the absence of gene looping, promoter directionality is compromised and longer uaRNA fragments are detected in the cell. We showed that Clp1 is essential for both the gene looping and the promoter directionality of the *CHA1* gene in budding yeast [215]. To determine the generality of the role of Clp1 in the termination of uaRNA synthesis, and thus the maintenance of promoter directionality in budding yeast, we mapped GRO-Seq reads of the mutant and the wild type cells grown at $37^{o}C$, to the annotated 5' end of genes. For this analysis, we choose only those genes whose 5' end was at least 700 bp from the neighboring genes. We therefore restricted our analysis to 1247 genes. We were unable to detect a distinct peak of uaRNA in the mutant at $37^{o}C$ in the metagene analysis. A small peak of uaRNA however was detected in about 10% of genes (Figures IV.5.A and IV.5.B). We believe that the inability to detect uaRNA signal in the mutant is due to the unstable nature of uaRNA. As mentioned before, uaRNA belongs to the category of ncRNAs called CUTs, which are extremely unstable and are immediately degraded by the yeast exosome. To understand the role of Clp1 in uaRNA synthesis, the GRO-Seq xperiments need to be repeated in Clp1 mutant strain that are deleted for the exosome component Rrp6. Deletion of Rrp6 will stabilize uaRNA in the Clp1 mutant. This strategy

was successfully used by the Proudfoot laboratory to determine the role of the Ssu72 component of the CPF complex in uaRNA synthesis in budding yeast [7].

Next, we examined the role of Clp1 in anti-sense transcription originating from the 3' end of genes. As mentioned above, we choose 634 genes for this analysis whose 3' ends were at least 700 bp away from the neighboring gene. Alignment of GRO-Seq reads to the 3' end of these genes revealed robust anti-sense transcription in the Clp1 mutant at 37°C (Figures IV.3 and IV.4). No such 3' end anti-sense transcript peak was observed either in the mutant at 25°C, or in the wild type strain at 37°C (Figures IV.3 and IV.4). It is difficult to say if the 3' initiated anti-sense transcription is the cause or the consequence of the defective termination. It is, however, known that a number of inducible genes in yeast carry a promoter structure for anti-sense transcription at their 3' end [219]. Such promoters are the mirror images of the canonical promoter at the 5' end of genes. These antisense promoters are repressed when the gene is transcriptionally active, but get activated when the gene is inactive [217]. Thus, a number of inducible promoters in yeast exhibit 3' end initiated anti-sense transcription under non-inducible conditions. Upon induction of the gene, however, anti-sense transcription is completely inhibited. It has been proposed that when such 3' end initiated anti-sense transcripts reach the promoter element at the 5' end of gene, they somehow repress initiation of sense transcripts by a mechanism that is not yet fully understood. On the basis of these observations, we propose that under normal conditions, when termination is efficient, the 3' initiated anti-sense transcription is repressed. When the termination is defective, for example in the Clp1 mutant at 37°C, 3' initiated anti-sense transcription is activated. When the polymerase engaged in anti-sense transcription reaches the promoter region

of the gene, it tends to repress the initiation of transcription, resulting in an overall decrease in the transcription of the gene.

Thus, both 5' end initiated and 3' end initiated anti-sense transcription tend to have an inhibitory effect on the sense transcription of the gene. An investigation into the mechanism of inhibition by anti-sense transcription will further contribute to our understanding of the biological role of anti-sense transcription in the cell.

IV.3.3 Clp1 regulates promoter-associated sense transcription

We have previously demonstrated that Clp1 affects reinitiation of transcription of the *CHA1* and *INO1* genes. A significant decrease in the density of RNAPII at the promoter region of both genes was observed in the Clp1 mutant at non-permissive temperature (Figure III.4). To determine if the Clp1 is required for promoter-associated sense transcription, we mapped the reads at the 5' end of genes in the Clp mutant at $25^{o}C$ and $37^{o}C$ as described above.The results show about a 27% decrease in the GRO-Seq signal in the mutant following a temperature shift to $37^{o}C$ (Figure IV.6). The mutant exhibited a decrease in promoter-associated polymerase even at $25^{o}C$ (Figure IV.7). These results strongly suggest that the role of Clp1 in transcription is not limited to the 3' end of the gene, but extends to the 5' end of genes as well on genomewide scale.

IV.3.4. Conclusion

Genomewide analysis of the role of Clp1 in transcription revealed several important findings. First, the data showed that Clp1 is a general termination factor. Second, Clp1 may be bringing about termination by facilitating pausing of polymerase near the poly(A) signal as the terminator-proximal peak of polymerase was not observed in the clp1 mutant. This pausing of the polymerase is considered a

prerequisite for the termination of transcription. Third, Clp1-mediated termination of transcription represses the 3'-initiated antisense transcripts. This finding is in accordance with the previous studies that showed inhibition of 3' initiated antisense transcription in the termination defective mutants. Fourth our data showed that Clp1 has an additional role in promoter-associated transcription. In the absence of a functional Clp1 protein, the density of transcriptionally active polymerase decreased by 27%. These results suggest that Clp1 is required for either initiation or reinitiation or both. Furthermore, Clp1 may also confer promoter directionality by limiting promoter initiated divergent transcription of upstream antisense RNA. Performing the GRO-Seq approach in an exosome mutant background will be required to investigate the role of this termination factor in inhibiting the 5'-initiated antisense transcription. Since Clp1 is essential for maintaining the integrity of the whole CFIA complex, these findings reflect the role of this complex rather than that of Clp1 alone.

APPENDIX A: EXPERIMENTAL PROCEDURES

A.1. CELL CULTURE

Cultures were started by inoculating 5 ml of YP-dextrose medium with colonies from a freshly streaked plate, and grown at 25°C overnight with constant shaking. Next morning, overnight grown cultures were diluted (1:100 dilution for the temperature-sensitive strains, and 0.5:100 dilution for the wild type strains) to an appropriate volume and grown to OD_{600}~0.4. The dilution was done in the appropriate synthetic complete-drop out medium. Induction was done for 2 hrs at 25 °C before shifting the cells to 37 °C for another 2 hours for the deactivation step. Usually, this takes the cells to OD_{600} of about 0.7-0.8. At this stage, the cells are ready for processing for RT-PCR, 3C, ChIP, or TRO assays.

A.2. TRANSCRIPTION RUN-ON ASSAY (TRO)

Transcription run-on (TRO) assay was performed by the modification of protocols described in Birse et al., 1997 and Hirayoshi and Lis, 1999 [229,230]. For *CHA1*, WT and *clp1-769-5* cells were grown in 100 ml of synthetic complete medium containing ammonium sulfate until A_{600} reached 0.4. Cells were centrifuged and resuspended in 100 ml of synthetic media containing serine and threonine (1 g/l each) and induced for 2 hours at 25°C. 50 ml of the cultures were centrifuged and resuspended in 50 ml of pre-warmed (37°C) serine and threonine containing medium and deactivation was done at 37°C for 120 minutes. The cell pellet obtained from 50 ml of liquid culture was washed with 10 ml cold TMN buffer (10 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 100 mM NaCl) and resuspended in 940 µl of DEPC (Diethylpyrocarbonate)-treated cold water. To the cell suspension, 60 µl of 10% sarkosyl was added and incubation performed on ice for 25

min to permeabilize the cells. Permeabilized cells were recovered by a low-speed centrifugation (1.2xg, 6 minutes) and directly used in the run on transcription assay. Elongation of transcripts initiated *in vivo* was resumed by resuspending cells in 120 µl of 2.5X reaction buffer (50 mM Tris-HCl pH 7.5, 500 mM KCl, 80 mM $MgCl_2$, 5 mM DTT), 45 µl of NTPs/RNase inhibitor mix (10 mM each of CTP, ATP, and GTP and 300 units of RNase Inhibitor), and 7 µl of [α-^{32}P]-UTP (3000 Ci/mmol, 10 µCi/µl). The reaction mix was incubated at 30°C for 2 minutes to allow transcript elongation. The reaction was stopped by adding 1 ml of cold TMN buffer and quickly spun at low speed. The recovered pellet was resuspended in 350 µl of Trizol. About 250 □l of acid-washed glass beads were added and the cells were lysed by vigorous shaking for 5 minutes on an agitator at room temperature. After lysis, tubes were spun for 5 minutes at 13800xg. To the recovered supernatant, 700 µl of Trizol and 200 µl of Chloroform were added and the samples were vigorously shaken on a vortexer, left on the bench for 5 minutes, and centrifuged at high speed for 10 minutes.

To isolate RNA, the supernatant was extracted twice with phenol/chloroform (pH 4.2). Labeled RNA was precipitated by adding 0.1 volumes of 10 M LiCl, 0.1 volumes of yeast tRNA (80 mg/ml) and 2.5 volumes of 100% ethanol. The mix was incubated at −20°C for 20 minutes followed by centrifugation at maximum speed for 15 minutes. The RNA pellet was resuspended in 60 µl of DEPC-treated water and denatured by adding 5 µl of 2 M NaOH followed by incubation on ice for 5 minutes. The NaOH was then neutralized by adding 12 µl of sodium acetate/acetic acid mix (0.3 M sodium acetate pH 5.2 and 0.5 µl of glacial acetic acid) and boiling the contents for 5 minutes.

In parallel, DNA probes of about 200-300 bp each in length, spanning the desired regions of the *CHA1* gene, including the upstream and downstream regions, were obtained by PCR amplification (See Fig. 1A for the position of probes). 10 µg of probe was denatured by boiling in 0.1 N NaOH and 1 mM EDTA for 10 minutes to form single stranded DNA. The heat-denatured probes were then slot-blotted a ZETA-probe GT membrane (*BIO-RAD*), according to manufacturer's instructions. Adsorbed DNA was crosslinked to the membrane by baking at 80°C in a vacuum oven for 30 minutes. The membrane was then prehybridized with 10 ml of hybridization solution (0.5M potassium phosphate pH 7.2, 7% SDS) at 55°C for at least 30 minutes. The denatured RNA in hybridization solution from the step described above was added to the prehybridized membrane. Labeled RNA was allowed to hybridize to the probe for 18-24 hours at 55°C in a rotator. After hybridization, the membrane was washed twice with 20 ml of a solution containing 0.1% SDS and 1XSSC for 7 minutes at 55°C, and twice with 20 ml of a solution containing 0.1% SDS and 0.1XSSC for 7 minutes at 55°C. After drying, the membrane was exposed to X-ray film overnight in an autoradiography cassette and the films were developed in a Kodak M35A X-OMAT system. All TRO signals were quantified using the GEL LOGIC 200 (KODAK) system and normalized with respect to the 18S control.

A.3. CHROMATIN IMMUNOPRECIPITATION (ChIP)

ChIP was performed as described in [120]. Primers used for ChIP-PCR are described appendix C. RNAPII ChIP was performed using anti-Rpb3 antibodies obtained from Santa Cruz (Cat# sc-101614). For ChIP analysis of CFI subunits Clp1, Rna14, Rna15 and Pcf11, a Myc-tag was inserted at the carboxy-terminus of each

subunit, and ChIP was performed using anti-Myc antibodies obtained from Upstate

Figure A.1 Chromosome Conformation Capture (CCC) approach.
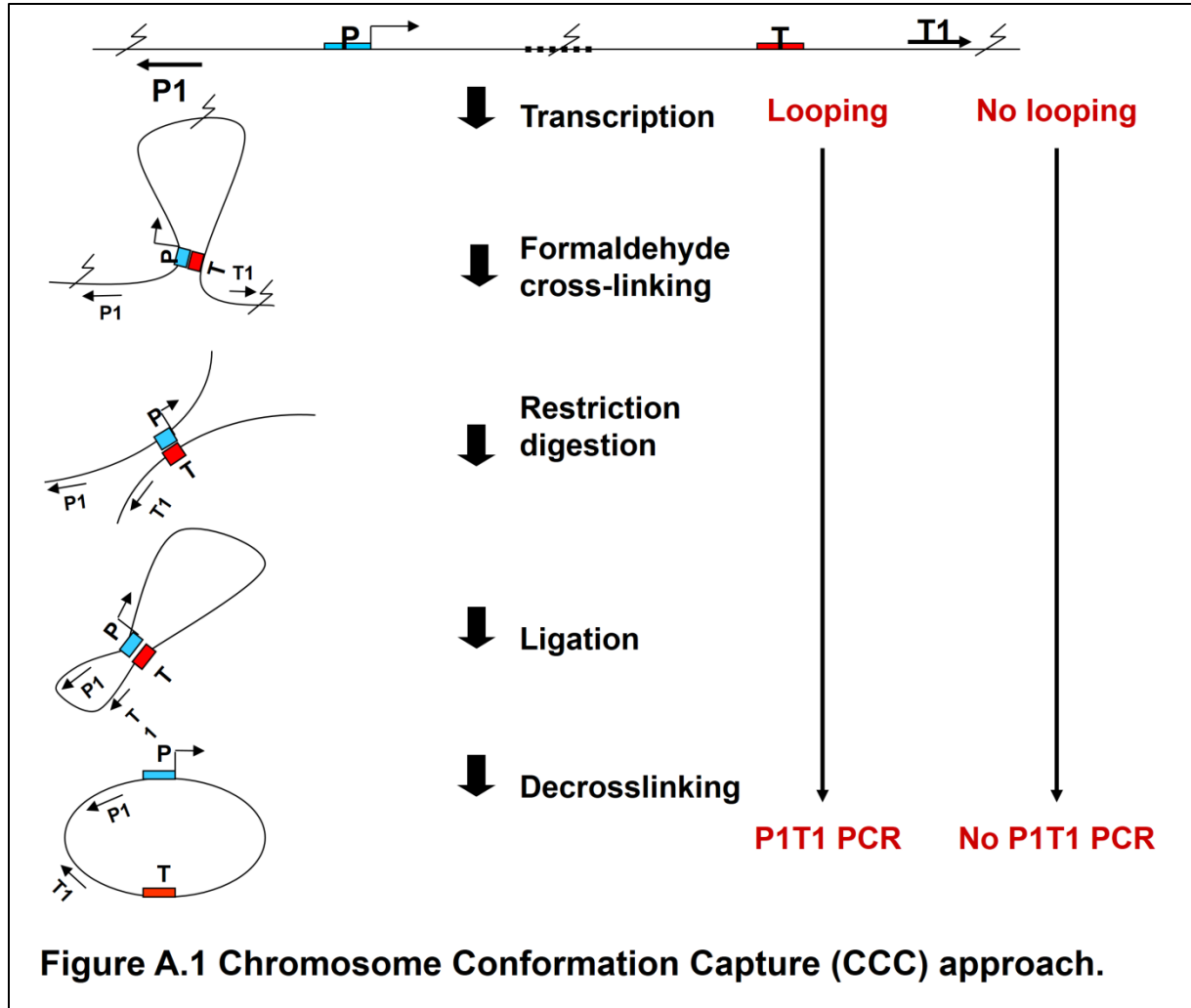
Biotechnology (Cat# 06-549). ChIP of TFIID was performed using anti-TBP antibodies

obtained from Santa Cruz (Cat# sc-33736). ChIP analysis of TFIIB was carried out

using anti-Myc antibodies in a strain with a C-terminus Myc-tagged TFIIB. For ChIP of

TFIIF, TFIIE and TFIIH, strains were constructed with a TAP-tag inserted at the

carboxy-terminus of Tfg2, Tfa2, and Ccl1 subunits respectively, and ChIP was

performed using IgG-Sepharose beads.

Crosslinking, cell lysis and isolation of chromatin was done as described in [120].

