2014

Extra-Transcriptional Effects of Chromatin Bound RNA Polymerase III Transcription Complexes

Asawari Korde

Louisiana State University and Agricultural and Mechanical College, akorde1@lsu.edu

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Recommended Citation

https://digitalcommons.lsu.edu/gradschool_dissertations/274

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contactgradetd@lsu.edu.
EXTRA-TRANSCRIPTIONAL EFFECTS OF CHROMATIN BOUND RNA POLYMERASE III TRANSCRIPTION COMPLEXES

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

The Department of Biological Sciences

by

Asawari Korde
B.S., University of Mumbai, India, 2005
M.S., University of Mumbai, India, 2007
August 2014
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. David Donze, who has made enormous contribution to the completion of my dissertation. His guidance, motivation, enthusiasm, and immense knowledge have encouraged me to develop my scientific abilities and to continue my career in research. It is rare to find such qualities as excellent teacher, brilliant researcher and a great human-being rolled into one person. I feel really lucky to have him as my mentor.

I would like to thank my excellent committee members Dr. Anne Grove, Dr. Patrick DiMario and Dr. Craig Hart for their continuous support, willingness to help, and insightful advices. My special thanks to Dr. Hart, who provided the opportunity to rotate in his lab and to learn basic fly molecular work. I thank Dr. James H. Oard (dean’s representative) for his valuable time and comments. Thanks to Dr. Maheshi Dassanayake and Dr. Dong-Ha Oh for their prompt help and advice on the RNA-seq project. Further, I would like to thank Dr. Alvaro at the Roy. J. Carver biotechnology center, at University of Illinois, who did an excellent job in sequencing our RNA samples. I am thankful to Dr. Kim for his timely advice and permission to use RT-PCR machine in his lab.

I would like to acknowledge the financial and academic support of Louisiana State University and its staff. I am particularly thankful to Ms. Prissy Milligan, administrative specialist (Graduate Studies) who makes the life of graduate students so much easier.

Thanks to my senior lab members, Dr. Kristy Brumfield and Dr. Richard Kleinschmidt for teaching me basic lab techniques. My special thanks to Kimberly Leblanc for being a wonderful lab-mate and friend. With her, surely, it was the most memorable time in the lab. Thanks to my co-author Jessica Rosselot, for her initial important work in identifying readthrough transcript. I am thankful to my group members Chance Nowak and Qing Wang for their major contribution in
RNA-seq experiment. I would like to thank Chance and other lab members, Kyle Kern and Dr. Raphyel Rosby for updating me with current science. My special thanks to other friends in the Department of Biological sciences including, Hana Kim for her generous support in RT-PCR technique, Wesley Frey for commenting on my most exciting and most confusing experimental results during my Ph.D. and Himanshu Raje for his valuable time and encouragement during my general examination.

My parents and brother have given me their unequivocal support throughout, as always, for which my mere expression of thanks likewise does not suffice. I would like to extend my thanks to my in-laws for their patience and encouragement. Above all, I would like to express the deepest appreciation to my husband, Tushar Kale for his enormous support, great patience at all times and even active participation in thesis formatting.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................................................................................. ii

LIST OF TABLES ................................................................................................................................................ v

LIST OF FIGURES .............................................................................................................................................. vi

ABSTRACT ......................................................................................................................................................... viii

CHAPTER 1  INTRODUCTION ............................................................................................................................. 1

CHAPTER 2  INTERGENIC TRANSCRIPTIONAL INTERFERENCE IS BLOCKED BY RNA POLYMERASE III TRANSCRIPTION FACTOR TFIIB IN SACCHAROMYCES CEREVISIAE* ................................................. 29

CHAPTER 3  FUNCTIONAL ANALYSIS OF ETC4 SITE AT RAD2-TNA1 LOCUS ........... 60

CHAPTER 4  GLOBAL ANALYSIS OF EFFECTS OF RNA POLYMERASE III TRANSCRIPTION COMPLEXES ON RNA POLYMERASE II TRANSCRIBED LOCI ........................................................................ 79

CHAPTER 5  DISCUSSION .................................................................................................................................. 95

REFERENCES .................................................................................................................................................... 115

APPENDIX: AUTHORIZATION FOR USE OF PUBLISHED MATERIAL ......................... 136

VITA ................................................................................................................................................................. 137
LIST OF TABLES

Table 2.1  Oligonucleotides used in this study ............................................................... 33
Table 2.2  Plasmids used in this study ........................................................................... 36
Table 2.3  *S. cerevisiae* strains generated in this study ................................................. 37
Table 3.1  Plasmids used in this study ........................................................................... 66
Table 3.2  Strains used in this study .............................................................................. 66
Table 3.3  Oligonucleotides used in this study ............................................................... 66
Table 4.1  Strains used in this study .............................................................................. 84
Table 4.2  Commands used for read alignment ............................................................. 87
Table 4.3  Commands used for DEseq analysis ............................................................... 89
### LIST OF FIGURES

| Figure 1.1. | Schematic of nucleosome assembly and histone modifications in eukaryotes. | 4 |
| Figure 1.2. | Chromatin remodeling protein families conserved from yeast to human. | 11 |
| Figure 1.3. | Schematic of typical promoters for RNA polymerase I, II and III. | 16 |
| Figure 1.4. | Transcription of a tRNA gene. | 18 |
| Figure 1.5. | Assembly of RNA polymerase III transcription complexes. | 19 |
| Figure 1.6. | Extra-transcriptional functions of Pol III transcription complexes bound sites in *S. cerevisiae*. | 22 |
| Figure 1.7. | Schematic for chromatin boundary activity. | 27 |
| Figure 2.1. | Strategy for RT-PCR quantitation of *SUT467-ATG31* extended transcripts. | 40 |
| Figure 2.2. | Mutation of the tV(UAC)D tRNA gene upstream of ATG31 results in read-through of the intergenic SUT467 transcript. | 42 |
| Figure 2.3. | Pol III transcription factors are required to block SUT467 readthrough transcription. | 45 |
| Figure 2.4. | *S. cerevisiae* Chromosome IV annotation and locations of 5’ ends of normal, intermediate, and long ATG31 transcripts. | 47 |
| Figure 2.5. | Binding of the TFIIIB complex is associated with blocking of SUT467 readthrough transcription. | 48 |
| Figure 2.6. | Genetic factors involved in tDNA chromatin boundary function have minimal effects on blocking of Pol II progression through tV(UAC)D. | 50 |
| Figure 2.7. | ATG31 translation is inhibited from the extended transcript under normal growth conditions. | 52 |
| Figure 2.8. | Mutations in tV(UAC)D inhibit function of ATG31 protein. | 53 |
| Figure 3.1. | Conserved B-box sequence across ETC sites. | 61 |
| Figure 3.2. | Predicted position of ETC4 conserved B-box sequence with respect to RAD2 and Sum1p binding sites. | 63 |
| Figure 3.3. | Sum1p mediated repression of TNA1 gene. | 63 |
LIST OF FIGURES (CONTINUED)

Figure 3.4. Hypothesis-TFIIIC bound ETC4 site acts as a barrier to spread of repression. .... 65
Figure 3.5. In absence of ETC4, survival of cells is significantly reduced upon UV exposure. ................................................................. 72
Figure 3.6. Northern analysis of RAD2 transcription. ................................................................. 72
Figure 3.7. Replacement of RAD2 ORF by URA3 coding sequence. ......................................... 73
Figure 3.8. Increase in mRNA levels indicates derepression of URA3 in SUM1 deleted mutants. ........................................................................ 74
Figure 3.9. Growth assays for rad2::URA3 strains suggest function of Ura3p is affected in etc4∆ mutants..................................................... 75
Figure 3.10. Deletion of ETC4 did not affect transcription of RAD2........................................... 76
Figure 4.1. Schematic of gene expression process. ................................................................. 80
Figure 4.2. Quality assessment of reads.................................................................................. 86
Figure 4.3. Representation of DESeq analysis........................................................................ 91
Figure 4.4. Visualization of potential extra-transcriptional effects on representative genes. ... 93
Figure 5.1. Binding sites for chromatin-associated proteins detected by Venters et al. at tV(UAC)D gene (A) and ETC4 locus (B). ................................................................. 101
Figure 5.2. Model for blocking intergenic transcriptional interference at tV(UAC)D locus. . 106
Figure 5.3. Functional categories of differentially expressed genes........................................ 108
Figure 5.4. Categorization of mis-expressed Pol II-transcribed loci. ........................................ 109
ABSTRACT

Transcription by RNA polymerase III (Pol III) requires sequential assembly of Pol III-specific transcription factors. At the tRNA gene, the TFIIC complex recognizes and specifically binds at intragenic promoter elements A-box and B-box and aids the assembly of TFIIB to upstream of the transcriptional start site. Upon binding, Pol III is recruited near start sites and transcription of tRNA genes is initiated. Apart from transcription of a gene, these bound Pol III complexes influence transcription, chromatin state and genome organization of neighboring RNA polymerase II (Pol II)-transcribed genes. Such effects are known as extra-transcriptional effects of Pol III complex.

Our study provides evidence of a unique “extra-transcriptional” activity of assembled Pol III transcription complexes at the tRNA gene that blocks progression of intergenic RNA polymerase II transcription. We demonstrated that the Pol III transcription complex bound to the tRNA gene upstream of the Saccharomyces cerevisiae ATG31 gene protects the ATG31 promoter against readthrough transcriptional interference from the upstream non-coding intergenic SUT467. The protection is predominately mediated by binding of the TFIIB complex. Failure to block this readthrough resulted in compromised ATG31 translation. Given the recent discovery of widespread pervasive transcription in yeast, protection of neighboring genes from intergenic transcriptional interference may be a key extra-transcriptional function of assembled RNA polymerase III complexes.

Our data from RNA-seq analysis demonstrated genome-wide effects of DNA bound Pol III complexes on neighboring chromosomal loci, by comparing expression profiles from tfc6 under-expressing mutants and wild-type S. cerevisiae strains. Reduced TFIIC occupancy in mutant strains altered Pol II derived transcripts and displayed 5’ or 3’ extension of protein-coding
genes, readthrough from non-coding transcripts and increase in the transcription of genes near the potential TFIIIC binding sites, including tRNA genes and putative ETC sites. Interestingly, not all genes in the vicinity of TFIIIC binding sites were transcriptionally mis-regulated, suggesting variable strength of influence on Pol II transcripts by TFIIIC bound sites. Finally, as observed in SUT467-ATG31 readthrough, we anticipated translation defects in 5’ or 3’ extended transcripts in mutants. Overall these genome-wide results suggest much complex regulatory role of Pol III transcription factors bound sites than previously anticipated.
CHAPTER 1
INTRODUCTION

Regulation of gene expression is vital during the life of a cell. Besides housekeeping genes that are constitutively active, various genes are only activated at specific developmental stages or under particular environmental conditions. Many regulatory mechanisms exist during cellular processes like replication, transcription and translation that control gene activity; therefore these mechanisms are necessary for proper functioning of a cell. Intrinsic properties of chromatin play a role in regulating each of these mechanisms.

Chromatin structure

In eukaryotic cells, long linear strands of DNA are folded into a confined nucleus. In order to achieve this higher level of compaction, DNA is coiled around highly basic proteins known as histones that neutralize the negative charge on DNA. There are 14 contact points between core histone proteins and the DNA duplex that contribute to the stable histone-DNA complex, and together with other associated non-histone proteins are referred to as chromatin. The nucleosome is the unit of chromatin that comprises a histone octamer made up of two molecules each of the histones H2A, H2B, H3, and H4, which binds and wraps about 1.7 turns (~147 bp) of DNA (Figure 1.1 part-A). The nucleosomes are separated from each other by short stretches of linker DNA (~18 bp in *S. cerevisiae* and up to ~60 bp in human). In multicellular eukaryotes, this linker DNA can be associated with histone H1 protein, which is not a part of core histone octamer (JANSEN and VERSTREPN 2011).

Nucleosome mapping studies have revealed that some nucleosomes are organized and highly positioned; that is, they occur at almost the same position within a population of cells. Positioning of nucleosomes partly depends on the underlying DNA sequence, for instance high percentage of AT content correlates with less nucleosome occupancy (JANSEN and VERSTREPN 2011).
2011), and is in part dictated by the presence of other DNA bound proteins. One consequence is that occupancy of nucleosomes can make the underlying DNA inaccessible to other proteins. The accessibility of the genetic information stored in DNA is connected to mechanisms that either work on chromatin or are influenced by chromatin. There are two main enzymatic activities that regulate chromatin access: chromatin modifying complexes that modify histones and chromatin remodeling complexes that recognize histone modifications and accordingly help reorganize nucleosomes. These activities in turn change nucleosome positioning by altering DNA-protein interactions in chromatin complexes. This results in open or closed states chromatin that eventually allows or prevents binding of regulatory proteins during various cellular processes including DNA replication and transcription. Broadly, depending on these states, chromatin is categorized into heterochromatin or euchromatin. Heterochromatin is generally a condensed structure containing closely spaced nucleosomes, is generally transcriptionally inactive, and can exist at many repetitive sequences and late replicating genes. Whereas, euchromatin is loosely packed with relatively fewer nucleosomes, is transcriptionally active and contains the majority of the genes, both actively transcribed and quiescent. Both heterochromatin and euchromatin are present in an alternating pattern of multi-gene blocks that allows coordinate co-regulation of groups of genes within the structural domains. The dynamic nature of chromatin allows change in these patterns during development of a cell. Some chromosomal regions such as centromere, telomere and mating loci in yeast remain condensed throughout the cell cycle, and are referred as constitutive heterochromatin. However, facultative heterochromatic regions change in response to gene activities.
**Chromatin modifications**

Chromatin modifications can occur either on DNA or on histones and are either associated with transcriptional activation or repression. Some modifications are reversible, and create additional layers of flexibility beyond the DNA sequence level.

DNA methylation directly acts on DNA and is widely conserved among higher eukaryotes. Methylation of cytosine residues within CpG islands on gene promoters is a primary epigenetic event that acts to suppress gene expression (Bird et al. 2002). DNA methylation accounts for the specific repression of genes in differentiated cells and for the stable silencing of transposable elements (Lippman et al. 2004). The correlation between DNA methylation and gene silencing has been extensively documented by a large body of evidence. Gene silencing can be distinguished from promoter-specific gene repression in that it acts in a regional and gene independent manner. It leads to transcriptional inactivation of larger regions of DNA by heterochromatin to inhibit access for DNA binding proteins or factors of the transcriptional machinery. Hence, chromatin plays an important role in determining the transcriptional status of a gene.

**Histone modification**

Other chromatin modifying events target globular core domains of histones as well as N- or C-terminal tails. These tails protrude from the nucleosome, exposing 20-35 residues (Luger et al. 1997). A variety of histone posttranslational modifications have been studied including acetylation, methylation, phosphorylation, ubiquitination, sumoylation, citrullination and ADP-ribosylation, and some of these important modifications are represented in Figure 1.1 part-B.

**Histone acetylation and deacetylation**

Histone acetylation and deacetylation are the most extensively studied histone modifications. During acetylation, an acetyl group is introduced to the positively charged ε-amino
group of lysine residues within the core histone protein or its N-terminal tail by the histone acetyltransferase enzymes. This reaction neutralizes the interaction of the N-termini of histones with the negatively charged DNA. As a consequence, the condensed chromatin is transformed into a more relaxed structure that is associated with greater levels of gene transcription (Grunstein 1997). This relaxation can be reversed by histone deacetylase enzymes which remove the acetyl groups, promoting the condensation of chromatin. Therefore, acetylation is generally associated with transcriptionally active chromatin and deacetylation with inactive chromatin (Grunstein 1997; Struhl 1998; Graff and Tsal 2013). Besides transcriptional regulation, histone acetylation is also involved in processes including replication and nucleosome assembly, higher-order chromatin packing and interactions of non-histone proteins with nucleosomes (Millar and Grunstein 2006).

Figure 1.1. Schematic of nucleosome assembly and histone modifications in eukaryotes. Adapted from (Xu et al. 2013).

A) Nucleosome consists of ~150 base pairs of DNA wound around a histone octamer, which is composed of two molecules of each: H2A (green), H2B (orange), H3 (brown) and H4 (blue) histone proteins. B) Major histone modifications occur on the protruding N-terminal tails but some of the histone modifications occur on C-terminal tails as well as globular domains of histone proteins. Key for histone modifications: Acetylation (A), methylation (M), phosphorylation (P) and ubiquitination (U).
There are several families of histone acetyltransferases. Gcn5 is a member of GNAT superfamily and involved in transcription initiation. It commonly modifies H3K4 but when associated with native complexes such as SAGA complex (in yeast), preferentially modifies a broad range of lysines on histone H3 and H2B (MILLAR and GRUNSTEIN 2006). Sas3, a member of the MYST family, modifies particularly histone H4 and H2A (BIRD et al. 2002).

Esa1 and Gcn5 are responsible for widespread acetylation of the region around the PHO5 gene, and thereby regulates PHO5 in yeast. MOF protein in Drosophila (Esa1 homologue) acetylates H4K16 on male X-chromosomes and increases transcription, commonly known as dosage compensation. The mammalian Esa1 homologue, TIP60, is involved with many aspects of the DNA damage repair pathway. Another acetyltransferase lies within the TAFII250 subunit of the basal transcription factor TFIIID complex (GRANT 2001; ROTH et al. 2001).

Histone deacetylases act as corepressors. A well-known HDAC in yeast is Rpd3p which is responsible for decetylating coding regions of transcriptionally active genes to prevent aberrant transcription initiation (LEE and SHILATIFARD 2007). In yeast, deacetylation of H4K16 is important for Sir protein mediated heterochromatin propagation. Sir2 is the nicotinamide adenine dinucleotide (NAD⁺) dependent histone deacetylase that sequentially deacetylates nucleosomes. The deacetylated histone tails then provide a platform for Sir3 and Sir4 interactions, which then recruit additional Sir2 to propagate the process to adjacent nucleosomes. As Sir proteins spread via the sequential deacylation mechanism that is independent of DNA sequence, they generate a specialized chromatin structure that is restrictive to transcription known as silenced chromatin. Silenced chromatin in yeast is found at the telomeres and cryptic mating-type loci, HMLα and HMRa (DONZE 2003; RUSCHE et al. 2003). Hst1 (Homology to Sir Two) is another NAD⁺ dependent deacetylase but unlike Sir2, it does not normally spread. It is part of the SUM1
complex (Sum1/Rfm1/Hst1) that represses over fifty genes involved in sporulation, NAD\(^+\) biosynthesis, and \(\alpha\)-cell identity (Bedalov et al. 2003; Zill and Rine 2008). In SIR2 mutants, overexpression of Hstp was found to restore silencing at HMR (Wang et al. 2002).

**Histone methylation**

Core histones can be methylated at amino groups of lysine and arginine residues only, but methylation is most commonly observed on lysines of H3 and H4 histone proteins (for example, H3K4, H3K9, H3K27, H4K20, H3K36 and H3K79). Lysine is able to be mono-, di-, or trimethylated by replacing each hydrogen of its NH\(_3\)\(^+\) group with a methyl group (Strahl et al. 1999; Strahl and Allis 2000). Arginine is able to be mono- or dimethylated with a free NH\(_2\) and NH\(_2\)\(^+\) group. Methylation of an arginine residue requires a complex including protein arginine methyltransferase (PRMT) while lysine requires a specific histone methyltransferase (HMT). The lysine-specific transferases are further broken down into whether or not they have a SET domain or a non-SET domain (Zhang and Reinberg 2001).

Methylated histones can either repress or activate transcription by blocking or encouraging DNA access to transcription factors. (Rice et al. 2003) Common methylation sites that are associated with gene activation include H3K4me2/3 and H3K79me3 whereas in higher eukaryotes, methylation of H3K9me2/3, H3K27me2/3, and H3K20me3 are associated with repression (Gilbert, S. F. 2010). Different degrees of residue methylation can confer different functions. For instance, monomethylated H4K20 (H4K20me1) is involved in the compaction of chromatin and therefore transcriptional repression. However, when dimethylated, it provides a platform for the binding of proteins involved in the repair of double-stranded DNA breaks (Zhang and Reinberg 2001). Histone methylation in *Saccharomyces cerevisiae* was long thought to be limited to H3K4, H3K36 and H3K79, catalyzed by the Set1, Set2 and Dot1 methyltransferases, respectively.
(Millar and Grunstein 2006). It is notable that these modifications have all been found to correlate with transcriptional activity, while the methylated marks that are absent in budding yeast (H3K9, H3K27 and H4K20) are associated with repression (Pokholyk et al. 2005). This picture has very recently been slightly altered by the findings of also H2BK37, H3R2, H3K42 and H4K20 as targets of methylation in budding yeast (Edwards et al. 2011; Gardner et al. 2011; Hyland et al. 2011), as well as indications that low levels of methylated H3K9 might be present (Garcia et al. 2007). However, in all these cases the responsible enzymes are as of yet unknown.

The activities of histone methyltransferases are counter balanced by the activity of histone demethylases that remove methyl groups from histone proteins. A tight regulation of both activities is necessary; loss of regulation leads to continuous gene expression that can result in cancer. Owing to this reversible nature, many anti-cancer therapies are targeted at methylation process (Albert and Helin 2010).

Histone phosphorylation

Histone H3 phosphorylation is highly conserved among eukaryotes from yeast to human, and has been extensively studied for many years. Phosphorylation of S10, T11 and S28 of histone H3 were found to be associated with chromosome condensation and segregation and conversely with H3 acetylation thus involved in transcriptional activation (Rossetto et al. 2012). Additionally, phosphorylation of H4S1 and H2BS10 was also demonstrated to occur during meiosis in yeast and apoptosis-induced chromatin condensation (Ahn et al. 2005; Rossetto et al. 2012). Phosphorylation of serine 139 of H2AX variant histone in mammalian cells and serine 129 of H2A in yeast create a specific signaling platform for recruitment and retention of DNA damage repair and signaling factors hence play key role in DNA damage response (Rossetto et al. 2012). It is not clear how H1 phosphorylation affects chromatin condensation during interphase and
mitosis. A number of studies indicate that interphase phosphorylation is involved in chromatin relaxation (Lever et al. 2000; Roth et al. 2001; Contreras et al. 2003); however, in metaphase chromosomes, H1 is hyperphosphorylated, and it has been shown that H1 hyperphosphorylation is required to maintain metaphase chromosomes in their condensed state (Th'ng et al. 1994).

**Histone ubiquitination**

Histone H2A K119 and H2B at K123 (in yeast) or K120 (in vertebrates) are ubiquitylated by histone ubiquitin ligases (Cao and Yan 2012). Monoubiquitylation of H2A and H2B have been involved in gene silencing and transcription activation, respectively. These ubiquitylated histones can interfere with chromatin compaction and therefore facilitate assembly of the DNA repair machinery on the DNA damage sites (Moyal et al. 2011). H2A monoubiquitylation is involved in silencing of X chromosome in female mammals (De Napolès et al. 2004; Fang et al. 2004). H2B ubiquitylation is necessary for reassembly of nucleosomes and restoration of the chromatin structure during transcription elongation (Xiao et al. 2005) and is also required for chromatin boundary integrity and spreading of other histone modifications. In general, histone ubiquitylation is implicated in transcriptional regulation by acting as a platform for other histone modifications; for example, the ubiquitylation of yeast H2B Lys 123 is required before the methylation of H3 that together contribute in telomeric gene silencing (Zhang et al. 2003). Like other histone modifications, monoubiquitylation of histones H2A and H2B is reversible. The ubiquitin modification can be removed by ubiquitin specific peptidases known as deubiquitylating enzymes (DUBs).

**Other histone modifications**

The small ubiquitin-related modifier (SUMO) functions in a manner similar to ubiquitin in that it is bound to target proteins as part of a post-translational modification and involved in variety
of cellular processes for instance, sumoylation of H4 has been found to have role in transcriptional repression (SHIIO and EISENMAN 2003).

Citrullination is a post-translational modification of histones where arginine is converted into citrulline, by removal of an imine. This process is catalyzed by deiminases, known as peptidyl arginine deiminases (PADs). These enzymes are largely restricted to bacteria, fungi and vertebrates. In human, there are five varieties of PADs and their activity is specific to the substrate. PAD4 can citrullinate nuclear substrates, including the histones H2A, H3, and H4, as well as the histone acetyltransferase p300 (WANG et al. 2004). In response to DNA damage, arginine 3 residue of histone H4 undergoes citrullination by the PAD4 mediated pathway (TANIKAWA et al. 2012). PAD4 also converts histone arginine (Arg) and mono-methyl arginine residues to citrulline. Recent studies have found that histone citrullination can be considered as an offset of histone arginine methylation and found to be associated with gene repression (WANG et al. 2004).

Another posttranslational modification, ADP-ribosylation, transfers ADP-ribose moiety from nicotinamide adenine dinucleotide (NAD⁺) onto acceptors such as amino acids. In vertebrates, arginine residues on histones are found to be ribosylated by the activity of arginine ADP-ribosyltransferases (OKAZAKI and MOSS 1999) and this action can be reversed by ADP-ribosylarginine hydrolases (TAKADA et al. 1994).

The combination of these histone modifications constitute a code (the ‘histone code’) that alters DNA-protein interactions that are important in diverse biological processes including gene regulation, DNA repair and chromosome condensation (JENUWEIN and ALLIS 2001). This code is deciphered by the readers, proteins that contain binding motifs specific for each modification: for example, chromodomains specifically recognize methylated residues, while bromodomains bind acetylated residues. These reader protein-containing complexes known as effectors recognize
posttranslational histone modifications and interpret the signal, eventually leading to chromatin remodeling (ROSSETTO et al. 2012).

**Chromatin modulators**

Chromatin remodeling factors are ATP-dependent chromatin binding protein complexes having the ability to either move or remove nucleosomes along a particular DNA sequence or to create a state of altered histone-DNA interaction (BECKER and HORZ 2002; SIF 2004). Principally, all chromatin remodeling complexes use the energy of ATP hydrolysis to loosen the contact between DNA and histones. Depending on their activities, these complexes are categorized in different classes and each of them has a subunit containing a conserved ATPase domain. ATPases are DNA translocases that are capable of the directional movement of DNA and therefore permitting the exposure of the DNA to regulatory factors (AALFS and KINGSTON 2000).

