Yeast Ssu72 Protein as a Mediator of Gene Loop Formation at Bidirectional Promoters

Kevin J. Lopez[a] and Jeffrey O. Henderson[a]

Research over the past ten years has revealed that bidirectional promoters are a common feature of eukaryotes from yeast to metazoans. Gene loops are three-dimensional structures that inhibit divergent transcription at bidirectional promoters. Ssu72, a multifunctional protein phosphatase, has been identified as a key component in directing gene loop formation. Molecular genetic characterization first identified Ssu72 as a component of the transcription initiation complex through an interaction with TFIIB. More recently, Ssu72 has been shown to inhibit divergent transcription by simultaneously binding to both promoter and terminator elements. Here, we describe the path of Ssu72 research from its discovery as a transcriptional regulatory protein to its role in the maintenance of transcriptional memory and repression of noncoding RNAs (ncRNAs). Finally, we consider the role of Ssu72 and gene loops in human health.

Keywords: BRCA1, Gene Loops, Gene Regulation, Saccharomyces cerevisiae, Ssu72, Transcription Regulation

Introduction

The central dogma of molecular biology relies on the assumption that the transfer of information during transcription is an efficient process; one with a carefully balanced mechanism that generates both protein encoding and non-coding RNA. Under the category of eukaryotic ncRNA fall many classes of RNAs such as rRNA, tRNA, splicing snRNA, rRNA modifying snoRNA, and regulatory RNAs like miRNA, stRNA, and siRNA. The fact that ncRNAs may have a variety of functional effects on protein synthesis and activation makes the production and inhibition of ncRNAs a topic of interest. Recent research has been published (Tan-Wong *et al.*, 2012) suggesting the three-dimensional shape (loop) of a gene during transcription plays a role in blocking the production of ncRNAs. Analysis of yeast Ssu72 protein has been instrumental in the discovery of this new gene loop function. By interacting with both promoter and terminator elements, this protein restricts divergent transcription of bidirectional promoters, thereby inducing transcription in the sense direction to form mRNA and repressing transcription in the antisense direction, which produces ncRNAs. Due to the relative structural simplicity of the yeast *Saccharomyces cerevisiae*, researchers have extensively studied the microorganism to gain insight on eukaryotic cell biology. O'Sullivan and coworkers (2004) demonstrated that gene loops were not only formed but actively involved in transcription. These authors used chromatin immunoprecipitation (ChIP) and chromosome conformation capture (3C) to reveal the existence of gene loops due to the position of RNA Polymerase II (RNAPII) on a *GAL1::FMP27* galactose inducible reporter construct in non-induced *S. cerevisiae*.

The ChIP analysis of *GAL1::FMP27* was carried out with an antibody specific for RNAPII on yeast grown in galactose, glucose or raffinose. The ChIP profile obtained from the galactose inducing conditions demonstrated large amounts of RNAPII complexed throughout the entire reporter gene. The ChIP profile obtained from the glucose repressive conditions exhibited almost no sign of RNAPII complexed whatsoever. Unexpectedly, the ChIP profile obtained from the raffinose non-inducing conditions displayed considerable amounts of RNAPII bound to the promoter and terminator regions but not the open reading frame (ORF). This discovery led these workers to hypothesize that the promoter and terminator somehow existed in close spatial proximity under non-inducing conditions.

In 3C, chromatin is cross-linked by formaldehyde, fragmented by restriction enzymes, diluted to prevent random molecular interactions, ligated to form intramolecular phosphodiester bonds, and finally, ligation products are identified by PCR. Formaldehyde fixation cross-links proteins to other proteins and to DNA, which results in overall cross-linking of physically touching segments throughout the genome. The ligation products demonstrate spatial proximity of restriction fragments that would not be connected linearly (Dekker *et al.*, 2002). The fact that the promoter and terminator regions were found to be ligated under non-inducing conditions strongly supported the existence of a gene loop. Furthermore, the inferred presence of a gene loop was found under inducing conditions but not in repressing conditions, indicating that the gene loop conformation is not a consequence of DNA curvature but is specific to the transcriptional status of the gene. Therefore, gene loops may be a common feature of promoting efficient transcription elongation (O'Sullivan *et al.*, 2004).