A.4. CHROMOSOME CONFORMATION CAPTURE ANALYIS (3C)

3C experiments were performed exactly as described previously [120]. The primers used for 3C analysis are shown in supplemental Table S2. A 50 ml cell culture was grown as described above. Cells were formaldehyde crosslinked for 15 minutes at 25°C. The crosslinked crude chromatin was digested with restriction endonuclease(s) (Alu1 for *INO1*; NlaIV and Alu1 for *CHA1*). After restriction digestion, the reaction volume was diluted by 7.5 fold to minimize intermolecular ligation in the next step. Ligation reactions were performed at room temperature for 90 minutes. The crosslinks were reversed by incubating at 65°C overnight. DNA was extracted with phenol-chloroform followed by ethanol precipitation. 500 ng of DNA was used as template in the PCR using the P1-T1 divergent primer pair as indicated in Figures II.3 A, III.12.A and D. Control PCR products were generated using a convergent primer pair (F2-R1). PCR and detection of products were performed exactly as described in [120]. Each experiment was performed with at least four independently grown cultures. The P1-T1 PCR signals are normalized with respect to F2-R1 PCR signals. The main steps in the CCC approach are shown in Figure A.1.

A.5. TRANSCRIPTION ANALYSIS (RT-PCR)

Isolation of total RNA and transcription analysis was performed by RT-PCR using oligo-dT primer at the reverse transcription step as described previously (El Kaderi et al 2009). The RT-PCR primers are shown in appendix C. A minus-RT control (without reverse transcriptase) was always performed to ensure that the RT-PCR signal was not coming from contaminating DNA. The RT-PCR results were normalized with respect to

the 18S rRNA control that is transcribed by RNAP I and requires a different set of transcription factors.

A.6. STRAND-SPECIFIC RT-PCR

Strand-specific RT-PCR was performed to distinguish between sense and anti-sense transcripts. Total RNA for this procedure was extracted using Trizol reagent. The cell pellet was resuspended in 500 µl of Trizol. Acid-washed glass beads (about 250 µl) were added to the cell suspension. Cells were lysed by vigorous shaking for 10 minutes on an agitator at 4^oC. Whole cell lysate was recovered by puncturing the bottom of the tube with a 22-guage needle, placing it on the top of a 15 ml pre-chilled centrifuge tube and centrifuging at 300xg for 2 minutes. The filtrate was transferred into a chilled 1.5 ml microfuge tube and 500 µl more Trizol reagent was added. After adding 200 µl of chloroform, tubes were vigorously agitated and left on the bench for 5 minutes. The tubes were then centrifuged at high speed for 10 minutes. The supernatant was extracted two times with an equal volume of phenol/chloroform (pH 4.3), followed by an extraction with chloroform only. RNA was precipitated using 0.1 volumes 10 M LiCl and 3 volumes cold ethanol in the presence of glycogen as a carrier. The precipitated RNA was collected by centrifugation at 14220xg on a table-top centrifuge for 15 minutes. The air-dried RNA pellet was resuspended in 50 µl of DEPC-treated water and the concentration was estimated using a spectrophotometer.

Strand specific RT-PCR was now performed as described in [120]. 1µg of RNA was used to make cDNA using strand-specific primers for *CHA1* as shown in Figure 2A. Primers A_s, B_s, C_s and D_s were used to reverse-transcribe sense mRNA, while A_{as}, B_{as} and C_{as} primers were used for reverse transcription of anti-sense transcripts. This was

followed by PCR amplification of cDNA for regions W, X, Y and Z using primer pairs A_{as}-B_s, A_{as}-A_s, B_{as}-C_s and C_{as}-D_s respectively. A minus-RT control (without reverse transcriptase) was always performed to ensure that the strand-specific RT-PCR signal was not due to contaminating DNA in the RNA preparation. RT-PCR results were normalized with respect to the 18S rRNA control that is transcribed by RNAP I and requires a different set of transcription factors.

A.7. IMMUNOPRECIPITATION-WESTERN BLOT

Immunoprecipitation was performed using C-terminal Myc-tagged strains for TFIIB, Pcf11, Rna14, Rna15, and clp1 proteins (NAH20, NAH22, NAH21, NAH25, and NAH29, respectively) constructed in the clp1 mutant background. The Myc-tag was amplified from pFA6a-13-myc-trp and inserted at the C-terminus of these subunits by one-step PCR-based chromosome modification.

A.7.1 GROWING CELLS AND OPTAINIG CHROMATIN

Freshly streaked plates were used to inoculate 5 ml of YP-Dextrose media and the cells were grown for 6 hours with shaking at 25°C. These cultures were diluted 1:100 in 300 ml of YPD and left to grow at 25°C with shaking till A_{600} reached 0.8. The cultures were then split into two halves of 150 ml each and centrifuged at 3000 rpm for 5 minutes. Cell pellet from each half was resuspended in 150 ml of pre-warmed (37^oC) YPD medium. For deactivation, the cells in the pre-warmed media were incubated at 37^oC for 2 hours. The other cell pellet from 150 ml culture was left at 25^oC for 2 hours. This step normally takes cells to an A_{600} of ~1.2. Equal O.D. units of the cultures were pelleted by centrifugation at 3000 rpm for 5. Upon the deactivation period, the cells were pelleted and washed once with 50 ml of cold water and pelleted again. The cell pellets

were flash frozen in liquid nitrogen, thawed by resuspending in 500 µl of IP lysis buffer

supplied with1mM PMSF).   The cells were lysed for 1 hour using ~ 300 µl of glass

beads and vigorous shacking at 4°C. The cell lysates were then centrifuged at 4°C

(14000 rpm for 15 minutes). The supernatants were subjected to immunoprecipitation

using anti Myc antibodies conjugated beads and western blots were conducted as in

Medler et al., 2011.

A.8. GRO-Seq

A.8.1 PREPARATION OF CELLS

Nuclear run-on reaction and the library construction for GRO-Seq were

(A)   (B)   (C)

without UTP   1 µM UTP   1 µM Br-UTP

1   ~10   ~7.5

A.2. TRO assay   for testing the efficiency of Br-UTP incorporation   during run-on reaction. The probes are PCR amplified fragments from a constitutively expressed gene, *ASC1*. The run-on buffer was supplied with either 1 µM of UTP (B), 1 µM of bromo-UTP (C), or left without adding any UTP (A). The reaction was performed in the presence of 50 µCi of [α-^{32}P]-CTP (3000 Ci/mmol). TRO signals were quantified using GEL LOGIC 200 (KODAK) system . The numbers under the gel pictures represent the relative intensities of the bands.

performed as described in Birse et al., 1997 and Core et al., 2012 with some

modifications. The wild type (W303-1a) and the *clp1 769-5* mutant cells were grown at

$25^{\circ}C$ in 100 ml of Yeast extract-Peptone-Dextrose (YPD) medium till A_{600} reached 0.4.

The cultures were then split into two halves of 50 ml each; transferred to 50 ml sterile

tubes; and centrifuged at 3000 rpm for 5 minutes. Cell pellet from one tube was

resuspended in 50 ml of pre-warmed ($37^{\circ}C$) YPD medium. For deactivation, the cells in

the pre-warmed media were incubated at $37^{\circ}C$ for 2 hours. The other cell pellet from 50

ml culture was left at $25^{\circ}C$ for 2 hours. This step normally takes cells to an A_{600} of ~ 0.8.

Equal O.D. units of the cultures were pelleted by centrifugation at 3000 rpm for 5

minute at $4^{\circ}C$. Cell pellets were washed with 10 ml of ice-cold TMN buffer (10 mM Tris-

HCl pH 7.5, 5 mM $MgCl_2$, 100 mM NaCl); left in the buffer for 10 minutes; centrifuged at

3000 rpm for 5 minutes; and then resuspended in 940 µl of DEPC

(Diethylpyrocarbonate)-treated ice-cold water. Chilling cells to ~ $4^{\circ}C$ at this stage is

essential for stopping any residual transcription. The cell suspensions were transferred

to a pre-chilled eppendorf tubes, and 60 µl of 10% sarkosyl was added to the cell

suspension. The tubes were tightly sealed with parafilm and placed inside a 50 ml tubes

backed with ice. The samples were incubated for 25 minutes on a nutator at $4^{\circ}C$ to

permeabilize cells. Sarkosyl is an anionic detergent that permeabilizes yeast cells and

prevents any new initiation events by inhibiting the PIC assembly. Sarkosyl, however,

does not interfere with the stability of preassembled PIC, and the catalytic activity of the

polymerase molecules. The permeabilized cells were recovered by a low-speed

centrifugation at 3600 rpm for 6 minutes at $4^{\circ}C$. The supernatant was quickly and

thoroughly aspirated out to remove the endogenous nucleotides. The remaining pellet

was then kept in ice till use in the run on transcription reaction.

A.8.2. NUCLEAR RUN-ON REACTION

100bp   **Total RNA**

← 28S

← 18S

← 5S /tRNA

A.3. The ratio of the intensities of the of 28S rRNA to the 18S rRNA bands is ~ 2:1 indicating that the RNA is completely intact
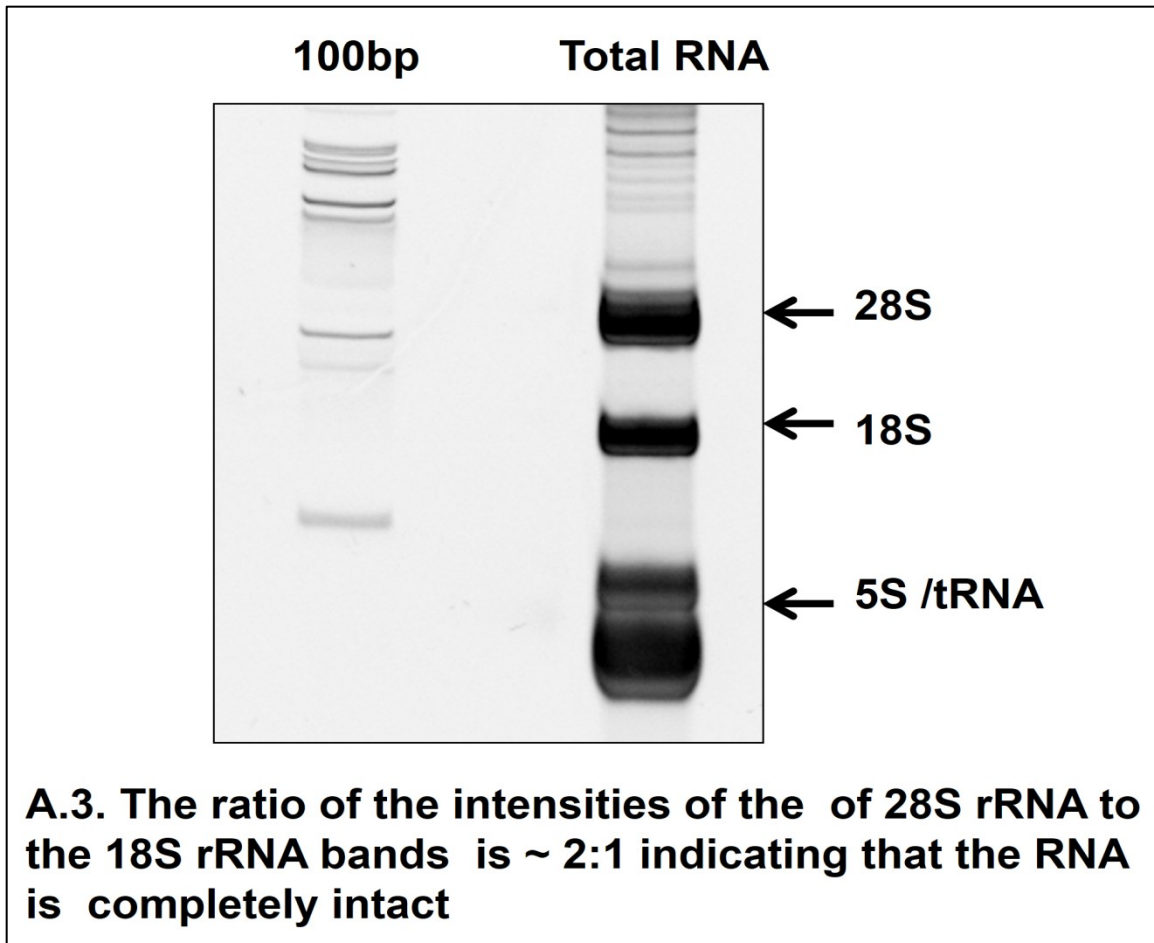
During the run-on reaction, the transcription events initiated *in vivo* are allowed to resume elongation by supplying the permeabilized cells with the nucleotides containing buffer, and incubating the reaction at 30^OC. In the run-on buffer UTP was replaced with the bromo-UTP nucleotides. Therefore, all the new transcription is done by RNA polymerase molecules that were already engaged in transcription before harvesting the cells and all the newly synthesized transcripts will contain Br-UTP nucleotide. For each run-on reaction, the cell pellet was thoroughly resuspended in 150 µl of run-on reaction buffer (50 mM Tris-HCl pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 2 mM DTT, 0.5 mM each of (ATP, CTP, GTP and BrUTP) and 5 µl RNase Inhibitor. The reaction mixture was

incubated in 30°C water bath for 5 minutes with gentle inversion of tubes every 2 minutes to allow elongation of the transcripts. The reactions were immediately stopped by adding 500 µl of ice-cold Trizol reagent (Ambion). The efficiency of Br-UTP incorporation was assessed by performing a TRO assay for *ASC1* (Figure A.2)

A.8.3. EXTRACTION OF TOTAL RNA

#	Sample ID	User name	Date and Time	Nucleic Acid Conc.	Unit	A260	A280	260/280	260/230	Sample Ty
1	after measy	Lab	3/29/2013 1:00 PM	96.1	ng/ul	2.403	1.095	2.19	1.77	RNA