The first identified chromatin remodeler was the SWI/SNF (Switch/Sucrose non fermentable) complex of *S. cerevisiae* (HIRSCHHORN et al. 1992), which is found in eukaryotes and its homologue (RapA) in bacteria. Nucleosome displacement by Swi2/Snf2 occurs by sliding or tracking nucleosomes along the DNA allowing binding of transcription factors, thereby affecting transcription (DECristofaro et al. 2001). Similarly, RSC (remodels the structure of chromatin), a member of the SWI/SNF family is associated with regulatory regions of genes and binding of RSC reduces nucleosome occupancy within these regions (JANSEN and VERSTREPEN 2011). ATPase activity of ISWI (imitation SWI) complex is similar to SWI/SNF but additionally contains SANT domain that allows this remodeling complex to interact with histones (BOYER et al. 2004). The Swr1-Complex (SWR-C) inserts Htz1 (*S. cerevisiae* homologue of H2A.Z) that destabilizes nucleosome at transcriptional start site and promotes gene activation (KROGAN et al. 2003).
Common families of chromatin remodeling complexes and their domains are schematically represented in Figure 1.2.

![Chromatin remodeling protein families](image)

Figure 1.2. Chromatin remodeling protein families conserved from yeast to human. Adapted from (Xu et al. 2013).
The ATPase domain includes conserved N-terminal DExx and a C-terminal HELICc subdomains. In addition to this, the SWI/SNF family contains an HSA domain for actin binding, and via its bromodomain acetylated histone tails are recognized. Similarly, SANT and SLIDE domains of ISWI family are responsible for histone binding.

**Eukaryotic transcription**

The complexity of eukaryotes utilizes a variety of mechanisms to ensure precise regulation of gene expression. Transcription is the first step in gene expression and is the major target of regulation. Eukaryotic transcription is an elaborate nuclear process that uses genetic information stored in compact DNA regulatory elements to replicate genomic sequences into RNA. The compartmentalization of eukaryotic cell separates transcription from cytoplasmic processes. Transcription can be divided into three sequential phases: (i) initiation, (ii) elongation, and (iii) termination. The transcriptional machinery that catalyzes these intricate phases is comprised of one of the three multi-subunit RNA polymerases- RNA polymerase (Pol) I, II and III and their associated transcription factors.

Initiation of transcription requires binding of transcription machinery including general transcription factors and RNA polymerase enzyme complexes at gene promoter sequences. The regions at 5’ and 3’ ends of most actively transcribed genes are nucleosome depleted (referred as nucleosome free regions-NFR), making the promoter sequences accessible to binding of regulatory
proteins (JANSEN and VERSTREPN 2011). Transcription activation occurs when positive
regulatory proteins called activators bind enhancer sequences (or in case of yeast, upstream
activating sequence-UAS) that are located distal to the target gene. Conversely, when negative
regulatory proteins called repressors bind gene promoter sequences (for example, operators), they
interfere with RNA polymerase binding and effectively prevent transcription.

Among three classes of RNA polymerases, RNA polymerase I (Pol I) is devoted to the
transcription of the large, tandemly repeated, ribosomal RNA genes encoding the 18S, 5.8S, and
25S (in yeast) or 28S (in human) rRNAs that form the catalytic core of ribosomes (WHITE 2005).
The Pol I promoter consists of upstream promoter element (UPE) and core promoter region (Figure
1.3 part-A). The UPE-binding factor (UBF-in human) or upstream activation factor (UAF-in yeast)
bind at UPE whereas SL1 complex comprising TBP recognizes core promoter region. Binding of
SL1 complex and the Pol I-specific TBP-associated factors (TAFs) recruit Pol I enzyme complex
directly to the promoter and remain bound to the DNA to support multiple rounds of transcription
(PAULE and WHITE 2000).

RNA polymerase II is responsible for transcription of protein coding messenger RNA
(mRNA) genes and non-coding RNA genes like small nuclear RNA (snRNA), small nucleolar
RNA (snoRNAs) and microRNA (miRNAs). Transcription by Pol II is one of the most extensively
studied mechanisms in eukaryotes. The initiation of Pol II transcription requires the pre-initiation
complex (PIC) that contains the Pol II enzyme complex and its specific general transcription
factors which assemble at a core promoter region of a gene. The core promoters consist of
combinations of sequence elements that may or may not include a TATA-box, initiator (Inr) region
around the +1 or CAP site, downstream promoter elements (DPE) and upstream element (Figure
1.3 part-B). Transcription factor TFIID includes TATA-binding protein (TBP) and TBP-associated
factors (TAFs) that are involved in promoter recognition and transcription activation. During the assembly of a pre-initiation complex, TFIID specifically recognizes the TATA-box located upstream of the transcription start site which then recruits TFIIB to join TFIID. Before entering a PIC, a Pol II enzyme complex and TFIIF are bound together, which are recruited by TFIIB at the transcription start site. Finally, to complete PIC assembly, Pol II recruits TFIIE which has ATPase and kinase activities and then TFIIH is recruited for its helicase activity (LEE and YOUNG 2000; WOYCHIK and HAMPSEY 2002). After initiation, the C-terminal domain of Pol II complex (CTD) is subjected to phosphorylation for further RNA elongation. The CTD is an important site for binding of chromatin remodeling complexes, and contributes to the recruitment of splicing, termination, capping and polyadenylation complexes onto a nascent RNA transcript.

Besides transcription of protein-coding genes, Pol II generates large numbers of non-coding transcripts with mostly unknown functions. Recent tiling array hybridization studies (DAVID et al. 2006; XU et al. 2009) revealed more widespread expression of genetic information than anticipated based on serial analysis of gene expression (SAGE) (VELCULESCU et al. 1995). These genomic expression analyses also revealed rapid degradation of many heavily transcribed intergenic transcripts by the nuclear exosome (WYERS et al. 2005; XU et al. 2009).

In yeast, strong Pol II promoters generate divergent non-coding short transcripts from the upstream nucleosome free region of strong promoters that contribute to pervasive and intergenic transcription. Two such intergenic non-coding transcript types have been found in S. cerevisiae - cryptic unstable transcripts (CUTs) and stable unannotated transcripts (SUTs) (XU et al. 2009). Recent studies have shown that even though both CUTs and SUTs are arising from the same types
of bidirectional promoters, their degradation pathways differ. *CUTs* are found to be accumulated during vegetative growth of a cell in absence of the nuclear-specific catalytic subunit of the exosome complex, Rrp6. Normally, *CUTs* are terminated by a Nrd1-Nab3-Sen1 mediated mechanism which also recruits the nuclear exosome for rapid and efficient degradation (Vasileva et al. 2008). However, some *CUTs* are reported to be degraded by cytoplasmic RNA decay pathways (Thompson and Parker 2007).

Unlike *CUTs*, *SUTs* are only partially susceptible to Rrp6-dependent degradation and therefore detected even in the presence of functional nuclear exosome (Xu et al. 2009; Xu et al. 2011). However, the abundance of *SUTs* fluctuates in wild type yeast depending on the environmental conditions. Primarily, these relatively stable *SUTs* are affected by cytoplasmic RNA decay pathways that include NMD (nonsense mediated decay) and Xrn1-dependent 5′ to 3′ exonucleolytic degradation (Marquardt et al. 2011). Xrn1 is the general cellular exonuclease and another class of Xrn1-sensitive transcripts have recently been found by strand-specific RNA sequencing. These intergenic transcripts are referred as Xrn1-sensitive unstable transcripts (*XUTs*) and are found to be Pol II dependent, polyadenylated and the majority of them are antisense to open reading frames. The meiotic non-coding transcripts were observed by Lardenois et al., 2011 in non-fermenting and non-respiring cells and were named meiotic unannotated transcripts (*MUTs*). Some *MUTs* were found to be antisense to coding regions and, similar to *CUTs*, the levels of *MUTs* were affected by Rrp6 in mitotic cells (Lardenois et al. 2011).

The third class of RNA polymerases, RNA polymerase III (Pol III) transcribes structural and catalytic RNAs, including transfer RNA (tRNA), 5S ribosomal RNA (5S rRNA), U6 small nuclear spliceosomal RNA (SNR6), snR52 small nucleolar RNA, cytoplasmic 7SL RNA and the

Pol III transcripts are untranslated and generally short, rarely exceeding 200 base pairs in length. Pol III gene promoters have peculiar features in that they mostly reside within the transcribed region of genes, and are referred to as intragenic control regions (ICR). ICRs are composed of essential blocks of sequences separated by less conserved regions. Transcription factors specific to Pol III recognize the essential blocks and bind at ICRs to recruit Pol III enzyme complexes for the transcription initiation. Assembly of transcription machinery on Pol III genes is mainly influenced by the structure of promoters and this structure varies with different types of promoter (Figure 1.3 part-C).

The type 1 promoter is present only in 5S ribosomal RNA genes and contains A-box and C-box sequences which are bound by TFIIIC transcription factor and TFIIIA binds intermediate element (IE). Type 2 promoter contains intragenic A-box and B-box sequences, and includes all tRNA genes (further discussed in detail). Type 3 promoter in the S. cerevisiae (Sc) SNR6 gene is a hybrid and consists of intragenic A-box, upstream TATA-box sequences and B-box is situated downstream of the transcription termination site (consecutive thymine residues). In human (Hs) U6 gene, A-box and B-box elements are absent in type 3 promoters, but like Pol II genes, these promoters contain the upstream TATA-box sequence Distal Sequence Element (DSE) and Proximal Sequence Element (PSE). In yeast, 5S RNA genes are arranged in tandem with the 35S rRNA genes where there are arrays of 100-200 copies of these genes. In contrast, the 276 tRNA genes are generally distributed throughout the genome (Percudani et al. 1997; Hani and Feldmann 1998).
Figure 1.3. Schematic of typical promoters for RNA polymerase I, II and III. Adapted from (TEICHMANN et al. 2010).

A) Pol I promoter consists of UPE or upstream promoter element (deep blue) and the core promoter region (orange) comprising transcriptional start site indicated by +1. B) Pol II promoter is composed of various elements which act as binding sites for transcription machinery. The upstream element (pink) binds an activating transcription factor. BRE (green) is TFIIB recognition element. The TATA-box (red) is AT rich sequence and specifically recognized by TBP or TATA-binding protein. The initiator-Inr (purple) consists of the CAP or start site. The DPE or downstream promoter element (light blue) is located downstream from the CAP site. C) Pol III transcribes target genes with divergent promoter structures. Both type 1 and type 2 genes exhibit intragenic promoter elements. 5S rRNA genes contain type 1 promoter, in which, A-box and C-box are recognized by TFIIIC and TFIIIA binds intermediate element. Type 2 promoter is found in tRNA genes which is comprised of A-box and B-box which are specifically recognized by TFIIIC during transcription. In S. cerevisiae, type 3 genes contain a hybrid promoter which has the intragenic A-box and B-box situated downstream to the gene. Unlike other Pol III promoters, in S. cerevisiae (Sc) SNR6 gene type 3 promoter contains TATA-box upstream of the start site. Whereas Human U6 gene contains type 3 promoter without intragenic promoter elements it is similar to Pol II promoters that consist of upstream TATA-box sequence Distal Sequence Element (DSE) and Proximal Sequence Element (PSE). The termination signal for Pol III genes is a stretch of consecutive thymine residues, indicated by TTTT and the numbers indicating the length of the gene in base pairs are mentioned at the end of each gene.
This study focuses on transcription of tRNA genes and the associated transcription machinery (Figure 1.4). Transcription of Pol III genes also begins with the step-wise assembly of a pre-initiation complex (PIC). In *S. cerevisiae*, a typical tRNA gene transcription machinery consists of three multimeric complexes: TFIIIC (6 subunits), transcription factors TFIIIB (3 subunits) and the Pol III enzyme (17 subunits), required for transcription initiation and for promoter recognition, respectively (Geiduscheck and Kassavetis 2001). As described earlier, tRNA genes consist of a type 3 promoter which contains *A*-box and *B*-box intragenic elements, each around 10-12 base pairs long. The *A*-box corresponds to the D loop of the tRNA molecule, and lies ~20 bp downstream of the transcription start site. The *B*-box corresponds to the T loop of the tRNA molecule and is located 30–90 bp downstream of the *A*-box. The *B*-box consensus sequence, GGTTCGANYCY contains a highly conserved cytosine residue (underlined), and serves as a high affinity binding site for Pol III specific transcription factor TFIIIC. The *A*-box is essential for proper positioning of TFIIIC for directing recruitment of another Pol III transcription factor, TFIIIB, upstream of the transcription start site (Chaussivert et al. 1995; Geiduscheck and Kassavetis 2001). The distance between *A*-box and *B*-box varies among tRNA genes, and the distance can alter efficiency of TFIIIC binding. The optimal distance between the *A*-box and *B*-box for TFIIIC binding in vitro transcription is 30–60 bp (Cannon et al. 1986; Baker et al. 1987).

In *S. cerevisiae*, TFIIIC consists of six subunits which are organized into two globular domains, τA (Tfc1, Tfc4 and Tfc7) and τB (Tfc3, Tfc6 and Tfc8). During transcription of tRNA genes, the primary step is the recognition and specific binding of TFIIIC at *A*-box and *B*-box via τA and τB domains, respectively (Geiduscheck and Kassavetis 2001). Binding of TFIIIC facilitates recruitment of TFIIIB complex which is composed of three subunits, TBP, Brf1 and Bdp1. TBP is the only subunit of the basal factors which is not dedicated solely to Pol III
Figure 1.4. Transcription of a tRNA gene. Adapted from (DONZE 2012)

Transcription of tRNA gene begins with recognizing and binding of RNA polymerase III specific transcription factor TFIIIC at B-box embedded in the gene. This facilitates binding of another transcription factor TFIIIB at the upstream of transcriptional start site. Finally, RNA polymerase III enzyme complex is recruited at transcriptional start site that can initiate tRNA gene transcription as it is used by all three RNA polymerases (Hu et al. 2002). The N-terminal part of Brf1 is paralogous to Pol II transcription factor TFIIIB and the Brf1 C-terminus is associated with TBP and Bdp1 binding, which is necessary for transcriptional activity of TFIIIB. Overall, in absence of DNA, these TFIIIB subunits are loosely associated and are sequentially assembled by TFIIIC. Brf1 and Bdp1 contact Tfc4 and Tfc8 respectively and the latter interaction connects TBP. The subsequent binding of subunits leads to bending of DNA for extremely stable TFIIIB-DNA interaction that recruits a multimeric (17 subunits) Pol III enzyme complex at the transcriptional start site (CHAUSSEVERT et al. 1995; LIAO et al. 2006; CIESLA and BOGUTA 2008). In S. cerevisiae, TFIIIC–TFIIIB–DNA initiation complex is recognized by six Pol III-specific subunits of the enzyme complex- C82, C53, C37, C34, C31 and C17. Whereas, the two large subunits C160 and C128 along with C25 and C11 are homologous to either Pol I, Pol II or both. AC19 and AC40 Pol III subunits are common to only Pol I whereas, the remaining five subunits -ABC27, ABC23,
ABC14.5, ABC10α and ABC10β are shared between all three classes of polymerases in yeast (HUANG and MARAIA 2001).

The Pol III enzyme complex accurately and efficiently recognizes a simple run of T (thymine) residues as a termination signal; even in absence of other transcriptional factors (SCHRAMM and HERNANDEZ 2002). After the initial round of transcription, a stable Pol III machinery on the tRNA genes more rapidly engages in subsequent transcription cycles without being released, in a process called facilitated re-initiation or recycling (DIECI et al. 2013). The tRNA gene is constantly transcribed by Pol III machinery to generate tRNA adapter molecules for protein synthesis. Therefore, it appears that Pol III transcription machinery is always bound at these genes during interphase (Figure 1.5 part-A).

Figure 1.5. Assembly of RNA polymerase III transcription complexes. Adapted from (TEICHMANN et al. 2010)
A) At tRNA gene TFIIIC (6 subunits), TFIIIB (3 subunits) and RNA polymerase III enzyme complex (17 subunits) are always assembled. B) A typical ETC site containing extended B-box sequence is a binding site for TFIIIC complex but it is not occupied by RNA polymerase III enzyme complex.

In S. cerevisiae, data from the genome-wide occupancy of Pol III components have revealed several new Pol III associated loci, beside known Pol III-transcribed genes (MOQTADERI and STRUHLL 2004). A locus, called ZOD1 (for zone of disparity), is considered as a functional Pol III promoter that consists of conserved A-box and B-box sequences which are bound by complete Pol III transcription machinery. Yet, discrete and functional RNA species from ZOD1 are not
detected hence it is thought to be transcriptionally inactive (ROBERTS et al. 2003; MOQTADERI and STRUHL 2004). Moreover, the other loci called ETC (for extra TFIIC) in Saccharomyces species consist of extended conserved B-box but not A-box sequences which is still sufficient for TFIIC binding (Figure 1.5 part-B). Moreover, these regions are non-transcribed since none of the ETC sites have measurable Pol III occupancy. Recent studies have revealed only TFIIC occupancy at ETC1-ETC8 and ETC10 sites (MOQTADERI and STRUHL 2004; NAGARAJAVEL et al. 2013). However, the iYGR033C region (designated as ETC9 in this study) has been found to be occupied by TFIIC and TFIIB, but not by the Pol III enzyme complex (GUFFANTI et al. 2006). Presence of conserved sequences among different Saccharomyces species suggested biologically meaningful function of these ETC sites (Moqtaderi & Struhl, 2004). Later, structural importance of these sites has been demonstrated in many recent studies (further explained in detail). In Schizosaccharomyces pombe, sequences similar to ETC sites, referred to as chromosome organizing clamp (COC) sites have been detected near the nuclear periphery, are bound by TFIIC and thought to be involved in chromosome organization (NOMA et al. 2006).

**Extra transcriptional functions of RNA polymerase III complex bound regions**

While Pol III and most of its transcription factors are thought to be dedicated to transcription of Pol III genes, emerging studies have shown that both partial and complete chromosomally bound Pol III transcription complexes can have effects on chromatin state and even on genome organization. For instance, Pol III transcribed tRNA genes not only function as transcription units for generating tRNA adapter molecules but also serve as the potential sites for other extra-transcriptional or product-independent roles (CLELLAND and SCHULTZ 2010; DONZE 2012). Similarly, other B-box containing sites (for example, ETC in S. cerevisiae) where partial or
complete Pol III complex assembly is found, also contribute to the extra-transcriptional functions. The overview of extra-transcriptional functions is represented in Figure 1.6.

**Direction of Ty element integration**

Sandmeyer and colleagues first identified the extra-transcriptional functions of Pol III-transcribed genes in yeast. The upstream region of Pol III genes was found to be a target for insertion of Ty3 retrotransposons in budding yeast (Chalker & Sandmeyer, 1990). Later, *in vitro* integration assays revealed the requirement of intact promoter elements and the Pol III transcription factors TFIIIB and TFIIIC for the insertion of Ty3 near transcription start site of transcriptionally competent *SUP2* tRNA gene. Binding of Pol III at *SUP2*, however was found to be inhibitory for the integration (Connolly & Sandmeyer, 1997; Kirchner, Connolly, & Sandmeyer, 1995). Further studies showed that TFIIIB is indispensable for Ty3 integration at the U6 gene (*SNR6*) and TFIIIC is not essential, but the presence of TFIIIC directs position-specific integration to the *SNR6* proximal initiation site (Yieh, Hatzis, Kassavetis, & Sandmeyer, 2002).

Another retrotransposon Ty1 also targets Pol III–transcribed genes, but the pattern of insertions is very different. In *S. cerevisiae*, the integration of Ty1 element occurs in a range of about 75–700 bp upstream of the transcription start site of diverse Pol III genes target as opposed to Ty3 integration just a few base pairs upstream of Pol III genes (Devine & Boeke, 1996). Pol III gene targeting by Ty1 requires an intact Pol III promoter at the target site (Devine & Boeke, 1996) and ATP-dependent chromatin remodeling factor ISW1 which is targeted to tRNA genes. Further, the N-terminal domain of TFIIIB subunit Bdp1 is essential for integration site selection (Bachman, Gelbart, Tsukiyama, & Boeke, 2005).
Figure 1.6. Extra-transcriptional functions of Pol III transcription complexes bound sites in *S. cerevisiae*. Modified from (DONZE 2012).

tRNA genes (orange box) are persistently occupied by Pol III transcription machinery (pink, yellow, and light green ovals). As a consequence of this, they are involved in transcription-independent functions such as directing Ty element insertion (A), overriding nucleosome positioning signal (B), nucleosome phasing (C) chromosomal organization sites (D), pausing replication fork propagation (E), mediating silencing of neighboring Pol II-transcribed gene (F) and functioning as chromatin boundaries (G).

Overriding nucleosome positioning

Intrinsic properties of DNA, including its sequence and the assembly of nearby chromatin bound proteins affect nucleosome positioning (JANSEN and VERSTREPN 2011). Morse *et al.*, (1992) showed that a mutated and transcriptionally incompetent tRNA gene lost the ability to
affect nucleosome positioning when a nucleosome positioning signal was incorporated near the gene. Whereas in the same condition, transcriptionally active tRNA gene bound with Pol III transcription machinery was able to supersede the nucleosome positioning signal (MORSE et al. 1992). Furthermore, nucleosome positioning data from DNA sequencing showed that binding Pol III complexes at nucleosome free regions of tRNA genes and ETC sites were responsible for displacing the nucleosome formation (MAVRICH et al. 2008; XU et al. 2009).

**Nucleosome phasing**

Recent nucleosome occupancy data have shown that tRNA genes act as nucleosome phasing signals, perhaps due to the stability, high occupancy and relatively fixed position of TFIIIC-TFIIIB complex. The ‘bootprints’ of TFIIIC-TFIIIB complexes also revealed that nucleosome positioning is disrupted up to ~1kb from tRNA genes in both directions. (NAGARAJAHEL et al. 2013).

**Chromosomal organization sites**

Structural maintenance of chromosome proteins (SMCs) are components of a variety of complexes that are central to the organization and segregation of chromosomes. The Smc1 and Smc3 dimer forms the core of the cohesin complex, which mediates sister-chromatid cohesion. In addition to the Smc1–Smc3 dimer, the cohesin complex contains Scc1 and Scc3 subunits. The Smc2 and Smc4 dimer forms the core of condensin, a protein complex that facilitates chromosome condensation in preparation for mitotic segregation (ARAGON et al. 2013).

D’Ambrosio et al. (2008) discovered that the condensin binding sites are closely associated with most of the genes that are transcribed by RNA polymerase III or the regions where Pol III transcription complexes are bound (for example ETC sites in S. cerevisiae) (D’AMBROSIO et al. 2008). Functional condensin proteins are required for the clustering of tRNA genes as well as
silencing of an adjacent Pol II gene. These findings suggested that at the loci where the Pol III transcription complexes are assembled, condensin interacts with these complexes (such as TFIIIC) and together act as a structural element in three-dimensional folding of chromosomes (HAEUSLER et al. 2008). Recent genome-wide footprint analysis of TFIIIC-TFIIIB in S. cerevisiae have indicated a close association of TFIIIC and centromeric nucleosomes perhaps via condensin that is thought to be important for 3-dimensional organization of nucleus (NAGARAJAVEL et al. 2013).

**Replication fork pausing sites**

Progression of replication forks appear to stall at tRNA genes with assembled transcription machinery, only when transcription and replication are orientated in opposite directions through these genes (DESHPANDE and NEWLON 1996). They showed that a temperature sensitive mutation in the large subunit of RNA the Pol III enzyme complex that affected transcription initiation, but not the assembly of TFIIIC and TFIIIB, also abolished replication pausing activity. That suggested the requirement of direct interaction between Pol III transcription machinery and advancing replication fork for arresting replication at the tRNA genes. Further they also proposed that at least to some extent, the accumulation of supercoiling between the progressing replication fork and transcription is responsible in stalling replication at the tRNA gene (DESHPANDE and NEWLON 1996). In S. pombe, tRNA genes (particularly tRNAGLU and sup3-e ) have been found to be capable of arresting replication forks moving in both orientations with respect to the tRNA gene transcription (MCFARLANE and WHITEHALL 2009).

**Pol II transcription inhibition or tRNA gene position effect**

Pol III transcribed tRNA genes can have negative effects on neighboring Pol II gene transcription, and are referred to as tRNA gene position effects (BOLTON and BOEKE 2003) or tRNA mediated gene silencing (KENDALL et al. 2000). Position effect was first observed by Kinsey
et al, 1991 when the SUP2 tRNA gene transcription complex was found to be affecting the transcription of upstream Pol II sigma transcript. Mutation of the B-box of SUP2 causing inactivation of SUP2 transcription increased transcription from the nearby sigma promoter 9-fold (Kinsey and Sandmeyer 1991). Similarly transcription of a Ty1 retrotransposon was affected when it was inserted upstream of Pol III transcribed SUP2 and SNR6 (U6 gene in S. cerevisiae) (Bolton and Boeke 2003). Previous studies in our lab also demonstrated increased CBT1 transcription upon TRT2 tRNA gene deletion, indicating tRNA position effect on CBT1 at its natural locus (Simms et al. 2004). Transfer RNA gene mediated (tgm) silencing (the suppression of neighboring Pol II gene transcription), is fundamentally different from other forms of transcriptional silencing in yeast. The silencing is independent of the tRNA gene orientation and does not involve simple blockage of RNA Pol II upstream activator sites. Instead, it is dependent on transcription of the tRNA gene, since mutations in the Pol III promoters and conditional mutations in Pol III enzyme complex) alleviate tgm silencing (Kendall et al. 2000). The mechanism of silencing near tRNA genes is found to be associated with the clustering of the tRNA genes near the nucleolar periphery (Wang et al. 2005).

**Chromatin boundaries**

Structurally and functionally discrete domains of chromatin- euchromatin and heterochromatin are separated by chromatin boundaries. There are two types of boundary elements. The first type, in which boundaries act as ‘barriers’ against self-propagating heterochromatin, prevent encroachment of silenced chromatin into transcriptionally active euchromatin regions (Figure 1.7 part-B). In S. cerevisiae, the telomeres and the cryptic mating-type loci (HML and HMR) represent silenced chromatin domains. At the HMR locus propagation of silenced chromatin is restricted by a naturally occurring tRNA gene that acts as a chromatin
barrier. The tRNA genes within the pericentric repeat elements of *S. pombe* act similarly, serving as barriers to pericentric heterochromatin (Noma *et al.* 2006). The second type of chromatin boundary has the ability to block the action of distal enhancer (or upstream activation sequence – UAS in yeast) on a promoter and hence acts as an insulator to prevent inappropriate activation of a gene (Figure 1.7 part-A). This enhancer-blocking action occurs only when the insulator is located between the enhancer/UAS and a promoter of a target gene. Our previous lab results showed insulator activity at *TRT2* tRNA, and *ETC4* sequences prevented the UAS from activating GAL promoters when inserted between UAS and *GAL10* or *GAL1* genes (Simms *et al.* 2008). *ETC6* may also act as an insulator which auto-regulates the activation of neighboring *TFC6* gene from its own UAS (Kleinschmidt *et al.* 2011).