Ssu72 Interactions

The *SSU72* gene is located from base pair 229,094 to base pair 229,714 on chromosome XIV of *S. cerevisiae* (Sun & Hampsey, 1996). The structure of Ssu72 contains a central five-stranded β-sheet bracketed by helices (Xiang *et al.*, 2010). The function of the 206 amino acid protein was first described by Sun and Hampsey (1996) as they sought to identify factors that were important for transcription initiation of RNAPII. While isolating and characterizing suppressors of defects in general transcription initiation factor TFIIB (which is encoded by gene *SUA7*), Sun and Hampsey identified three revertants, which they designated, *ssu71*, *ssu72*, and *ssu73* (for suppressors of *sua7*, genes 1, 2, and 3). These revertants were specific to the *sua7-1* allele having a glutamic acid-62-

lysine (E62K) substitution in TFIIB causing an aberrant initiation pattern and cold sensitive (cs^-) growth defects. Both *ssu71* and *ssu73* suppressors were not only noted to compensate for the cs^- phenotype but also to restore the normal initiation pattern in the presence of *sua7-1 cs⁻*. This observation stands to reason when one comes to the realization that *SSU71* is *TFG1* (the gene that encodes the largest subunit of TFIIF), and *SSU73* is *RPB9* (the gene that encodes a subunit of RNAPII). In contrast to *ssu*71 and *ssu*73, *ssu*72-1 was noted to be an enhancer of the *sua7-1 cs⁻* phenotype which stimulated a synthetic heat sensitive phenotype and caused the start selection site of *ADH1* and *CYC1* genes to shift further downstream from the normal sites. After genetic and molecular analysis, Sun and Hampsey (1996) concluded that *SSU72* encodes a novel protein that might affect the assembly of the transcription pre-initiation complex (PIC). Furthermore, mutational analysis demonstrated that *SSU72* is essential for cell viability.

Ssu72, like most proteins in the human proteome, is quite versatile. We will focus on its role in gene loops, but it is worth noting that the protein also has a function in the resolution of sister chromatid arm cohesion (Kim *et al*., 2010). Kim and colleagues' budding yeast two-hybrid assay identified Ssu72 as a Rad21-binding protein. Depletion or inactivation of Ssu72 led to premature resolution whereas overexpression caused resistance to resolution. This study also showed that Ssu72 counteracts the phosphorylation of SA2 cohesion subunit and dephosphorylates SA2 itself. Ssu72, therefore, has a role in proper chromosome alignment and segregation during mitosis. These observations demonstrate the importance of this multifunctional protein in the yeast cell cycle.

Transcription Initiation Interactions

The assembly of the PIC of RNAPII relies on the interaction of general transcription factors. TFIIB plays an important role in this assembly by mediating the interaction the TFIID complex, composed of TATA Binding Protein (TBP) and TBP-associated factors (TAF) with the TFIIF-RNAPII complex. Sun and Hampsey (1996) made the initial inference of Ssu72's role in transcription initiation from how an *ssu72* mutant enhanced a TFIIB start site selection defect. Wu and coworkers (1999) then demonstrated direct binding of Ssu72 to TFIIB by GST pull-down. Additionally, Ssu72 was determined to have phosphatase activity with a phosphatase motif that hydrolyzes ρ-nitro-phenyl phosphate (Ganem *et al*., 2003 and Meinhart *et al*., 2003 as cited in Hampsey *et al*., 2011). This activity was then demonstrated to catalyze the removal of phosphate from Ser5-P of the heptad repeated sequence (Tyr_1-Ser_2-Pro_3-Thr_4-Ser_5-Pro_6-Ser_7) of the carboxyl-terminal domain (CTD) of RNAPII, prepping it for initiation (Hampsey *et al*., 2011). Xiang, Manley, and Tong (2012) have recently confirmed that Ssu72 has Ser7 phosphatase activity as well. Using the crystal structure of the N-terminal domain of human symplekin, human Ssu72, and a 10-mer Ser7-P peptide, these authors noticed that the Ser7 peptide is bound to the Ssu72 active site with its backbone running in the opposite direction as that of the Ser5 peptide. Ssu72 also has a 4000-fold lower phosphatase activity upon the Ser7 peptide than it does on the Ser5 peptide. Collectively, these experiments have solidified the evidence that Ssu72 plays a role in transcription initiation.