A.4. Purity of the total RNA preparation. A260/A280 is 2.19 which indicate the purity of the sample
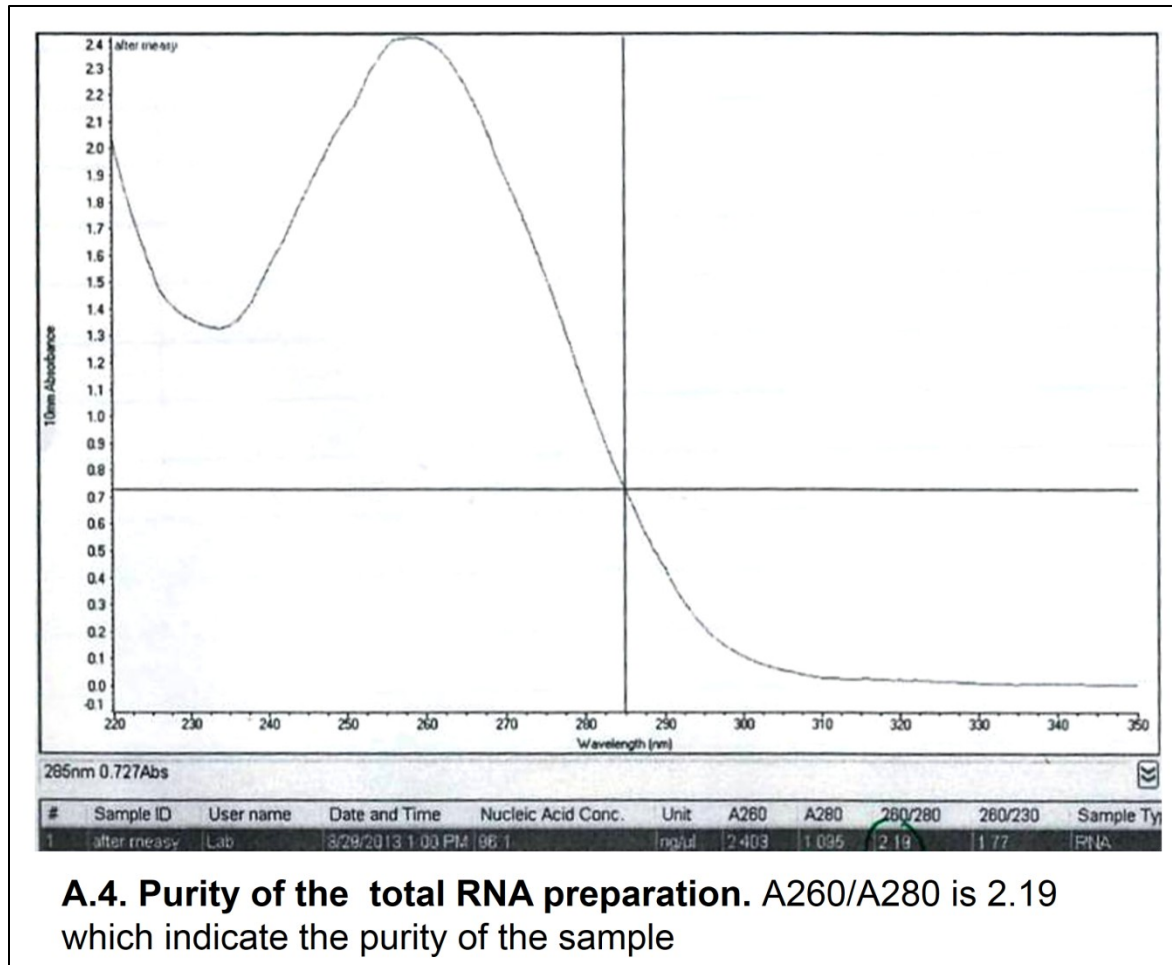
About 250 µl of acid-washed glass beads were added to the run-on reaction and the cells were lysed by vigorous agitation at room temperature for 5 minutes. The samples were recovered by puncturing the bottom of the tubes using a red hot 22g needle; placing them in 15 ml falcon tubes; and spinning at 1500 rpm for 2 minutes at 4°C. Recovered filtrate was transferred to pre-chilled 1.5 ml eppendorf tubes. An
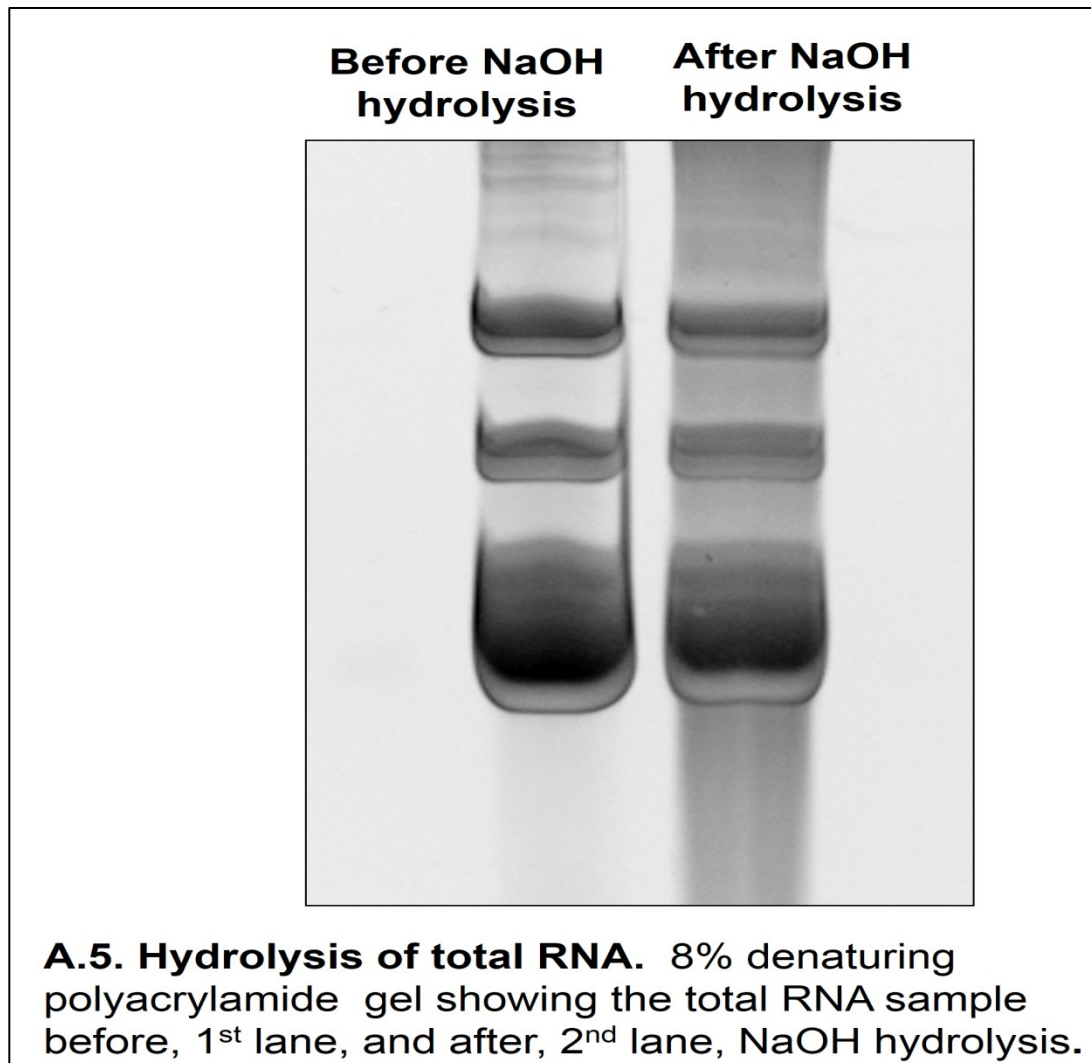
additional 500 µl of Trizol and 200 µl of chloroform were added to the tubes, and the samples were vigorously agitated. After incubation at room temperature for 5 min, samples were vigorously shaken again followed by centrifugation at 14000 rpm for 10 minutes. The upper aqueous phase was carefully transferred to another tube while taking care that the precipitated DNA in the interphase is left untouched. Three consecutive acid phenol-chloroform extractions followed by a chloroform only extraction were performed to further purify the RNA. Total RNA was precipitated in 0.3 M NaCl and 3 volume of ice-cold 100% ethanol. 2 µl of glycogen was added as a carrier. The mixture was incubated at −20°C for 60 min followed by centrifugation at 13200 rpm for 30 minutes at 4°C. RNA pellet was washed once with 1 ml of ice-cold 75% ethanol and centrifuged for another 10 minutes. The resultant RNA pellet was air dried for 5 minutes and resuspended in 55 µl of DEPC-treated water. RNA integrity was assessed by running 5 µl of each RNA sample on a 8% denaturing polyacrylamide gel (8% polyacrylamide, 7M urea, 1X TBE buffer) (Figure A.3.). The RNA purity was also assessed by measuring the RNA absorbance at 260 nm using nanodrop spectrophotometer (Figure A.4).

A.8.4 RNA HYDROLYSIS

To improve resolution, the RNA samples were subjected to a partial hydrolysis using NaOH. RNA hydrolysis was performed by adding 5 µl of 1N NaOH (f.c. 500 mM), mixing thoroughly and placing in ice for 20 minutes. NaOH was neutralized by adding 30 µl of 1M Tris-HCl (pH 6.6) (Figure A.5.). Samples were purified using Rneasy kit (Qiagen) to remove any unincorporated BrUTP nucleotides and to change buffer composition. Samples were eluted from the Rneasy columns in a final volume of 100 µl

of DEPC-treated water. The RNA samples were incubated at 65°C for 5 minutes and immediately placed on ice till ready for binding to the BrUTP antibody conjugated beads (Santa Cruz).

A.8.5 IMMUNOPURIFICATION OF NASCENT RNA

A.5. Hydrolysis of total RNA. 8% denaturing polyacrylamide gel showing the total RNA sample before, 1st lane, and after, 2nd lane, NaOH hydrolysis.

The nascent nuclear run on transcripts (NRO-RNA) represent a small fraction of the total RNA isolated in the preceding step. The Br-UTP nucleotides incorporated in the nascent transcripts serve as the affinity tag for the purification nascent RNA. In parallel, 100 µl slurry of BrdUTP antibody-conjugated beads were prepared for each sample. The beads were transferred to a clean eppendorf tube and washed three times

each with 500 µl of the binding buffer (0.25X SSPE buffer, 1mM EDTA, 0.05% Tween, 37.5 mM NaCl). The washed beads were resuspended in 500 µl of blocking buffer (1X binding buffer, 0.1 PVP (polyvinylpyrrolidone),1 µg/ ml Ultra Pure BSA) and the tubes

A.6. GRO-Seq library construction. 8% denaturing polyacrylamide gels showing different stages of the libraries construction. From left, 100bp ladder, nascent RNA immunopurified using BrdUTP antibody conjugated beads, and the nascent RNA after rRNA depletion. (4) 1.5% agarose gel showing the final library sent for sequencing
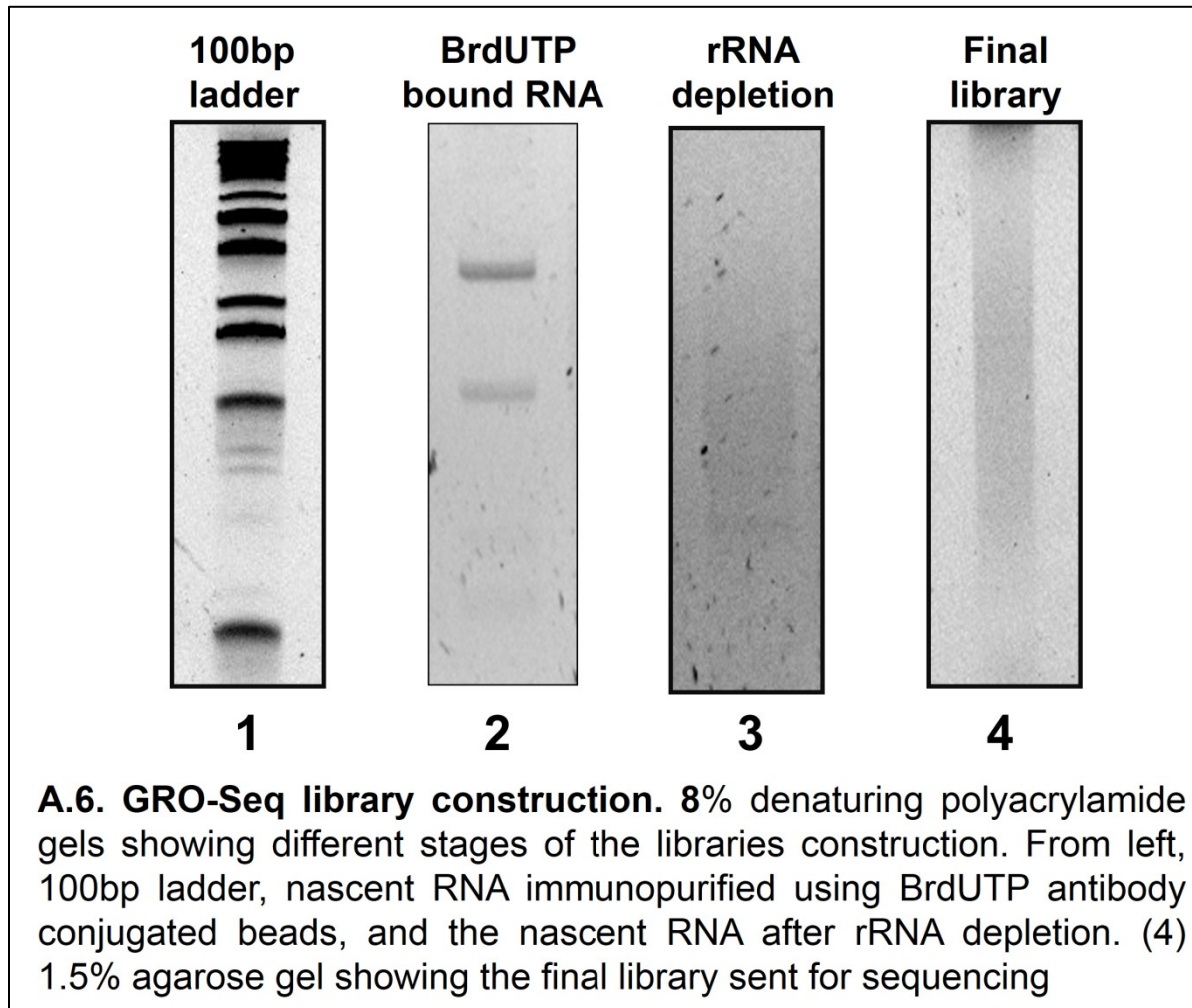
were incubated at 4°C on a nutator for two hours. The blocked beads were washed twice with the binding buffer and resuspended in another 400 µl of the same buffer. The RNA samples prepared previously were bound to the beads at 4°C nutator for one hour. Upon binding, the beads were washed consecutively with 500 µl of binding buffer, low salt buffer (0.2X SSPE buffer, 1mM EDTA, 0.05% Tween), and high salt buffer (0.25X SSPE buffer, 1mM EDTA, 0.05% Tween,100 mM NaCl) once each. This was followed two washes each of 500 µl TET buffer (10 mM Tris-HCl pH 7.5, 1mM EDTA. 0.5%

Tween). All the washes were done on nutator for three minutes each. All the centrifugation steps between the washes were at 1500 rpm for 1 minute each, with incubation of tubes on ice for 30 seconds before aspirating out the wash buffer. The nascent RNA transcripts were eluted two times each with 125 µl and one time with 250 µl of elution buffer (20 mM DTT, 150 mM NaCl, 50 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.1% SDS) for a total elution time of 10 minutes in 42°C water bath. RNA was precipitated in 0.3 M NaCl and 3 volume of ice-cold 100% ethanol and 2 µl of glycogen as a carrier. The samples were incubated at −20°C for 60 min followed by centrifugation (13200 rpm for 30 minutes) at 4°C. RNA pellets were washed with ice-cold 75% ethanol and centrifuged for another 10 minutes. The pellets were air dried for 5 minutes and resuspended in 12 µl of DEPC-treated water. 2 µl of each sample was run in 8% denaturing polyacrylamide gel to check the binding of samples to the BrdUTP antibody conjugated beads (Figure A.6).

A.8.6. rRNA DEPLETION

GRO-Seq libraries were constructed using the ScriptSeq^{TM} complete Kit (epicenter). The nascent, purified RNA samples were first depleted of rRNA. The depletion of rRNA was performed as recommended by the Ribo-Zero™ Magnetic Kit (epicentre). 5 µg of each RNA sample was combined with 10 µl of rRNA removal solution and 4 µl of reaction buffer in a 40 µl reaction volume. The mixture was gently mixed and incubated in 68 °C water bath for 10 minutes followed by incubation at room temperature for 5 minutes. In parallel, 225 µl (per sample) of magnetic beads suspension was processed. The bead suspension was transferred to a 1.5 ml RNase-free tube and placed in a magnetic stand for 2 minutes. After aspirating out the clear

phase, the tubes were removed from the stand and the beads were washed two times

each with 225 µl of RNase-Free water. Each time, the tubes were placed in the

Grow cells under the desired condition
⇩
Permeabilize cells and perform the run-on reaction
⇩
Total RNA extraction
⇩
Nascent RNA selection by Immunopurification
⇩
rRNA depletion
⇩
cDNA synthesis and 5' terminal tagging
⇩
3'- terminal tagging
⇩
Di-tagged cDNA purification
⇩
PCR amplification and bar-coding (for multiplexing)
⇩
Final Library purification
⇩
Illumina HiSeq

Figure A.7 A flow chart for the main steps in the GRO-Seq protocol

magnetic stand for 2 minutes before removing the supernatant. The washed beads

were resuspended thoroughly in 65 µl of Resuspension Solution supplied with 1 µl of

RiboGuard RNase Inhibitor. The previously treated RNA mixture was added to the

processed beads and they were mixed immediately by quick pipetting followed by brief

and gentle agitation. The mix was incubated at room temperature for 5 minutes,

agitated briefly, and incubated in 50°C water bath for 5 minutes. The tubes were immediately placed in the magnetic stand for 2 minutes before the transfer of the supernatant (~ 90 µl of rRNA-depleted sample) to a new RNase-Free tube. The volume of the rRNA depleted samples was adjusted to 180 µl with RNase-Free water. Samples were ethanol precipitated by adding 18 µl of 3M sodium acetate, 2 µl of glycogen, and 600 µl of ice-cold 100% ethanol, mixed thoroughly and placed in -20°C for 2 hours. Tubes were centrifuged at 13200 rpm for 30 minutes at 4°C, and pellets were washed once with ice-cold 75% ethanol and centrifuged for another 10 minutes. RNA pellets were air dried for 5 minutes and resuspended in 10 µl of DEPC-treated water each. Aliquots of the rRNA-depleted samples were run in 8% denaturing polyacrylamide gel to assess the depletion efficiency (Figure A.6.).