In this study, *S. cerevisiae* is used as a model system to analyze extra-transcriptional functions of Pol III complexes at naturally bound loci. For the past several decades *S. cerevisiae* has been a key model system for eukaryotic molecular genetic research because its basic cellular mechanics of replication, recombination, cell division and metabolism are generally conserved between yeast and larger eukaryotes, including mammals.

Apart from the well-defined genetic composition, there are other unique characteristics ascribed to *S. cerevisiae* such as ability to transform synthetic DNA directly into yeast allowing easy production of altered forms of proteins, ability to recombine an exogenous DNA (with partial homologous segments) directly to specific genomic locations, and viability of both haploid and diploid forms of yeast. The ease of genetic manipulation of yeast allows its use for analyzing the mechanisms involved in chromatin remodeling and gene regulation.
Figure 1.7. Schematic for chromatin boundary activity. Panel A depicts insulator activity where it prevents inappropriate activation of genes from enhancers. Panel B shows blocking of heterochromatin propagation by chromatin barrier, thereby keeping the following gene transcriptionally active.

In the overview of this dissertation, the second chapter demonstrates the unique extra-transcriptional function of RNA polymerase III complexes (predominantly the TFIIIB complex) assembled at the tRNA gene upstream of ATG31, demonstrating the ability to block cryptic intergenic transcriptional interference by RNA polymerase II. When TFIIIB binding was weakened, readthrough of intergenic transcription was observed that inhibited the translation of ATG31 and affected the function of Atg31p in autophagy. Therefore intergenic transcription blocking was found to be an important regulatory mechanism to protect the neighboring gene from transcriptional interference.

In the third chapter, another locus RAD2-TNA1 was analyzed for examining a hypothesized repression blocking-insulator type function of ETC4 site that is located between RAD2 and TNA1 genes. The potential insulator type ability of ETC4 at its natural location was predicted to be necessary for preventing inappropriate repression of RAD2 gene from the upstream Sum1-Hst1 repressor.
The fourth chapter describes initial steps involved in RNA-seq analysis for determining the genome-wide effects of extra-transcriptional functions of assembled Pol III transcription complexes on neighboring genomic loci. The preliminary computational data comparison between wild type strains and TFC6 down-regulated mutants (exhibiting reduced TFIIIC binding) showed 3’ or 5’ extension of some of the Pol II transcripts and/or differential expression of Pol II transcribed genes which are in the vicinity of either tRNA genes or ETC sites.

Finally, chapter five contains thorough discussion of the results of all the above studies and the consequences of those results. New strategies for studying extra-transcriptional effects of Pol III transcription complex bound sites are also discussed in this chapter.
Introduction

In eukaryotes, the process of transcription is divided among three RNA polymerases- RNA polymerase I, II, and III. In the yeast *Saccharomyces cerevisiae*, RNA polymerase III (Pol III) transcribes a variety of small RNAs, including transfer RNA (tRNA), 5S ribosomal RNA (5S rRNA), U6 spliceosomal RNA, snR52 small nucleolar RNA, 7SL RNA, and the RNA component of RNase P. Assembly of the transcription machinery on Pol III genes is mainly determined by the structure of the promoter. A unique feature of most Pol III promoters is the presence of internal control regions (ICRs) that are composed of conserved sequences separated by more variable regions. The most common promoter arrangement used by Pol III is the class II promoter, found mainly in tRNA genes (tDNAs). Class II promoters consist of the conserved intragenic *A-box* and *B-box* sequences that are bound by the transcription factor complex TFIIIC (PASCALI and TEICHMANN 2012; ACKER et al. 2013).

In yeast, the entire Pol III transcription machinery bound to tDNAs consists of three multimeric protein complexes: the transcription factors TFIIIC (6 subunits) and TFIIIB (3 subunits), which are required for promoter recognition and preinitiation complex formation, and the 17 subunit Pol III enzyme (GEIDUSCHEK and KASSAVETIS 2001; HUANG and MARAIA 2001; ACKER et al. 2013). The initial step in the transcription of tDNAs in yeast is the binding of the TFIIIC complex to the *A-* and the *B-boxes.*

*Previously Published in *Genetics Society of America*
Asawari Korde, Jessica M. Rosselot and David Donze (2013). “Intergenic Transcriptional Interference Is Blocked by RNA polymerase III Transcription Factor TFIIIB in *Saccharomyces cerevisiae*.” *Genetics Society of America.*
The 6 subunits of TFIIIC are organized into two globular domains, \( \tau A \) (Tfc1p, Tfc4p, and Tfc7p) and \( \tau B \) (Tfc3p, Tfc6p, and Tfc8p). \( \tau B \) specifically binds to the \( B \)-box with high affinity and favors \( A \)-box binding by \( \tau A \) (Geiduschek and Kassavetis 2001). The most currently refined \( B \)-box consensus sequence, GWT\( \underline{C} \)RANNC (Marck et al. 2006; Orioli et al. 2012) contains a highly conserved cytosine residue (underlined), and mutation of this cytosine compromises TFIIIC binding (Donze 2012). TFIIIC binding is required to recruit TFIIIB at most Pol III promoters. TFIIIB is composed of three proteins, TATA-binding protein (TBP), TFIIB-related factor (Brf1p) and \( B \)" (B-double prime, Bdp1p). Binding of TFIIIB forms an exceptionally kinetically stable TFIIIB–DNA complex (Cloutier et al. 2001), which then recruits the Pol III enzymatic complex and helps maintain it for multiple transcription cycles in a process called facilitated recycling (Dieci and Sentenac 1996; Ferrari et al. 2004).

While Pol III and its transcription factors are generally thought to be dedicated to transcription of Pol III target genes, emerging studies have shown that either partial or complete DNA-bound Pol III transcription complexes can have effects on transcription, chromatin state, and genome organization of neighboring Pol II genes. These so-called “extra-transcriptional” (Donze 2012) or “product independent” (Clelland and Schultz 2010) effects of Pol III complexes, mostly demonstrated in \( S. \) cerevisiae, include the following activities: targeting integration of Ty retroelements (Chalker and Sandmeyer 1990; Ji et al. 1993; Devine and Boeke 1996), displacement of nucleosomes (Morse et al. 1992), phasing of adjacent nucleosomes (Nagarajavel et al. 2013), position effect repression of adjacent Pol II promoters (Hull et al. 1994), chromatin boundary/insulator functions (Donze 2012), and pausing of replication forks (Deshpande and Newlon 1996; Sekedat et al. 2010). In some instances, the TFIIIC complex alone can mediate extra-transcriptional functions, as Extra-TFIIIC (ETC) sites (Moqtaderi and
STRUHL 2004), chromosomal loci that bind only TFIIIC without recruiting TFIIIB or Pol III, can act as insulators (SIMMS et al. 2008), can directly regulate Pol II promoters (KLEINSCHMIDT et al. 2011), and can tether chromosomal regions to the nuclear periphery (HIRAGA et al. 2012).

Using the S. cerevisiae model system, we have previously described multiple types of extra-transcriptional functions of the TRT2 tDNA at the STE6-CBT1 locus. In MATα cells, TRT2 serves as a barrier to prevent repression of the neighboring Pol II-transcribed CBT1 gene, whereas in MATa cells TRT2 exerts an apparent tRNA position effect, as deletion of TRT2 results in an increase in CBT1 gene transcription (SIMMS et al. 2004). This modest position effect (approximately threefold increase in CBT1 mRNA levels) was shown to be due in part to the tDNA acting as an insulator, as it prevents inappropriate activation of the CBT1 promoter by the Mcm1p transcription factor that binds to the nearby STE6 upstream activation sequence (UAS) (SIMMS et al. 2008).

Manual inspection of the S. cerevisiae genome reveals that about one-quarter of all tDNAs lie between divergently transcribed genes in the yeast genome and could potentially show a similar insulator effect. Given the modest insulator effect observed at the CBT1 locus, we investigated the ATG31-tV(UAC)D-SES1 locus anticipating a more robust effect, as genome-wide expression data indicate that SES1 is transcribed at considerably higher levels (~70-fold) than is ATG31 in rich media (HOLSTEGE et al. 1998; XU et al. 2009). Our reasoning was that transcription factors responsible for the high level activation of SES1 would more strongly activate ATG31 upon deletion of the tDNA. Surprisingly, when we performed Northern blot analysis on RNA from wild-type and tV(UAC)D deleted (referred to hereafter as tdnaΔ) strains, we found that ATG31 mRNA levels were not only increased, but that a longer transcript with an extended 5’-UTR (5’-untranslated region) replaced the normal transcript. We show here that this longer transcript is due
to readthrough of the noncoding stable unannotated transcript *SUT467* (Xu *et al.* 2009) and mutations that inhibit TFIIB complex assembly or stability allow readthrough. Progression of transcription from the upstream *SUT467* start site prevents normal *ATG31* transcriptional initiation, and the extended 5’-UTR inhibits translation of the *ATG31* coding sequence. Since *ATG31p* is required for autophagy, reduced translation results in compromised autophagy and fitness under nitrogen starvation conditions in strains exclusively expressing the extended transcript. This work identifies another novel extra-transcriptional function of tDNAs, the ability to block progression of cryptic intergenic transcription, preventing subsequent deleterious transcriptional interference of an adjacent promoter.

**Materials and methods**

Yeast cultures were grown in nutrient-rich YPD media (1% yeast extract, 2% peptone, and 2% dextrose) at 30°C on a rotary shaker unless otherwise noted. For induction of autophagy, cells were grown to mid-log phase (*A*600 = 0.7) in YPD, collected by centrifugation (3000 rpm 3-5 min), washed with water, then resuspended in nitrogen starvation media (1.7 g/liter yeast nitrogen base without amino acids and without ammonium sulfate, plus 2% dextrose). For Northern blot analysis of temperature-sensitive mutants, cultures were grown at 30°C to an OD$_{600}$ of 0.7 and then incubated at 37°C for 1 hour before RNA extraction.

Plasmid pDD1232 was created by cloning a 1.35-kb XhoI-SpeI cut *ATG31-SES1* intergenic fragment (PCR amplified with oligos DDO1281/-1282, which added an artificial XhoI site) into Bluescript SK+ (all oligos used are listed in Supporting Information, Table 2.1). Two-step PCR mutagenesis was performed using pDD1232 as template and T7 and T3 primers with mutagenic primers (DDO184/-1284; DDO183/-1285, respectively) to amplify the fragment containing the tdna deletion; this fragment was then cloned into Bluescript SK+ as above to create pDD1233.
Table 2.1 Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>tV(UAC)D</td>
<td></td>
<td>T3 primer</td>
</tr>
<tr>
<td>DDO-183</td>
<td>AATTAACCTCTACTAAAGGG</td>
<td>T7 primer</td>
</tr>
<tr>
<td>DDO-184</td>
<td>GTAATACGACTCCTAGTTAGGG</td>
<td>ATG3 intergenic SpeI</td>
</tr>
<tr>
<td>DDO-1281</td>
<td>GACCTTATCTATCCCTGCTGTC</td>
<td>SES1 intergenic XhoI</td>
</tr>
<tr>
<td>DDO-1282</td>
<td>CTTTTTTTCTTCAAGTCCTCCTCGAGCCTGTGGCTC</td>
<td>SES1 CDS KO PCR check</td>
</tr>
<tr>
<td>DDO-1283</td>
<td>CAGATGGGATAGACGATTTCTTC</td>
<td>tDNA SDM delete top</td>
</tr>
<tr>
<td>DDO-1284</td>
<td>GATATGTACTTTCAATTTATTATTTTACAGA</td>
<td>tDNA SDM delete bottom</td>
</tr>
<tr>
<td>DDO-1285</td>
<td>GGGATTAAGTCTTTTCAATGGGTTTCGTAAA</td>
<td>ATG3 CDS internal</td>
</tr>
<tr>
<td>DDO-1325</td>
<td>GAGATCCCAGTCCATCCAATTG</td>
<td>ATG31 CDS internal</td>
</tr>
<tr>
<td>DDO-1328</td>
<td>TATTTTGGATATTTGAGCCACCTTATTTATTAGG</td>
<td>ATG31-SES1 intergenic KO top</td>
</tr>
<tr>
<td>DDO-1329</td>
<td>GACAGAGACACCACCTCCTGTGGATATCTTGGAGACCTTCTTTACATTATAATAGCGACTCC</td>
<td>ATG31-SES1 intergenic KO bottom</td>
</tr>
<tr>
<td>DDO-1330</td>
<td>CAACACATATGTCCATCCTTTACAG</td>
<td>intergenic KO check</td>
</tr>
<tr>
<td>DDO-1391</td>
<td>CTTTACACGCCGAGATCCCCGTCGTTGGATC</td>
<td>tv(uac)d B-box delete top</td>
</tr>
<tr>
<td>DDO-1392</td>
<td>GATAAAAAGATTCTGATCCCAACCGACGGGA</td>
<td>tv(uac)d B-box delete bottom</td>
</tr>
<tr>
<td>DDO-1393</td>
<td>CTTTACACGCCGGAAGATCCCCGAGTGGCAACCTCGTGAATCTTTATTTAC</td>
<td>tv(uac)d B-box mut. Top</td>
</tr>
<tr>
<td>DDO-1394</td>
<td>GATAAAAAGATTCTGATCCCAACCGAGGTTTGCAACCTCGGCTGTAAG</td>
<td>tv(uac)d B-box mut. Bottom</td>
</tr>
<tr>
<td>DDO-1468</td>
<td>CAAATTTATTCGAGTCCGATACAAAAAGATTC</td>
<td>tv(UAC)D flip BamHI end</td>
</tr>
<tr>
<td>DDO-1469</td>
<td>GTGGATTATCTCTGAATTTCTTTTACATTATAATAGCG</td>
<td>tv(UAC)D flip EcoRI end</td>
</tr>
<tr>
<td>DDO-1474</td>
<td>CGCCGTGTAACGGCCGAGTCTTTGAATCTAGATTGGACCGGGAAATTGAC</td>
<td>tv(uac)d A-box mut. Top</td>
</tr>
<tr>
<td>DDO-1475</td>
<td>GTCAATTTCAATGGGTTTCGGTGCCATTAGCAGCTTATTTTAATAGCGACCGGGGGTGA</td>
<td>tv(uac)d A-box mut. Bottom</td>
</tr>
<tr>
<td>DDO-1489</td>
<td>AATTCCACTCTCCTTTTTAATGGCTCCACTCAGGA</td>
<td>ETC4 EcoRI end</td>
</tr>
<tr>
<td>DDO-1490</td>
<td>GATCCTCTTTTCTGTATATAACCATTCAGACTCCCGTGGAGGCAATACAAAAAGGGGAAGTGG</td>
<td>ETC4 BamHI end</td>
</tr>
</tbody>
</table>
(Table 2.1 continued)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDO-1534</td>
<td>AATATTTATAGCTCAGGATCCAAAAGACATGACGATATAATGTAC</td>
<td>ETC9 PCR BamHI end</td>
</tr>
<tr>
<td>DDO-1535</td>
<td>AATATTTATAGCTCAGAATTCAAGAAAACACGGATAATTCACCAGGTC</td>
<td>ETC9 PCR EcoRI end</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Northern Analysis probes</strong></td>
<td></td>
</tr>
<tr>
<td>DDO-1369</td>
<td>TAATACGACTCATAATAGTACGGGAATGGGAGAGCATTG</td>
<td>ATG31 CDS T7</td>
</tr>
<tr>
<td>DDO-1370</td>
<td>GAATGTACAGTTACTGTTTATG</td>
<td>ATG31 CDS</td>
</tr>
<tr>
<td>DDO-1404</td>
<td>TAATACGACTCATAATAGCACAAGAGTGGTGCTTCTG</td>
<td>SUT467 T7</td>
</tr>
<tr>
<td>DDO-1366</td>
<td>CATGCTTGACGCTTATATACCC</td>
<td>SUT467</td>
</tr>
<tr>
<td>DDO-1371</td>
<td>TAATACGACTCATAATAGTACATGGACTGACTGGAGCTT</td>
<td>SES1 CDS T7</td>
</tr>
<tr>
<td>DDO-1372</td>
<td>GTGGGACATCAACACATTTATAC</td>
<td>SES1 CDS</td>
</tr>
<tr>
<td>ACT1 5'</td>
<td>ATGGGATCTCTGAGTGTGTTCTACGCG</td>
<td>ACT1 CDS</td>
</tr>
<tr>
<td>ACT1 3'</td>
<td>TAATACGACTCATAATAGGCAACTCTCAATTCTGAGAGG</td>
<td>ACT1 CDS T7</td>
</tr>
<tr>
<td>5'-RACE, ChIP and qRT-PCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DDO-59</td>
<td>CATACTCGAAGGGTGTGGG</td>
<td>tN(GUU)C upstream, Chr. III</td>
</tr>
<tr>
<td>DDO-60</td>
<td>GATTTTTCCATTCGCCATGC</td>
<td>tN(GUU)C downstream, Chr. III</td>
</tr>
<tr>
<td>DDO-402</td>
<td>ATGGGATTCTGAGTGTGCTGC</td>
<td>ACT1 ATG/intron qPCR</td>
</tr>
<tr>
<td>DDO-403</td>
<td>CAAAACGCTTGGATGGGAAAC</td>
<td>ACT1 internal qPCR</td>
</tr>
<tr>
<td>DDO-1284</td>
<td>GATATGTACTCATTTATATTTATCTTTTACGAAACCATTGGAAATGTCAATTTATCC</td>
<td>Extended transcript specific primer for qRT-PCR (same as tDNA∆)</td>
</tr>
<tr>
<td>DDO-1364</td>
<td>GATATGTACTCATTTATATTTATTTATATTTTAC</td>
<td>Long transcript RACE inner</td>
</tr>
<tr>
<td>DDO-1527</td>
<td>GACGTTTATCTTGGATATGAC</td>
<td>Long transcript RACE outer</td>
</tr>
<tr>
<td>DDO-1541</td>
<td>CATAAACAGTAACTGTAAACATTTC</td>
<td>ATG31 RACE inner</td>
</tr>
<tr>
<td>DDO-1542</td>
<td>TCTCTTCTATCGACTCTGAGGC</td>
<td>ATG31 RACE outer</td>
</tr>
<tr>
<td>DDO-1555</td>
<td>GTTCCTTGGATTGCAAGGCATAACC</td>
<td>ATG31 intermediate</td>
</tr>
<tr>
<td>DDO-1576</td>
<td>CAATGTATTTACGGCTAATAAG</td>
<td>ATG31 intergenic ChIP</td>
</tr>
<tr>
<td>DDO-1577</td>
<td>TTTCGACATAATAATAGGCAGAC</td>
<td>ATG31 intergenic ChIP</td>
</tr>
<tr>
<td>DDO-1606</td>
<td>CTTCATAATTATTATCTTTTACGAAAC</td>
<td>qPCR bridge primer</td>
</tr>
</tbody>
</table>

34
Site-directed mutagenesis was performed on pDD1232 to create pDD1248 (tdna B-boxΔ), pDD1249 (tdna B-box mutant), pDD1261 (tdna A-box mutant), and pDD1260 (tdnaΔ::EcoRI–BamHI linker), using DDO1391/-1392, DDO1393/-1394, DDO1474/-1475, and DDO1466/-1467 primer sets, respectively. The B-box mutant had the invariant cytosine and following guanine bases changed to GC, and the A-box mutant scrambled the entire consensus. Plasmids pDD1262 (flipped orientation of the tDNA) and pDD1272 (tdnaΔ::ETC9) were created by cloning EcoRI-BamHI-digested PCR-amplified fragments using DDO1468/-1469 (flip), or DDO1534/-1535 (ETC9),
respectively, into EcoRI-BamHI-digested pDD1260. Yeast genomic DNA was used as PCR template. Plasmid pDD1263 (tdnaΔ::ETC4) was constructed by directly ligating complementary oligonucleotides (DDO1489/-1490) containing EcoRI-BamHI overhangs into pDD1260. All plasmids were confirmed by Sanger sequencing and are listed in Table 2.2.

Yeast strains were generated from wild-type S. cerevisiae W303-1a; genotypes of all strains used and generated in this study are given in Table 2.3. Parent tdnaΔ::URA3 (DDY4605–4607) strains were created by amplifying URA3 with primers DDO1279/-1280 containing homology to the flanking region of tv(UAC)D, and then this DNA was transformed into wild-type DDY3 followed by selection of Ura+ colonies and PCR identification of homologous recombinants. Linearized tdna mutant plasmids were digested with XhoI and SpeI, individually transformed into a tdnaΔ::URA3 strain, and 5-fluoroorotic acid (5-FOA) resistant colonies were isolated. Recombinants were identified by PCR and verified by DNA sequencing of the product. Yeast strains for chromatin immunoprecipitation (ChIP) were created by crossing existing BRF1-3X-FLAG (DDY1495) and TFCl-3X-FLAG (DDY3860) strains to DDY4607, and then FLAG-tagged Ura+ progeny were backcrossed to each tdna mutant.

Table 2.2 Plasmids used in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pYM6</td>
<td>9XMyc-klTRP1</td>
</tr>
<tr>
<td>pDD 1232</td>
<td>ATG31-SES1 intergenic</td>
</tr>
<tr>
<td>pDD 1233</td>
<td>ATG31-SES1 intergenic tv(uac)dΔ</td>
</tr>
<tr>
<td>pDD 1248</td>
<td>ATG31-SES1 intergenic tv(uac)d b-boxΔ</td>
</tr>
<tr>
<td>pDD 1249</td>
<td>ATG31-SES1 intergenic tv(uac)d b-box point mutant</td>
</tr>
<tr>
<td>pDD 1260</td>
<td>ATG31-SES1 intergenic tv(uac)dΔ::EcoRI-BamHI linker</td>
</tr>
<tr>
<td>pDD 1261</td>
<td>ATG31-SES1 intergenic tv(uac)d a-box mutant</td>
</tr>
<tr>
<td>pDD 1262</td>
<td>ATG31-SES1 intergenic tv(UAC)D flip</td>
</tr>
<tr>
<td>pDD 1263</td>
<td>ATG31-SES1 intergenic tv(uac)dΔ::ETC4</td>
</tr>
<tr>
<td>pDD 1262</td>
<td>ATG31-SES1 intergenic tv(uac)dΔ::ETC9</td>
</tr>
</tbody>
</table>
### Table 2.3 S. cerevisiae strains generated in this study