Interactions in Termination

That Ssu72 is a phosphatase allows it to hold multiple roles in different mechanisms. This enzymatic capability endows Ssu72 with the ability to be part of initiation-elongation transition and in the elongation-termination stage. Due to this capability, Ssu72 becomes an essential component of transcription, and integral to mRNA production. The phosphorylation of the CTD heptad repeat changes as RNAPII advances along the DNA template. As noted above, hypophosphorylated RNAPII is recruited into the PIC. Phosphorylation of Ser5 by kin28, the kinase subunit of TFIIH, initiates transcription and during elongation CTD kinase subunit 1 (Ctk1) phosphorylates Ser2 (Reyes-Reyes & Hampsey, 2007). The action of these kinases combined with the phosphatase action of regulator of transcription 1 (Rtr1) and F cell production 1 (Fcp1) on Ser5 and Ser2, respectively, results in a complexly orchestrated interplay between the kinases and the phosphatases forming a dual gradient of CTD modification. Bataile and coworkers (2012), through ChIP and gene microarray studies, suggest that CTD serines are differentially phosphorylated along genes in a prescribed pattern during the transcription cycle. Interestingly, Ser5-P becomes more prevalent when RNAPII is towards the 5' end and Ser2-P when RNAPII is towards the 3' end of the gene (Keuhner *et al*., 2011).

Dephosphorylation of Ser5 was proposed to be carried out by one of the two RNAPII phosphatases, Ssu72 and/or Rtr1. Rtr1 is specific to the 5' end of genes and contributes dephosphorylation in that region exclusively. ChIP analysis revealed that depletion of Ssu72 lead to termination with a polymerase that contained abnormally high amounts of phosphorylated Ser5 (Bataile *et al*., 2012). This work indicated that Ssu72 targets P-Ser5 at the 3' end of genes, which is on the complete opposite end of transcription (linearly speaking). The same degron-*ssu72* yeast stain also demonstrated a similar change in P-Ser7 phosphorylation. Further research determined that Ssu72 is a dual phosphatase that targets both Ser5 and Ser7 prior to termination (Zhang *et al*., 2012). Ess1 catalyzes the cis-trans isomerization of Pro6 forming a P-Ser5-Pro6 isomer which facilitates rapid phosphorylation by Ssu72 (Werner-Allen *et al*., 2011). Posttermination erasure of Ser5 and Ser7 by Ssu72 enables the complete dephosphorylation of Ser2 by Fcp1, hypophosphorylating RNAPII in preparation for a new round of initiation (Zhang *et al*., 2012).

As the exploration of eukaryotic transcription termination advanced, researchers determined that Ssu72 also has a role in the polyadenylation (poly(A)) of maturing transcripts. Ssu72 was found to be an essential element of the yeast pre-RNA 3'-end processing complex that couples processing with transcription termination. Binding of 3'-end processing factors to both the RNAPII CTD and RNA makes RNAPII more susceptible to pausing (Hobson *et al*., 2012). Studies from the Martinson laboratory (2010) show that transcription through a poly(A) signal reduces the rate of RNAPII elongation and causes RNAPII pausing downstream of the poly(A) site. This type of pausing supports the model in which the cleavage and poly(A) factor (CPF) is recruited to

the elongation complex by interactions with the RNAPII body. CPF is only able to interact with the body of RNAPII efficiently through Ssu72. CPF is the first protein to bind to the signaling region near the cleavage site of the pre-mRNA, to which the poly(A) tail will be added by polyadenine polymerase. When the AAUAAA sequence that the CPF recognizes is transcribed, CPF binds to this site and produces RNAPII pausing. Cleavage factor IA (CFIA) binds to a second downstream signaling region, located at the portion of the pre-mRNA that is cleaved before polyadenylation. This second downstream signaling region is GU-rich and is required for efficient processing. In this model, CPF will then bind to CFIA, releasing its hold of the RNAPII body, and will accompany CFIA to the CTD, where CFIA will interact with the CTD to lead to CPF mediated cleavage and release of paused RNAPII (Hobson *et al.*, 2012).