A.8.7. LIBRARY PREPARATION

Construction of the final library was done in six sequential steps. First, 2 µl of the rRNA depleted samples were fragmented in a 12 µl reaction volume containing 1 µl of RNA Fragmentation Solution and 2 µl of cDNA Synthesis Primer. The components were mixed carefully and incubated in thermocycler for 5 minutes at 85°C before placing on ice. Second, the fragmented RNA samples were reverse transcribed and 5' tagged. This was done by combining 3 µl of cDNA Synthsis PreMix, 0.5 µl of 100mM DTT and 0.5 µl StarScrip Reverse Transcriptase in a 0.2 ml tube, mixing carefully and adding the mix to the fragmented RNA sample and mixing the whole sample properly. The tubes were placed in a thermocycler and incubated at 25°C for 5 minutes followed by 42°C for 20 minutes. The tubes were cooled to 37°C before adding 1 µl of Finishing Solution and incubating at 37°C for 10 minutes. The tubes were incubated at 95°C for 3 minutes to

stop the reaction. The tubes were then incubated at 95°C for 3 minutes, cooled and then kept at 25°C. Third, the cDNA samples were 3' terminal–tagged by thoroughly mixing 7.5 μl Terminal Tagging Premix and 0.5 μl of DNA polymerase in a 0.2 ml tube and adding the mix to the cDNA samples from the previous step and mixing all well. The reactions were incubated in a thermocycler for 15 minutes at 25°C followed by 3 minutes at 95°C after which the reactions were cooled to 4°C. Fourth, the terminal tagged cDNA samples were purified using Qiagen MinElute PCR purification kit and eluted with 25 μl of the provided elution buffer. Normally, the sample volume recovered from this elution step is 22.5 μl. Fifth, all the purified cDNA samples were used in the final PCR amplification reactions. Briefly, in a 50 μl reaction volumes, 22.5 μl samples were combined with 5 μl of advantage buffer, 1 μl of advantage polymerase, 1 μl of the provided Forward PCR primer, 1 μl of the ScriptSeq Index PCR primer, for barcoding, (epicentre) and 1 μl of dNTP mix. For sample multiplexing, ScriptSeq Index PCR primer #1, #6, and #12 were used for barcoding the wildtype cells, Clp1mutant cells grown at the permissive temperature (25°C), and Clp1 mutant cells grown at the restrictive temperature (37°C), respectively. In a thermocycler, the reactions were incubated at 95°C for 1 minute followed by 15 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 68°C for 3 minutes. The reactions were then incubated at 68°C for 7 minutes for the final extension step. Finally, the libraries were subjected to a final purification step using the Qiagen MinElute PCR purification kit. Each library was eluted in 30 μl elution buffer of which 3 μl were ran in a 1.5% agarose gel for the quality check and 1 μl was used for the optical density reading (Figure A.5.). A flow chart for the main steps in the GRO-Seq protocol is shown in Figure A.6.

A.9. QUANTIFICATION

The quantification was performed as described in [214].

APPENDIX B: STRAINS

B.1. STRAINS USED IN CHAPTER II

Strain	Genotype	Reference
H-144 (W303 1A)	MATa leu2-3 can1-100 ura3-1 ade2-1 his3-11,15 trp1-1	
WZ8 (hrp1-5)	MATα cup1Δ ura3 his3 trp1 lys2 ade2 leu2 hrp1::HIS3[pRS315-hrp1-L205S (LEU2	Kuehner and Brow, 2008
H-264 (pap1-1)	MATα ade1/ade2 lys2 ura3-52 pap1-1	Claire Moore, 2006
H-261 (rna14-1)	MATα ura3-1 trp1-1 ade2-1 leu2-3,112 his3-11,15 rna14-1	Claire Moore, 2002

B.2. STRAINS USED IN CHAPTER III

Strain	Genotype	Reference
By4733	MATa his3Δ200 trp1Δ63 leu2Δ0 met15Δ0 ura3Δ0	
Clp1- 769-5	MATa ura3Δ0 leu2Δ0 his3Δ1lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1ts::URA3	Ben-Aroya S, et al. (2008)
SAM53	BY4733, MATa his3Δ200 trp1Δ63 leu2Δ0 met15Δ0 ura3Δ0 CLP1-Myc(trp)	Medler et al., 2011
NAH20	MATa ura3Δ0 leu2Δ0 his3Δ1 lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1-ts::URA3.SUA7(TFIIB)-Myc-KMX	This study
NAH21	MATa ura3Δ0 leu2Δ0 his3Δ1 lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1-ts::URA3 Rna14-Myc-KMX	This study
NAH22	MATa ura3Δ0 leu2Δ0 his3Δ1 lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1-ts::URA3 Pcf11-Myc-KMX	This study
NAH25	MATa ura3Δ0 leu2Δ0 his3Δ1 lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1-ts::URA3 Rna15-Myc-KMX	This study
NAH26	MATa ura3Δ0 leu2Δ0 his3Δ1lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1ts::URA3 Trp1 Δ (KMX)	This study
NAH29	MATa ura3Δ0 leu2Δ0 his3Δ1lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1ts::URA3 Trp1 Δ Clp1-Myc (TRP)	This study
NAH31	MATa ura3Δ0 leu2Δ0 his3Δ1 lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1-ts::URA3 CCL1-TAP (TRP)	This study
NAH32	MATa ura3Δ0 leu2Δ0 his3Δ1 lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1-ts::URA3 TFA2-TAP (TRP)	This study
NAH33	MATa ura3Δ0 leu2Δ0 his3Δ1 lys2Δ0can1Δ::LEU2-MFA1pr::His3 clp1-ts::URA3 TFG2-TAP (TRP)	This study
NAH36	MATa his3Δ200 trp1Δ63 leu2Δ0 met15Δ0 ura3Δ0 TFA2-TAP (TRP)	This study
NAH37	MATa his3Δ200 trp1Δ63 leu2Δ0 met15Δ0 ura3Δ0 CCL1-TAP (TRP)	This study
NAH38	MATa his3Δ200 trp1Δ63 leu2Δ0 met15Δ0 ura3Δ0 TFG2-TAP (TRP)	This study

B.3. STRAINS USED IN CHAPTER IV

Strain	Genotype	Reference
By4733	MATa his3Δ200 trp1Δ63 leu2Δ0 met15Δ0 ura3Δ0	
clp1- 769-5	MATa ura3Δ0 leu2Δ0 his3Δ1lys2Δ0 can1Δ::LEU2-MFA1pr::His3 clp1ts::URA3	Ben-Aroya S, et al. (2008)

B.4. STRAIN CONSTRUCTION

SAM53, which contained the Myc-tag at the carboxy-terminal of Clp1 in BY4733 strain background, was constructed by transforming the parental strain with the PCR product amplified from pFA6-13Myc-TRP1. The temperature-sensitive mutant *clp1-769-5* was kindly provided by Dr. Philip Hieter. Strains NAH20, NAH21, NAH22, NAH31, NAH32 and NAH33 were derived from the temperature sensitive *clp1-769-5* strain by adding either the Myc or the Tap-tag at the carboxy terminus of an initiation factor or a termination factor. Strains NAH20 (Myc-tagged TFIIB), NAH21 (Myc-tagged Rna14), NAH22 (Myc-tagged Pcf11) and NAH25 (Myc-tagged Rna15), which contained the Myc-tag at the carboxy-terminus of the indicated factor, were constructed by transforming the *clp1-769-5* strain with the PCR product amplified from pFA6-13Myc-KanMX6. For TAP-tagging of the general transcription factors, first the temperature-sensitive *clp1-769-5* strain was made *trp1⁻* by replacing *TRP1* with a *KanMX* cassette that was PCR amplified from pUG6. Next a TAP-tag was inserted at the carboxy-terminus of TFIIH subunit Ccl1 (NAH31), TFIIF subunit Tfg2 (NAH33) and TFIIE subunit Tfa2 (NAH32) by transforming the *clp1-769-5-(trp1)* strain with the TAP-cassette amplified from plasmid pBS1479. NAH29 was constructed by deleting the *TRP1* gene from the temperature sensitive *clp1-769-5* strain by one step gene replacement using a PCR product

containing *KanMX* amplified from pUG6. *KanMX* was then excised out utilizing the CRE recombinase activity and the resultant strain was transformed with the PCR product amplified from pFA6-13Myc-TRP1 to insert the Myc-tag at the carboxy-terminus of clp1. Strains NAH36 (TAP-tagged Tf2a), NAH37 (TAP-tagged Ccl1), and NAH38 (TAP-tagged Tfg2), which contained the Myc-tag at the carboxy-terminus of the indicated factor in BY4733 strain background, was constructed by transforming the parental strain with the PCR product amplified from pFA6-13Myc-TRP1.

APPENDIX C: PRIMERS

C.1. *CHA1* TRO-PRIMERS

Name	Sequence
CHA1 A	GATAGCCTCTTGCGACCTTATT
	CTTAACAGGAGCCGCCCAT
CHA1 B	GCCCCAGCGGAAATGTAA
	CATTCATATTTCAAGAAAAATTGTG
CHA1 C	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 D	AATTCAAAAGGACGGTAAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 E	GTTGGTGGAGGTGGTTTATACA
	TCTGGTGTTGTATTTGCGAGC
CHA1 F	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 G	GGAAGAAGCGTTGGATAGCAT
	CGTTTTGGATATGTTGATGCTTAC
CHA1 H	GCACAGAATTTGTATAAAGGGG
	GCTTTTCTTCACTTAGTAAGGATTAA
CHA1 I	GTTCCGTAATAATCTTCCCAGC
	CTGGGGTCTTCATTTGTGTCA

C.2. ChIP-PRIMERS

INO1 RNAPII-ChIP

Name	Sequence
INO1 A	GAAATATGCGGAGGCCAAG
	GGAGGTGATTGGAGCAATATTATC
NO1 B	GCTTGTTCTGTTGTCGGGTTC
	TCTTCGTAACTACAGCATTTTCG
NO1 C	GTATTAAACCGGTCTCCATTGC
	CCGACGGGCTTCATATATTTG
NO1 D	GATATCCAGAATTTCAAAGAAGAAAC
	TATTCTGCGGTGAACCATTAATATAG
NO1 E	CTCATTTCAACGACTCTCTTTTC
	ATGTTAAGTATATGTATTGATGGAAGG
NO1F	GGTAGATGCGAGAAAGTGCTG
	CTTCTTTCTCGTCCTCCTCCT
CHA1 A	GCCCCAGCGGAAATGTAA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC

CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CCCCTTTATACAAATTCTGTGC
CHA1 E	GCACAGAATTTGTATAAAGGGG
	GCTTTTCTTCACTTAGTAAGGATTAA

CHA1- TFIIB –ChIP

Name	Sequence
CHA1 A	GATAGCCTCTTGCGACCTTATT
	CATTCATATTTCAAGAAAAATTGTG
CHA1 B	AATTCAAAAGGACGGTAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CCCCTTTATACAAATTCTGTGC

INO1 TFIIB-ChIP

Name	Sequence
INO1 A	GCTTGTTCTGTTGTCGGGTTC
	GGAGGTGATTGGAGCAATATTATC
INO1 B	GATATCCAGAATTTCAAAGAAGAAAC
	TATTCTGCGGTGAACCATTAATATAG
INO1 C	GTATTAAACCGGTCTCCATTGC
	CCGACGGGCTTCATATATTTG
INO1 D	CTCATTTCAACGACTCTCTTTTTC
	GCACTTTCTCGCATCTACCTCA

CHA1-TBP –ChIP

Name	Sequence
CHA1 A	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CGTTTTGGATATGTTGATGCTTAC

INO1 TBP-ChIP

Name	Sequence
INO1 A	GCTTGTTCTGTTGTCGGGTTC
	GGAGGTGATTGGAGCAATATTATC
INO1 B	GATATCCAGAATTTCAAAGAAGAAAC
	TATTCTGCGGTGAACCATTAATATAG
INO1 C	GTATTAAACCGGTCTCCATTGC

	CCGACGGGCTTCATATATTTG
INO1 D	CTCATTTCAACGACTCTCTTTTTC
	GCACTTTCTCGCATCTACCTCA

CHA1-TFIIE-ChIP

Name	Sequence
CHA1 A	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CCCCTTTATACAAATTCTGTGC

INO1 TFIIE-ChIP

Name	Sequence
INO1 A	GAAATATGCGGAGGCCAAG
	GGAGGTGATTGGAGCAATATTATC
INO1 B	GATATCCAGAATTTCAAAGAAGAAAC
	TATTCTGCGGTGAACCATTAATATAG
INO1 C	GTATTAAACCGGTCTCCATTGC
	CCGACGGGCTTCATATATTTG
INO1 D	CTCATTTCAACGACTCTCTTTTTC
	GCACTTTCTCGCATCTACCTCA

CHA1- TFIIH-ChIP

Name	Sequence
CHA1 A	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	GGAAAAAATCAATACTAGCAAAATA

INO1 TFIIH-ChIP

Name	Sequence
INO1 A	GCTTGTTCTGTTGTCGGGTTC
	TCTTCGTAACTACAGCATTTTCG
INO1 B	TATTCTGCGGTGAACCATTAATATAG
	GTATTAAACCGGTCTCCATTGC
INO1 C	GTATTAAACCGGTCTCCATTGC

	CCGACGGGCTTCATATATTTG
INO1 D	GACAAAGAGGCAATAGTTCAAAAG
	CTCATTTCAACGACTCTCTTTTTC

CHA1-TFIIF-ChIP

Name	Sequence
CHA1 A	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	GGAAAAAATCAATACTAGCAAAATA

INO1-TFIIF-ChIP

Name	Sequence
INO1 A	GAAATATGCGGAGGCCAAG
	GGAGGTGATTGGAGCAATATTATC
INO1 B	GATATCCAGAATTTCAAAGAAGAAAC
	TATTCTGCGGTGAACCATTAATATAG
INO1 C	GTATTAAACCGGTCTCCATTGC
	CCGACGGGCTTCATATATTTG
INO1 D	CTCATTTCAACGACTCTCTTTTTC
	GCACTTTCTCGCATCTACCTCA

CHA1-Clp1p-Myc-ChIP

Name	Sequence
CHA1 A	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CCCCTTTATACAAATTCTGTGC

INO1-Clp1p-Myc-ChIP

Name	Sequence
INO1 A	GAATATTGAACTTATTTAATTCACATGG
	GGAGGTGATTGGAGCAATATTATC
INO1 B	GTATTAAACCGGTCTCCATTGC
	CCGACGGGCTTCATATATTTG
INO1 C	GATATCCAGAATTTCAAAGAAGAAAC