<table>
<thead>
<tr>
<th>Name</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDY3</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3Δ-1</td>
</tr>
<tr>
<td>DDY232</td>
<td>MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 rpc3Δ-236 hmrΔ</td>
</tr>
<tr>
<td>DDY246</td>
<td>MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 rpc160Δ::HIS3 p-rpc160-112</td>
</tr>
<tr>
<td>DDY261</td>
<td>MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 tfc3-G349E hmrΔ</td>
</tr>
<tr>
<td>DDY416</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 hmrΔ brf1Δ::HIS3 p-brf1 II.6</td>
</tr>
<tr>
<td>DDY420</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 hmrΔ brf1Δ::HIS3 p-brf1 II.6</td>
</tr>
<tr>
<td>DDY947</td>
<td>MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 sas2Δ::TRP1</td>
</tr>
<tr>
<td>DDY1376</td>
<td>MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 nhp6a:URA3 nhp6b:HIS3</td>
</tr>
<tr>
<td>DDY1495</td>
<td>MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 BRF1:3XFLAG:KanMX nhp6b::HIS3</td>
</tr>
<tr>
<td>DDY1631</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 rsc2Δ::TRP1</td>
</tr>
<tr>
<td>DDY1676</td>
<td>MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 rpd3::LEU2</td>
</tr>
<tr>
<td>DDY2236</td>
<td>MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 HMR-ADE2 hzt1Δ::KanMX</td>
</tr>
<tr>
<td>DDY2509</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 bdf1Δ::HIS3</td>
</tr>
<tr>
<td>DDY3860</td>
<td>MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 TFC1:3XFLAG:KanMX</td>
</tr>
<tr>
<td>DDY4300</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tcf6 promoter mutant 3</td>
</tr>
<tr>
<td>DDY4607</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)dΔ::URA3</td>
</tr>
<tr>
<td>DDY4624</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)dΔ</td>
</tr>
<tr>
<td>DDY4625</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1</td>
</tr>
<tr>
<td>DDY4652</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)dΔ</td>
</tr>
<tr>
<td>DDY4653</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)dΔ</td>
</tr>
<tr>
<td>DDY4764</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 atg31Δ::TRP1</td>
</tr>
<tr>
<td>DDY4769</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)d B-boxΔ</td>
</tr>
<tr>
<td>DDY4816</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(UAC)D flip</td>
</tr>
<tr>
<td>DDY4817</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)d A-box mutant</td>
</tr>
<tr>
<td>DDY4819</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)dΔ::ETC4</td>
</tr>
<tr>
<td>DDY4901</td>
<td>MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 tv(uac)dΔ::URA3 BRF1:3XFLAG:KanMX</td>
</tr>
<tr>
<td>DDY4904</td>
<td>MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 tv(uac)dΔ::URA3 TFC1:3XFLAG:KanMX</td>
</tr>
<tr>
<td>DDY4970</td>
<td>MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)dΔ::ETC9</td>
</tr>
<tr>
<td>DDY5003</td>
<td>MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 tv(uac)dΔ::ETC9 BRF1:3XFLAG:KanMX</td>
</tr>
<tr>
<td>DDY5006</td>
<td>MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 tv(uac)dΔ::ETC9 TFC1:3XFLAG:KanMX</td>
</tr>
<tr>
<td>DDY5010</td>
<td>MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 ydl156Δ::KanMX</td>
</tr>
<tr>
<td>Name</td>
<td>Genotype</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DDY5012</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 ATG31-9X-myc::TRP1}$</td>
</tr>
<tr>
<td>DDY5014</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)dΔ ATG31-9X-}$</td>
</tr>
<tr>
<td></td>
<td>$\text{myc::TRP1}$</td>
</tr>
<tr>
<td>DDY5018</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 tv(uac)d B-boxΔ ATG31-9X}$</td>
</tr>
<tr>
<td></td>
<td>$\text{myc::TRP1}$</td>
</tr>
<tr>
<td>DDY5020</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 tv(uac)d B-box pt. mutant}$</td>
</tr>
<tr>
<td>DDY5044</td>
<td>$\text{MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 pho13::URA3}$</td>
</tr>
<tr>
<td></td>
<td>$\text{pho8Δ60::HIS3 atg8Δ::TRP}$</td>
</tr>
<tr>
<td>DDY5046</td>
<td>$\text{MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 pho13::URA3}$</td>
</tr>
<tr>
<td></td>
<td>$\text{pho8Δ60::HIS3}$</td>
</tr>
<tr>
<td>DDY5051</td>
<td>$\text{MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 pho13::URA3}$</td>
</tr>
<tr>
<td></td>
<td>$\text{pho8Δ60::HIS3}$</td>
</tr>
<tr>
<td>DDY5072</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 pho13::URA3}$</td>
</tr>
<tr>
<td></td>
<td>$\text{pho8Δ60::HIS3 tv(uac)dΔ}$</td>
</tr>
<tr>
<td>DDY5074</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 pho13::URA3}$</td>
</tr>
<tr>
<td></td>
<td>$\text{pho8Δ60::HIS3}$</td>
</tr>
<tr>
<td>DDY5078</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 pho13::URA3}$</td>
</tr>
<tr>
<td></td>
<td>$\text{pho8Δ60::HIS3 tv(uac)d B-boxΔ}$</td>
</tr>
<tr>
<td>DDY5081</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 pho13::URA3}$</td>
</tr>
<tr>
<td></td>
<td>$\text{pho8Δ60::HIS3 tv(uac)d B-box pt. mut.}$</td>
</tr>
<tr>
<td>SG154.2</td>
<td>$\text{MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 bar1Δ SCC2::scc2-}$</td>
</tr>
<tr>
<td></td>
<td>$\text{D730V::HYG}$</td>
</tr>
<tr>
<td>ROY1032</td>
<td>$\text{MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 HMRΔI smc1-2::LEU2}$</td>
</tr>
<tr>
<td>ROY1060</td>
<td>$\text{MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 HMRΔI smc3-1::LEU2}$</td>
</tr>
<tr>
<td>ROY1063</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 HMRΔI scc1-73::TRP1}$</td>
</tr>
<tr>
<td>ZFY155</td>
<td>$\text{MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 pho13::URA3}$</td>
</tr>
<tr>
<td></td>
<td>$\text{pho8Δ60::HIS3}$</td>
</tr>
<tr>
<td>DDY4908</td>
<td>$\text{MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 TFC1:3XFLAG:KanMX}$</td>
</tr>
<tr>
<td>DDY4917</td>
<td>$\text{MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 tv(uac)dΔ::ETC4}$</td>
</tr>
<tr>
<td></td>
<td>$\text{TFC1:3XFLAG:KanMX}$</td>
</tr>
<tr>
<td>DDY4920</td>
<td>$\text{MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 tv(uac)d B-box point mutant}$</td>
</tr>
<tr>
<td></td>
<td>$\text{TFC1:3XFLAG:KanMX}$</td>
</tr>
<tr>
<td>DDY4925</td>
<td>$\text{MATa his3-11,15 leu2-3,112 lys2Δ trp1-1 ura3-1 tv(uac)d B-box point mut}$</td>
</tr>
<tr>
<td>DDY4935</td>
<td>$\text{MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 tv(uac)d flip}$</td>
</tr>
<tr>
<td></td>
<td>$\text{BRF1:3XFLAG:KanMX}$</td>
</tr>
<tr>
<td>DDY4938</td>
<td>$\text{MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 BRF1:3XFLAG:KanMX}$</td>
</tr>
<tr>
<td>DDY4943</td>
<td>$\text{MATa his3-11,15 leu2-3,112 trp1-1 ura3-1 tv(uac)d A-box mutant}$</td>
</tr>
<tr>
<td></td>
<td>$\text{BRF1:3XFLAG:KanMX}$</td>
</tr>
</tbody>
</table>
The 9X-Myc epitope tag was amplified from a TRP1 marked cassette (Knop et al. 1999) using DDO1462/-1463 and transformed into wild-type and tdna mutants. ATG31-9X-myc Trp⁺ homologous recombinants were identified by PCR (DDO1419/-1464) and confirmed by Western blotting. ATG31 and ATG8 knockout strains were constructed by standard yeast homologous recombination. For autophagy induction alkaline phosphatase assays, pho13Δ pho8Δ60 strains were created by crossing a pho13Δ::URA3 pho8Δ60::HIS3 strain (kindly provided by Daniel Klionsky) to wild-type and tdna mutants.

RNA isolation and Northern blotting were performed as described (Simms et al. 2004). Most Northern results were verified with three (but at least two) independently isolated mutant strains; Table 2.3 lists only the specific strains shown in the figures. Primers used to amplify Northern probes are listed in Table 2.1. 5’-RACE analysis was performed on RNA isolated from the DDY420 brf1 II.6 mutant using the First Choice RLM-RACE kit (Ambion/Life Technologies, AM1700). Individual clones were sequenced by standard Sanger sequencing and mapped to the S. cerevisiae genome on the Saccharomyces Genome Database at http://www.yeastgenome.org (Cherry et al. 2012).

Quantitative RT-PCR was performed as follows: First-strand cDNA was synthesized from 0.5 mg of total RNA after DNase treatment (RQ1 DNase, Promega M6101). Synthesis was extended from long transcript specific primer DDO1284 using ProtoScript M-MuLV first-strand cDNA synthesis kit (NEB E6300S). Quantitative reverse transcription PCR (qRT-PCR) was
performed on 1:4 diluted cDNA using primers DDO1606/-1555 and Sybr Green super mix (Bio-Rad 170-8882) with 60⁰C annealing temperature. Results were normalized to amplicons from ACT1 control primers (DDO402+403). Reactions were run and analyzed using a Bio-Rad MyiQ as described (Kim et al. 2011) and examined by agarose gel electrophoresis to verify that only the predicted PCR products were amplified. The primers were designed to specifically amplify readthrough transcripts; the strategy is described and illustrated in Figure 2.1.

**Figure 2.1.** Strategy for RT-PCR quantitation of *SUT467-ATG31* extended transcripts. In order to quantify extended transcripts while minimizing amplification of trace amounts of chromosomal DNA remaining after DNase treatment, a bridge primer method was designed. First strand cDNA was synthesized with readthrough specific primer DDO-1284, which has homology to normally untranscribed regions on either side of the tDNA (green and orange boxes), but omitting the transcribed tDNA sequence. This oligo can only base pair with RNA that is transcribed from upstream through the tDNA, presumably looping out the tRNA sequence within the extended transcript. After first strand synthesis, qPCR was performed with the bridge primer DDO-1606, which is homologous to 21 base pairs of the green sequence, and only 5 base pairs within the orange sequence. By having only 5 contiguous base pairs of homology to the orange region, amplification of residual genomic DNA was minimized.
Chromatin immunoprecipitation was performed as described previously (Rusche et al. 2002). Anti-FLAG epitope antibody was purchased from Sigma (F1804). Primers DDO1527/-1555 or DDO1576/-1577 were used to amplify desired regions surrounding $tV(UAC)D$. For Western analysis, yeast minilytsates were prepared by glass bead lysis of log-phase cultures directly in lysis buffer (50 mM Tris pH 7.5, 1% SDS, 5 mM EDTA, 14.3 mM $\beta$-mercaptoethanol, 1 mM PMSF, 2 mg/ml leupeptin and pepstatin). Pellets from 5 ml YPD culture at A600 1.0 were resuspended in 200 $\mu$l lysis buffer and then vortexed with glass beads at 4 for 10 min. Lysis buffer, 100 $\mu$l, was added, and the mixture was boiled for 3 min, cooled on ice, and centrifuged at 4°C to remove cell debris. To 100 $\mu$l of clarified minilytsate, an equal amount of 2X SDS-PAGE loading buffer was added, and after boiling, 15 ml was loaded on 12% acrylamide protein gel. Proteins were transferred to Millipore Immobilon membrane by semidry transfer and incubated in blotto (10% TBS/10% SDS/5% dry milk) for 1 hour. Primary Myc antibody (c-Myc 9E10, Santa Cruz Biotechnology), anti-mouse Ig-horseradish peroxidase secondary antibody (GE healthcare) were used for Western analysis. Immuno-star Western chemiluminescent kit (Bio-Rad) was used to detect the secondary antibody.

The alkaline phosphatase pho8Δ60 assay was performed as described (Klionsky 2007). The cell survival assay was adapted from (Kabeya et al. 2007). Yeast strains were grown in YPD rich media to $A_{600} = 1.0$ (10$^7$ cells/ml). Cells were harvested by centrifugation, washed once with distilled water, and resuspended at $10^7$ cells/ml in media lacking nitrogen. Cells were incubated for 6 days at 30°C on rotatory shaker, and every other day 200 $\mu$l of culture dilutions was plated on YPD plates in triplicate. Plates were incubated at 30°C for 48 hour and survival rate was obtained by counting resulting colonies.
Results

Mutation of the tDNA upstream of ATG31 results in readthrough of SUT467

Figure 2.2 part-A depicts the ATG31–SES1 locus, showing the location of the tV(UAC)D tDNA, and the extent of transcripts normally produced from the region. To test the initial hypothesis that tV(UAC)D might act as an insulator by preventing promiscuous activation of ATG31 by regulatory elements associated with the strong promoter of the neighboring divergent SES1 gene, we created tV(UAC)D deleted (tdnaΔ) mutant strains. Northern blot analysis using a

Figure 2.2. Mutation of the tV(UAC)D tRNA gene upstream of ATG31 results in readthrough of the intergenic SUT467 transcript.
(A) Schematic of the ATG31–SES1 locus on S. cerevisiae chromosome IV. Colored arrows indicate known annotated transcripts: ATG31, black; SUT467, blue. The overlapping red arrow represents the extended readthrough transcript. (B) Northern blot analysis of ATG31 expression in wild-type and tdnaΔ strains reveals the extended transcript. The ATG31 coding sequence probe hybridized to RNA of 800 bp in wild-type strains (black arrow) and to RNA of 1200 bp after tDNA deletion (red arrow). The SUT467 probe hybridized to the predicted 300-bp transcript in wild-type cells (blue arrow) and to the same 1200-bp extended transcript in tdnaΔ strains. The normal ATG31 transcript was absent in tdnaΔ strains. Each pair of lanes contained total RNA from independent wild-type and mutant strains. (C) B-box deletion (B-boxΔ) or mutation of the invariant cytosine in the B-box (B-box mut) also resulted in extended readthrough transcription. Strains used were: (B) DDY4625 and DDY3 (wild-type); DDY4653 and 4624 (tdnaΔ); (C) DDY3 (wt); DDY4652 (tdnaΔ); DDY4769 (B-boxΔ); and DDY4925 (B-box mut).
probe homologous to the $ATG31$ coding sequence (Figure 2.2 part-B, left) showed not only an apparent slight increase in the level of $ATG31$ mRNA in the tdna∆ strains (compared to $ACT1$ controls), but also an increase in the length of the transcript by 400 nucleotides (shifting from 800 bases in wild-type strains to 1200 bases), with apparent absence of the normal length mRNA. Recent tiling array and RNA-seq studies have identified widespread pervasive and intergenic transcription in eukaryotic cells, and in yeast this often appears to occur as bidirectional transcription from strong promoters (Neil et al. 2009; Xu et al. 2009). Inspection of data from these studies indicated that the $SUT467$ intergenic transcript initiates upstream of the $SES1$ promoter and terminates near the tDNA (Figure 2.2 part-A); therefore, we hypothesized that the extended $ATG31$ transcript in the tdna∆ strain was a readthrough $SUT467$ transcript that interferes with the production of the normal $ATG31$ transcript.

To confirm that the extended transcript in tdna∆ strains was due to readthrough of $SUT467$ and not a 3’ extension, we repeated the Northern analysis using a probe specific for the transcribed $SUT467$ RNA sequence (Figure 2.2 part-B, right). The results showed that this probe hybridized to RNA of 300 bases in wild-type strains, consistent with previous annotations of $SUT467$. The tdna∆ strains showed a longer 1200-base transcript, the same length as when using the $ATG31$ coding-sequence probe. We concluded that in the absence of the tDNA at this region, $SUT467$ readthrough occurs and interferes with normal $ATG31$ transcription initiation, producing only the observed extended RNA. Since our gross deletion of the tDNA sequence removed 90 bp of chromosome IV, we confirmed this readthrough effect by creating strains that either had only the B-box sequence of the tDNA deleted or contained a mutation in the invariant cytosine residue in the B-box. Both of these mutations of the tDNA were expected to result in loss of TFIIIC binding and inhibition of Pol III complex assembly. Northern blot analysis with either probe shown in
Figure 2.2 part-C confirmed that in each of these mutant backgrounds, complete readthrough of SUT467 occurred as in the tdnaΔ strains.

The slightly shorter transcript observed in the tdnaΔ strain compared to the B-boxΔ and point mutant strains is also consistent with the long transcript being a readthrough from upstream of the tDNA, as this reflects the 90-bp deletion. Also observed were shorter RNAs hybridizing to only the SUT467 probe, which appear to terminate between the tDNA and ATG31.

**TFIIIB binding is correlated with blocking of SUT467 readthrough**

Our previous studies of heterochromatin barrier and insulator function of tDNAs have shown that assembled TFIIIC alone can block the spread of silencing from the HMR locus and can insulate a UAS from a promoter (Simms *et al.* 2008). To determine which components of the Pol III complex are required to block SUT467 progression, we analyzed ATG31 transcripts in various temperature-sensitive strains compromised for Pol III complex function and formation when pulsed at the nonpermissive temperature before RNA extraction. Figure 2.3 part-A shows the results of Northern blot analysis of these strains. Temperature-sensitive mutations in RNA Polymerase III subunit genes RPC31 and RPC160 affect transcription initiation and elongation, respectively (Dieci *et al.* 1995; Thuillier *et al.* 1995). These mutants had relatively little effect on the ability of the tDNA to block progression of SUT467 (lanes 2 and 3 compared to wild type in lane 1) as evidenced by a minimal alteration of the ratio of normal to extended transcripts. In contrast, mutations in the TFIIIB subunit encoding BRF1 gene (brf1-II.6 and -II.9) that impair interactions of Brf1p with TBP (Andrau *et al.* 1999) showed a major shift to the longer extended transcript (lanes 4 and 5), with relatively little normal length RNA. Interestingly, these mutants also showed an intermediate length ATG31 transcript that initiates just upstream of the tDNA coding sequence (see 5’-RACE analysis below).
Mutations involving TFIIIC also resulted in readthrough transcription. RNA isolated from the temperature-sensitive, DNA-binding defective tfc3 G349E mutant strain (LEFEBvre et al. 1994) showed apparent equal amounts of both normal and long ATG31 transcripts, with a small relative amount of the intermediate transcript (Figure 2.3 part-A, lane 6). A strain harboring a mutation in the TFC6 promoter that results in reduced expression of Tfc6p and slow growth (Kleinschmidt et al. 2011) showed a similar pattern (lane 8), shifted a bit more to the long and intermediate transcripts. These results demonstrate that loss of TFIIIC function also results in readthrough SUT467 transcription, but this could be due to loss of TFIIIB assembly in the absence of full TFIIIC activity.

Figure 2.3. Pol III transcription factors are required to block SUT467 readthrough transcription. (A) Northern analysis of temperature-sensitive mutants of the Pol III complex was performed as in Figure 2.2, except that each culture was shifted from 30°C to 37°C for 1 hour prior to RNA extraction. Extended ATG31 transcripts are most prominent in TFIIIB and TFIIIC subunit mutants, which also express an intermediate length ATG31 transcript. Strains used in lanes 1-8 were DDY3 (wt); DDY232 (rpc31-236); DDY246 (rpc162–112); DDY416 (brf1 II.9); DDY420 (brf1 II.6); DDY261 (tfc3 G349E); DDY3 (wt); and DDY4300 (tfc6 promoter mutant). (B) 5’- RACE analysis of extended and intermediate ATG31 transcripts. 5’- RACE was performed to map transcriptional start sites (TSS) for the various transcripts observed in the Northern blot analysis of the brf1 II.6 mutant. Colored solid boxes represent the range of alternative TSS, which were observed in three distinct clusters.

To verify that the transcription start site of our readthrough transcript initiates in the region of the annotated SUT467 transcriptional start site (TSS) and to map the TSS of the observed intermediate transcript, we performed 5’-RACE analysis on RNA isolated from the brf1-III.6
mutant, because it contains all three transcripts as detected by Northern blotting. As shown schematically in Figure 2.3 part-B, 5'-RACE ends that correspond to the annotated ATG31 mRNA, and within a 94-nucleotide range that overlaps the annotated SUT467 TSS, were mapped. The intermediate transcript was found to begin very close to the beginning of the tRNA coding sequence. The exact Saccharomyces Genome Database chromosome IV coordinates corresponding to each individually mapped 5'-RACE end are listed in Figure 2.4.

To further assess the mechanistic requirements of each Pol III transcription factor in preventing Pol II readthrough transcription, we constructed yeast strains specifically modified at the ATG31 upstream tDNA locus and analyzed the long vs. short RNA phenotypes. Inverting the orientation of the tDNA had no effect, as no extended ATG31 mRNA was detected (Figure 2.5A, lane 1). Mutation of the A-box within the tDNA or replacement of the tDNA with the ETC4 site resulted in only the extended transcript being produced (Figure 2.5 part-A, lanes 2 and 3). Each of these replacements was expected to bind TFIIIC, but not be able to efficiently recruit TFIIIB or Pol III. Interestingly, replacing the tDNA with the tDNA remnant upstream of the TIM21 gene, recently referred to as ETC9 (NAGARAJAVEL et al. 2013), was sufficient to block readthrough transcription (Figure 2.5 part-A, lane 4). This tDNA remnant has previously been shown to bind both TFIIIC and TFIIIB, but not the Pol III enzymatic complex (GUFFANTI et al. 2006).

These results suggest that recruitment of TFIIIB is the critical step that prevents readthrough transcription of SUT467, as binding of TFIIIC alone at the ETC4 site is not sufficient to block Pol II progression. Mutation of the A-box has been demonstrated to impair TFIIIB assembly (HUIBREGTSE and ENGELKE 1989), and this mutation also allows readthrough. These interpretations assume that each of these sequences used to replace the tDNA have the same in vivo binding characteristics at the ATG31 locus as they do in their native chromosomal locations.
Figure 2.4. *S. cerevisiae* Chromosome IV annotation and locations of 5’ ends of normal, intermediate, and long *ATG31* transcripts. Annotations are as listed in the *Saccharomyces* Genome Database, and the predicted TFIIIB footprint was inferred from published data as described and referenced in the main text. Normal transcripts are shown as black arrows, with extended transcripts in green and red. Transcription start sites (TSS) were mapped by 5’-RACE as described in Materials and Methods. RNA was from the brf1 mutant strain DDY420, as all three transcripts are present in this mutant. First strand cDNA synthesis was random primed and amplified by nested PCR using adaptor inner and outer primers (as described in Ambion First-Choice RLM RACE kit) with nested DDO-1541/1542 for the normal *ATG31* mRNA, and DDO-1527/1364 for the intermediate and long transcripts. Since DDO-1364 lies in the normally untranscribed region between *ATG31* and the tDNA, only cDNA copied from extended transcripts are amplified. The coordinates are listed for each individual sequenced 5’-RACE clone, and their locations are marked on the map with arrowheads using the same color-coding scheme (coordinates as of October 11, 2013).

To confirm such assumptions regarding the presence or absence of each transcription factor complex at these sequences when moved to the *ATG31* locus, we crossed a 3X-FLAG-epitope-tagged *BRF1* allele into each of these mutants. ChIP results using anti-FLAG antibody shown in Figure 2.5 part-B demonstrate that the tDNA flip and *ETC9* alleles are strongly enriched for TFIIIB at levels comparable to wild-type tDNAs, while insertions unable to block readthrough transcription (*A-box* mut, *ETC4*, and *B-box* mut) had significantly reduced TFIIIB ChIP signals,
comparable to background signals observed in the no antibody control panels. Primers amplifying a separate control tDNA on chromosome III showed similar enrichment in each of the samples, indicating that equivalent amounts of ChIP DNA were added to each PCR reaction.

We also created strains containing a 3X-FLAG-epitope-tagged TFC1 allele to assess the binding of TFIIIC at modified ATG31 loci. The results in Figure 2.5 part-C show that TFIIIC but not TFIIIB is associated with the ETC4 insertion, and both TFIIIC and TFIIIB are bound at the ETC9 insertion and at the wild-type tDNA locus. The control tDNA again showed equivalent

![Figure 2.5](image-url)

Figure 2.5. Binding of the TFIIIB complex is associated with blocking of SUT467 readthrough transcription.

Strains were constructed to recruit the entire Pol III complex, TFIIIB and TFIIIC, or TFIIIC alone to the ATG31–SES1 intergenic region. Each construct was tested for the ability to block readthrough and for binding of Pol III transcription factor complexes to the ectopic locations. (A) Schematic of the modified ATG31 loci and Northern blot of each strain using the ATG31 probe. Lane 1, DDY4816 (tDNA flip); lane 2, DDY4817 (A-box mut); lane 3, DDY4819 (ETC4 replacement); lane 4, DDY4970 (ETC9 replacement); and lane 5, DDY4925 (B-box mut). Replacement of the tDNA by ETC4, or mutating the A-box or B-box, resulted in the presence of the extended transcript (red labels). However, inversion of the tDNA sequence or replacement with the ETC9 sequence still blocked readthrough (black labels). (B and C) Confirmation of expected Pol III transcription factor binding in the above mutants by chromatin immunoprecipitation. Each tDNA mutant strain was crossed to strains containing either BRF1-3X-FLAG or TFC1-3X-FLAG alleles, and then subjected to ChIP analysis using anti-FLAG antibody. (B) The absence of TFIIIB upstream of ATG31 in the A-box mutant, ETC4 replacement, and B-box mutant correlates with the presence of the extended transcript, suggesting that TFIIIB binding is required to block readthrough. Strains used (left to right) were DDY4935, -4943, -4949, -5003, and -4946. (C) ChIP analysis of BRF1-3XFLAG and TFC1-3XFLAG strains demonstrates that TFIIIC but not TFIIIB is bound in ETC4 replacement strains, indicating that TFIIIC binding alone cannot block readthrough transcription. Strains used (left to right) were DDY4938, -5003,-4949, -3860, -5006, and -4917.
levels of enrichment in the ChIP samples. These results demonstrate an association of TFIIIB binding with blocking of SUT467 transcription. Importantly, and contrary to results seen in our earlier tDNA heterochromatin blocking studies (Simms et al. 2008), TFIIIC binding alone to ETC4 is not sufficient to block cryptic transcript readthrough.

**Mutations in genes affecting tDNA heterochromatin barrier function have minimal impact on transcript blocking**

Previous studies on the heterochromatin barrier activity of tDNAs revealed the involvement of other chromatin-associated proteins in this extra-transcriptional function (Donze et al. 1999; Donze and Kamakaka 2001; Jambunathan et al. 2005; Braglia et al. 2007). To assess the potential role of these tDNA associated proteins in blocking readthrough transcription, we performed Northern blot analysis (using the ATG31 coding sequence probe) on RNA isolated from a number of these mutants. The results in Figure 2.6 part-A showed that each of these mutants contain mostly normal-length ATG31 transcripts; however, low levels of readthrough are apparent in some strains, most obvious in nhp6 (lane 3) and smc3 (lane 12) mutants in the particular blot shown. However, the intensity of these signals was relatively weak and was often difficult to consistently distinguish from background in different blots.

To confirm these apparent low levels of readthrough, we used readthrough transcript-specific primers to develop an RT-PCR assay to measure differences in the relative levels of the long transcript compared to a wild-type strain. The inset in Figure 2.6 part-B shows an inverted ethidium-stained gel image that verifies that the primers specifically amplified the readthrough cDNA, as the B-box mutant strain showed significantly higher levels of RT-PCR product than the wild-type strain. There also appears to be a low level of readthrough in the wild-type strain, which is consistent with a genome-wide transcriptome analysis that identified a single readthrough clone overlapping this locus (Miura et al. 2006). Quantitative RT-PCR was performed on the same
RNA samples shown in the Northern blot in Figure 2.6 part-A, and those that showed a significant increase in the long ATG31 transcript relative to the wild-type parent are shown in Figure 2.6 part-B. The B-box mutant strain measured ~80-fold more readthrough transcript than wild type in this assay, while other mutants were confirmed to have modest (ranging from 2- to 15-fold) yet detectable increased levels of the long transcript as suggested by the Northern analysis.

![Figure 2.6](image_url)

Figure 2.6. Genetic factors involved in tDNA chromatin boundary function have minimal effects on blocking of Pol II progression through tV(UAC)D.

(A) RNA from strains containing mutations that weaken tDNA boundary function were analyzed by Northern blotting using the ATG31 probe. Strains in lanes 1-9 were DDY3, -947, -1376, -2236, -2058, -2509, -1631, -1676, and -5010; lane 10, SG154.2; lanes 11-13, ROY1032, -1060, and -1063; lane 14, DDY4925. (B) RT-PCR analysis of readthrough transcription also shows only minimal effects.

**The extended ATG31 transcript is not efficiently translated**

Given the extended 5'-UTR present on the long ATG31 transcript, we next asked to what extent translation of the ATG31 protein was affected by readthrough SUT467 transcription. We created ATG31-9X-myc epitope-tagged strains in wild-type and tDNA mutant backgrounds and then analyzed ATG31 protein expression by Western blotting. In each strain producing the long transcript, we observed a drastic reduction in Atg31p levels, almost undetected under normal
growth conditions (Figure 2.7 part-B left panel, lanes 3-8). Atg31p is required for autophagy in yeast.

Autophagy is a conserved cellular response that recycles cellular components upon nutrient limitation and during normal regulated molecular turnover and involves the formation of autophagosome vesicles that capture and degrade macromolecules after fusion with other membrane bound vesicles (Reggiori and Klionsky 2013; Stanley and Adolphs 2013).

Nutrient starvation not only induces autophagy but also arrests cells at early G1 phase. Tor (Target of rapamycin), a phosphatidylinositol kinase-related kinase, is known to be involved in the signaling pathway from nutrient starvation to G1 arrest in yeast (Barbet et al. 1996). There are two types of Tor in yeast- Tor1p and Tor2p. The immunosuppressive drug rapamycin is the inhibitor of Tor2p and causes cycle arrest in yeast at early G1 phase (Heitman et al. 1991). Due to this cellular stress, autophagy is induced in rapamycin treated cells. Northern analysis on tdna mutants under rapamycin induced autophagy condition (Noda and Ohsumi 1998) showed appearance of normal ATG31 transcript (Figure 2.7 part-A, right panel, lanes 14-17). Moreover, Atg31p was also detected in tdna mutants upon autophagy induction but the protein levels were significantly reduced as compared to wild-type under the same growth condition (Figure 2.7 part-B Right panel, lanes 14-17).