RNAPII pausing appears to be an important intermediate step to termination but pausing alone does not induce termination. What has been proposed to lead to termination is something called the "torpedo" model (described in Hobson *et al.*, 2012). To competently release paused RNAPII not only are CPF and CFIA necessary, but 5'-3' ribonuclease RNA trafficking protein (Rat1) in complex with its activating cofactor Rat-interacting 1 (Rai1) and its CTD-interacting partner regulator of Ty1 transposition 103 (Rtt103) are also essential. Rat1 is recruited to the 3' end of the gene through interactions of the CTD with the 3'end processing factors and Rtt103 (Hobson *et al.*, 2012). Proteins that interact with phosphorylated Ser2 in the CTD and poly(A) site RNA elements recruit Rat1 to RNAPII where the creation of an unprotected 5'end by the 3' endoribonuclease, or by other cleavage events downstream, allows Rat1 to begin to rapidly "chew" its way towards the RNA exit channel on RNAPII. Collision of Rat1 with RNAPII then promotes termination. The 3' end processing complex introduces Rat1 while pausing RNAPII (which allows Rat1 to "catch" up with the polymerase) and ultimately these carefully choreographed actions lead to the destabilization of RNAPII and its release. In the Torpedo model, the role of Ssu72 is significant because it promotes recruitment of protein 1 of CFI (Pcf11; a component of the CFIA), a reader of the CTD code that promotes RNA cleavage and Rat1 recruitment. The association of Pcf11/Rtt103 requires dephosphorylation of Ser5-P and Ser7-P by Ssu72 in order to associate with the 3' elongation complex (Zhang *et al.*, 2012). It is through its many interactions in transcription that one begins to understand why Ssu72 is not only essential for cell viability in *S. cerevisiae,* but also highly conserved among eukaryotic organisms (He *et al.*, 2003).

Gene Loops

Michael Hampsey has been one of the leading researchers involved in the discovery and study of gene loops. Hampsey's lab used two experiments to determine whether gene loops were static structures of the genome or dynamic functions whose formation coincided with transcription. Hampsey and colleagues first used an *rpb1-1* mutant which encoded for an altered form of RNAPII. This altered RNAPII was temperature dependent and would stall transcription when the temperature was shifted from 24ºC to 37ºC. The looping signal for both *BUD3* and *SEN1* was diminished for the mutants when the temperature shifted to 37ºC while the increase in temperature had no effect on the looping of the wild-type strains. Hampsey's second experiment mirrored O'Sullivan's 3C procedure with inducible *GAL1* promoter versions of *BUD3* and *SEN1*. The analysis revealed an increase in the looping signal when the carbon source was shifted from glucose to galactose, demonstrating that gene loops are structures that form in a transcription dependent manner (reviewed by Hampsey *et al.*, 2011).

In 2007, Singh and Hampsey further extended O'Sullivan and coworkers research on gene loops by demonstrating the presence of gene loops in *BLM10, SAC3, GAL10,* and *HEM3* genes. With the use of ChIP, Singh and Hampsey were able to provide evidence that gene loops were not idiosyncratic to long genes (FMP27, BUD3 and SEN1) but could very well be a general characteristic of RNAPII transcribed genes in *S. cerevisiae*. These experiments also demonstrated that gene loop existence was based on the genetic interaction between TFIIB (*sua7-1*) and Ssu72. *sua7-1* mutants (e.g. E65K) diminished looping at all genes tested (*SEN1, BLM10, SAC3, GAL10,* and *HEM3*). Furthermore, these authors demonstrated that the looping is TFIIB-dependent in a manner independent of its role in transcription; thus, the *sua7-1* mutants had no effect on transcription levels of the genes, but only loop formation. This data was confirmed by El Kaderi and colleagues (2009) who demonstrated that gene looping does not occur during activator-independent increase in transcription for both *MET16* and *INO1*; therefore, transcription is not dependent on gene looping. These workers also contributed to gene loop research by providing five additional insights. (1) Gene loops accompany activated transcription and require a transcription activator. (2) Activators do not directly interact with the 3' end of genes during transcription. (3) Activators interact with TFIIB when a gene is activated and is in a loop conformation. (4) TFIIB cross-links to the distal ends of the gene in an activator-dependent manner; and (5) physical interaction of TFIIB with the 3' end of genes requires Rna15, a termination factor.