	TATTCTGCGGTGAACCATTAATATAG
INO1 D	CTCATTTCAACGACTCTCTTTTTC
	ATGTTAAGTATATGTATTGATGGAAGG

CHA1-Rna14-TAP-ChIP

Name	Sequence
CHA1 A	GCCCCAGCGGAAATGTAA
	CATTCATATTTCAAGAAAAATTGTG
CHA1 B	AATTCAAAAGGACGGTAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GTTGGTGGAGGTGGTTTATACA
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CGTTTTGGATATGTTGATGCTTAC

INO1-Rna14-TAP-ChIP

Name	Sequence
INO1 A	GCTTGTTCTGTTGTCGGGTTC
	GGAGGTGATTGGAGCAATATTATC
INO1 B	GATATCCAGAATTTCAAAGAAGAAAC
	TATTCTGCGGTGAACCATTAATATAG
INO1 C	GTATTAAACCGGTCTCCATTGC
	CCGACGGGCTTCATATATTTG
INO1 D	CTCATTTCAACGACTCTCTTTTTC
	GCACTTTCTCGCATCTACCTCA

CHA1-Rna15-TAP-ChIP

Name	Sequence
CHA1 A	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CCCCTTTATACAAATTCTGTGC

INO1-Rna15-TAP-ChIP

Name	Sequence
INO1 A	GAAATATGCGGAGGCCAAG
	GGAGGTGATTGGAGCAATATTATC
INO1 B	GATATCCAGAATTTCAAAGAAGAAAC
	TATTCTGCGGTGAACCATTAATATAG
INO1 C	GTATTAAACCGGTCTCCATTGC
	CCGACGGGCTTCATATATTTG

INO1 D	CTCATTTCAACGACTCTCTTTTTC
	GCACTTTCTCGCATCTACCTCA

CHA1-Pcf11-TAP-ChIP

Name	Sequence
CHA1 A	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CGTTTTGGATATGTTGATGCTTAC

INO1-Pcf11-TAP-ChIP

Name	Sequence
CHA1 A	GCGATGAGATAAGATAAAAGGGA
	GATTACCGATTCCTCTACTTTTGA
CHA1 B	AATTCAAAAGGACGGTAAAAGAT
	AAGGGATGAACATAAATGGGC
CHA1 C	GGTGGAAACGAATGGATGTC
	TCTTAGTGTTGTAACCCAAATGC
CHA1 D	GGAAGAAGCGTTGGATAGCAT
	CGTTTTGGATATGTTGATGCTTAC

CHA1-Strand-Specific RT-PCR primers

Name	Sequence
A_{as}	CGAGTACTAATCACCGCGAAC
B_{as}	AATTCAAAAGGACGGTAAAAGAT
C_{as}	GGAAGAAGCGTTGGATAGCAT
A_s	AAGAGAAACTGTATAAACATTTTCC
B_s	TCTCTTGTCTATCCAGCACTTAAAA
C_s	AAGGGATGAACATAAATGGGC
D_s	TGCTATCCAACGCTTCTTCC

C.3. CCC-PRIMERS

CHA1-CCC

Name	Sequence
CHA1P1	GATTACCGATTCCTCTACTTTTGA
CHA1T1	GTAAGCATCAACATATCCAAAACG
CHA1 F	AATTCAAAAGGACGGTAAAAGAT
CHA1 R	AAGGGATGAACATAAATGGGC

INO1-CCC

Name	Sequence
INO1 P1	GAACCCGACAACAGAACAAGC
INO1 T1	GTTGAGGTAGATGCGAGAAAGTG
INO1 F	GATATCCAGAATTTCAAAGAAGAAAC
INO1 R	TATTCTGCGGTGAACCATTAATATAG

MET16-CCC

Name	Sequence
MET16 P1	TTTGCTGGCCTTAGTTTTGATC
*MET16*T1	GGAAGATGGAAGGGCAAGG
INO1 F	GATATCCAGAATTTCAAAGAAGAAAC
INO1 R	TATTCTGCGGTGAACCATTAATATAG

C.4. RT-PCR PRIMERS

CHA1-RT-PCR

Name	Sequence
CHA1 A	AATTCAAAAGGACGGTAAAAGAT
CHA1 B	AAGGGATGAACATAAATGGGC
18 S F	GGAATAATAGAATAGGACGTTTGG
18 S R	GTTAAGGTCTCGTTCGTTATCG

MET16-RT-PCR

Name	Sequence
MET16-A	CATTTGGTTTGACTGGCTTGG
MET16-B	TCGTACTTGTCATCATCTTTCTCC
18 S F	GGAATAATAGAATAGGACGTTTGG
18 S R	GTTAAGGTCTCGTTCGTTATCG

INO1-RT PCR

Name	Sequence
INO1 A	GATATCCAGAATTTCAAAGAAGAAAC
INO1 B	TATTCTGCGGTGAACCATTAATATAG
18 S F	GGAATAATAGAATAGGACGTTTGG
18 S R	GTTAAGGTCTCGTTCGTTATCG

C.5. Strain-making primers

Name	Sequence
5′ TFIIB-Myc-F2	TTGCTAATGGTGTAGTGTCTTTGGATAACTTACCGGGCGTTGAAAAGAAACGGATCCCCGGGTTAATTAA
3'TFIIB-Myc-R1	CACGAGTACCCGTGCTTCTTGTTCCTATAATTTACTGTTTTATCACTTCAGAATTCGAGCTCGTTTAAAC

5′F1-RNA14-Myc-tag	TTTTAAATGATCAAGTAGAGATTCCAACAGTTGAGAGCAC CAAGTCAGGTCGGATCCCCGGGGTTAATTAA
3′R1-RNA14- Myc-tag	AGATGTGTTGGTATAAATATTCATATATACCTATTTATTAAC GTAATGTTAGAATTCGAGCTCGTTTAAAC
5′F1- PCF11-Myc-tag	CTAATAGTGGCAAGGTCGGTTTGGATGACTTAAAGAAATT GGTCACAAAACGGATCCCCGGGGTTAATTAA
3′R1- PCF11-Myc-tag	TAATATAATATATAGTTATTAAATTTAAATGTATATATGCAG TTCTGCTCGAATTCGAGCTCGTTTAAAC
5′ F2-RNA15-HA-tag	CTATTTGGGACTTAAAACAAAAGCATTAAGGGGAGAATTT GGTGCATTTCGGATCCCCGGGGTTAATTAA
3'RNA15-HA-tag	ATCATTGCGGAACCGCATTTTTTTTTTGTATTTTTGCCTCC CTAGTTTCAGAATTCGAGCTCGTTTAAAC
5′TFA2-TAP-C	TTACTAACACTCATATGACCGGTATCTTGAAAGATTATTCC CATAGAGTATCCATGGAAAAGAGAAG
3'TFA2-TAP-C	CAGTCTCTTTAACCTAATATGCAAACGAAAATGATTTAATC AAAACAACCTACGACTCACTATAGGG
5′TFG2-TAP-C	GAGACGCGGAGGCTGACTTGGAAGATGAAATAGAAATGG AAGATGTCGTTTCCATGGAAAAGAGAAG
3'TFG2-TAP-C	CTCAAGAAACTGCGTAAATATAAAATTAATGAAGAAAATCT GATTGTCAATACGACTCACTATAGGG
5′ CCL1-C-TAP	AGTTGAATGGAGAAGATACTTCGTCCACCGTTGAGAAAAA GCAAAAAACATCCATGGAAAAGAGAAG
3' CCL1-C-TAP	CTTAATCTATATATATATAAAACAGAAACCTACGGTAAC AGAGCTGTTTACGACTCACTATAGGG
5′TRP1-KMX	TATTGAGCACGTGAGTATACGTGATTAAGCACACAAAGGC AGCTTGGAGTCAGCTGAAGCTTCGTACGC
3'TRP1-KMX	TGCAGGCAAGTGCACAAACAATACTTAAATAAATACTACTC AGTAATAACGCATAGGCCACTAGTGGATCTG

C.6. ScriptSeq Index PCR-GRO-Seq

Name	Sequence
Index # 1	Index 1 5′-ATCACG-3′
Index # 6	Index 6 5′-GCCAAT-3′
Index # *12*	Index 12 5′-CTTGTA-3′

APPENDIX D: MEDIA

YEAST EXTRACT-PEPTON-DEXTROSE (YPD) medium (1 liter)

Component	Quantity	Notes
Yeast extract	10 g	
Peptone	20 g	
Dextrose	20 g	• 100 ml of 20% stock-add after autoclaving
Agar	20 g	• For plates only
NaOH	1 pellet	• For plates only

INOSITOL DROP-OUT MEDIUM (1 liter)

Component	Quantity	Notes
Ammonium Sulfate	5 g	
Vitamin Stock	1 ml	• Of 1000X stock solution
Trace Elements Stock	1 ml	• Of 1000X stock solution
Salt Mix	1.7 g	
Inositol drop-out amino acid Mix	230 mg	
Dextrose	20 g	• 100 ml of 20% stock-add after autoclaving

TRACE ELEMNTS STOCK (1000X; 100 ml)-FOR INOSITOL DROP-OUT MEDIA

Component	Quantity	Notes
Boric acid	50mg	• Autoclave
Copper sulfate	4 mg	• Store in a dark bottle at 4°C
Potassium iodide	10 mg	
Ferric chloride	20 mg	
Manganese sulfate	40 mg	
Sodium molybdate	20 mg	
Zinc sulfate	40 mg	

VITAMIN STOCK (1000X; 100 ml)- FOR INOSITOL DROP-OUT MEDIA

Component	Quantity	Notes
Biotin	2 mg	- Autoclave - Store in a dark bottle at $4^{o}C$
Calcium pantothenate	200 mg	
Folic acid	0.2 mg	
Niacin	40 mg	
β-Aminobenzoic acid	20 mg	
Pyridoxine hydrochloride	40 mg	
Riboflavin	20 mg	
Thiamin hydrochloride	40 mg	

SALT MIX- FOR INOSITOL DROP-OUT MEDIA

Component	Quantity	Notes
Potassium phosphate monobasic	85 g	
Potassium phosphate dibasic	15 g	
Magnesium sulfate	50 g	
Sodium chloride	10 g	
Calcium chloride	10 g	

AMINO ACID MIX- FOR INOSITOL DROP-OUT MEDIA

Component	Quantity	Notes
Adenine hemisulfate	40 mg	
Histidine	20 mg	
Leucine	60 mg	
Lysine	30 mg	

Methionine	20 mg	
Tryptophan	40 mg	
Uracil	20 mg	

INOSITOL STOCK (100 X; 100 ml)

Component	Quantity	Notes
Inositol (for plus inositol medium)	1 g	• 1 ml/ liter of inositol drop-out medium

METHIONINE DROP-OUT MEDIUM (1 liter)

Component	Quantity	Notes
Yeast nitrogenous base	6.7 g	• without amino acids
methionine drop-out mix	1 g	
Agar	20 g	• For plates only
NaOH	1 pellet	• For plates only
Dextrose	20 g	• 100 ml of 20% stock-add after autoclaving

METHIONINE DROP-OUT MIX- FOR METHIONINE DROP-OUT MEDIA

Component	Quantity	Notes
Adenine	2.5 g	
L-arginine	1.2 g	
L- asparatic acid	6.0 g	
L- glutamic acid	6.0 g	
L-Histidine	1.2 g	
L-leucine	3.6 g	
L-lysine	1.8 g	

L-phenylalanine	3.0 g	
L-tryptophan	2.4 g	
L-tyrosine	1.8 g	
L-valine	9.0 g	
Uracil	1.2 g	

AMMONIUM SULFATE MEDIUM (1 liter)-FOR *CHA1* RPRESSION

Component	Quantity	Notes
Yeast nitrogenous base	1.7 g	• Without amino acids • Without ammonium sulfate
Ammonium sulfate	5 g	
Amino acid mix	230 mg	
Dextrose	20 g	• 100 ml of 20% stock-add after autoclaving

SERINE/ THREONINE MEDIUM (1 liter)-FOR *CHA1* ACTIVATION

Component	Quantity	Notes
Yeast nitrogenous base	1.7 g	• Without amino acids • Without ammonium sulfate
L-serine	1 g	
L-threonine	1 g	
Amino acid mix	230 mg	
Dextrose	20 g	• 100 ml of 20% stock-add after autoclaving

AMINO ACID MIX FOR *CHA1* MEDIA

Component	Quantity	Notes
Adenine hemisulfate	40 mg	
Histidine	20 mg	

Leucine	60 mg	
Lysine	30 mg	
Methionine	20 mg	
Tryptophan	40 mg	
Uracil	20 mg	

TRYPTON DROP-OUT MEDIUM (1 liter)

Component	Quantity	Notes
Yeast nitrogenous base	6.7 g	• without amino acids
Trypton drop-out amino acid mix	1 g	
Agar	20 g	
NaOH	1 pellet	
Dextrose	20 g	• 100 ml of 20% stock-add after autoclaving

TRYPTON DROP-OUT MIX

Component	Quatity	Notes
Adenine	2.5 g	
L-arginine	1.2 g	
L- asparatic acid	6.0 g	
L- glutamic acid	6.0 g	
L-Histidine	1.2 g	
L-leucine	3.6 g	
L-lysine	1.8 g	
L-methionine	1.2 g	

L-phenylalanine	3.0 g	
L-tyrosine	1.8 g	
L-valine	9.0 g	
Uracil	1.2 g	

G418 PLATES (KMX-MEDIUM) -1 liter

Component	Quantity	Notes
Yeast nitrogenous base	10.0 g	• without amino acids
Peptone	20.0 g	
Agar	20. 0 g	
Dextrose	20 g	• 100 ml of 20% stock-add after autoclaving
G418	1.0 ml	• Of 400 mg/ml

2XYT MEDIUM-1 liter

Component	Quantity	Notes
Yeast extract	10.0 g	• without amino acids
Tryptone	16.0 g	
NaCl	5. 0 g	
Agar	20.0 g	• For plates only

APPENDIX E: BUFFERS AND SOLUTIONS

STOCK SOLUTIONS

Reagent	Molarity/ concentration/ percentage	Notes
Tris-HCl- pH 8.0	1.0 M	• Adjust pH using HCl
EDTA pH 7.0 to 8.0	0.5 M	• Adjust pH using NaOH
NaCl	5.0 M	• Autoclave
KCl	2.0M	• Autoclave
SDS	10%	• Filter sterilize
$CaCl_2$	1.0 M	• Autoclave
$MgCl_2$	1.0 M	• Autoclave
PEG (Mw 4000)	50 %	• Filter sterilize
LiOAc	1.0 M	• Filter sterilize
Glycine	2.5 M	• Autoclave
Ammounium acetate	7.5 M	• Autoclave
NaOAc pH 5.2	3.0 M	• Adjust pH using glacial acetic acid
Glycerol	50 %	• Autoclave
Tergitol	10 %	• Autoclave
Triton X-100	10 %	• Filter sterilize
LiCl	5.0 M	• Autoclave
HEPES pH 7.9	1.0 M	• Adjust the pH using KOH • Filter sterilize
Sodium deoxycholate	10%	• Filter sterilize
KOH	10.0 M	• Autoclave

Dextrose	20 %	• Autoclave
PMSF	100 mM	• Don't autoclave • Keep at 4°C
Glycogen	20 mg/ ml	• Filter sterilize
DTT	1.0 M	• Filter sterilize
Ethedium bromide	10.0 mg / ml	• Don't autoclave • Keep at 4°C
Ammonium acetate	7.5 M	• Autoclave
TE	10X	• 100 mM Tris.HCl pH 8.0 • 10 mM EDTA
TAE	50 X	• 2.0 M Tris-acetate • 50 mM EDTA
TBS	10X	• 100 mM Tris-HCl pH 8 • 2M NaCl

AGAROSE GEL ELECTROPHORESIS BUFFER (1X TAE)

Component	Concentration	Notes
Tris-acetate	40 Mm	• Autoclave
0.5 M EDTA pH 8.0	1 mM EDTA	• Autoclave