Since Atg31p is required for the formation of autophagosomes upon nitrogen starvation (Kabeya et al. 2007), we tested the efficiency of this response using two well-characterized assays to measure autophagy induction. We first created a series of strains that produce the readthrough transcript and contain the pho8Δ60 and pho13Δ alleles. These mutations reduce background alkaline phosphatase levels, and the phosphatase activity of the precursor Pho8Δ60 protein can be activated only if it is proteolytically processed during autophagy (Klionsky 2007). Since ATG31
is required for these events, we reasoned that the reduction in Atg31p levels due to the extended 5’-UTR would result in reduced processing of Pho8Δ60p upon induction of autophagy and therefore reduced levels of alkaline phosphatase activity upon shifting cells to nitrogen starvation conditions.

Figure 2.7. ATG31 translation is inhibited from the extended transcript under normal growth conditions.
A) Northern analysis with Myc-tagged tdna mutants produced only extended ATG31 transcript under normal growth conditions (lanes 3-8). After induction of autophagy (rapamycin treatment), along with wild type, tdna mutants also showed appearance of normal length ATG31 transcript (lanes 14-17), except tdnaΔ (right panel lanes 12 and 13). B) Atg31p was detected in wild type-Myc tagged strains by using 9E10 Myc antibody. In tdnaΔ and other tdna mutants that produced only the extended transcript, ATG31 translation was significantly inhibited (lanes 3-8) as compared to wild-type (lanes 1 and 2). Under induced condition, increased translation of ATG31 was observed in tdna mutants except tdnaΔ (lanes 14-17). Lanes 9 and 18 shows undetected ATG31 transcript (in A) and Atg31p (in B) in atg31Δ strain which was used as a negative control.

Figure 2.8 part-A shows this to be the case, as strains producing the readthrough transcript showed significantly reduced induction of phosphatase activity compared to a pho13Δ pho8Δ60 strain producing only normal ATG31 mRNA (wt in Figure 2.8 part-A). Complete deletion of ATG31 or ATG8 in control strains severely reduced starvation-induced phosphatase activity as
expected. These results confirm that induction of autophagy is compromised in strains that predominately produce the long ATG3 transcript.

As a second assay for the efficiency of autophagy induction, we tested the viability of yeast cells producing the long transcript when placed under autophagy-inducing conditions. Previous studies have shown that complete deletion of ATG3 results in reduced survival of cells undergoing nitrogen starvation due to inhibition of autophagy (KABEYA et al. 2007). When we tested a strain containing the B-box mutation in the tDNA (producing the long transcript), we found an intermediate level of survival compared with wild-type and atg31Δ strains (Figure 2.8 part-B). The loss of survival of tdna mutant is not due to effects on the neighboring SES1 gene, as wild-type and tdna mutant strains show equivalent expression levels of SES1 when analyzed by Northern blotting (Figure 2.8 part-C).

Figure 2.8. Mutations in tV(UAC)D inhibit function of ATG31 protein.
A) Inhibition of autophagy induction in yeast expressing the extended transcript as measured by the Pho8Δ60 alkaline phosphatase assay. Strains used were: DDY5051 (wt control), DDY5072 (tdnaΔ), DDY5078 (B-boxΔ), DDY5081 (B-box mut), DDY5044 (atg8Δ), and DDY5046 (atg31Δ). B) Production of the extended transcript is associated with reduced survival under nitrogen starvation conditions. Strains used were: DDY5012 (wt), DDY5081 (B-box mut), and DDY4764 (atg31Δ). C) Northern blot analysis of SES1 in wild-type and tdnaΔ strains show no significant difference in SES1 mRNA levels. Mutation of tDNA does not affect SES1 expression. Treatment with rapamycin induces autophagy and results in decreased SES1 mRNA levels, and this downregulation is also unaffected when the tDNA is deleted.
This result suggests that readthrough of the cryptic transcript reduces Atg31p translation to a level that compromises fitness of the cells during nitrogen starvation.

**Discussion**

The results described in this study demonstrate that stable RNA polymerase III transcription factor complexes containing TFIIIB assembled at tDNAs have the capacity to block the progression of intergenic transcription by RNA polymerase II. High-throughput microarray and sequencing technologies have led to the identification of much more diversity in transcriptomes, from prokaryotes to humans, than was previously appreciated (Core *et al.* 2008; Dornenburg *et al.* 2010; Wei *et al.* 2011). In *S. cerevisiae*, such pervasive transcripts include the cryptic unstable transcripts (CUTs), stable unannotated transcripts (SUTs), Xrn1-sensitive unstable transcripts (XUTs), and meiotic unannotated transcripts (MUTs) (Wyers *et al.* 2005; Xu *et al.* 2009; Lardinois *et al.* 2011; Van Dijk *et al.* 2011). Additionally, alterations in the prevalence of intergenic transcripts and transcript start and end sites have been observed under different growth and stress conditions (Xu *et al.* 2009; Waern and Snyder 2013). Since the vast majority of these cryptic transcripts and transcript isoforms have unknown functions, it has been speculated that they may represent inherent sloppiness of the transcriptional process, referred to as “transcriptional noise” (Struhl 2007).

Where functions of such pervasive transcription have been identified in *S. cerevisiae*, it appears that it is not necessarily the RNA produced but the act of transcription itself that leads to the observed function. The short noncoding *SRG1* transcript inhibits *SER3* expression by transcriptional interference and promoter occlusion mechanisms, as the path of the *SRG1* transcript overlaps transcription factor binding sites within the *SER3* promoter (Martens *et al.* 2004; Martens *et al.* 2005). In this case, the *SRG1* transcript terminates near the beginning of *SER3,
while mutation of the tDNA at SUT467 results in uncontrolled readthrough all the way to the end of ATG31. Similar cis-linked mechanisms may be at work at other yeast loci where noncoding transcription appears to block initiation or elongation of ADH1 (Bird et al. 2006), IMD2 (Kuehner and Brow 2008), URA2 (Thiebaut et al. 2008), FLO11 (Bumgarner et al. 2009), PHO84 (Camblong et al. 2007), and IME4 (Hongay et al. 2006; Gelfand et al. 2011). There is evidence that there are also trans-effects of the noncoding RNA product regulating the PHO84 locus (Camblong et al. 2009). Additionally, full repression of yeast GAL genes (Houseley et al. 2008), IME1 (Van et al. 2012), and again PHO84 (Camblong et al. 2007) requires chromatin modifications associated with ongoing noncoding transcription. While more instances are likely yet to be identified, this handful of yeast genes has incorporated intergenic transcription into their regulatory programs and generally appears to use it as a means of repression. However, we show in this study that if left unchecked, progression of noncoding transcription can have negative consequences on neighboring gene expression, resulting in reduced fitness of cells. This result demonstrating a cryptic transcript-blocking activity of bound Pol III complexes can be added to the list of extra-transcriptional effects of the RNA Polymerase III system.

Our results presented here demonstrate that the tDNA upstream of ATG31 protects against such repressive transcriptional interference effects. Our data are consistent with a model in which TFIIIB, as part of the Pol III complex associated with the tV(UAC)D tDNA, serves as a physical impediment to elongating RNA Pol II initiating at the SUT467 transcriptional start site. In the absence of TFIIIB, nearly complete readthrough by Pol II occurs to produce an extended SUT467-ATG31 RNA transcript. This transcript is not efficiently, if at all, translated into Atg31 protein, as scanning ribosomes (Kozak 2005) attaching at the 5' end of the extended transcript would encounter start and stop codons before reaching the ATG31 start codon. This readthrough transcript
appears to be both capped and polyadenylated, as the 5'-RACE protocol includes a phosphatase treatment before decapping and adaptor ligation, ensuring that only capped 5' ends are mapped, and the long transcript is enriched in Northern analysis of poly(A)-purified RNA (A. Korde, unpublished data).

The small amount of Atg31p we detect in our Western blots likely results from low levels of normal ATG31 transcripts that are undetectable in Northern blots from cells grown in rich media. Extracts from autophagy induced tdna mutants show a slight increase in protein levels, along with detectable normal ATG31 transcripts in Northern blots of RNA isolated from the same cultures (Right panels of Figure 2.7 part-A and part-B). This suggests that under conditions that induce ATG31, limited normal initiation is slightly enhanced, but protein levels are still lower than in wild-type cells.

Previous work from our lab and others has shown that certain extra-transcriptional effects associated with tDNAs can be mediated by binding of the TFIIIC complex alone. Propagation of silencing at the HMR mating locus can be blocked by replacing the tDNA downstream of the HMR-I silencer with an ETC site, and insertion of an ETC site between UASG and GAL10 insulates the promoter from Gal4p activation (Simms et al. 2008). Heterochromatin boundary activity of TFIIIC-only containing complexes is also observed in Schizosaccharomyces pombe (Noma et al. 2006; Scott et al. 2006). Additionally, the ETC6 site within the TFC6 promoter may modulate transcription by an insulator-like mechanism (Kleinschmidt et al. 2011). In this case of preventing readthrough of intergenic transcription, the binding of TFIIIC alone is clearly not sufficient. While TFIIIC binds to B-box sequences in vitro with extremely high affinity (Lefebvre et al. 1994; Jourdain et al. 2003), this binding is somehow tempered by passage of the Pol III enzymatic complex during transcription of the internal control element regions.
On the other hand, after recruitment of TFIIIB by TFIIIC, the tightly bound TFIIIB complex appears to be fixed, as in vitro experiments have shown that TFIIIB-DNA complexes are resistant to high salt and heparin treatments (Kassavetis et al. 1990; Kassavetis et al. 1995). The fully assembled TFIIIB complex also is thought to be “kinetically trapped” (Cloutier et al. 2001), with a half-life on the order of a full yeast cell cycle, and fully assembled TFIIIB likely persists at tDNAs until regulated release during mitosis or stationary phase (Fairley et al. 2003; Roberts et al. 2003). Such characteristics of TFIIIB are consistent with our results that suggest that formation of this complex is the major impediment to cryptic transcript readthrough by SUT467. These results are also compatible with earlier in vitro studies that demonstrated the ability of Pol III to transcribe through assembled TFIIIC but not assembled TFIIIB (Bardeleben et al. 1994).

Our results suggest that TFIIIC yields to Pol II in a similar manner as it does to Pol III, since replacing the tDNA with ETC4 allowed readthrough of Pol II even though chromatin immunoprecipitation analysis revealed that TFIIIC was bound to the ectopic ETC4 site (Figure 3 part-C). While we have not mapped the exact 3’ end of the SUT467 transcript, the annotated end mapped by tiling array analysis (Xu et al. 2009) places it within 20 bp of the 5’extent of the expected TFIIIB footprint at this tDNA (schematically depicted in Figure 2.4). Estimation of this 5’ end of the TFIIIB footprint is based on earlier in vitro footprinting studies (Kassavetis et al. 1989) and a recent global “bootprinting” analysis of in vivo bound Pol III transcription factors (Nagarajavel et al. 2013) . The location of the 3’ end of SUT467 is consistent with TFIIIB being a transcriptional roadblock that is resistant to displacement by transcribing Pol II.

A curious sidelight to this study is the appearance of the intermediate length transcript in TFIIIB and TFIIIC mutants, but not in tDNA mutants. This is most likely initiated by Pol II, as in
brf1, tfc3, and tfc6 mutants, the tDNA terminator sequence is still present, so it is unlikely that this is a Pol III transcript. We speculate that in these mutants, TFIIB binding still occurs, but is unstable, and perhaps dissociation of the Brf1p and Bdp1p subunits occurs before loss of TBP at the site. Such a lingering TBP might then recruit factors necessary to then subsequently recruit Pol II immediately upstream of the tDNA. Alternatively, the Pol III complex may mask a cryptic Pol II promoter, which is revealed in a subset of cells containing mutations in the Pol III transcription factors.

There is mounting evidence that a much larger fraction of genomes is transcribed than was previously appreciated. While RNA degradation pathways generally keep most of these transcripts at low levels (Wolin et al. 2012), it has become clear that the act of intergenic transcription can have significant effects on neighboring genes. Due to such observations, one must consider how mutation of a specific genomic locus may affect expression of nearby genes in addition to the targeted gene when assigning the actual cause of observed phenotypes (Wei et al. 2011). To assess the global nature of RNA Pol III extra-transcriptional effects, we are conducting RNA-seq analysis of wild-type vs. Tfc6p under-expressing mutant strains. Previous studies have been conducted to determine the global effects of Pol III deficiencies (Conesa et al. 2005), but the RNA was analyzed by coding sequence microarray, which could not detect effects involving intergenic transcription. Inspection of preliminary RNA-seq results suggests that when the Pol III complex is globally compromised, several tDNA proximal genes may be affected as described here, and in other possibly unique ways (Q. Wang, A. Korde, and C. Nowak, unpublished results). This type of result also raises the question of how to interpret phenotypes due to mutations that may globally affect intergenic transcription (which may be relevant in mutants of other DNA and chromatin binding proteins), or as shown here for mutation at a specific locus, as unchecked cryptic transcription can
lead to unexpected and even detrimental misexpression of downstream genes. While a subset of pervasive transcription products themselves may be noise, multiple mechanisms must exist to keep secondary effects of their production in check.
CHAPTER 3
FUNCTIONAL ANALYSIS OF ETC4 SITE AT RAD2-TNA1 LOCUS

Introduction

In *Saccharomyces cerevisiae*, genome-wide surveys for RNA polymerase III (Pol III) transcription complex occupancy revealed eight intergenic loci which were found to be occupied by the TFIIIC complex but not by TFIIIB or Pol III, thus referred as ‘Extra-TFIIIC’ (ETC) sites (MOQTADERI and STRUHL 2004). These non-transcribed TFIIIC bound loci are mostly situated between divergently transcribed RNA polymerase II (Pol II) genes (MOQTADERI and STRUHL 2004). Sequence analysis of ETC sites revealed that they are comprised of a conserved B-box sequence, and also contain an additional conserved 10-base 3’ extension (Figure 3.1). This extended B-box consensus is conserved among four yeast species: *S. cerevisiae*, *S. mikatae*, *S. bayanus*, and *S. paradoxus*, suggesting an important biological function (MOQTADERI and STRUHL 2004).

Thousands of sites were identified in human (MOQTADERI et al. 2010) and mouse (CARRIERE et al. 2012) that are occupied by TFIIIC and are situated between closely spaced divergently transcribing Pol II genes, suggesting similarity with yeast ETCs. In *Schizosaccharomyces pombe*, TFIIIC binding sites were found to be predominantly associated with the nuclear periphery, possibly mediating three-dimensional organization of the fission yeast genome. Therefore these TFIIIC bound loci were named chromosome-organizing-clamp (COC) sites in *S. pombe* (NOMA et al. 2006). COC sites are found between divergently transcribed genes and majority of them were within a few hundred base pairs of promoters of Pol II transcribed genes, but the effect of these sites on Pol II-transcribed genes is unclear. (KIRKLAND et al. 2013). Additionally, in *S. pombe* inverted repeat regions at cryptic mating type loci were also occupied by TFIIIC and found to possess heterochromatin barrier function (NOMA et al. 2006).
Figure 3.1. Conserved B-box sequence across ETC sites. The B-box consensus (surrounded by blue box) is derived from 274 *S. cerevisiae* tRNA genes and found to be conserved at ZOD1 and eight ETC loci in different *Saccharomyces* species. ETC4 is highlighted with red. Neighboring RNA polymerase II transcribed genes are specified in the parenthesis. Nucleotides in bold are identical across the four yeasts *S. cerevisiae*, *S. bayanus*, *S. mikatae*, and *S. paradoxus*. The size of the letter in the sequence indicates its degree of identity across the contributing sequences. The cytosine residue (in green box) is highly conserved in all B-box consensus across the four different species under study. Adapted from (MOQTADERI and STRUHL 2004).

In *S. cerevisiae*, functional analyses of some of the ETC sites have revealed chromatin boundary function. Our previous lab studies have shown that the ETC6 site upstream of the TFC6 gene on chromosome IV exhibits enhancer blocking properties (KLEINSCHMIDT et al. 2011) as does ETC4 when ectopically placed between the GAL gene and the upstream activation sequence (UAS) (SIMMS et al. 2008). In addition to insulator activity, heterochromatic barrier function was also observed when ETC4 (SIMMS et al. 2008) or ETC2 (VALENZUELA et al. 2009) were used in reporter constructs. When a 90 bp fragment containing the ETC4 sequence was cloned between GAL10 and its UAS, cells were unable to grow on minimal media containing galactose, indicating insulator type function of TFIIIC bound at ETC4 ectopic sites (SIMMS et al. 2008). Similarly, when the same
ETC4 fragment along with ADE2 reporter gene were artificially placed downstream of the HMR locus, heterochromatin propagation was blocked due to ETC4 barrier function. However, the mutation of the conserved B-box of ETC4 inhibits TFIIIC binding; therefore cloning of the mutated ETC4 sequence between GAL10 and its UAS or at HMR could not function as an insulator to gene activation or barrier to heterochromatin spread, respectively (Simms et al. 2008).

For tRNA genes, flanking sequences of the gene plays an important role in boundary activities (Donze and Kamakaka 2001), perhaps by stabilizing the binding of transcription factors at the gene promoter. Therefore we were interested in examining the activity of ETC4 at its natural chromosomal location that is between divergently transcribing RAD2 and TNA1 genes (Figure 3.2 part-A and part-B). ETC4 is conserved in other Saccharomyces species (Figure 3.2 part-C). In Saccharomyces cerevisiae, Rad2p is involved in nucleotide excision repair (NER). Together with other Rad proteins, Rad2p binds DNA lesions that are caused due to UV exposure or chemical crosslinking. As the first members of DNA damage response, they cause unwinding of the surrounding DNA duplex, followed by targeting incisions on both sides of the damaged DNA, which releases a 25-30 bases of damaged fragment (Prakash and Prakash 2000).

TNA1 encodes for high affinity nicotinic acid plasma membrane permease enzyme (Llorente and Dujon 2000). Low levels of extracellular nicotinic acid activates transcription of TNA1, whereas at higher concentration of nicotinic acid, TNA1 is repressed by Sum1/Hst1 complex (Bedalov et al. 2003). Repression activity of Hst1-Sum1 complex was previously observed on the TNA1 gene (Bedalov et al. 2003) and we also confirmed it by detecting increase in expression of TNA1 in sum1Δ mutants (Figure 3.3).
Figure 3.2. Predicted position of ETC4 conserved B-box sequence with respect to RAD2 and Sum1p binding sites. Adapted from *Saccharomyces* genome database.

A) Sum1p binding sites are located upstream of divergently transcribed RAD2 and TNA1 genes on Chromosome VII. Coordinates are obtained from *Saccharomyces* genome database (SGD). B) The schematic for the region from RAD2 (green) to the nearest Sum1p binding site (Blue) is depicted. Red solid box indicated predicted position of ETC4 site. Numbers presented below denote distances in base pairs between specified regions. C) B-box consensus at ETC4 (red box) and Sum1p binding sequence (blue box) are conserved across other *Saccharomyces* species such as *S. mikatae* and *S. paradoxus*. Gray region and stars indicate sequence identity among three Saccharomyces species. At the bottom, arrow denotes transcriptional start site for RAD2 gene.

Figure 3.3. Sum1p mediated repression of TNA1 gene.

Northern analysis shows derepression of TNA1 in sum1Δ mutants. Actin mRNA levels were used as input control. Strains were grown in minimal media containing all amino acids. Key: WT1-DDY3, WT2-DDY4, sum1Δ mutant 1-DDY 4233 and sum1Δ mutant 2-DDY 4234.
Sum1p binding sites were identified within the intergenic region between TNA1 and RAD2 by chromatin immunoprecipitation, followed by microarray analysis (ChIP on ChIP) (Harbison et al. 2004; MacIsaac et al. 2006) (Figure 3.3). SUM1 encodes a sequence-specific DNA-binding repressor protein that binds to the operators (middle sporulation elements, or MSEs) of middle-sporulation genes (Xie et al. 1999) in mitotic cells. The SUM1 gene had previously been identified by a dominant mutation (SUM1-1), that has single missense mutation in SUM1 gene, changing threonine to isoleucine at the C-terminus of the protein (Chi and Shore 1996). Later, the role of Sum1p was identified in regulation of replication and non-meiotic genes that encode enzymes in the de novo pathway for NAD\(^+\) biosynthesis (Bedalov et al. 2003). While Sum1p binding to the MSE element represses the genes (Pierce et al. 2003), binding to the silencer is necessary for silencing at HML (Irlbacher et al. 2005).

Upon binding to specific sequences, both Sum1-1 as well as Sum1 recruit Hst1p. Hst1p (homologue of Sir2p) is a NAD\(^+\)-dependent histone deacetylase (Hickman and Rusche 2007) that deacetylates tails of H3 and H4 (McCord et al. 2003), thereby it is responsible for the repressive function. When associated with Sum1p, Hst1p causes Sir2-independent transcriptional repression of target genes. In sir2\(\Delta\) S. cerevisiae mutants, Sum1-1p/Hst1p complex suppresses silencing defects and spreads heterochromatin at the HMR silent mating locus (Klar et al. 1985; Laurensen and Rine 1991). Following recruitment to the HMR silencers, Sum1-1p and Hst1p spread across the HMR domain, deacetylating the histones and thereby mediating repression (Lynch et al. 2005). Therefore, expecting a similar mechanism at RAD2-TNA1 locus, we were interested to see if TFIIIC bound ETC4 site has any role in protecting RAD2 gene by preventing the spread of Sum1p/Hst1p-mediated repression.
In this study, \textit{RAD2} was used as a reporter gene to investigate the potential boundary function of \textit{ETC4} at its natural location. We hypothesized that TFIIIC bound at the \textit{ETC4} site acts as a repression-blocking barrier which prevents repression of \textit{RAD2} gene from Hst1p/Sum1p complex (Figure 3.4 part-A). Therefore, deletion of \textit{ETC4} would compromise TFIIIC binding and result in loss of the barrier activity (Figure 3.4 part-B).

![Figure 3.4](image)

**Figure 3.4.** Hypothesis: TFIIIC bound \textit{ETC4} site acts as a barrier to spread of repression
A) Under high concentrations of nicotinic acid, \textit{TNA1} gene (yellow) is repressed by Sum1p-Hst1p complex (light and dark blue beads). In presence of \textit{ETC4} site (Red box) bound TFIIIC complex (orange) would prevent repression of \textit{RAD2} (green) by Sum1p-Hst1p complex. B) Further, when \textit{ETC4} is deleted, due to absence of TFIIIC, Sum1p-Hst1p mediated repression would affect the transcription of \textit{RAD2}.

**Materials and methods**

Plasmids used in this study are listed in Table 3.1. All yeast strains used in this study are isogenic to W303-1a and listed in Table 3.2. Oligonucleotides used in this study are listed in Table 3.3.
Table 3.1 Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDD 638</td>
<td>pRS416 with URA3 marker</td>
</tr>
<tr>
<td>pDD 1154</td>
<td>ETC4 sequence in TOPO vector</td>
</tr>
<tr>
<td>pDD 1194</td>
<td>RAD2-TNA1 intergenic etc4Δ</td>
</tr>
</tbody>
</table>

Table 3.2 Yeast stains used in this study

<table>
<thead>
<tr>
<th>Strains</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDY 3</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1</td>
</tr>
<tr>
<td>DDY 3863</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 etc4Δ::URA3</td>
</tr>
<tr>
<td>DDY 4091</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 etc4Δ box B</td>
</tr>
<tr>
<td>DDY 4233</td>
<td>MATa ade2-1 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 sum1Δ::LEU2</td>
</tr>
<tr>
<td>DDY 4234</td>
<td>MATa ade2-1 his3-112 leu2-3,112 lys2Δ sum1Δ::URA3 trp1-1 ura3-1</td>
</tr>
<tr>
<td>DDY 4385</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 etc4Δ</td>
</tr>
<tr>
<td>DDY 4432</td>
<td>MATa ADE2-1 his3-11 leu2-3,112 LYS2 trp1-1 ura3-1 sum1Δ::LEU2 etc4Δ</td>
</tr>
<tr>
<td>DDY 4457</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 rad2[ORF]Δ::URA3[ORF]</td>
</tr>
<tr>
<td>DDY 4535</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 rad2[ORF]Δ::URA3[ORF] etc4Δ</td>
</tr>
<tr>
<td>DDY 4537</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 rad2[ORF]Δ::URA3[ORF] etc4Δ</td>
</tr>
<tr>
<td>DDY 4538</td>
<td>MATa ADE2-1 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 rad2[ORF]Δ::URA3[ORF] sum1Δ::LEU2 etc4Δ</td>
</tr>
<tr>
<td>DDY 4540</td>
<td>MATa ADE2-1 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 rad2[ORF]Δ::URA3[ORF] sum1Δ::LEU2 etc4Δ</td>
</tr>
<tr>
<td>DDY 4572</td>
<td>MATa ADE2 his3-11 leu2-3,112 LYS2 trp1-1 ura3-1 rad2[ORF]Δ::URA3[ORF] sum1Δ::LEU2</td>
</tr>
<tr>
<td>DDY 4575</td>
<td>MATa ADE2 his3-11 leu2-3,112 LYS2 trp1-1 ura3-1 rad2[ORF]Δ::URA3[ORF] sum1Δ::LEU2</td>
</tr>
<tr>
<td>DDY 4577</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 rad2[ORF]Δ::URA3[ORF]</td>
</tr>
</tbody>
</table>

Table 3.3 Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Oligos</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDO 402</td>
<td>ATGGATTCTGAGGTTGCTGC</td>
<td>ACT1 RT-PCR control</td>
</tr>
<tr>
<td>DDO 403</td>
<td>CAAAACGCTTGGATGGAAAC</td>
<td>ACT1 RT-PCR control</td>
</tr>
<tr>
<td>DDO 701</td>
<td>GCAACAAATTACATGGAGTC</td>
<td>TNA1 northern probe top</td>
</tr>
</tbody>
</table>
Construction of strains

Creating etc4 yeast mutants

Plasmid pDD 1154 was subjected to site directed mutagenesis using DDO 1129 and DDO 1130 to generate pDD 1194 containing 23bp deleted from ETC4 (etc4Δ). EcoRI digested fragment from pDD 1194 containing mutated ETC4 was inserted into yeast DDY 3863 by homologous recombination to create etc4Δ mutants (DDY 4385). SUM1 deleted (DDY 4233 and DDY 4234) strain were obtained from Laura Rusche. The double mutants etc4Δ sum1Δ (DDY 4432) were created by crossing etc4Δ yeast strain (DDY 4385) with sum1Δ strain (DDY 4233).
Creating *rad2::URA3* strains

*URA3* ORF was PCR amplified from pDD 638 using DDO 1286 and DDO 1272 which included approximately 300 bp of homology on each side of the *URA3* coding sequence. The *URA3* fragment was then used to replace *RAD2* coding sequence in wild-type (DDY 3), *etc4Δ* (DDY 4385) and *etc4Δ sum1Δ* (DDY 4432) to create DDY 4457 (or DDY 4577), DDY 4535-4537 and DDY 4538-4540 respectively by homologous recombination.