Gene Loop Function in Transcriptional Memory

El Kaderi and colleagues' 2009 paper on transcription initiation and termination concludes with an incipient speculation of the function of gene loops, something that has now been tagged "transcriptional memory." They discuss how the presence of a gene loop may facilitate termination-assisted transfer of RNA RNAPII from the terminator to the promoter during the second and subsequent rounds of transcription. Immediately after transcription initiation, RNAPII clears the promoter and TFIIB dissociates from the initiation complex leaving behind many components of the initiation complex referred to as the "scaffold." Scaffold-based reinitiation, contrary to *de novo* initiation, requires very few components be recruited to form the PIC. Loop formation, via a pioneer round combined with Ssu72-catalyzed dephosphorylation of the CTD, augments the rate of reinitiation.

The effect of transcriptional memory was demonstrated and published a couple months later by Laine, Singh, Krishnamurthy and Hampsey (2009). Through ingenious experimental methods, Laine and colleagues were able to demonstrate gene looping is, in fact, associated with

transcriptional memory. Through ChIP, the researchers discovered that the *GAL10* gene loop was present up to 4 hours after the cell environment had been transposed to a glucose repressive state. Further study strengthened the transcriptional memory hypothesis through reactivation kinetics of *GAL10*; 3C analysis revealed that looping occurred coincident with galactose induction in the wild type strain but was defective in the *sua7-1* mutant E62K. The kinetics of induction of the *sua7-1* mutants was essentially identical to the kinetics of the initial activation of the wild type (>1 hour). However, the kinetics of reactivation in the wild-type strains was observed to reach maximum transcript levels in 2 minutes, at least 30 times faster than the mutants to reach maximum transcription levels during reactivation. The definite difference in kinetics is strong evidence that the transcriptional memory phenomenon correlates with gene looping. It was also demonstrated by Tan-Wong and coworkers (2009) that the memory gene loops interact with the nuclear pore complex through association with myosin-like protein 1 (Mlp1) . An *mlp1Δ* strain does not maintain the memory gene loop and consequently loses transcriptional memory (Xu *et al*., 2009).

Bidirectional Promoters

With the modern advances in molecular biology and bioinformatics scientists have been able to dig deeper and collect more data on how the human body functions at the molecular level. The decade-long project ENCODE (Encyclopedia of DNA Elements) has found that 80% of the human genome serves a biochemical purpose with over 90% of the entire human DNA participating in transcription (Preker *et al*., 2008). A portion of the transcriptome had previously gone unnoticed due to the RNA exosome nuclear surveillance machinery that degrades aberrant ncRNAs synthesized by RNAPII. To eliminate RNA degradation, Preker and colleagues depleted the exonucleolytic RNA exosome activity via a mutation in ribosomal RNA-processing protein 6 (Rrp6), an associate-protein to the nine-protein core exosome ring complex which is part of the catalytically active exosome subunits that act to degrade ncRNAs. The depletion of exosome activity revealed novel ncRNAs and cryptic unstable transcripts (CUTs) which are usually rapidly degraded by the RNA surveillance pathway. Through deeper investigation, Xu and colleagues (2009) were able to gain insights on the origin of CUTs. Transcripts of a *Δrrp6* mutant were identified, and those with confidently mapped 5' ends were analyzed for transcript start sites (TSS) with the intent to find out whether the transcripts had hallmarks. Nucleosome-free promoter regions (5'-NFRs), which facilitate transcription by allowing RNAPII to bind to DNA, are hallmarks of gene promoters. Analysis of the transcripts of the *Δrrp6* mutant indicated that all classes of transcripts (open reading frame transcripts, stable annotated transcripts (SUTs) and CUTs) displayed deficiency of nucleosomes upstream of the TSS. Additionally, 63% of the non-overlapping divergent transcript pairs involving at least one unannotated transcript demonstrated no intermediate nucleosome (Xu *et al*., 2011). This data modifies our perspective on how the genome is transcribed. The data suggest that bidirectionality is an inherent feature of promoters which must be in part controlled by the exosome nuclear surveillance machinery.