SOLUTIONS FOR YEAST GENOMIC DNA EXTRACTION

Reagent	Composition	Notes
Lysis buffer	2% Triton X-100 100 mM NaCl 10 mM Tris-HCl pH 8.0 10 mM EDTA 1% SDS	

SOLUTIONS FOR LiOAc/DMSO YEAST TRANSFORMATION

Reagent	Composition	Notes
LiAOAc buffer	0.1 M LiAOAc 10 mM Tris-HCl(pH=8.0) 1 mM EDTA	

PEG solution	50 % w/v PEG (M.W. = 4000) 0.1 LiAOAc 10 mM Tris-HCl (pH=8.0) 1 mM EDTA	• Filter sterilize
DMSO	100 %	

SOLUTIONS FOR PLASMID MINIPREP

Solution	Composition	Notes
Solution I	50 mM Dex 10 mM EDTA 25 mM Tris-HCl pH 8.0	
Solution II	0.1 N NaOH 1% SDS	
Solution III	30 ml 5M KOAc 5.75 ml glacial HOAc 14.25 ml H2O	• Store at – 20 ^{o}C

YEAST CELL WASH

component	Concentration	Notes
Wash buffer I	1X TBS	• Autoclave
Wash buffer II	1XTBS 1% Triton X-100	• Autoclave

CHROMATIN IMMUNOPRECIPITATION (ChIP) BUFFERS AND SOLUTIONS

FA-LYSIS BUFFER

Reagent	Concentration	Notes
HEPES-KOH pH 7.9	50 mM	• Store at -20 oC
NaCl	140 mM	
EDTA	1 mM	
Triton X-100	1 %	
Sodium Deoxycholate	0.1 %	

PMSF	1 mM	
SDS	0.07 %	

FA-LYSIS BUFFER + 500 mM NaCl

Reagent	Stock Concentration	Volume added
HEPES-KOH pH 7.9- 8.0	50 mM	• Store at - 20
NaCl	500 mM	
EDTA pH 8.0	1 mM	
Triton X-100	1 %	
Sodium Deoxycholate	0.1 %	
PMSF	1 mM	
SDS	0.07 %	

ChIP WASH BUFFER

Reagent	Concentration	Notes
Tris-HCl pH 7.5 to 8	10 mM	• Store at -20 oC
LiCl	250 mM	
Triton X-100	0.5 %	
EDTA pH 8.0	1 mM	
Sodium Deoxycholate	0.5 %	
SDS	0.1 %	

ChIP ELUTION BUFFER

Reagent	Concentration	Notes
Tris-HCl pH 7.5 to 8.0	50 mM	• Store at room temperature

SDS	1 %	
EDTA pH 8.0	10 mM	

REVRESE TRANSCRIPTION PCR (RT-PCR) BUFFERS AND SOLUTIONS
HIGH TE BUFFER

Reagent	Concentration	Notes
Tris-HCl pH 7.5	50 mM	• Store at RT
EDTA	20 mM	

RNA-LYSIS BUFFER

Reagent	Concentration	Notes
Tris-HCl pH 8.0	80 mM	
$CaCl_2$	10 mM	
β-mercatoethanol	10 mM	
VCR (Shake well)	10 mM	

CHROMOSOME CONFORMATION CAPTURE SOLUTION
TM BUFFER

Component	Concentration	Notes
Tris HCl pH 7.5- 8.0	10 mM	
$MgCl_2$	5 mM	

TRANSCRIPTION RUN-ON ASSAY SOLUTIONS AND BUFFERS

Reagent	Composition	Notes
20X SSC	3 M NaCL 300mM Na3CitrateX2H2O	• Adjust pH to 7.0 using HCl
Sarkosyl	10%	
Boiling solution	0.4 N NaOH 1 mM EDTA	

Hybridization solution	0.5M potassium phosphate pH 7.2 7% SDS	
Membrane wash I	0.1% SDS 1% SSC	
Membrane wash II	0.1% SDS 0.1% SSC	
2.5 XRun-on buffer	50 mM Tris-HCl pH 7.5 500 mM KCl 80 mM MgCl$_2$ 5 mM DTT	
NTPs/RNase inhibitor mix	10 mM each of CTP, ATP, and GTP 300 units of RNase Inhibitor 7 µl of [α-^{32}P]-UTP (3000 Ci/mmol, 10 µCi/µl	
TMN buffer	10 mM Tris-HCl pH 7.5 5 mM MgCl$_2$ 100 mM NaCl	
NaOAc/HOAc mix	0.3 M sodium acetate pH 5.2 0.5 µl of glacial acetic acid	
LETS buffer	0.1M LiCl 0.2% SDS 10 mM EDTA 10 mM Tris-HCl pH 7.5	

IMMUNOPRECIPITATION ASSAY BUFFERS AND SOLUTIONS

Component	Concentration	Notes
IP lysis buffer	10% glycerol 20 mM Tris-HCl pH 8.0 50 mM KCl 0.5mM EDTA 1 mM MgCl2 0.1% TritonX-100 1 mM PMSF (add directly before use)	• Autoclave • Keep at 4°C
30% Acrylamide:Bis Solution	1 % Bisacrylamide 29% Acrylamide	
4% stacking gel	125 mM Tris-Cl, pH 6.8 0.1% SDS 5% Acrylamid mix 0.1 Ammonium persulfate	• Keep at 4°C
Electrode buffer	25 mM tris 250 mM glycine	

	0.1 SDS	
5X laemeli buffer	250 mM Tris HCl pH 6.8 50 % Glycerol 10% SDS 2.8 M β-mercaptoethanol 0.1% Bromophenol blue	
Transfer Buffer	20% Methanol 24 mM Tris-base 192 mM Glycine	

REFERENCES

1. Roeder RG, Rutter WJ (1969) Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature 224: 234-237.

2. Seila AC, Core LJ, Lis JT, Sharp PA (2009) Divergent transcription: a new feature of active promoters. Cell Cycle 8: 2557-2564.

3. Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Munster S, et al. (2009) Bidirectional promoters generate pervasive transcription in yeast. Nature 457: 1033-1037.

4. Arigo JT, Eyler DE, Carroll KL, Corden JL (2006) Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol Cell 23: 841-851.

5. van Dijk EL, Chen CL, d'Aubenton-Carafa Y, Gourvennec S, Kwapisz M, et al. (2011) XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. Nature 475: 114-117.

6. Lardenois A, Liu Y, Walther T, Chalmel F, Evrard B, et al. (2011) Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6. Proc Natl Acad Sci U S A 108: 1058-1063.

7. Tan-Wong SM, Zaugg JB, Camblong J, Xu Z, Zhang DW, et al. (2012) Gene loops enhance transcriptional directionality. Science 338: 671-675.

8. Wilson RC, Doudna JA (2013) Molecular mechanisms of RNA interference. Annu Rev Biophys 42: 217-239.

9. Chinen M, Tani T (2012) Diverse functions of nuclear non-coding RNAs in eukaryotic gene expression. Front Biosci (Landmark Ed) 17: 1402-1417.

10. Edwards AM, Kane CM, Young RA, Kornberg RD (1991) Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. J Biol Chem 266: 71-75.

11. Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, et al. (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. Science 288: 640-649.

12. Krishnamurthy S, Hampsey M (2009) Eukaryotic transcription initiation. Curr Biol 19: R153-156.

13. Armache KJ, Kettenberger H, Cramer P (2003) Architecture of initiation-competent 12-subunit RNA polymerase II. Proc Natl Acad Sci U S A 100: 6964-6968.

14. Runner VM, Podolny V, Buratowski S (2008) The Rpb4 subunit of RNA polymerase II contributes to cotranscriptional recruitment of 3' processing factors. Mol Cell Biol 28: 1883-1891.

15. Bataille AR, Jeronimo C, Jacques PE, Laramee L, Fortin ME, et al. (2012) A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. Mol Cell 45: 158-170.

16. Cramer P, Bushnell DA, Kornberg RD (2001) Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science 292: 1863-1876.

17. Kornberg RD (2007) The molecular basis of eucaryotic transcription. Cell Death Differ 14: 1989-1997.

18. Fuda NJ, Ardehali MB, Lis JT (2009) Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature 461: 186-192.

19. Woychik NA, Hampsey M (2002) The RNA polymerase II machinery: structure illuminates function. Cell 108: 453-463.

20. Hahn S (2004) Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol 11: 394-403.

21. Zawel L, Reinberg D (1992) Advances in RNA polymerase II transcription. Curr Opin Cell Biol 4: 488-495.

22. Lu H, Zawel L, Fisher L, Egly JM, Reinberg D (1992) Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature 358: 641-645.

23. Murakami K, Elmlund H, Kalisman N, Bushnell DA, Adams CM, et al. (2013) Architecture of an RNA Polymerase II Transcription Pre-Initiation Complex. Science.

24. He Y, Fang J, Taatjes DJ, Nogales E (2013) Structural visualization of key steps in human transcription initiation. Nature 495: 481-486.

25. Shandilya J, Roberts SG (2012) The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim Biophys Acta 1819: 391-400.

26. Juven-Gershon T, Hsu JY, Theisen JW, Kadonaga JT (2008) The RNA polymerase II core promoter - the gateway to transcription. Curr Opin Cell Biol 20: 253-259.

27. Murakami K, Gibbons BJ, Davis RE, Nagai S, Liu X, et al. (2012) Tfb6, a previously unidentified subunit of the general transcription factor TFIIH, facilitates dissociation of Ssl2 helicase after transcription initiation. Proc Natl Acad Sci U S A 109: 4816-4821.

28. Myers LC, Kornberg RD (2000) Mediator of transcriptional regulation. Annu Rev Biochem 69: 729-749.

29. Bushnell DA, Westover KD, Davis RE, Kornberg RD (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science 303: 983-988.

30. Liu X, Bushnell DA, Silva DA, Huang X, Kornberg RD (2011) Initiation complex structure and promoter proofreading. Science 333: 633-637.

31. Ni Z, Saunders A, Fuda NJ, Yao J, Suarez JR, et al. (2008) P-TEFb is critical for the maturation of RNA polymerase II into productive elongation in vivo. Mol Cell Biol 28: 1161-1170.

32. Wade JT, Struhl K (2008) The transition from transcriptional initiation to elongation. Curr Opin Genet Dev 18: 130-136.

33. Svejstrup JQ (2004) The RNA polymerase II transcription cycle: cycling through chromatin. Biochim Biophys Acta 1677: 64-73.

34. Yudkovsky N, Ranish JA, Hahn S (2000) A transcription reinitiation intermediate that is stabilized by activator. Nature 408: 225-229.

35. Pei Y, Schwer B, Shuman S (2003) Interactions between fission yeast Cdk9, its cyclin partner Pch1, and mRNA capping enzyme Pct1 suggest an elongation checkpoint for mRNA quality control. J Biol Chem 278: 7180-7188.

36. Adelman K, Marr MT, Werner J, Saunders A, Ni Z, et al. (2005) Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. Mol Cell 17: 103-112.

37. Nechaev S, Adelman K (2011) Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. Biochim Biophys Acta 1809: 34-45.

38. Kessler MM, Zhelkovsky AM, Skvorak A, Moore CL (1995) Monoclonal antibodies to yeast poly(A) polymerase (PAP) provide evidence for association of PAP with cleavage factor I. Biochemistry 34: 1750-1759.

39. Huang Y, Carmichael GG (1996) Role of polyadenylation in nucleocytoplasmic transport of mRNA. Mol Cell Biol 16: 1534-1542.

40. Bernstein P, Peltz SW, Ross J (1989) The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. Mol Cell Biol 9: 659-670.

41. Sachs A (1990) The role of poly(A) in the translation and stability of mRNA. Curr Opin Cell Biol 2: 1092-1098.

42. Proudfoot NJ (1989) How RNA polymerase II terminates transcription in higher eukaryotes. Trends Biochem Sci 14: 105-110.

43. Mandel CR, Tong L (2007) How to get all "A"s in polyadenylation. Structure 15: 1024-1026.

44. Connelly S, Manley JL (1988) A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. Genes Dev 2: 440-452.

45. Whitelaw E, Proudfoot N (1986) 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 5: 2915-2922.

46. Russo P (1995) Saccharomyces cerevisiae mRNA 3' end forming signals are also involved in transcription termination. Yeast 11: 447-453.

47. Zaret KS, Sherman F (1982) DNA sequence required for efficient transcription termination in yeast. Cell 28: 563-573.

48. Birse CE, Minvielle-Sebastia L, Lee BA, Keller W, Proudfoot NJ (1998) Coupling termination of transcription to messenger RNA maturation in yeast. Science 280: 298-301.

49. Proudfoot NJ (2011) Ending the message: poly(A) signals then and now. Genes Dev 25: 1770-1782.

50. Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription. Cell 108: 501-512.

51. Proudfoot N (2000) Connecting transcription to messenger RNA processing. Trends Biochem Sci 25: 290-293.

52. Buratowski S (2005) Connections between mRNA 3' end processing and transcription termination. Curr Opin Cell Biol 17: 257-261.

53. Lykke-Andersen S, Jensen TH (2007) Overlapping pathways dictate termination of RNA polymerase II transcription. Biochimie 89: 1177-1182.

54. Richard P, Manley JL (2009) Transcription termination by nuclear RNA polymerases. Genes Dev 23: 1247-1269.

55. Proudfoot NJ (1986) Transcriptional interference and termination between duplicated alpha-globin gene constructs suggests a novel mechanism for gene regulation. Nature 322: 562-565.

56. Eggermont J, Proudfoot NJ (1993) Poly(A) signals and transcriptional pause sites combine to prevent interference between RNA polymerase II promoters. EMBO J 12: 2539-2548.

57. Prescott EM, Proudfoot NJ (2002) Transcriptional collision between convergent genes in budding yeast. Proc Natl Acad Sci U S A 99: 8796-8801.

58. Steinmetz EJ, Ng SB, Cloute JP, Brow DA (2006) cis- and trans-Acting determinants of transcription termination by yeast RNA polymerase II. Mol Cell Biol 26: 2688-2696.

59. Minvielle-Sebastia L, Keller W (1999) mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. Curr Opin Cell Biol 11: 352-357.

60. Gilmartin GM (2005) Eukaryotic mRNA 3' processing: a common means to different ends. Genes Dev 19: 2517-2521.

61. Mandel CR, Bai Y, Tong L (2008) Protein factors in pre-mRNA 3'-end processing. Cell Mol Life Sci 65: 1099-1122.

62. Millevoi S, Vagner S (2010) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. Nucleic Acids Res 38: 2757-2774.

63. Osborne BI, Guarente L (1989) Mutational analysis of a yeast transcriptional terminator. Proc Natl Acad Sci U S A 86: 4097-4101.

64. Hyman LE, Moore CL (1993) Termination and pausing of RNA polymerase II downstream of yeast polyadenylation sites. Mol Cell Biol 13: 5159-5167.

65. Sato K, Ito R, Baek KH, Agarwal K (1986) A specific DNA sequence controls termination of transcription in the gastrin gene. Mol Cell Biol 6: 1032-1043.

66. Connelly S, Manley JL (1989) A CCAAT box sequence in the adenovirus major late promoter functions as part of an RNA polymerase II termination signal. Cell 57: 561-571.

67. Glover-Cutter K, Kim S, Espinosa J, Bentley DL (2008) RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. Nat Struct Mol Biol 15: 71-78.

68. Ashfield R, Enriquez-Harris P, Proudfoot NJ (1991) Transcriptional termination between the closely linked human complement genes C2 and factor B: common termination factor for C2 and c-myc? EMBO J 10: 4197-4207.

69. Russo P, Sherman F (1989) Transcription terminates near the poly(A) site in the CYC1 gene of the yeast Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 86: 8348-8352.