**Growth media**

All yeast cells were grown in nutrient rich YPD (1% yeast extract, 2% peptone and 2% glucose) medium (unless otherwise mentioned) at 30°C on a rotary shaker. Strains were grown in YPD for UV sensitive assays. For measuring colony sizes and growth curve experiment, minimal media (YMD) containing 2% glucose and 1.7g/l of yeast nitrogen base with all amino acids except uracil (-ura) with or without agar was used for growing *RAD2* replaced *URA3* strains.

**UV sensitivity assay**

For preliminary UV sensitivity assay, isolated colonies from each wild-type (DDY 3), *sum1Δ* (DDY 4233), *etc4Δ* (DDY 4385) and *etc4Δ sum1Δ* (DDY 4432) were grown overnight in 4 ml liquid YPD. Next day, optical densities (O.D) of the overnight cultures were determined by UV spectrophotometer to estimate the amount of cells (O.D 1.0 = 10⁷ cells/ml). For each strain, about 10³ cells/ml were achieved by serially diluting respective overnight cultures. 100 µl of those 10³ dilutions were plated on each YPD plate containing agar to obtain approximately 100 cells per plate. After plating, half the number of plates for each strain were exposed to ultraviolet light (5000 μJ) using UV stratalinker to obtain approximately 50% survival for wild-type (LD50). UV-exposed plates were immediately covered in foil to maintain dark condition during DNA repair. All plates including control plates (unexposed to UV) were incubated for 2 days at 30°C. Number of colonies
on triplicate plates for each ‘control’ and ‘UV-exposed’ were counted to calculate percentage survival for a particular strain.

**Northern analysis**

For isolation of total yeast RNA, cells were grown in YPD at 30°C up to \( A_{600} = 1.0 \). For cycloheximide-mediated RAD2 induction, cells were grown in YPD media up to \( A_{600} = \sim 0.8 \) and then cycloheximide was added as per 0.1 mg/ml concentration. These treated cultures were allowed to grow for another 1 hour. Pellets were washed with DEPC water and re-suspended in extraction buffer (50mM NaOAc, 10 mM EDTA, pH to 5.0 with HOAc) and 1% SDS before storing at -20°C. Total RNA was extracted by hot acid/phenol at 65°C and purified with buffered phenol/chloroform, then precipitated with 100 % ethanol. RNA pellets were washed with 70% ethanol, dried and re-suspended in DEPC treated water. To prepare samples for northern analysis, 10ug total RNA for each sample was dried using SpeedVac DNA concentrator and re-suspended in 10ul RNA formaldehyde loading buffer with ethidium bromide. Each RNA sample was heated for 15 minutes at 65°C and was resolved on a MOPS/formaldehyde/agarose gel and blotted to Zeta–Probe (Bio-Rad) membrane by wicking in transfer buffer (0.01N NaOH + 3M NaCl). After UV crosslinking, the Zeta-probe membrane was incubated in ULTRAhyb (Ambion/Life Technology) for a pre-hybridization step. Probes were generated by PCR amplification of ORFs from TNA1, RAD2 or URA3 using DDO 701+702, DDO 1225+1226 and DDO 1229+1230, respectively; including the T7 RNA polymerase promoter attached to the downstream primer. The PCR products were \(^{32}\)P-\(\alpha\)-UTP-labeled in a T7 polymerase *in vitro* transcription reaction at 37°C for 30min to 1hr., and then the radioactive probe was filtered to remove unincorporated UTP before adding to the filter.
membrane. Hybridization was carried out overnight at 65°C. Membranes were washed twice in 2X SSC for 5min each in 65°C, then washed twice in 0.1X SSC for 15min each in 65°C. Zeta-probe membrane was then placed under phosphor screen in exposure cassette for 2hr to overnight and scanned using Typhoon scanner (LSU Genomics Facility).

**Growth assays**

Overnight culture of *RAD2* replaced *URA3* strains in YMD-ura were used of growth assay in YMD-ura. For obtaining growth curve, starting with O.D 0.1, 25ml culture was grown in liquid YMD-ura and optical densities were measured at interval of 1 hour up to 8 hours and the last reading was taken after 16 hours. Every hour optical density was measured using spectrophotometer at 600 nm wavelength. For estimating growth by colony sizes, after serial dilution about 100 cells plated on YMD-ura agar plates for each strain and they were allowed grow for two days. For each strain, diameter of about 35 colonies were measured using ImageJ software and the averages were compared to estimate the overall colony size variation.

**RT-PCR**

First strand of cDNA was synthesized from 0.5µg of total RNA with gene specific primer DDO 1611 using ProtoScript M-MuLV first strand cDNA synthesis kit (Biolabs- NEB # E6300S). Total RNA were used to make control samples which did not contain reverse transcriptase enzyme (hence No RT controls) 1µl of cDNA or No RT control was used as a template and using gene-specific primers (DDO 1612 and DDO1608) cDNA were amplified for 35 cycles by *Taq* polymerase. PCR products were resolved on 1% agarose gel to visualize the difference in the intensities of *RAD2* mRNA levels.
Results

Deletion of ETC4 affects UV resistance

In order to investigate the role of TFIIIC bound ETC4 site in potentially protecting RAD2 from Sum1p/Hst1p-mediated repression, we created etc4Δ mutant yeast by integrating mutagenized ETC4 fragment in wild-type as well as sum1Δ strains by homologous recombination. In response to DNA damaging conditions such as UV exposure, RAD2 expression increases to produce more Rad2p protein for nucleotide excision repair which is necessary for cell survival (Siegele et al. 1989). We conducted a UV-sensitivity assay and compared survival of etc4Δ mutants with wild-type after exposure to ultraviolet light (Figure 3.5). Survival of sum1Δ mutants were comparable to wild-type survival, but etc4Δ mutants showed reduced survival after UV exposure, suggesting partial repression of RAD2.

To detect effects of Sum1p-Hst1p mediated repression on RAD2 transcription, we performed Northern analysis on etc4Δ and etc4Δ sum1Δ double mutants under induced and normal growth conditions. RAD2 transcription was induced by adding DNA damaging agent, cycloheximide, in the growth cultures. For each strain total RNA was extracted from induced and normal (control) yeast cells. Unfortunately, on the northern blot, we could not detect discrete mRNA bands using RAD2 probe, instead we observed a smear. However, actin mRNA were detected after re-probing the same blot with actin probe, which suggested specific instability of RAD2 mRNA. (Figure 3.6).

Replacement of RAD2 ORF with URA3 ORF as a marker gene

Considering the possibility of degradation or intrinsic instability of RAD2 mRNA, we replaced RAD2 ORF by coding sequence of URA3 marker gene in wild-type as well as in mutants, leaving the endogenous RAD2 promoter intact. The schematic for construction of these strains is
Figure 3.5. In absence of ETC4, survival of cells is significantly reduced upon UV exposure. Left plot: Number of colonies counted for strains under control (No UV exposure) and UV (exposed to 5000) conditions are represented. Cells from each strain were grown in YPD media and serially diluted to obtain ~100-200 colonies on control plates. For UV treatment, after plating on YPD agar, plates were exposed to 5000 µJ of UV light. Right plot: Percentage survival is plotted that suggests ETC4 deletion increases UV sensitivity and thereby reduces cell survival. Colony counts are representing average of colonies from at least 3 plates for each strain and error bar denotes their standard deviations. Key: WT- DDU3, sum1Δ- DDU 4233, etc4Δ- DDU 4385 and etc4Δ sum1Δ- DDU 4432.

Figure 3.6. Northern analysis of RAD2 transcription. Total RNA from WT (DDY3), etc4Δ (DDY 4385) and etc4Δ sum1Δ (DDY 4432) were extracted from cultures growing under normal growth condition as well as after cycloheximide treatment. Northern blot represents undetected mRNA bands and steady mRNA levels for actin are depicted in the bottom panel.
shown in Figure 3.7. From the resulting rad2::URA3 strains, URA3 transcription was determined by Northern analysis to assess any repressive effect on URA3 marker gene transcription in the background of ETC4 deleted condition (Figure 3.8). As expected, sum1Δ mutants showed higher mRNA levels as compared to wild-type indicating derepression of URA3 in absence of Sum1 protein. Whereas, the etc4Δ mutant showed no change in URA3 transcription when compared to wild-type mRNA levels.

Figure 3.7. Replacement of RAD2 ORF by URA3 coding sequence.
~1.5 Kb URA3 coding sequence (orange box) was inserted at the place of RAD2 ORF (green box) while keeping endogenous RAD2 promoter intact. The replacement was achieved by homogenous recombination between RAD2 specific flanking sequences in URA3 construct. Construction of control rad2::URA3 as well as mutant strains is explained in the flow chart.

As a second test for URA3 expression driven from the endogenous RAD2 promoter, we performed a growth assay on rad2::URA3 strains, in liquid minimal media lacking uracil. As expected rad2::URA3 etc4Δ mutants showed less growth in uracil dropout media, indicating drastically reduced efficiency of producing functional Ura3p. Contradicting the results from
Northern analysis using URA3 probe indicate increase in transcription in sum1Δ (lanes 3 and 4) and etc4Δ sum1Δ double (lanes 7 and 8) mutants. Actin (ACT1) mRNA levels represents loading control. Key: Lane 1-DDY4457, lane 2- DDY 4577, lane 3- DDY 4572, lane 4- DDY 4575, lane 5- DDY 4535, lane 6- DDY 4537, lane 7- DDY 4538 and lane 8- DDY 4540.

Northern analysis, rad2::URA3 sum1Δ showed the highest growth and despite of the high levels of URA3 transcription detected from etc4Δ sum1Δ double mutants, they were moderately efficient in producing functional Ura3p for cell growth in minimal media lacking uracil (Figure 3.9 part-A). Similar results were observed when colony sizes were determined by measuring diameter of colonies for each rad2::URA3 strain. The mean colony sizes are plotted as histograms in Figure 3.9 part-B.

Consistent with the growth curve, the etc4Δ mutant showed reduced colony size on the minimal media without uracil. Therefore, both growth assays suggested the impaired expression of Ura3p in ETC4 deleted strains, which could be due to repression by Sum1p-Hst1p complex. However, overall URA3 transcription levels detected on Northern blot did not correlate with results from growth assays.

As we were unable to detect RAD2 mRNA on northern blots, we synthesized cDNA from RAD2 mRNA by reverse transcription (RT). To eliminate the genomic DNA being used as a
Figure 3.9. Growth assays for rad2::URA3 strains suggest function of Ura3p is affected in etc4Δ mutants.
A) Growth rate is plotted for each rad2::URA3 strain as a measure to detect ability to produce functional Ura3 protein for cell growth. Cells were allowed to grow in liquid minimal media without uracil and growth is measured by determining cell densities at every hour up to 8 hours. For each condition two separate strains were examined. Key: #1 rad2::URA3- DDY 4457, #2 rad2::URA3- DDY 4577, #1 rad2::URA3 sum1Δ- DDY 4572, #2 rad2::URA3 sum1Δ- DDY 4575, #1 rad2::URA3 etc4Δ- DDY 4535, #2 rad2::URA3 etc4Δ- DDY 4538, #1 rad2::URA3 etc4Δ sum1Δ- DDY 4537, #2 rad2::URA3 etc4Δ sum1Δ- DDY 4540. B) Colony growth was determined by plating approximately 35-40 colonies on minimal media lacking uracil. After 2 days of incubation, average colony size was calculated by measuring diameter (in millimeters) of more than 30 colonies for each strain. The graph is representing mean colony sizes for rad2::URA3 (DDY 4577), rad2::URA3 sum1Δ (DDY 4575), rad2::URA3 etc4Δ (DDY 4535) and rad2::URA3 etc4Δ sum1Δ (DDY 4540).
template during RT PCR, during cDNA synthesis, we designed a gene-specific primer which had 22 bp non-specific sequence at 5’ end. The cDNA template was subjected to RT PCR and to amplify only synthesized cDNA, we used a set of gene-specific primers, in which the downstream primer was complementary to the non-specific sequence. After RT-PCR, mRNA levels were compared between etc4 mutants and wild-type strain that indicated no significant change in transcription levels but this could be due to saturation. (Figure 3.10). Therefore, quantitative RT PCR is required for conclusive results.

Figure 3.10. Deletion of ETC4 did not affect transcription of RAD2. RT-PCR of etc4Δ mutants (lanes 8 and 9) when compared to wild-type (lane 7), did not show any detectable change in RAD2 mRNA levels when gene specific primers (DDO 1608 and DDO 1612) were used. No RT controls (lanes 10-12) did not show any amplification, confirming absence of residual genomic DNA in the PCR. Lanes 1-3 show amplification of actin from cDNA and No RT controls (lanes 4-6) using DDO 402 and DDO 403. Strains used: wild-type DDY3 (lanes 1, 4, 7 and 10), etc4 mutants- DDY 4091 (lanes 2, 5, 8 and 11) and DDY 4385 (lanes 3, 6, 9 and 12).

**Discussion**

In addition to Pol III-transcribed genes, Pol III transcription complexes are also bound at non-transcribed loci in the genome. Global analysis of occupancy of these complexes revealed *ETC* sites which are found to be occupied by TFIIIC but not Pol III. The extended *B-box* sequence present at *ETC* sites was found to be essential for TFIIIC binding (MOQTADERI and STRUHL 2004). Previously, many extra-transcriptional functions have been ascribed to TFIIIC bound loci (DONZE 2012). Previous lab studies have demonstrated heterochromatin barrier as well as enhancer-
blocking insulator type activities of the \textit{ETC4} sequence, and when \textit{ETC4} was engineered at ectopic sites it was occupied by TFIIIC but not TFIIIB (Simms \textit{et al.} 2008).

In this study, boundary function of \textit{ETC4} at its natural location was investigated. However, we could not conclusively determine \textit{ETC4} function at \textit{RAD2-TNA1} locus. Gene-specific repression (Xie \textit{et al.} 1999) activity as well as silencing spreading ability (Rusche and Rine 2001) have been attributed to Hst1p associated complexes. Therefore, the rationale of the study was based on the assumption that similar to the Sum1p-Hst1p mediated spreading of silencing at cryptic mating locus, Sum1p-Hst1p complexes which are bound upstream of \textit{RAD2} affect \textit{RAD2} expression. However, our results could not detect any consistent effect of \textit{SUM1} deletion on the transcription of \textit{URA3} marker gene which was used to replace \textit{RAD2} ORF. On the other hand, growth assays showed increase in the ability to produce functional Ura3p in \textit{rad2::URA3 sum1\Delta} mutant, suggested possibility of Sum1p mediated repression. At the same time, for \textit{rad2::URA3 etc4\Delta sum1\Delta} double mutants, high levels of \textit{URA3} transcription were not consistent with the results observed in growth assays. Considering the important role of nucleosome positioning in transcription, we cannot eliminate the possibility that the deletion of \textit{ETC4} and absence of bound Sum1p at upstream binding sites may have positively affected transcription of downstream Pol II-transcribed gene (\textit{RAD2} or \textit{URA3}). Based on these results, reduced activity of Rad2p (UV sensitivity) and even Ura3p (growth assays) have been observed in \textit{etc4} mutants but apparently are not due to reduction of mRNA levels of \textit{RAD2} or \textit{URA3}. Therefore, we speculate that deletion of \textit{ETC4} that abolishes TFIIIC binding at the upstream of \textit{RAD2} possibly influences selection of \textit{RAD2} transcriptional start sites and hence might affect translation levels and/or protein function. However, to uncover the precise role of \textit{ETC4} at \textit{RAD2-TNA1}, further experimentation is required. The following experiments are intended to be performed in future.
• Detection of altered protein or inhibited translation:
  o Amplification of 9X-Myc fragment using primers containing sequence homology for RAD2 downstream sequence.
  o Transformation of wild-type and etc4Δ mutants with 9X-Myc fragment.
  o Selection of positive transformants by PCR.
  o Protein extraction from epitope tagged strains and western analysis.

• Results of western analysis will reveal the effect of ETC4 deletion on translation, which could possibly alter the RAD2 5’ UTR due to utilization of alternate transcription start sites.

• 5’ RACE will be performed on total RNA extracted from wild-type and etc4Δ mutants to map the altered transcriptional start sites.
CHAPTER 4
GLOBAL ANALYSIS OF EFFECTS OF RNA POLYMERASE III TRANSCRIPTION COMPLEXES ON RNA POLYMERASE II TRANSCRIBED LOCI

Introduction

In a multicellular organism almost all cells contain nearly exact copies of chromosomal DNA. In spite of this, cells can have very distinct appearances, functions and ability to respond to extracellular stimuli. These differences are possible because cells make use of different stretches of DNA, called genes, as templates to build functional cellular products (proteins or regulatory RNA transcripts) by differentially regulating gene expression (Figure 4.1).

In the first step of gene expression, known as transcription, the information stored in the DNA in the form of A (adenine), T (Thymine), G (Guanine), and C (Cytosine) sequences, is used to create ribonucleic acid molecules (RNA). RNA is synthesized using one of the DNA strands as a template and has the same chemical structure except that thymine is replaced by uracil (U). Some RNA molecules can be the end product themselves, and others can in turn be used as a template for the creation of other molecules, proteins, in a process called translation. The RNAs that are used as a template for proteins are known as messenger RNA (mRNA) and the ones that are directly involved in regulatory mechanism but do not generate proteins are classified as non-coding RNAs (ncRNAs).

The collection of RNAs produced from the set of genes which are transcribed in any specific cellular condition is known as the transcriptome. Multi-level regulatory mechanisms acting on transcription and translation processes are responsible for generating multiple variants of an RNA molecule transcribed from the same gene that results in production of different functional protein products. For instance, many eukaryotic genes can produce more than one variant of mRNA because of alternative splicing, RNA editing, or alternative transcription initiation and
termination sites, adding more complexity to the cellular transcriptome. Therefore to interpret this level of complexity and to reveal the molecular constituents of cells during a specific cellular condition, transcriptome analysis is necessary. Moreover, unlike other cellular products, RNA

![Diagram of gene expression process](image)

**Figure 4.1.** Schematic of gene expression process.
On the top, three separate genes in the genome are shown. The middle gene is zoomed in for showing the details of its three exons and two introns. During the transcription process, the gene is used as a template for mRNA synthesis. On the left, resulting transcript is produced using a regular transcription start site (solid arrow at 5' end of the gene) whereas the right side transcript is a result of beginning of transcription from the alternate start site (dotted arrow at 5' end of the gene). During the post-transcriptional event, pre-mRNA is spliced into mature mRNA. Alternate splicing is observed when exons are excluded. The matured mRNA are then translated into proteins (bead-like structure at the bottom). Differently spliced transcripts generally give rise to proteins with different functions.

samples can be more easily and reproducibly measured in a high-throughput manner with a variety of current technologies (MARIONI et al. 2008; ELLIOTT et al. 2009).

Many technologies have been used over the years for the purpose of measuring gene expression including hybridization-based techniques and sequence-based techniques. The older complementary hybridization-based microarray technique contains thousands of short single stranded DNA molecules called probes, which are attached to fixed locations on a glass or polymer slide. The extracted mRNA are reverse transcribed into single stranded complementary DNA
(cDNA) and each molecule is labelled with a fluorescent dye. These labelled cDNA are then allowed to hybridize with the complementary probe attached to the microarray surface. The expression can then be estimated by the optical measurement of the amount of fluorescence coming from each probe (ALLISON et al. 2006). Although microarrays are a powerful and relatively inexpensive technology, they present several limitations. For example, prior knowledge about a probe sequence is necessary, occurrence of high background noise due to partial complementarity of cDNA to the probes and this method restricts the detection of differential expression of only the specific gene target (probe) between RNA samples.

Unlike microarray methods, sequence-based approaches directly determine the cDNA sequence. Initial approaches of cDNA sequencing with expressed sequence tags (EST) was generally low throughput and not quantitative. New high throughput tag-based methods including serial analysis of gene expression (SAGE) (HARBERS and CARNINCI 2005), cap analysis of gene expression (CAGE) (KODZIUS et al. 2006) and massively parallel signature sequencing (MPSS) (BRENNER et al. 2000) provide precise, ‘digital’ gene expression levels. However, all these approaches are based on expensive Sanger sequencing technology, and short tags cannot be uniquely mapped to the reference genome, thus some portions of a transcriptome usually remained unstudied.

In recent times, sequencing of RNA (RNA-seq) has emerged as the preferred technology for the simultaneous measurement of transcript sequences and their abundance. RNA-seq is also called “a revolutionary tool for transcriptomics” (WANG et al. 2009) since this new high throughput sequencing method has enabled both mapping and quantifying transcriptomes. Compared to other RNA measuring technologies such as quantitative PCR (qPCR) and microarrays, RNA-seq has higher throughput and lower background noise. RNA-seq can measure the expressions of tens of
thousands of genes simultaneously in a single experiment in few days. It also generates ‘digital’ results as opposed to ‘analog’ signals generated in microarray experiments (KAPUR et al. 2008).

Since RNA-seq facilitates sequencing of all of the RNA transcripts in a cell, it provides the ability to look at alternative gene spliced transcripts, post-transcriptional changes, gene fusion, mutations like single nucleotide polymorphism (SNP) and changes in gene expression (MAHER et al. 2009) or even a ribosomal profiling (INGOLIA et al. 2012). RNA-seq can also be used to determine exon/intron boundaries and to confirm or revise previously annotated 5’ and 3’ transcript boundaries. Nagalakshmi et al. (2008) have successfully applied RNA-seq to map the transcribed regions of the yeast genome such as 5’ and 3’UTR (untranslated region) boundaries, alternative initiation codons, 3’ heterogeneity and occurrence of many overlapping genes in Saccharomyces cerevisiae (NAGALAKSHMI et al. 2008). In addition to S. cerevisiae, RNA-seq has been applied to Schizosaccharomyces pombe (WILHELM et al. 2008), Arabidopsis thaliana (LISTER et al. 2008) and human cells (MORTAZAVI et al. 2008). In this study we have applied high throughput RNA-seq approach to analyze the genome-wide effects of extra-transcriptional functions of RNA polymerase III complexes on neighboring genomic loci in Sacchraomyces cerevisiae.

Among three types of RNA polymerases, RNA polymerase III (Pol III) transcribes a variety of small RNA, including transfer RNA (tRNA), 5S ribosomal RNA (5S rRNA), U6 small nuclear RNA, snR52 small nucleolar RNA, cytoplasmic 7SL RNA and RNA component of RNase P. Transcription of these genes requires sequential assembly of Pol III specific transcription factors TFIIIA (only for 5S rRNA), TFIIIC and TFIIIB along with the Pol III enzyme complex.

While Pol III and most of its transcription factors are thought to be dedicated to transcription of Pol III genes, emerging studies have shown that both partial and complete chromosomally bound Pol III transcription complex can have effects on chromatin state and even on genome organization.

82
For instance, Pol III transcribed tRNA genes not only function as transcription units for transfer RNA but also serve as the potential sites for other extra-transcriptional roles (DONZE 2012). Previous global analysis of Pol III transcription complex occupancy have revealed, apart from tRNA genes, other chromosomal locations such as Extra-TFIIC (ETC) sites. These sites are either bound by only TFIIC complex (at ETC1-8 and ETC 10) or by both TFIIC and TFIIB (at ETC9) but not Pol III. Further sequence analysis showed that ETC sites contain conserved B-box consensus sequences which are essential for TFIIC binding. These bound transcription factors, even in absence of Pol III enzyme, contribute to the extra-transcriptional functions.

Previous studies have demonstrated extra-transcriptional effects of Pol III transcription complex bound loci that have influenced neighboring Pol II-transcribed genes (DONZE 2012). In one of the cases, autoregulation of the TFC6 gene promoter was demonstrated to be governed by binding of TFIIC transcription factor at the upstream ETC6 site. This TFIIC binding regulates the expression of TFC6 gene which encodes one of the subunits of TFIIC (Tfc6p). It was observed that mutation in the promoter region of TFC6 gene upstream of ETC6 down-regulated TFC6 transcription and in turn TFIIC binding was reduced at known potential binding loci, including ETC6 (KLEINSCHMIDT et al. 2011). On account of this and many other studies on the specific loci, we were interested in performing global analysis for identifying effects of occupancy of Pol III transcription complexes (partial or complete) on nearby chromatin regions. RNA-seq of thisTfc6p under-expressing mutant followed by expression analysis revealed many differentially expressed loci, and visualization software enabled detection of 5’ or 3’ extensions of previously annotated genes.
**Materials and methods**

*Saccharomyces cerevisiae* strains used in this study are listed in Table 4.1.

Table 4.1 Strains used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDY 3</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1</td>
</tr>
<tr>
<td>DDY 3630</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 ETC6 wild-type (TFC6-ESC2)</td>
</tr>
<tr>
<td>DDY 4300</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 TFC6 promoter mutant #39</td>
</tr>
<tr>
<td>DDY 4301</td>
<td>MATa ADE2 his3-11 leu2-3,112 lys2Δ trp1-1 ura3-1 TFC6 promoter mutant #39</td>
</tr>
</tbody>
</table>

**Preparation of RNA sample for RNA-seq analysis**

Total RNA were extracted from wild type (DDY 3 and DDY 3630) and *TFC6* down-regulated mutants (DDY 4300 and DDY 4301) by acid-phenol extraction technique (as described in Chapters 2 and 3). The extracted total RNA were subjected to DNase treatment using Promega RQ1 DNase kit (M6101) to obtain DNA free total RNA. 50 µg of total RNA was incubated for 20 min at 37 °C with 10X DNase buffer and DNase RQ enzyme. The enzymatic reaction was stopped by adding stop buffer (provided in the kit) and RNA was purified using phenol-chloroform and precipitated by absolute ethanol. The extracted DNase treated total RNA pellet then re-dissolved in nuclease-free water to a concentration of 1µg/µl. The quality of RNA was checked on 1% formaldehyde denaturing gel.