The existence of bidirectional promoters demands investigation of the mechanism that regulates the transcription of the coding mRNAs because it is unclear whether the production of CUTs has any adaptive value or if their production is merely a byproduct of the promoter's intrinsic characteristics. It has been estimated that 50% of *S. cerevisiae* promoters are bidirectional in nature (Castelnuovo & Stutz, 2013). Bidirectional promoters have also been identified in mammals, including humans, albeit at a lower incidence than in yeast (reviewed by Seila *et al*., 2009). Divergent transcription suggests that initiation by one polymerase would induce the formation of a PIC in the opposite direction on the same gene. Analysis of high-throughput sequencing data from short RNA samples of mouse ES cells and genome-wide nuclear run-on assays from human fibroblasts identified a much higher density of RNA mapping to the 5' coding-region of most genes than throughout the downstream regions. These authors also found almost equal amounts of antisense RNAs upstream of the TSS for a majority of the genes tested. From this data, it was proposed that divergent transcription from a bidirectional promoter is common in most active genes but without significant elongation by the antisense RNAPII.

Additionally, ChIP-sequencing data for RNAPII and analysis of histone modifications demonstrated that RNAs both downstream and upstream of the TSS overlap locations of bound RNAPII. Thus, the polymerases were enclosed by nucleosomes that indicate active transcription initiation, such as histone 3 triply methylated on lysine 4 (H3K4me3) and H3 and H4 acetylation, suggesting the formation of a nucleosome-free region. Alternatively, marks of transcription elongation such as H3K79me2, H3K36me3 and H2Bub are found only in the sense direction. Therefore, these investigators concluded that polymerase occupancy due to initiation was nearly equivalent in both directions, but the polymerase moving antisense did not productively elongate (Seila *et al*., 2009).

Elucidation on the role of bidirectional promoters has been an area of active research the past few years (reviewed in Grzechnik *et al*., 2014). By understanding events at bidirectional promoters, one can gain insight on the order and regulation of the transcriptome. Divergent gene pairs are mainly co-expressed, coupling protein-coding genes involved in the same process allows for coordinated temporal and environmental responses. Non-coding transcription can have a repressive effect by removing bound transcription factors from the protein coding region or by recruiting the Polycomb repressive complex-2 which catalyzes histone trimethylation (Wei *et al*., 2011). Furthermore, antisense transcription of one gene (*GAL80*) can repress the transcription of a tandem gene (*SUR7*) which has a promoter in the sense direction upstream of the first gene (Xu *et al*., 2011). Work on convergent transcription using FLAG and HA-tagged RNAPII and affinity chromatography demonstrated how the converging polymerases were not able to transcribe a full gene in the manner that RNAPII with only one elongation complex was able to transcribe a full gene (Wei *et al*., 2011). These researchers also explained how the depletion of long intergenic non-coding RNAs in humans decreases the expression of distal protein-coding genes. The question

remained: What mechanism controls the production and consequently the regulation of these ncRNAs? In November of 2012, Tan-Wong and coworkers published an article in *Science* suggesting that the three-dimensional shape (loop) of a gene during transcription plays a role in blocking the production of ncRNAs.

Gene Loops Function in Transcription Directionality and ncRNA Repression

To test whether intragenic looping enhances the directionality of transcription, Proudfoot's group (Tan-Wong *et al*., 2012) examined the expression of divergently transcribed ncRNA at the *FMP27* locus in *S. cerevisiae*. They compared the effect of mutating *RRP6* and *SSU72* alone or together on the genomic profile of coding and ncRNAs. Wild-type, *ssu72-2*, *Δrrp6*, and double *ssu72-2Δrrp6* strains were grown at the semi-permissive condition of 32°C. Total RNA was hybridized to strand-specific *S. cerevisiae* tiling arrays. From these arrays, it was noted that *ssu72-2* mutation alone or in combination with *Δrrp6* gave rise to both CUTs, due to the loss of the exosome activity from *Δrrp6,* and many additional ncRNAs. The new cryptic transcripts were labeled Ssu72-restricted transcripts (SRTs) and were identified as running in a divergent orientation from bidirectional promoters. By focusing on pairs of tandem genes where the promoter of the downstream gene is far from the terminator region of the upstream gene, these authors concluded that the enhanced synthesis of the ncRNA originates from the downstream bidirectional promoter, rather than the upstream, poly (A) site.