70. Core LJ, Waterfall JJ, Lis JT (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322: 1845-1848.

71. Zhao J, Hyman L, Moore C (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev 63: 405-445.

72. Gross S, Moore C (2001) Five subunits are required for reconstitution of the cleavage and polyadenylation activities of Saccharomyces cerevisiae cleavage factor I. Proc Natl Acad Sci U S A 98: 6080-6085.

73. Kessler MM, Zhao J, Moore CL (1996) Purification of the Saccharomyces cerevisiae cleavage/polyadenylation factor I. Separation into two components that are required for both cleavage and polyadenylation of mRNA 3' ends. J Biol Chem 271: 27167-27175.

74. Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature 415: 141-147.

75. Kuehner JN, Pearson EL, Moore C (2011) Unravelling the means to an end: RNA polymerase II transcription termination. Nat Rev Mol Cell Biol 12: 283-294.

76. Noble CG, Beuth B, Taylor IA (2007) Structure of a nucleotide-bound Clp1-Pcf11 polyadenylation factor. Nucleic Acids Res 35: 87-99.

77. Gordon JM, Shikov S, Kuehner JN, Liriano M, Lee E, et al. (2011) Reconstitution of CF IA from overexpressed subunits reveals stoichiometry and provides insights into molecular topology. Biochemistry 50: 10203-10214.

78. Barilla D, Lee BA, Proudfoot NJ (2001) Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 98: 445-450.

79. Hammell CM, Gross S, Zenklusen D, Heath CV, Stutz F, et al. (2002) Coupling of termination, 3' processing, and mRNA export. Mol Cell Biol 22: 6441-6457.

80. Brodsky AS, Silver PA (2000) Pre-mRNA processing factors are required for nuclear export. RNA 6: 1737-1749.

81. Nino CA, Herissant L, Babour A, Dargemont C (2013) mRNA Nuclear Export in Yeast. Chem Rev.

82. Luo W, Johnson AW, Bentley DL (2006) The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. Genes Dev 20: 954-965.

83. Kim M, Vasiljeva L, Rando OJ, Zhelkovsky A, Moore C, et al. (2006) Distinct pathways for snoRNA and mRNA termination. Mol Cell 24: 723-734.

84. Zhang Z, Fu J, Gilmour DS (2005) CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3'-end processing factor, Pcf11. Genes Dev 19: 1572-1580.

85. Kyburz A, Sadowski M, Dichtl B, Keller W (2003) The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. Nucleic Acids Res 31: 3936-3945.

86. Takagaki Y, Manley JL, MacDonald CC, Wilusz J, Shenk T (1990) A multisubunit factor, CstF, is required for polyadenylation of mammalian pre-mRNAs. Genes Dev 4: 2112-2120.

87. Minvielle-Sebastia L, Preker PJ, Wiederkehr T, Strahm Y, Keller W (1997) The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. Proc Natl Acad Sci U S A 94: 7897-7902.

88. Paushkin SV, Patel M, Furia BS, Peltz SW, Trotta CR (2004) Identification of a human endonuclease complex reveals a link between tRNA splicing and pre-mRNA 3' end formation. Cell 117: 311-321.

89. Haddad R, Maurice F, Viphakone N, Voisinet-Hakil F, Fribourg S, et al. (2012) An essential role for Clp1 in assembly of polyadenylation complex CF IA and Pol II transcription termination. Nucleic Acids Res 40: 1226-1239.

90. Ramirez A, Shuman S, Schwer B (2008) Human RNA 5'-kinase (hClp1) can function as a tRNA splicing enzyme in vivo. RNA 14: 1737-1745.

91. Ghazy MA, Gordon JM, Lee SD, Singh BN, Bohm A, et al. (2012) The interaction of Pcf11 and Clp1 is needed for mRNA 3'-end formation and is modulated by amino acids in the ATP-binding site. Nucleic Acids Res 40: 1214-1225.

92. Holbein S, Scola S, Loll B, Dichtl BS, Hubner W, et al. (2011) The P-loop domain of yeast Clp1 mediates interactions between CF IA and CPF factors in pre-mRNA 3' end formation. PLoS One 6: e29139.

93. Medler S, Al Husini N, Raghunayakula S, Mukundan B, Aldea A, et al. (2011) Evidence for a complex of transcription factor IIB with poly(A) polymerase and cleavage factor 1 subunits required for gene looping. J Biol Chem 286: 33709-33718.

94. Chen J, Moore C (1992) Separation of factors required for cleavage and polyadenylation of yeast pre-mRNA. Mol Cell Biol 12: 3470-3481.

95. Logan J, Falck-Pedersen E, Darnell JE, Jr., Shenk T (1987) 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 U S A 84: 8306-8310.

96. Kerppola TK, Kane CM (1991) RNA polymerase: regulation of transcript elongation and termination. FASEB J 5: 2833-2842.

97. Rosonina E, Kaneko S, Manley JL (2006) Terminating the transcript: breaking up is hard to do. Genes Dev 20: 1050-1056.

98. Kireeva ML, Komissarova N, Kashlev M (2000) Overextended RNA:DNA hybrid as a negative regulator of RNA polymerase II processivity. J Mol Biol 299: 325-335.

99. Komissarova N, Becker J, Solter S, Kireeva M, Kashlev M (2002) Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. Mol Cell 10: 1151-1162.

100. West S, Proudfoot NJ, Dye MJ (2008) Molecular dissection of mammalian RNA polymerase II transcriptional termination. Mol Cell 29: 600-610.

101. Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, et al. (2006) Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell 24: 735-746.

102. Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A (2008) The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat Struct Mol Biol 15: 795-804.

103. Sun ZW, Hampsey M (1996) Synthetic enhancement of a TFIIB defect by a mutation in SSU72, an essential yeast gene encoding a novel protein that affects transcription start site selection in vivo. Mol Cell Biol 16: 1557-1566.

104. Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, et al. (2010) The genetic landscape of a cell. Science 327: 425-431.

105. Milgrom E, West RW, Jr., Gao C, Shen WC (2005) TFIID and Spt-Ada-Gcn5-acetyltransferase functions probed by genome-wide synthetic genetic array analysis using a Saccharomyces cerevisiae taf9-ts allele. Genetics 171: 959-973.

106. Collins SR, Kemmeren P, Zhao XC, Greenblatt JF, Spencer F, et al. (2007) Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. Mol Cell Proteomics 6: 439-450.

107. Wu WH, Pinto I, Chen BS, Hampsey M (1999) Mutational analysis of yeast TFIIB. A functional relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with TFIIB. Genetics 153: 643-652.

108. Jensen TH, Boulay J, Olesen JR, Colin J, Weyler M, et al. (2004) Modulation of transcription affects mRNP quality. Mol Cell 16: 235-244.

109. Wilmes GM, Bergkessel M, Bandyopadhyay S, Shales M, Braberg H, et al. (2008) A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. Mol Cell 32: 735-746.

110. Fiedler D, Braberg H, Mehta M, Chechik G, Cagney G, et al. (2009) Functional organization of the S. cerevisiae phosphorylation network. Cell 136: 952-963.

111. Holbein S, Wengi A, Decourty L, Freimoser FM, Jacquier A, et al. (2009) Cordycepin interferes with 3' end formation in yeast independently of its potential to terminate RNA chain elongation. RNA 15: 837-849.

112. Hang M, Smith MM (2011) Genetic analysis implicates the Set3/Hos2 histone deacetylase in the deposition and remodeling of nucleosomes containing H2A.Z. Genetics 187: 1053-1066.

113. Rodriguez-Navarro S, Fischer T, Luo MJ, Antunez O, Brettschneider S, et al. (2004) Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. Cell 116: 75-86.

114. Calvo O, Manley JL (2001) Evolutionarily conserved interaction between CstF-64 and PC4 links transcription, polyadenylation, and termination. Mol Cell 7: 1013-1023.

115. He X, Khan AU, Cheng H, Pappas DL, Jr., Hampsey M, et al. (2003) Functional interactions between the transcription and mRNA 3' end processing machineries mediated by Ssu72 and Sub1. Genes Dev 17: 1030-1042.

116. Ganem C, Devaux F, Torchet C, Jacq C, Quevillon-Cheruel S, et al. (2003) Ssu72 is a phosphatase essential for transcription termination of snoRNAs and specific mRNAs in yeast. EMBO J 22: 1588-1598.

117. Tu J, Song W, Carlson M (1996) Protein phosphatase type 1 interacts with proteins required for meiosis and other cellular processes in Saccharomyces cerevisiae. Mol Cell Biol 16: 4199-4206.

118. Estruch F, Cole CN (2003) An early function during transcription for the yeast mRNA export factor Dbp5p/Rat8p suggested by its genetic and physical interactions with transcription factor IIH components. Mol Biol Cell 14: 1664-1676.

119. Sanders SL, Jennings J, Canutescu A, Link AJ, Weil PA (2002) Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. Mol Cell Biol 22: 4723-4738.

120. El Kaderi B, Medler S, Raghunayakula S, Ansari A (2009) Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. J Biol Chem 284: 25015-25025.

121. Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, et al. (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature 440: 637-643.

122. Graumann J, Dunipace LA, Seol JH, McDonald WH, Yates JR, 3rd, et al. (2004) Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast. Mol Cell Proteomics 3: 226-237.

123. Drysdale CM, Jackson BM, McVeigh R, Klebanow ER, Bai Y, et al. (1998) The Gcn4p activation domain interacts specifically in vitro with RNA polymerase II holoenzyme, TFIID, and the Adap-Gcn5p coactivator complex. Mol Cell Biol 18: 1711-1724.

124. Lee KK, Sardiu ME, Swanson SK, Gilmore JM, Torok M, et al. (2011) Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. Mol Syst Biol 7: 503.

125. Pascual-Garcia P, Govind CK, Queralt E, Cuenca-Bono B, Llopis A, et al. (2008) Sus1 is recruited to coding regions and functions during transcription elongation in association with SAGA and TREX2. Genes Dev 22: 2811-2822.

126. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, et al. (2000) A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature 403: 623-627.

127. Ezeokonkwo C, Ghazy MA, Zhelkovsky A, Yeh PC, Moore C (2012) Novel interactions at the essential N-terminus of poly(A) polymerase that could regulate poly(A) addition in Saccharomyces cerevisiae. FEBS Lett 586: 1173-1178.

128. Breitkreutz A, Choi H, Sharom JR, Boucher L, Neduva V, et al. (2010) A global protein kinase and phosphatase interaction network in yeast. Science 328: 1043-1046.

129. De Wever V, Reiter W, Ballarini A, Ammerer G, Brocard C (2005) A dual role for PP1 in shaping the Msn2-dependent transcriptional response to glucose starvation. EMBO J 24: 4115-4123.

130. Dantonel JC, Murthy KG, Manley JL, Tora L (1997) Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. Nature 389: 399-402.

131. Wong J, Nakajima Y, Westermann S, Shang C, Kang JS, et al. (2007) A protein interaction map of the mitotic spindle. Mol Biol Cell 18: 3800-3809.

132. Ansari A, Hampsey M (2005) A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. Genes Dev 19: 2969-2978.

133. Kim M, Krogan NJ, Vasiljeva L, Rando OJ, Nedea E, et al. (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature 432: 517-522.

134. Sikorski TW, Ficarro SB, Holik J, Kim T, Rando OJ, et al. (2011) Sub1 and RPA associate with RNA polymerase II at different stages of transcription. Mol Cell 44: 397-409.

135. Calvo O, Manley JL (2005) The transcriptional coactivator PC4/Sub1 has multiple functions in RNA polymerase II transcription. EMBO J 24: 1009-1020.

136. Nedea E, He X, Kim M, Pootoolal J, Zhong G, et al. (2003) Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. J Biol Chem 278: 33000-33010.

137. O'Sullivan JM, Tan-Wong SM, Morillon A, Lee B, Coles J, et al. (2004) Gene loops juxtapose promoters and terminators in yeast. Nat Genet 36: 1014-1018.

138. O'Reilly D, Greaves DR (2007) Cell-type-specific expression of the human CD68 gene is associated with changes in Pol II phosphorylation and short-range intrachromosomal gene looping. Genomics 90: 407-415.

139. Singh BN, Hampsey M (2007) A transcription-independent role for TFIIB in gene looping. Mol Cell 27: 806-816.

140. Perkins KJ, Lusic M, Mitar I, Giacca M, Proudfoot NJ (2008) Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. Mol Cell 29: 56-68.

141. Tan-Wong SM, French JD, Proudfoot NJ, Brown MA (2008) Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. Proc Natl Acad Sci U S A 105: 5160-5165.

142. Yun K, So JS, Jash A, Im SH (2009) Lymphoid enhancer binding factor 1 regulates transcription through gene looping. J Immunol 183: 5129-5137.

143. Laine JP, Singh BN, Krishnamurthy S, Hampsey M (2009) A physiological role for gene loops in yeast. Genes Dev 23: 2604-2609.

144. Tan-Wong SM, Wijayatilake HD, Proudfoot NJ (2009) Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. Genes Dev 23: 2610-2624.

145. Yue X, Schwartz JC, Chu Y, Younger ST, Gagnon KT, et al. (2010) Transcriptional regulation by small RNAs at sequences downstream from 3' gene termini. Nat Chem Biol 6: 621-629.

146. Yadon AN, Singh BN, Hampsey M, Tsukiyama T (2013) DNA looping facilitates targeting of a chromatin remodeling enzyme. Mol Cell 50: 93-103.

147. Jash A, Yun K, Sahoo A, So JS, Im SH (2012) Looping mediated interaction between the promoter and 3' UTR regulates type II collagen expression in chondrocytes. PLoS One 7: e40828.

148. Le May N, Fradin D, Iltis I, Bougneres P, Egly JM (2012) XPG and XPF endonucleases trigger chromatin looping and DNA demethylation for accurate expression of activated genes. Mol Cell 47: 622-632.

149. Henriques T, Ji Z, Tan-Wong SM, Carmo AM, Tian B, et al. (2012) Transcription termination between polo and snap, two closely spaced tandem genes of D. melanogaster. Transcription 3: 198-212.

150. Crevillen P, Sonmez C, Wu Z, Dean C (2013) A gene loop containing the floral repressor FLC is disrupted in the early phase of vernalization. EMBO J 32: 140-148.

151. Dieci G, Sentenac A (1996) Facilitated recycling pathway for RNA polymerase III. Cell 84: 245-252.

152. Cochet-Meilhac M, Nuret P, Courvalin JC, Chambon P (1974) Animal DNA-dependent RNA polymerases. 12. Determination of the cellular number of RNA polymerase B molecules. Biochim Biophys Acta 353: 185-192.

153. Kimura H, Tao Y, Roeder RG, Cook PR (1999) Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Mol Cell Biol 19: 5383-5392.

154. Lee TI, Wyrick JJ, Koh SS, Jennings EG, Gadbois EL, et al. (1998) Interplay of positive and negative regulators in transcription initiation by RNA polymerase II holoenzyme. Mol Cell Biol 18: 4455-4462.

155. Ishihama A (2000) Functional modulation of Escherichia coli RNA polymerase. Annu Rev Microbiol 54: 499-518.

156. Borggrefe T, Davis R, Bareket-Samish A, Kornberg RD (2001) Quantitation of the RNA polymerase II transcription machinery in yeast. J Biol Chem 276: 47150-47153.

157. Hampsey M, Singh BN, Ansari A, Laine JP, Krishnamurthy S (2011) Control of eukaryotic gene expression: gene loops and transcriptional memory. Adv Enzyme Regul 51: 118-125.

158. Kundu S, Peterson CL (2010) Dominant role for signal transduction in the transcriptional memory of yeast GAL genes. Mol Cell Biol 30: 2330-2340.

159. Moabbi AM, Agarwal N, El Kaderi B, Ansari A (2012) Role for gene looping in intron-mediated enhancement of transcription. Proc Natl Acad Sci U S A 109: 8505-8510.