**Preparation of RNA-seq libraries at sequencing center**

RNA-seq library preparation and Illumina sequencing were performed at the Roy J. Carver Biotechnology center at the University of Illinois (http://www.biotech.uiuc.edu/htdna). The RNA-seq libraries were prepared with Illumina's TruSeq Stranded RNA Sample Prep kit that generated 5’ to 3’ strand-specific libraries.
Results and discussion

Quality check of reads

Millions of reads (total 196,295,402) were generated by high throughput sequencing of DDY 3, DDY 3630, DDY 4300 and DDY 4301 RNA-seq libraries. The FASTQ sequence files obtained for each strain were subjected to FastQC software tool to ensure the quality of reads. FastQC tool aims to provide a quality check report, which can spot problems that originate either in the sequencer or in the starting library material. The analysis in FastQC was performed by a series of analysis modules such as per base sequence quality, per base sequence content, GC content, sequence length distribution and also generated a list of overrepresented sequences. As an example, Figure 4.2 part-A represents results obtained from some of the FastQC analysis modules. Most sequencers generate a quality check report as part of their analysis pipeline, for instance, Illumina uses CASAVA 1.8 software tool and the base quality-scores (Sanger score) obtained from CASAVA 1.8 (ASCII) for each FASTQ file are shown in Figure 4.2 part-B. High quality score (above 28) indicate good quality of reads due to successful sequencing.

Read alignment

After the quality check, reads from FASTQ files were aligned to W303 reference genome for S. cerevisiae by using Bowtie2 which generated SAM (sequence alignment or map format) output files. Bowtie2 is a very powerful aligning tool that has a diverse array of applications, for example, mapping of RNA transcripts back to a known genome, or to enumerate the number of sequences mapped back to longer regions such as a gene in the desired genome. The SAM files were then converted to BED files by using program written in perl language. For strand specificity reads BED files were sorted into Crick strand-specific and Watson strand-specific files.
Figure 4.2. Quality assessment of reads.

A) Some of the important analysis modules provided by FastQC online software for DDY3 FASTQ file are shown in different panels. Per base quality is represented in BoxWhisker type plot where each yellow box represents the inter-quartile range (25-75%). The central red line is the median value. The upper and lower whiskers represent the 10% and 90% points. The background of the graph divides the y axis into very good quality calls (green), calls of reasonable quality (orange), and calls of poor quality (red). All bases show high quality since the quality number (on y axis) is above 28 (Green background). Per Base GC Content is plotted for the GC content of each base position in a file. GC bias which is changing in different bases indicating an overrepresented sequences. The duplication levels plot is indicating the relative number of sequences with different degrees of duplication. DDY3 RNA-seq has substantial non-unique sequences since the duplication is almost 93%. The peak in the sequence length distribution plot shows length of the sequences in this case 100 bp. The per sequence quality score is showing quality of subset of the sequences in a FASTQ file. The X-axis denotes the quality score which is above 38, indicating good quality of sequences. B) Base quality score (sanger score) for each of the RNA-seq FASTQ file is generated by CASAVA 1.8 tool. Red boxes indicate inter-quartile range and the black line is the median value. All scores are above 28, indicating good quality of bases in each read generated.
The GenomeCoverageBed module from Bedtools was used to obtain a coverage after alignment for the genome file. Generally coverage reflects the average number of times a given region has been sequenced by independent reads. For Bedtools, the input genome file contained chromosome-specific data such as chromosomes and their lengths. The resulting BedGraph files were the outcome of genome-wide coverage and contained chromosome number, start and end chromosome coordinates for the coverage level, followed by the coverage level itself. All the commands for read alignment are listed in Table 4.2.

### Table 4.2 Commands for read alignment.

<table>
<thead>
<tr>
<th>Jobs</th>
<th>Commands</th>
</tr>
</thead>
<tbody>
<tr>
<td>To build W303 genomic_ref_index</td>
<td>Reference_strain_W303$ bowtie2-build -f W303_MPG_2012_ALAV00000000.fsa W303_genomic_ref_index</td>
</tr>
<tr>
<td>Bowtie2 alignment</td>
<td>$ bowtie2 --very-sensitive -k 1 --al output_file_name -x genomic_index_file_path -U fastq_file_path -S samoutput_filename.text</td>
</tr>
<tr>
<td>To convert SAM to BED with perl program</td>
<td>$sam_bed.pl Enter SAM file to select sequences: samoutput_file_name</td>
</tr>
<tr>
<td>To create crick strand file</td>
<td>grep +$ bed_filename &gt; crick_strand.bed</td>
</tr>
<tr>
<td>To create watson strand file</td>
<td>grep -v +$ bed_filename &gt; watson_strand.bed</td>
</tr>
<tr>
<td>BED to BEDGRAPH for crick strand</td>
<td>genomeCoverageBed -bg -i crick_strand.bed -g sequence_length_input_file_path &gt; crick_sorted.bedgraph</td>
</tr>
<tr>
<td>BED to BEDGRAPH for watson strand</td>
<td>genomeCoverageBed -bg -i watson_strand.bed -g sequence_length_input_file_path &gt; watson_sorted.bedgraph</td>
</tr>
</tbody>
</table>

### Visualization of RNA-seq results

Integrative Genomics Viewer (IGV) was used as a visualization tool for exploring RNA-seq data. After normalizing the read count for each strain (wild-type or mutant), transcription start sites as well as occurrence of extended transcripts in the mutant strains were visualized. RNA polymerase II (Pol II) transcribed genes within the vicinity of tRNA genes or predicted **ETC** sites
showed 3’ or 5’ extension of some of the transcripts or difference in their transcription level in mutants (DDY 4300 and DDY 4301) as compared to those in wild-type (DDY3 and DDY3630), indicating potential extra-transcriptional effects of Pol III complex binding sites on neighboring Pol II genes.

**Analysis of differential expression**

The preliminary results from IGV suggested changes in the expression levels of a subset of yeast genes. Therefore, to understand global effects of reduced TFIIC occupancy on the genome expression profile, we determined differentially expressed genes and their distribution in the genome. In order to detect extension of transcripts, we determined expression for regions upstream and downstream to the gene. To identify statistically significant differentially expressed transcriptional units in these specific regions, we categorized nearby regions of a gene into gene promoter (upstream) and gene end (downstream). Differential expression was statistically evaluated by DESeq software that was implemented in ‘R’ language for analyzing mapped short read counts obtained from RNA-seq. All commands for DESeq analysis are listed in Table 4.3.

A combined count data text file is the prerequisite for DESeq analysis (ANDERS and HUBER 2010) that also serves as a proxy for the magnitude of gene expression since transcripts of greater abundance in the cell usually have more reads generated from RNA libraries. A combined count data file was created by extracting count data from the mapped reads (SAM files) for each RNA sample and then combining count data for all samples in the gene expression study. Therefore, the resulting count data file contained contig names and number of reads from all the RNA samples (wild-type, and mutants) that were mapped to the respective contigs in the reference genome.
Table 4.3 Commands for DEseq analysis.

<table>
<thead>
<tr>
<th>Jobs</th>
<th>Commands</th>
</tr>
</thead>
</table>
| To install of DESeq within R | source ("http://www.bioconductor.org/biocLite.R")
| | biocLite ("DESeq") |
| To load the DESeq library | library ("DESeq") ← |
| To open counts file in a table with rows and header | > countsTable <- read.table("read_count_4_strains.txt",header=TRUE, row.names=1) |
| To check/read table | > head(countsTable) |
| To convert the table into R data frame | > design <- data.frame(row.names = colnames(countsTable), condition = c("wt","wt","mut","mut")) |
| To assign conditions | > conds <- factor(c("wt","wt","mut","mut")) |
| To create a countDataSet | > data <- newCountDataSet(countsTable,conds) |
| To estimate the size factors | > data <- estimateSizeFactors (data) |
| For normalizing counts | > head (counts(data, normalized=TRUE)) |
| To estimate variance/dispersion | > data <- estimateDispersions(data) |
| To plot mean normalized counts Vs dispersion with fitted curve | > plotDispEsts(data) |
| To run the negative binomial test | > results <- nbinomTest(data, "wt","mut") |
| To save results with mean counts and p-value | > write.table(results, file="DESeq_results.txt", sep="\t", row.names=rownames(results), col.names=colnames(results), quote=F) |
| To plot of normalized mean counts Vs log2 fold change (MA-plot) | > plotDE <- function(results) (plot(results$baseMean, results$log2FoldChange, log="x", pch=20, cex=1, col=if else(results$sval<.05, "red", "gray"))) > plotDE(results) |
| To plot histogram for distribution of p-values | > hist(results$sval, breaks=100, col="skyblue", border="slateblue", main = "") |
| To visualize differential expression using heat maps | > datablind <- estimateDispersions(data) > vsd <- getVarianceStabilizedData( data ) > install.packages("gplots") |
| For analyzing first 100 most differentially expressed genes | > select <- order(results$sval) [1:100] > colors <- colorRampPalette(c("yellow","red"))(100) > heatmap( vsd[select,], col = colors, scale = "none" ) |
The combined count data text file (tab delineated) was read in ‘R’ in the form of data table with defined rows and header. The table was then converted into ‘R’ dataframe and conditions (‘wt’ for wild-type and ‘mut’ for mutant) were assigned to each column containing number of reads (counts) mapped to the respective chromosomes. The resulting new count data set file was the central structure for DESeq analysis. The counts in different samples (wild-type or mutant) were normalized by estimating size factor for all the library sizes. As a preliminary check for quality of generated data, we performed estimateDispersions function that performs three steps. First, it estimates a dispersion value for each gene, then it fits a curve through the estimates. Finally, it assigns to each gene a dispersion value. The level of dispersion is related to the biological variation seen in each treatment. The variance between counts is the sum of two factors, firstly, the level of variation between replicates of each treatment, and secondly an uncertainty measure based on the concentration of the counts. If the dispersion or biological variation is more, the difference between counts from each treatments (here wild type and mutant) should be bigger to be considered as a significant difference. Figure 4.3 part-A represents the relationship between the level of dispersion per gene and the mean of normalized counts for each gene.

Many studies have shown that the variance grows faster than the mean in RNA-seq data. This is known as ‘overdispersion’ (ANDERS and HUBER 2010). To overcome this, DESeq analysis comprises negative binomial distribution, which also accounts for flexible data-driven relationships between mean and variance to generate more balanced and accurate results. Unlike Poisson distribution, negative binomial distribution reflects both technical and biological variability by considering biological sampling variance correctly, which removes potential selection biases in the ‘hit list’ of differentially expressed genes.
Figure 4.3. Representation of DESeq analysis.

A) Scatter plot of dispersion of mean normalized counts (reads) shows dispersion values (black dots) and fitted values (red line). More dispersion value (variance) requires vast difference between expression of a gene in wild type and mutant to be called as a significant differential expression. B) MA-plot shows log2 fold change in the expression versus mean normalized counts for both wild type and reduced TFIIC occupancy conditions. The red color marks genes detected as differentially expressed at 5% false rate (FDR) based on negative binomial distribution. C) Histogram of p-values based on the negative binomial test represent distribution of p-values for the data. Differentially expressed genes contain extremely low p-values, otherwise values are spread within the range of zero to one. The p values from genes with very low counts, which take discrete values are accumulated at right side of the histogram. D) Heatmap is showing differential expression for 25 most highly expressed genes. At top right of the panel color key indicates the values (in terms of color) given for counts (reads) mapped to a particular gene. Yellow to red= low to high.
To verify differential expression between wild type and mutant conditions we performed negative binomial test that generated a dataframe with mean normalized counts for individual conditions (wild-type and mutant), p-values, difference in gene expression (fold change). The logarithmic values for differences in gene expression (log2fold change) are plotted against the mean normalized counts in Figure 4.3 part-B. Red dots represents differentially expressed genes that are significant at 5% false discovery rate (FDR). Distribution of p-values are represented as histograms (Figure 4.3 part-C). Differentially expressed genes contained low p-values, while for other genes p-values were spread uniformly over the range from zero to one. Very high p-values were assigned to genes with very low counts (at the right side of the histogram, in Figure 4.3 part-C).

The final quality assessment was performed by clustering and visualizing the DESeq data in the form of ‘heatmaps’ that verified accuracy of analysis for differentially expressed genes. The 25 most highly differentially expressed genes are shown in the following heatmap (Figure 4.3 part-D). In the heatmap, comparison of expression levels between wild-type and mutants for each gene or the surrounding region revealed significant differences.

Combining the results from IGV and DESeq analysis, we selected 9 differentially expressed (p-value < 0.5) representative genes that showed an extra-transcriptional effect of adjacent tRNA genes or potential ETC sites. Figure 4.4 represents the snapshots of these genes with the observed extra-transcriptional effects. We anticipated that the 5’ extension of genes under reduced TFIIC occupancy could be due to readthrough (in case of NUP2) as we observed at the ATG31-tV(UAC)D locus (Korde et al. 2014) or due to alternate transcription start sites (for TRM12, FAR3 and PMT7) that indicated the influence of nearby tRNA gene in selecting transcriptional start position.
Figure 4.4. Visualization of potential extra-transcriptional effects on representative genes. Based on the results from DESeq analysis and manual visualization by IGV tool, above genes were considered as representative genes. Each panel represents RNA-seq data from wild type DDY3 (gold color) and TFC6 mutant DDY 4300 (purple color). The location of each target gene is indicated by black bars and its name is mentioned at the bottom of the panel. Red and orange bars indicate position of nearest tRNA gene or potential ETC site, respectively. Some of the top differentially expressed genes in TFC6 mutant showed 5’ extension (A), 3’ extension (B) or increase in the mRNA level (C).

Whereas, the readthrough observed in TFC6 under-expressing mutants hinted at the possibility of Pol II transcription blocking function of adjacent tRNA gene. As compared to 5’ extension, 3’
extension of genes (such as PCL5 and DGF5) was found to be a rare scenario. Highly de-repressed genes such as SPO74, SRL4 and ARG8 under low TFIIIC occupancy suggested potential tRNA position effect from neighboring tRNA genes or from potential ETC sites (in case of SLR4 and ARG8). However, ARG8 and SRL4 are under the influence of Gcn4 regulation (CONESA et al. 2005) which likely contribute to the increased levels of mRNA under reduced TFIIIC binding condition. In addition to 5’ extension, PMT7 showed slight decrease in the mRNA level in mutants indicating positive effect of adjacent tRNA gene on the gene transcription.

The results obtained from the computational analysis were considered as suggestive of extra-transcriptional effects, and to validate these results individual experiments were designed for each representative locus. Previous studies have demonstrated that B-box mutation of tRNA genes or ETC sites compromises TFIIIC binding and hence abolishes extra-transcriptional effects on neighboring Pol II transcribed gene (DONZE and KAMAKAKA 2001; SIMMS et al. 2008; KLEINSCHMIDT et al. 2011; KORDE et al. 2014). Therefore to verify the extra-transcriptional effects of tRNA genes or predicted ETC sites on nearby representative Pol II transcribed genes, we mutated individual B-box sites. Considering the requirement of B-box for the activity of tRNA gene or ETC site, we anticipated that mutation in these regions would recreate the effects observed in the Tfc6p under-expressing mutant. The future experiments will involve following steps to validate results from RNAseq analysis. Mutation in tRNA gene or ETC sites at targeted loci (representative gene).

- Extraction of total RNA from mutants and their respective parent strains (controls).
- cDNA synthesis.
- Quantitative RT PCR or Northern analysis to confirm observed effects.
- Western analysis for detecting altered proteins from genes that showed extended transcripts.
CHAPTER 5
DISCUSSION

Genes are regulated at every step of their expression pathways from DNA to protein. Extensive research on gene regulation has revealed mechanisms much more complex than previously anticipated. In eukaryotes, genes within chromatin need to be available for various nuclear processes such as DNA replication, DNA repair or transcription. The orchestral work performed by several regulatory mechanisms effectively control gene expression. Chromatin associated events such as histone modifications and chromatin remodeling aid access the regulatory proteins to DNA that assists the transcription machinery to initiate gene expression. For most genes, the control of transcription is paramount since it is the most obvious, efficient and cost-effective stage of regulation for cells. In eukaryotes, transcription is mediated by three RNA polymerases, which are recruited near the transcription start site of the target gene by RNA polymerase-specific transcription factors. These transcription factors recognize conserved promoter regions to promote the assembly of pre-initiation complexes (PIC). RNA polymerase III (Pol III) requires sequential assembly of pre-initiation complexes for transcribing 5S rRNA, tRNA and other non-coding genes. Unlike a protein coding gene, a tRNA gene contains internal promoter sequences (A-box and B-box) which are the binding sites for Pol III specific transcription factor TFIIIC. Once bound to DNA, it assembles the initiation factor TFIIIB complex, which is responsible for targeting Pol III to the transcriptional start site, and then transcription initiates. These motifs within and upstream of a tRNA gene affect its transcription efficiency and function. Chromatin immunoprecipitation followed by microarray hybridization studies have strengthened the notion that the multimeric Pol III transcription complexes are persistently occupied on all yeast tRNA genes during active growth. Such a tight occupancy state at tRNA genes, which are
interspersed throughout all chromosomes, has a potential to exert a genome-wide influence on neighboring loci. Such effects have been found to be mostly independent of the transcription process, and are referred to as extra-transcriptional (described in chapter 1- Introduction) (CONEsA et al. 2005; SIMMS et al. 2008; DONZE 2012). In addition to Pol III transcribed genes such as tRNA genes, other chromosomal loci including ETC (extra-TFIIC) sites are found to be occupied by partial Pol III transcription machinery which also have similar influences on nearby chromatin or gene activities. This study mainly focused on the extra-transcriptional functions of Pol III transcriptional complexes.

Boundary elements organize eukaryotic chromatin into functionally distinct domains, euchromatin and heterochromatin, by preserving their structural integrity, and by preventing regulatory cross talk between different domains (CAPELSON and CORCES 2004; GASZNER and FELSENFELD 2006; VALENZUELA and KAMAKAKA 2006; LUNYAK 2008). Based on their function, boundary elements are typically characterized by two fundamental properties: (i) the ability to protect from chromosomal position effects by acting as barriers against the self-propagation of repressive chromatin (KELLUM and SCHEDL 1991; GDULA et al. 1996; DONZE and KAMAKAKA 2001; SIMMS et al. 2004; OKI and KAMAKAKA 2005) and (ii) the ability to insulate or block regulatory interactions between distal enhancers and proximal gene promoters (UDVARDY et al. 1985; RECILLAS-TARGA et al. 2002; SIMMS et al. 2008; RAAB and KAMAKAKA 2010; KLEINSCHMIDT et al. 2011). Some boundary elements are able to act both as chromatin barriers and enhancer blocking insulators (NOMA et al. 2001; LUNYAK et al. 2007).

In eukaryotes, three different mechanisms for boundary activity have been uncovered. First, fixed boundary elements consist of specific DNA sequences and their associated proteins, which establish boundaries with well-defined positions (UDVARDY et al. 1985; KELLUM and
Schedl 1991; Kellum and Schedl 1992; Chung et al. 1993; Piakaart et al. 1998). Second, there are variable boundary elements that do not occupy specific DNA sequences or genomic locations but maintain chromatin domains through opposing action of chromatin modifying enzyme complexes (Fourel et al. 2004; Kimura and Horikoshi 2004; Oki and Kamakaka 2005). Third, boundary activity can be observed at non-protein-coding transcriptional units which are bound by transcription factors, such as tRNA genes in yeast (Donze and Kamakaka 2001; Oki and Kamakaka 2005; Scott et al. 2006; Valenzuela and Kamakaka 2006).

In yeasts, Saccharomyces cerevisiae and Schizosaccharomyces pombe, tRNA genes can function as chromatin boundaries that restrict the spread of heterochromatin silencing (Donze et al. 1999; Donze and Kamakaka 2001; Noma et al. 2006; Scott et al. 2006) and mediate enhancer-blocking insulator activity in S. cerevisiae (Simms et al. 2008). However recruitment of the entire Pol III complex is not always necessary for boundary activity. TFIIIC bound ETC loci also act as boundaries to both repressed chromatin and enhancer activation (Simms et al. 2004; Valenzuela and Kamakaka 2006; Simms et al. 2008; Valenzuela et al. 2009). Stable occupancy by transcription factors is essential for efficient boundary function (Donze and Kamakaka 2001; Simms et al. 2004; Oki and Kamakaka 2005; Noma et al. 2006; Valenzuela and Kamakaka 2006; Simms et al. 2008). Recruitment of chromatin remodeling complexes can mediate nucleosome displacements, which enable binding of transcription factors at promoters.

**Nucleosome positioning and rearrangements**

Nucleosomes have the ability to block biologically relevant transcription factor binding sites, and action of chromatin remodeling factors transiently allows access to such sites (Li et al. 2007; Williams and Tyler 2007). Genome-wide studies found that nucleosome density at promoter regions is typically lower than that in the coding region (Bernstein et al. 2002; Lee et
Therefore, it seems that eukaryotic cells tend to position sequence-specific transcription factor binding sites within accessible regions. These regions are flanked on both sides by positioned nucleosomes (Yuan et al. 2005). A nucleosome-depleted region can by itself block the spread of silencing (Bi et al. 2004), suggesting that the creation of nucleosome free region may be an important step in boundary function. Transfer RNA genes and even ETC loci contain well positioned nucleosomes at their 5’ and 3’ flanking regions (Nagarajavel et al. 2013), with the site occupied by the Pol III complex being nucleosome free.

Nucleosomes can be displaced from promoter DNA by promoter-binding transcription factors, in combination with ATP-dependent nucleosome remodeling complexes and histone chaperones. Most remodeling factors can move nucleosomes along DNA by nucleosome sliding (Clapier and Cairns 2009). Nucleosome eviction is dependent on activity of chromatin remodelers which are required for TFIIIC binding and boundary activity (Valenzuela et al. 2009). RSC chromatin remodeler localizes to tRNA genes (Ng et al. 2002) where it evicts histones (Parnell et al. 2008) and our previous studies have shown that mutants in RSC complex affect tRNA-mediated heterochromatin barrier function (Jambunathan et al. 2005).

Our results indicate that the efficiency of blocking intergenic transcription involved binding of other accessory proteins which might have increased the stability of transcription factors or created a conducive environment by altering chromatin. Genetic analysis of heterochromatin barrier function revealed involvement of chromatin-associated proteins including chromatin remodelers, regulatory DNA-binding proteins and chromosomal structural proteins (Donze and Kamakaka 2001; Jambunathan et al. 2005). Our study demonstrated that loss of NHP6 and mutation in the some of the subunits of condensin complex (smc 3-1 and scc D730V) reduced intergenic transcription blocking activity at tV(UAC)D gene whereas mutation in genes encoding
other chromatin or DNA binding complexes such as SAS2, HTZ1, YTA7, BDF1, RSC2, RPD3 showed weak or no effect on interference blocking activity at tV(UAC)D.

Nhp6p provides transcriptional initiation fidelity to tRNA genes since its binding distinctively alters the interaction of TFIIC subunits with the A-box promoter element that provides more confined DNA placement of TFIIB and, as a further consequence, in the placement of Pol III (KASSAVETIS and STEINER 2006). Chromatin immunoprecipitation (ChIP) experiments showed reduced TBP binding at the Pol III transcribed-SNR6 gene, in nhp6a nhp6b mutant (ERIKSSON et al. 2004). Mutation in NHP6 genes have caused changes in transcriptional start sites at tRNA genes (KASSAVETIS and STEINER 2006) or compromised heterochromatin barrier function of tRNA gene (BRAGLIA et al. 2007).

The Smc proteins (KOSHLAND and STRUNNIKOV 1996) are a family of proteins required for chromosome condensation and cohesion. They associate with Sister chromatid cohesion (Scc) proteins to form heterodimeric complexes. Mutations in these chromosome structural complexes significantly affect boundary function (DONZE et al. 1999; MENEGHINI et al. 2003). Recruitment of cohesin complexes along with tRNA gene barrier were required for cell cycle progression, which was needed establishment of silencing at yeast HMR locus (LAZARUS and HOLMES 2011). Genome wide analysis also showed close association of cohesin complexes with TFIIC (D'AMBROSIO et al. 2008). Recent studies have further implicated the role of Pol III machinery in the recruitment of the condensin class of chromosome binding and organizing proteins. Multiple studies have shown by chromatin immunoprecipitation that the condensin subunits associate with Pol III genes and TFIIC-only binding sites via direct interaction between condensin subunits and TFIIC (D'AMBROSIO et al. 2008; HAEUSLER et al. 2008). Considering the close association of Nhp6 and Smc proteins with Pol III transcription factors, loss or mutation of these genes, might
have disturbed interaction and thereby stability of transcription factors at the \( tV(UAC)D \) tRNA gene.

Manual inspection of Pol III transcription associated proteins on *Saccharomyces* Genome Database, revealed binding sites for many other Pol II–specific transcription regulators including chromatin modulators such as Bur6, Spt6, Spt16, Hbt1, Isw2 and Rsc9 near \( tV(UAC)D \) (chapter 2) and Bur6, Reb1, Rsc9 and Rap1 at the *ETC4* (chapter 3) normal chromosomal locus (Figure 5.1). Presence of binding sites for chromatin remodelers and strong enrichment of Spt6p [at \( tV(UAC)D \)] and Rap1p (at *ETC4*) indicated close association of nucleosome re-arrangements and boundary function at these loci. Given that we detected weak readthrough transcription of *SUT467-ATG31* in *rsc2Δ* strains, and that binding of another component of RSC complex –Rsc9 has been detected at \( tV(UAC)D \) gene (SAHA *et al.* 2006). These results suggest potential RSC complex involvement in nucleosome eviction that may affect intergenic transcription blocking activity.