160. Maraia RJ, Kenan DJ, Keene JD (1994) Eukaryotic transcription termination factor La mediates transcript release and facilitates reinitiation by RNA polymerase III. Mol Cell Biol 14: 2147-2158.

161. Martin M, Cho J, Cesare AJ, Griffith JD, Attardi G (2005) Termination factor-mediated DNA loop between termination and initiation sites drives mitochondrial rRNA synthesis. Cell 123: 1227-1240.

162. Jansa P, Burek C, Sander EE, Grummt I (2001) The transcript release factor PTRF augments ribosomal gene transcription by facilitating reinitiation of RNA polymerase I. Nucleic Acids Res 29: 423-429.

163. Spitalny P, Thomm M (2008) A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. Mol Microbiol 67: 958-970.

164. Mapendano CK, Lykke-Andersen S, Kjems J, Bertrand E, Jensen TH (2010) Crosstalk between mRNA 3' end processing and transcription initiation. Mol Cell 40: 410-422.

165. Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, et al. (2008) Divergent transcription from active promoters. Science 322: 1849-1851.

166. Sigova AA, Mullen AC, Molinie B, Gupta S, Orlando DA, et al. (2013) Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. Proc Natl Acad Sci U S A 110: 2876-2881.

167. Preker P, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, et al. (2008) RNA exosome depletion reveals transcription upstream of active human promoters. Science 322: 1851-1854.

168. Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, et al. (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature 457: 1038-1042.

169. Flynn RA, Almada AE, Zamudio JR, Sharp PA (2011) Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. Proc Natl Acad Sci U S A 108: 10460-10465.

170. Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, et al. (2005) Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell 121: 725-737.

171. Jacquier A (2009) The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. Nat Rev Genet 10: 833-844.

172. Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, Libri D (2006) Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. Mol Cell 23: 853-864.

173. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, et al. (2007) High-resolution profiling of histone methylations in the human genome. Cell 129: 823-837.

174. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. Cell 130: 77-88.

175. Minsky N, Shema E, Field Y, Schuster M, Segal E, et al. (2008) Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. Nat Cell Biol 10: 483-488.

176. Marquardt S, Hazelbaker DZ, Buratowski S (2011) Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. Transcription 2: 145-154.

177. Brannan K, Kim H, Erickson B, Glover-Cutter K, Kim S, et al. (2012) mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. Mol Cell 46: 311-324.

178. Steinmetz EJ, Conrad NK, Brow DA, Corden JL (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. Nature 413: 327-331.

179. Mischo HE, Proudfoot NJ (2013) Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. Biochim Biophys Acta 1829: 174-185.

180. Morlando M, Greco P, Dichtl B, Fatica A, Keller W, et al. (2002) Functional analysis of yeast snoRNA and snRNA 3'-end formation mediated by uncoupling of cleavage and polyadenylation. Mol Cell Biol 22: 1379-1389.

181. Dheur S, Vo le TA, Voisinet-Hakil F, Minet M, Schmitter JM, et al. (2003) Pti1p and Ref2p found in association with the mRNA 3' end formation complex direct snoRNA maturation. EMBO J 22: 2831-2840.

182. Nedea E, Nalbant D, Xia D, Theoharis NT, Suter B, et al. (2008) The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes. Mol Cell 29: 577-587.

183. Ghazy MA, He X, Singh BN, Hampsey M, Moore C (2009) The essential N terminus of the Pta1 scaffold protein is required for snoRNA transcription termination and Ssu72 function but is dispensable for pre-mRNA 3'-end processing. Mol Cell Biol 29: 2296-2307.

184. Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA (2013) Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature 499: 360-363.

185. Ntini E, Jarvelin AI, Bornholdt J, Chen Y, Boyd M, et al. (2013) Polyadenylation site-induced decay of upstream transcripts enforces promoter directionality. Nat Struct Mol Biol 20: 923-928.

186. Kessler MM, Henry MF, Shen E, Zhao J, Gross S, et al. (1997) Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. Genes Dev 11: 2545-2556.

187. Patel D, Butler JS (1992) Conditional defect in mRNA 3' end processing caused by a mutation in the gene for poly(A) polymerase. Mol Cell Biol 12: 3297-3304.

188. Barnwal RP, Lee SD, Moore C, Varani G (2012) Structural and biochemical analysis of the assembly and function of the yeast pre-mRNA 3' end processing complex CF I. Proc Natl Acad Sci U S A 109: 21342-21347.

189. Chen S, Hyman LE (1998) A specific RNA-protein interaction at yeast polyadenylation efficiency elements. Nucleic Acids Res 26: 4965-4974.

190. Henry M, Borland CZ, Bossie M, Silver PA (1996) Potential RNA binding proteins in Saccharomyces cerevisiae identified as suppressors of temperature-sensitive mutations in NPL3. Genetics 142: 103-115.

191. Twu JS, Bretthauer RK (1971) Properties of a polyriboadenylate polymerase isolated from yeast ribosomes. Biochemistry 10: 1576-1582.

192. Hafe LA, Keller EB (1975) The polyadenylate polymerases from yeast. J Biol Chem 250: 1838-1846.

193. Lingner J, Kellermann J, Keller W (1991) Cloning and expression of the essential gene for poly(A) polymerase from S. cerevisiae. Nature 354: 496-498.

194. Thomas D, Barbey R, Surdin-Kerjan Y (1990) Gene-enzyme relationship in the sulfate assimilation pathway of Saccharomyces cerevisiae. Study of the 3'-phosphoadenylylsulfate reductase structural gene. J Biol Chem 265: 15518-15524.

195. Griac P, Henry SA (1999) The yeast inositol-sensitive upstream activating sequence, UASINO, responds to nitrogen availability. Nucleic Acids Res 27: 2043-2050.

196. Hahn S, Young ET (2011) Transcriptional regulation in Saccharomyces cerevisiae: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. Genetics 189: 705-736.

197. Spitz F, Furlong EE (2012) Transcription factors: from enhancer binding to developmental control. Nat Rev Genet 13: 613-626.

198. Mischo HE, Proudfoot NJ (2012) Disengaging polymerase: Terminating RNA polymerase II transcription in budding yeast. Biochim Biophys Acta.

199. Calvo O, Manley JL (2003) Strange bedfellows: polyadenylation factors at the promoter. Genes Dev 17: 1321-1327.

200. Wang Y, Roberts SG (2010) New insights into the role of TFIIB in transcription initiation. Transcription 1: 126-129.

201. Lykke-Andersen S, Mapendano CK, Jensen TH (2011) An ending is a new beginning: transcription termination supports re-initiation. Cell Cycle 10: 863-865.

202. Shandilya J, Wang Y, Roberts SG (2012) TFIIB dephosphorylation links transcription inhibition with the p53-dependent DNA damage response. Proc Natl Acad Sci U S A 109: 18797-18802.

203. Wang Y, Fairley JA, Roberts SG (2010) Phosphorylation of TFIIB links transcription initiation and termination. Curr Biol 20: 548-553.

204. Mukundan B, Ansari A (2011) Novel role for mediator complex subunit Srb5/Med18 in termination of transcription. J Biol Chem 286: 37053-37057.

205. Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, et al. (2006) Proteome survey reveals modularity of the yeast cell machinery. Nature 440: 631-636.

206. Dichtl B, Blank D, Ohnacker M, Friedlein A, Roeder D, et al. (2002) A role for SSU72 in balancing RNA polymerase II transcription elongation and termination. Mol Cell 10: 1139-1150.

207. Collins SR, Miller KM, Maas NL, Roguev A, Fillingham J, et al. (2007) Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. Nature 446: 806-810.

208. Zhang Y, Zhang M, Zhang Y (2011) Crystal structure of Ssu72, an essential eukaryotic phosphatase specific for the C-terminal domain of RNA polymerase II, in complex with a transition state analogue. Biochem J 434: 435-444.

209. Ben-Aroya S, Coombes C, Kwok T, O'Donnell KA, Boeke JD, et al. (2008) Toward a comprehensive temperature-sensitive mutant repository of the essential genes of Saccharomyces cerevisiae. Mol Cell 30: 248-258.

210. Petersen JG, Kielland-Brandt MC, Nilsson-Tillgren T, Bornaes C, Holmberg S (1988) Molecular genetics of serine and threonine catabolism in Saccharomyces cerevisiae. Genetics 119: 527-534.

211. Zawel L, Kumar KP, Reinberg D (1995) Recycling of the general transcription factors during RNA polymerase II transcription. Genes Dev 9: 1479-1490.

212. Ranish JA, Yudkovsky N, Hahn S (1999) Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. Genes Dev 13: 49-63.

213. Costa FF (2010) Non-coding RNAs: Meet thy masters. Bioessays 32: 599-608.

214. El Kaderi B, Medler S, Ansari A (2012) Analysis of interactions between genomic loci through Chromosome Conformation Capture (3C). Curr Protoc Cell Biol Chapter 22: Unit22 15.

215. Al Husini N, Kudla P, Ansari A (2013) A role for CF1A 3' end processing complex in promoter-associated transcription. PLoS Genet 9: e1003722.

216. Mayer A, Lidschreiber M, Siebert M, Leike K, Soding J, et al. (2010) Uniform transitions of the general RNA polymerase II transcription complex. Nat Struct Mol Biol 17: 1272-1278.

217. Xu Z, Wei W, Gagneur J, Clauder-Munster S, Smolik M, et al. (2011) Antisense expression increases gene expression variability and locus interdependency. Mol Syst Biol 7: 468.

218. Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, et al. (2011) A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. Cell 145: 622-634.

219. Murray SC, Serra Barros A, Brown DA, Dudek P, Ayling J, et al. (2012) A pre-initiation complex at the 3'-end of genes drives antisense transcription independent of divergent sense transcription. Nucleic Acids Res 40: 2432-2444.

220. Malik S, Durairaj G, Bhaumik SR (2013) Mechanisms of antisense transcription initiation from the 3' end of the GAL10 coding sequence in vivo. Mol Cell Biol 33: 3549-3567.

221. Hongay CF, Grisafi PL, Galitski T, Fink GR (2006) Antisense transcription controls cell fate in Saccharomyces cerevisiae. Cell 127: 735-745.

222. Camblong J, Iglesias N, Fickentscher C, Dieppois G, Stutz F (2007) Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. Cell 131: 706-717.

223. Berretta J, Pinskaya M, Morillon A (2008) A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. Genes Dev 22: 615-626.

224. Zhang DW, Mosley AL, Ramisetty SR, Rodriguez-Molina JB, Washburn MP, et al. (2012) Ssu72 phosphatase-dependent erasure of phospho-Ser7 marks on the RNA polymerase II C-terminal domain is essential for viability and transcription termination. J Biol Chem 287: 8541-8551.

225. Wei W, Pelechano V, Jarvelin AI, Steinmetz LM (2011) Functional consequences of bidirectional promoters. Trends Genet 27: 267-276.

226. Sun H, Wu J, Wickramasinghe P, Pal S, Gupta R, et al. (2011) Genome-wide mapping of RNA Pol-II promoter usage in mouse tissues by ChIP-seq. Nucleic Acids Res 39: 190-201.

227. Churchman LS, Weissman JS (2011) Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 469: 368-373.

228. Nagalakshmi U, Waern K, Snyder M (2010) RNA-Seq: a method for comprehensive transcriptome analysis. Curr Protoc Mol Biol Chapter 4: Unit 4 11 11-13.

229. Birse CE, Lee BA, Hansen K, Proudfoot NJ (1997) Transcriptional termination signals for RNA polymerase II in fission yeast. EMBO J 16: 3633-3643.

230. Hirayoshi K, Lis JT (1999) Nuclear run-on assays: assessing transcription by measuring density of engaged RNA polymerases. Methods Enzymol 304: 351-362.

ABSTRACT

AN INVESTIGATION INTO THE ROLE OF CFIA 3' END PROCESSING COMPLEX IN THE TERMINATION AND INITIATION/REINITIATION OF TRANSCRIPTION

by

NADRA AI-HUSINI

May 2014

Advisor: Dr. Athar Ansari

Major: Biology

Degree: Doctor of Philosophy

In budding yeast, as in higher eukaryotes, transcription of protein coding genes is executed by a highly specialized, conserved polymerase called RNA polymerase II (RNAPII). The transcription cycle of RNAPII has four major steps: initiation, elongation, termination, and reinitiation. The successful accomplishment of each of these steps requires a number of accessory factors. Many of these factors operate at multiple steps in the transcription cycle. The major focus of this study was to examine the function of Clp1, which is an RNA processing factor operating at the 3′ end of genes, in the transcription cycle. Clp1 is one of the four subunits of the CFIA 3′ end processing complex. It is the least investigated CFIA subunit. The role of the other three subunits of the CFIA complex in 3' end processing and termination of transcription is well documented.

Here we investigate the role of Clp1 in the initiation as well as the termination of transcription. We used a temperature-sensitive mutant of Clp1 to assess its function. We demonstrated a direct role for this factor in the termination of transcription of *CHA1.* We used three different approaches; TRO assay, RNAPII-ChIP assay, and strand

specific RT-PCR, to demonstrate the termination function of Clp1. In addition, we showed that Clp1 is also involved in the early steps of the transcription. Our results strongly suggest that Clp1 participates in promoter-associated transcription. We provide multiple lines of evidence in support of a role for Clp1 at the 5′ end of genes. First, the presence of Clp1 in the vicinity of the promoter region implies its involvement early in the transcription cycle. Second, the decrease in RNAPII density near the promoter without a parallel decrease in the level of the GTFs suggested a role for Clp1 in reinitiation of transcription. Third, an increase in 5' initiated antisense divergent transcripts in the Clp1 mutant supports a role for the factor in providing directionality to the promoter-bound polymerase. To assess the generality of the observed functions of Clp1, we investigated the role of Clp1 in the transcription cycle on a genomewide scale using GRO-Seq approach. Our results show that the number of transcriptionally active genes decreased by at least two-fold in the clp1mutant. The GRO-Seq results strongly suggest a genomewide function for Clp1 in the termination of transcription, and indicate that Clp1 is required for the pausing of RNAPII that is a pre-requisite for the termination of transcription. We also observed a dramatic increase in 3' initiated antisense transcription in the absence of a functional Clp1 protein.

Using the chromosome conformation capture approach, CCC, we observed a role for Clp1 in gene loop formation. We found a strong correlation between the Clp1 function in gene looping, and its role in promoter-associated transcription which implies gene looping as the means through which this factor is exerting its functions at the 5′ end of genes.

AUTOBIOGRAPHICAL STATEMENT

NADRA AL-HUSINI

EDUCATION

2008-2013- **Ph.D** in **Molecular Biology**, Wayne State University, Detroit, Michigan, USA

2006-2007-   **M.Sc** in Biotechnology, Wayne State University, Detroit, Michigan, USA.

1993-1997-   **B.Sc**. in Microbiology, Taiz University, Taiz, Yemen

PUBLICATIONS

Al Husini N, Kudla P, Ansari A. (2013) A role for CF1A 3' end processing complex in promoter-associated transcription. PLoS Genet. 2013 Aug; 9(8):e1003722.

Medler S, **Al Husini N**, Raghunayakula S, Mukundan B, Aldea A, Ansari A. (2011) Evidence for a complex of transcription factor IIB with poly(A) polymerase and cleavage factor 1 subunits required for gene looping.
J Biol Chem. 2011 Sep 30; 286(39):33709-18

AWARDS

2011   **Best poster award**. Biological Sciences Departmental Retreat. WSU, Detroit, Michigan, USA

2012   **Best Poster Award**. The Graduate Student Research Symposium (GSRD XVI)Medical School, WSU, Detroit, Michigan

2012   **Best poster award**. Biological Sciences Departmental Retreat. WSU, Detroit, Michigan, USA

2012   **Best Graduate Research Assistant** (GRA), Biological Sciences Department, WSU, Detrit, Michigan, USA

2012/2013   **Rumble Fellowship award.** Wayne State University, Detroit, Michigan, USA