The nucleosome forms a strong barrier to Pol II transcription *in vitro*, but yeast Pol III can transcribe through nucleosomes by mobilizing histones along the templates (CLARK and FELSENFELD 1992; STUDITSKY *et al.* 1994; STUDITSKY *et al.* 1995; STUDITSKY *et al.* 1997). TFIIS, an elongation factor for Pol II, facilitates transcription through templates containing assembled nucleosomes (KULISH and STRUHL 2001; KIREEVA *et al.* 2005). Restoration of nucleosomes by re-depositing histones back onto transcribed regions is required for masking cryptic promoters which can be recognized by transcription factors and lead to the generation of cryptic transcripts initiated from internal start sites within the body of yeast genes that may cause transcriptional interference (KAPLAN *et al.* 2003; MASON and STRUHL 2003; SCHWABISH and STRUHL 2004; SCHWABISH and STRUHL 2006).
Elongating RNA polymerases can cause the direct or in *cis* suppression of another transcriptional process, which is referred to as transcriptional interference (PALMER *et al.* 2011). Recent genomic expression studies in a number of organisms have confirmed the existence of pervasive transcripts (JOHNSON *et al.* 2005; DAVID *et al.* 2006; STEINMETZ *et al.* 2006; XU *et al.* 2009). These intergenic transcripts have been found to contribute to transcriptional interference, which recently has been recognized as a potentially widespread mechanism of gene regulation in organisms from viruses to microbes to metazoans. Regulation of sense and antisense transcripts of
IME4 gene is the classic example of cell specific regulation by transcriptional interference in *S. cerevisiae*. In addition to this, importance of intergenic transcription is detected in zinc homeostasis (Bird et al. 2006), entry into meiosis (Hongay et al. 2006) and variegated FLO11 expression (Bumgarner et al. 2009). In *S. cerevisiae*, bidirectional transcription from strong promoters also generate cryptic intergenic transcripts, such as CUTs, SUTs, MUTs, XUTs and NUTs, and these RNA molecules are rapidly degraded by the nuclear exosome or cytoplasmic degradation pathways under normal conditions. (Wyers et al. 2005; Xu et al. 2009; Lardenois et al. 2011; van Dijk et al. 2011; Schulz et al. 2013). SUTs are targeted by both nuclear (Rrp6) and cytoplasmic RNA decay pathways (5′ to 3′ decay by Xrn1). However, either of the two pathways degrade a distinct fraction of the individual SUTs, or the transcripts have relatively long half-lives, because transcripts are detected even when both pathways are functional (Marquardt et al. 2011). Therefore, transcription of SUTs can potentially exert interference to other transcription units, as evidenced in our study. SUT467 readthrough inhibited translation of ATG31 and reduced Atg31p affected fitness of cells under nitrogen starvation (Chapter 2).

Hence, in spite of regulatory roles for transcriptional interference that have been found across a variety of biological processes, if left unchecked, progression of noncoding transcription can have negative consequences on neighboring gene expression. Palmer et al., proposed two different mechanisms for transcription interference. First, interfering RNA polymerases prevent binding of transcription factors at a promoter by occlusion. Second, the promoter whose activity is dependent on the binding of transcription factors can be substantially inhibited by the dislodgement of the transcription factors due to the transcription activity from a neighboring strong promoter. It is feasible that slow-binding transcription factors might be effectively removed via frequent dislodgement by Pol II arriving from strong interfering promoter. Strong transcriptional
interference could also result when polymerases traverse a promoter which relies heavily on pre-initiation complex (PIC)-assisted re-initiation, where a complex of transcription factors stably associated throughout multiple rounds of transcription initiation (PalméR et al. 2011). This could be the case in our studies at tV(UAC)D, since SES1 UAS is a strong regulatory element that generates SUT467, and transcription elongation by Pol II appears to dislodge weakly bound TFIIIC and TFIIIB complexes in the mutants, allowing the unchecked Pol II progression to interfere with the downstream ATG31 promoter.

**Protein-protein interactions between Pol III transcription factors**

The previous barrier activities at cryptic mating loci suggested that a key step in the formation of a heterochromatin barrier is the stable binding of TFIIIC. The binding of TFIIIC was necessary for boundary function but subsequent binding of TFIIIB likely improved the probability of complex formation at the boundary (Simms et al. 2008). Boundary function appears to be critically dependent on the ability of factors to bind stably to their target sites on chromatin. Flanking sequences and chromatin bound factors near such promoters enhance the stability of transcription factor complexes, and thereby contribute to efficient boundary activity. (Simms et al. 2008; Valenzuela et al. 2009).

Binding to A-box and B-box promoter elements of tRNA gene is primarily mediated by subunits Tfc1p and Tfc3p (Kassavetis et al. 1989; Kassavetis et al. 1990; Bartholomew et al. 1991). Therefore, binding of TFIIIC complex was inhibited in tV(UAC)DA- and B-box mutants. Brf1p initiates TFIIIB assembly by means of interaction between its N-terminal region and Tfc4p of TFIIIC complex (Bartholomew et al. 1991; Khoo et al. 1994; Chaussivert et al. 1995; Moir et al. 1997). At certain promoters, Brf1p assures effective assembly, while the TBP subunit selects the precise DNA binding site (Joazeiro et al. 1996). The C-terminal domain of Brf1p
constitutes the binding site for the conserved SANT domain of Bdp1p, the third subunit of TFIIIB complex (GEIDUSCHEK and KASSAVETIS 2001). Bdp1p, interacts with Tfc8p, which further connects TBP and favors binding of TBP to DNA. Once bound to DNA, TFIIIB makes an extremely stable nucleoprotein complex (ANDRAU et al. 1999). TBP and Bdp1p mediate bending of DNA and protein-protein interactions stabilize TFIIIB-DNA complex due to steric obstructions to DNA escape and also confer resistance to dissociation by simple electrolytes and polyelectrolytes. (GROVE et al. 1999) C-terminal domain of Brf1p can simultaneously embrace TBP and form a platform able to interact with the Pol III-specific subunit C34 (FERNANDEZ-TORNERO et al. 2010). All three subunits of TFIIIB are required for Pol III recruitment, but direct interactions have only been identified in the case of Brf1p (BARTHOLOMEW et al. 1993; WERNER et al. 1993; KHOO et al. 1994). Hence, multiple interaction between subunits of TFIIIC and TFIIIB ensures assembly of Pol III complexes in which Brf1 plays an integral role.

Mutations in subunits of TFIIIC and TFIIIB complexes impair heterochromatin barrier function at HMR locus (DONZE and KAMAKAKA 2001). The temperature sensitive tfc3(G349E) mutant containing transition substitution of G to A resulting in Gly to Glu at amino acid position 349 exhibits significantly lower DNA binding affinity of the TFIIIC complex (LEFEBVRE et al. 1994; ARREBOLA et al. 1998). Mutations in C-terminal region of Brf1p (brf1 II.6 or brf1 II.9) reduce interaction with TBP, and therefore primarily affect the formation of TFIIIB-DNA complex (ANDRAU et al. 1999). Pol III subunit mutants (rpc31 and rpc160) are defective in Pol III binding and initiation of tRNA transcription (THUILLIER et al. 1995) but do not affect assembly of TFIIIC and TFIIIB.

At tV(UAC)D, Pol III complexes containing tfc3 and brf1 mutants could not block cryptic intergenic Pol II transcription, likely due to the defects described above. In the cases of A- and B-
mutants, mutations in the promoter region compromised binding of TFIIIC complex at the tRNA gene. These weak interactions might have diminished the assembly of Pol III transcription factors or might have at least increased their assembly time which could have impaired the formation of stable Pol III transcription complexes on DNA. However, mutation in Pol III enzyme subunits primarily affected transcription initiation at tRNA gene (Thuillier et al. 1995) but did not necessarily alter the assembly of TFIIIC and TFIIIB onto DNA. Therefore, in Pol III enzyme subunit mutants, a stable complex of transcription factors was expected to form at tRNA gene, which could resist dislodgement by cryptic Pol II transcription initiated from the upstream strong promoter (SES1 UAS).

When the tDNA was replaced with ETC sequences in order to determine minimal factor requirement for transcription-blocking mechanism, TFIIIC bound at an ectopic ETCA sequence was not sufficient for preventing readthrough, indicating an inability of blocking transcriptional interference by TFIIIC alone. Replacement of tRNA gene with ETCA where both TFIIIC and TFIIIB were bound, significantly blocked readthrough as compared to only TFIIIC bound ETCA. This indicated higher stability of TFIIIC-TFIIIB-DNA complex as compared to TFIIIC-DNA stability at ETC sites. However, as compared to complete Pol III transcription machinery bound at tRNA gene, TFIIIC-TFIIIB bound ETCA sequences were not fully efficient in blocking readthrough since northern analysis detected low levels of extended mRNA as a result of SUT467-ATG31 readthrough. Based on the Palmer et al model for transcriptional interference and results described in Chapter 2, a model for blocking of intergenic transcription at tV(UAC)D tRNA gene is represented in Figure 5.2.
Figure 5.2. Model for blocking intergenic transcriptional interference at tV(UAC)D locus.
A) In wild-type, stable occupancy of Pol III transcription complexes blocks readthrough from SUT467, producing normal ATG31 transcript (black solid arrow). B) Mutation in Tfc3 subunit of TFIIIC complex diminishes assembly of TFIIIC and TFIIIB on tRNA gene, allowing readthrough (solid red arrow). Occasional assembly may block readthrough by some polymerases, producing few normal ATG31 transcripts (dotted black arrow). Weakened assembly also allows the appearance of an intermediate transcript (Green arrow) initiating upstream of tRNA gene. C) Mutation in Brf1 subunit of TFIIIB complex affects TFIIIB assembly onto tRNA gene, allowing readthrough (solid red arrow). Occasional assembly may block readthrough, producing few normal ATG31 transcripts (dotted black arrow). Initiation of intermediate transcripts (Green arrow) also occurs in this situation. D) Mutation in Pol III subunits inhibits recruitment of the Pol III enzyme complex, but does not affect TFIIIC-TFIIIB-DNA binding, blocking readthrough and only allowing ATG31 normal transcripts. E) B-box mutation compromises binding of TFIIIC complex, thereby failure to assemble TFIIIB and Pol III, producing readthrough (solid red arrow). F) Weak TFIIIC binding at replaced ETC4 sequence due to absence of box A, producing readthrough (solid red arrow). G) At replaced ETC9 sequence, binding of TFIIIB stabilizes TFIIIC-TFIIIB complexes on tRNA gene, hence significantly blocking readthrough (dotted red arrow) and producing normal ATG31 transcripts.
Global effects of bound Pol III transcription complexes

RNA-seq analysis (Chapter 4) on wild-type and TFC6 under-expressing mutants that affected binding of TFIIC at all potential loci (KLEINSCHMIDT et al. 2011), have revealed a global picture of extra-transcriptional effects exerted by TFIIC bound loci. DESeq data have reported many differentially expressed chromosomal loci including Pol II-transcribed genes adjacent to tRNA genes or predicted ETC sites. The analysis was based on stringent statistical significance value (0.05) which potentially have eliminated false positive results. The first 100 most differentially expressed genes in the mutants were categorized based on FunCat annotation scheme, available at MIPS website (http://mips.helmholtz-muenchen.de/genre/proj/yeast/). Figure 5.3 represents the functional catalogue of differentially expressed genes. Reduced TFIIC binding due to TFC6 under-expression in mutants brings about a significant enrichment in transcripts of genes whose products are involved in central metabolism, cell cycle associated activities and transcription, to a lesser extent in interaction with the cellular environment. The functional categories associated with the differentially expressed genes from mutants were consistent with the results obtained by Conesa et al. while studying differential expression of genes under defective Pol III transcription (CONESA et al. 2005). Since the recognition and binding of TFIIC at specific promoter sequences is the initial and necessary step in the transcription of all Pol III-transcribed genes, reduced TFIIC occupancy likely decreased Pol III transcription globally. The functional categories indicate an interplay between components of Pol III transcription complexes and the regulation of cellular metabolism. Previous studies on a genomic scale have shown that in Pol III transcription mutants, many Gcn4p regulated genes are activated. Reduced transcription of initiator methionine tRNA induces GCN4 translation generating elevated levels of Gcn4p to activate transcription of target genes (Conesa et al., 2005). We speculate that similar mechanisms
are at play here, since our analysis did not show change in the transcription levels of *GCN4* when *TFC6* under-expressed mutants were compared with wild-types but the level of transcripts increased for the genes which are known to be under control of Gcn4p. After comparing with previously known Gcn4p regulated genes, we have found 11 genes (out of first 100) that were differentially expressed but under the control of Gcn4p induction. After importing our RNA-seq data into the Integrated Genome Viewer, manual inspection of differentially expressed genes revealed that 32 of the 100 top mis-regulated genes were in the vicinity of a tRNA gene. (Figure 5.4). However, not all tRNA gene proximal Pol II genes showed altered transcription effects. The plausible explanation would be influence of factors such as stability of core transcription factors and binding of chromatin associated proteins.

Figure 5.3. Functional categories of differentially expressed genes. The first 100 the most differentially expressed genes (reported by DESeq analysis) were categorized based on FunCat annotation scheme, available at MIPS website: (http://mips.helmholtz-muenchen.de/genre/proj/yeast/) Numbers in black represent count of genes in each category.
Figure 5.4. Categorization of mis-expressed Pol II-transcribed loci. First 100 differentially expressed loci were categorized based on presence of Pol III factors assembly sites such as tRNA gene or predicted ETC site in their vicinity. Further, taking into account of observed aberrant transcription changes, Pol to loci near tRNA gene or ETC sites are separated. In this pie chart, numbers represent count of genes.

The degree of occupancy of all the genomic tRNA genes by TFIIIC, TFIIIB and Pol III is variable across the genome (Kurjan et al. 1980; Moqtaderi and Struhl 2004; Moqtaderi et al. 2010; Oler et al. 2010), at least at the level of resolution of ChIP assays. Coding sequences outside the A- and B-box promoter elements as well as non-coding sequences flanking tRNA genes positively or negatively influence the extent of tRNA gene transcription (Raymond and Johnson 1987). DNA sequences upstream and downstream of various tRNA genes have been shown to affect the stability of binding of TFIIIC and TFIIIB (Sprague et al. 1980; Raymond et al. 1985; Joazeiro et al. 1996; Donze and Kamakaka 2001). It was found that 5’ flanking sequences surrounding the transcriptional start site of tRNA gene play a specific role in both determination of transcriptional efficiency and in fixing the precise site of initiation, since they provide a platform for TFIIIB binding (Fruscoloni et al. 1995). Although, TFIIIB does not require specific sequences for binding, recognizable sequence patterns underlying the binding region of TFIIIB occur upstream of tRNA genes in many eukaryotic genomes (Giuliodori et al. 2003), and these patterns may influence TFIIIB binding stability. Some tRNA genes efficiently recruit TFIIIC and
TFIIB at their normal sites in euchromatin, but when these genes are moved to the silenced HMR domain, they are unable to block the spread of silencing emanating from the HMR-E silencer.

Therefore, it is possible that in addition to the internal promoter sequences that are recognized by TFIIC, flanking sequences adjacent to TFIIC binding sites play a role in determining which TFIIC-bound promoter elements can function as chromatin boundaries, and flanking sequences may also affect the ability to block Pol II progression. The spacing between A- and B- promoter boxes differs in different tRNA genes. The nature of the complexes could be changed by having different geometries or stability of binding by one or more components. The topology or strength of the TFIIC-DNA interaction might subtly alter its interaction with other binding partners such as condensin (D’AMBROSIO et al. 2008; HAEUSLER et al. 2008). This indicates that even though all tRNA genes have conserved promoters which are recognized by core transcription factors, they possess different abilities to exert extra-transcriptional effects. This was consistent with the results obtained from tV(UAC)D locus, supporting the importance of overall chromatin environment for the extra-transcriptional activity of Pol III transcription complexes bound sites.

Consistent with the orientation-independent barrier function at HMR tRNA THR gene (DONZE and KAMAKAKA 2001) and at ETC9 locus (VALENZUELA et al. 2009), change in orientation of tV(UAC)D gene or of ectopic ETC9 did not affect the interference blocking ability. While investigating the blocking of transcriptional interference, the appearance of an intermediate length transcript in brfl and TFC6 promoter mutants and weakly in tfc3 mutant suggested the possibility of recruitment of Pol II at tV(UAC)D gene due to the weak assembly of Pol III transcription complexes. Interestingly, manual inspection of the tRNA gene upstream sequences revealed possible start sites for Pol II transcription ~10 bps upstream of tRNA start site that closely
mapped with 5’ ends of the intermediate transcript (Integrated Genome viewer-IGV and our 5’RACE analysis in (Chapter 2). During assembly of TFIIIB, subunits of TFIIIC complex reorient and possibly facilitate stable binding of TFIIIB at the upstream of tRNA gene. In intermediate transcript producing mutants, mutation in \textit{tfc3} or \textit{brf1} might have affected the assembly of TFIIIB which in turn have prevented recruitment of Pol III. Moreover, instability of binding and retention of the TFIIIB complex could have occasionally exposed cryptic binding sites for Pol II transcription which otherwise are hidden under assembled Pol III complexes. The N-terminus of Brf1p, which is not affected in either of the mutants is a paralog of TFIIIB, a Pol II transcription factor (BUSHNELL \textit{et al.} 2004). During Pol II driven transcription, TFIIIB binds to TBP within the TFIID complex and recruits Pol II at the transcriptional start site, which might have been the case in the production of intermediate length transcript. Another possibility that we speculated at \textit{tV(UAC)D} was C-terminus mutation in Brf1p caused inefficient interaction with TBP but possibly allowed interaction of TBP with DNA upstream of tRNA gene. This bound TBP might have recruited Pol II-specific factors such as TFIID followed by Pol II recruitment.

Another candidate mechanism for recruiting Pol II at the intermediate start site could be TFIIIC and its associated proteins, since at the IR insulator in \textit{S. pombe}, TFIIIC binding sites do not recruit Pol III, but instead Pol II appeared to be associated with TFIIIC (NOMA \textit{et al.} 2006). Even though Pol II recruitment at TFIIIC binding sites has not been demonstrated in \textit{S. cerevisiae}, interaction of Pol II by TFIIIC associated proteins as occurs in \textit{S. pombe} cannot be completely eliminated at \textit{tV(UAC)D} locus.

**Future directions**

Blocking of intergenic transcription by \textit{tV(UAC)D} tRNA gene required occupancy of Pol III transcription factors. The minimal factor requirement for this boundary function was the stable
assembly of TFIIIB. Previous studies have revealed that TFIIIC alone can function as a boundary element and binding of TFIIIC to DNA results in conformational changes that facilitate assembly of TFIIIB. Association of TFIIIB with DNA is much stronger than TFIIIC-DNA binding, since mild salt or heparin treatment can easily dissociate TFIIIC but cannot displace fully assembled TFIIIB. *In vitro* studies for Pol III transcription re-initiation have reported that once bound to DNA, extremely stable TFIIIB complexes can direct multiple rounds of Pol III transcription, even in absence of TFIIIC (Dieci *et al.* 2013). Considering the stability of TFIIIB-DNA complex, it is valid to ask a question if TFIIIB alone can create an efficient boundary to block cryptic transcription at \( tV(UAC)D \). There are two approaches to obtain assembly of TFIIIB in absence of TFIIIC. First, *in vitro*, TFIIIB complex can be assembled by adding purified TFIIIC complex and treatment with heparin can remove bound TFIIIC but not assembled TFIIIB. Second, the difficulty in assembling TFIIIB complex *in vivo* in absence of TFIIIC can be overcome by placing putative TATA sequence which can be recognized by the TBP subunit of TFIIIB. The results obtained by using these approaches would need further validation since both approaches might alter the natural chromosomal environment at \( tV(UAC)D \) gene.

Blocking of intergenic transcription did not seem to be severely affected by loss of chromatin-associated proteins previously implicated in heterochromatin barrier function. Therefore, it would be informative to experimentally test the involvement of individually detected \( tV(UAC)D \) gene- associated chromatin remodelers for their potential roles in the blocking of intergenic transcription.

It is evident that stability of Pol III transcription complexes is essential for boundary activity and is influenced by many factors, including sequences flanking the gene promoter and binding of chromatin bound proteins, which modulate nucleosome positioning and facilitate steady
binding of transcription factors. Moreover, it is possible that different types of boundary activities also require different extents of stability and associated chromosomal environment. It will be interesting to compare the \textit{HMR} tDNA\textsuperscript{Thr} and the \textit{ATG31} upstream tDNA to each other in both heterochromatin barrier and Pol II blocking activity assays.

Based on genome-wide detection of transcription factor binding sites by Venters \textit{et al.} (2011) a Rap1p binding site coincides with the \textit{ETC4} locus. This suggests that deletion of \textit{ETC4} might have altered binding of Rap1p, which is actively involved in transcriptional activation, repression as well as nucleosome displacement. Therefore, there is a possibility that loss of Rap1p binding might have affected the transcription of \textit{RAD2} in our \textit{etc4} mutant strains (Chapter 3). It would be necessary to determine the role of Rap1p in \textit{RAD2} transcription by examining mRNA levels under conditions of compromised \textit{RAP1} function. As described in chapter 3, \textit{ETC4} might have a role in the proper translational regulation of \textit{RAD2}, therefore future analysis of Rad2p levels may help explain the increase in UV sensitivity in our \textit{etc4} mutants.

The effects of nearby tRNA genes on differentially expressed loci (obtained from DESeq analysis) can be confirmed by molecular experiments that will involve construction of tRNA \textit{B-box} mutants and detection of transcript levels by qRT-PCR or northern analysis in the mutants. Analysis of such RNA levels in \textit{brf1} and \textit{tfc3} mutants using gene specific probes should determine a minimal transcription factor requirement for each representative altered characteristic.

Knowledge about the regulation of gene activity by physical landmarks such as insulators and barriers is an expanding area of research. This study strengthens the notion that tRNA genes and \textit{ETC} sites play integral transcription-independent roles in the regulation of many Pol II-transcribed loci, including protein-coding genes. Functional analysis of the \textit{tV(UAC)D} locus ascribed a novel function to tRNA genes and also highlighted the importance of stable assembly
of Pol III transcription factors for the effective boundary activity. Our RNA-seq analysis has provided a global picture that emphasizes an intricate role of these interspersed loci in the *S. cerevisiae* genome, and has added to the list of known extra-transcriptional functions of the RNA Polymerase III complex bound to chromosomes.
REFERENCES

AALFS, J. D., and R. E. KINGSTON, 2000 What does 'chromatin remodeling' mean? Trends

ACKER, J., C. CONESA and O. LEFEBVRE, 2013 Yeast RNA polymerase III transcription factors

AHN, S. H., W. L. CHEUNG, J. Y. HSU, R. L. DIAZ, M. M. SMITH et al., 2005 Sterile 20 kinase
phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in

209-220.

ALLISON, D. B., X. CUI, G. P. PAGE and M. SABRIPOUR, 2006 Microarray data analysis: from

ANDERS, S., and W. HUBER, 2010 Differential expression analysis for sequence count data.
Genome Biol 11: R106.

ANDRAU, J. C., A. SENTENAC and M. WERNER, 1999 Mutagenesis of yeast TFIIIB70 reveals C-
terminal residues critical for interaction with TBP and C34. J Mol Biol 288: 511-520.

ARAGON, L., E. MARTINEZ-PEREZ and M. MERKENSCHLAGER, 2013 Condensin, cohesin and the

ARREBOLA, R., N. MANAUD, S. ROZENFELD, M. C. MARSOLIER, O. LEFEBVRE et al., 1998
Tau91, an essential subunit of yeast transcription factor IIIC, cooperates with tau138 in

the binding of yeast transcription factor tau to two intragenic regions. Proc Natl Acad Sci
U S A 84: 8768-8772.

BARBET, N. C., U. SCHNEIDER, S. B. HELLIWELL, I. STANSFIELD, M. F. TUITTE et al., 1996 TOR
controls translation initiation and early G1 progression in yeast. Mol Biol Cell 7: 25-42.

BARDELEBEN, C., G. A. KASSAVETIS and E. P. GEIDUSCHEK, 1994 Encounters of Saccharomyces
cerevisiae RNA polymerase III with its transcription factors during RNA chain

and topography of RNA polymerase III in transcription complexes. Mol Cell Biol 13:
942-952.
BARTHOLOMEW, B., G. A. KASSAVETIS and E. P. GEIDUSCHEK, 1991 Two components of Saccharomyces cerevisiae transcription factor IIIB (TFIIIB) are stereospecifically located upstream of a tRNA gene and interact with the second-largest subunit of TFIIIC. Mol Cell Biol 11: 5181-5189.


DEVINE, S. E., and J. D. BOEKE, 1996 Integration of the yeast retrotransposon Ty1 is targeted to regions upstream of genes transcribed by RNA polymerase III. Genes Dev 10: 620-633.


MIURA, F., N. KAWAGUCHI, J. SESE, A. TOYODA, M. HATTORI et al., 2006 A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. Proc Natl Acad Sci USA 103: 17846-17851.


RAYMOND, G. J., and J. D. JOHNSON, 1987 The 5'-flanking sequence of yeast tRNA(Leu3) genes enhances the rate of transcription from stable pre-initiation complexes. Nucleic Acids Res 15: 9881-9894.


SEKEDAT, M. D., D. FENYÖ, R. S. ROGERS, A. J. TACKETT, J. D. AITCHISON et al., 2010 GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. Molecular systems biology 6: 353.

SEKINGER, E. A., Z. MOQTADERI and K. STRUHL, 2005 Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. Mol Cell 18: 735-748.


APPENDIX: AUTHORIZATION FOR USE OF PUBLISHED MATERIAL

Request

On Mar 24, 2014, I contacted Ann Delacey to request authorization for using my published material. Following is the response from Ann Delacey.

Permission

From: Ann Delacey <adelacey@copyright.com>
To: Asawari Korde <arkorde1@lsu.edu>

Date: Tue, Mar 25, 2014 at 4:38 PM

Subject: License and terms and conditions from Copyright Clearance Center

Dear Asawari,

I have attached your permission request and terms and conditions that govern it. Please contact me if you have any questions or concerns.

Best,
Ann DeLacey

Attachment:

This is a License Agreement between Asawari Korde ("You") and Genetics Society of America ("Genetics Society of America") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Genetics Society of America, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

| License Number | 3353236128023 |
| License date   | Mar 29, 2014  |
| Licensed content publisher | Genetics Society of America |
| Licensed content title | Genetics |
| Licensed content date | Dec 31, 1949 |
| Type of Use | Thesis/Dissertation |
| Requestor type | Author of requested content |
| Format | Electronic |
| Portion | chapter/article |
| Title or numeric reference of the portion(s) | Full article |
| Title of the article or chapter the portion is from | Intergenic transcription is blocked by RNA polymerase III |
| Editor of portion(s) | N/A |
| Author of portion(s) | Asawari Korde |
VITA

Asawari Korde is the daughter of Mr. Rajanikant Korde and Mrs. Anuradha Korde. She was born in Thane, India, in 1984, and did most of her schooling in Mumbai, India. Asawari graduated with a Bachelor of Science in biotechnology from KET’s V. G. Vaze College, Mumbai, India, in 2005 and a master’s degree in biotechnology from Institute of science, India, in 2007. She began her doctoral research in the spring of 2009 in the Department of Biological Sciences at Louisiana State University under the guidance of Dr. David Donze. Ms. Korde is expected to graduate with the degree of Doctor of Philosophy in biochemistry and molecular biology in August 2014.