Their array data demonstrated the presence of 605 SRTs and 1982 CUTs. To focus on promoter-associated ncRNAs (pncRNAs), the authors concentrated on 135 SRTs and 678 CUTs that were transcribed convergently between tandem open reading frames. Earlier work (Wei *et al*., 2011) using a genome-wide RNAPII occupancy profile of the *ssu72-2* mutants revealed a distinct peak upstream of the TSS, which was absent in the wild type, as well as an RNAPII accumulation over SRTs. These results established that the loss of Ssu72 promotes *de novo* initiation of SRTs positioned between tandem ORFs (pSRTs). Analysis of public genome wide data demonstrated that pSRT-associated promoters are particularly depleted of histone H4 acetylation (Tan-Wong *et al*., 2012). This implies that these promoters lie in regions where the nucleosome is in a repressed transcriptional state. The loss of Ssu72, consequently, seemed to relax histone acetylation leading to pSRT expression. Overall, these results suggest that gene looping decreases divergent transcription by a mechanism that involves histone H4 deacetylation by a histone deacetylase (HDAC). Loss of HDAC Rco1 (in the Rpd3S complex), a protein known to contribute to H4 deacetylation in a gene's 3' regions, also increased antisense transcription. The fact that the ncRNAs induced by the loss of Rco1 are derived from the 3' end classifies them as different ncRNA than SRTs, which are derived from divergent transcription from promoters. Castelnuovo and Stutz (2013) clarified that Rco1-restricted transcripts are linked and antisense to the upstream promoter TSS of ORFs in tandem pairs, as opposed to the downstream promoter from which SRTs originate. From these observations, the authors conclude that Ssu72 enforces promoter directionality with the caveat that deacetylation still holds an association with directionality as it is necessary for gene loop formation (Tan-Wong *et al*., 2012).

Taking this conclusion a step further Tan-Wong and coworkers (2012) wondered that if Ssu72 is required for gene-loop formation, whether other gene-loop-associated factors similarly act to restrict pncRNAs. Mutations in TFIIB and polyadenylation complex components, such as Pta1, Rna14, and Rna15 also increased divergently transcribed SRTs suggesting that intragenic looping has a direct role in regulating transcriptional directionality.

Tan-Wong and coworkers (2012) immediately sought to test their hypothesis by determining whether *cis* mutations that affect intragenic looping would also lead to changes in RNAPII directionality. They did so by replacing the poly(A) signal in the 3' untranslated region with an Rnt1 cleavage signal. This results in normal termination but blocked intragenic looping and polyadenylation. Implementing the change in two yeast genes and in a β–globin transgene in human embryonic kidney cells resulted in a three-fold increase of divergently transcribed ncRNAs suggesting that intragenic looping plays a conserved role in regulating transcriptional directionality. In the case of a bidirectional promoter the gene loop limits assembly of two PICs and, therefore, forces transcription in the sense direction by allowing for the maintenance of only the sense PIC (Rhee and Pugh, 2012).

Medical Perspectives on Repression of ncRNA by Gene Looping

Probably the most obvious relevancy of these new findings is that many processes that have been determined to occur in yeast cells have been proven to occur in human cells. The fact that these structures are homologous within the cells of humans gives biomedical researchers a new approach to maintaining the health of the human population. It is interesting to note that gene loops are not unique to yeast and metazoans. The HIV provirus forms a transcription-dependent gene loop between the long terminal repeat promoter and the 3' long terminal repeat poly(A) site; the proteins involved in forming this gene loop offer additional targets for HIV chemotherapy (Perkins *et al*., 2008). Of importance to the cancer research community, a gene looping model has been proposed for the transcriptional regulation of the tumor suppressor gene *BRCA1* in humans (Tan-Wong *et al*., 2008). Manipulation of these loops may lead to beneficial health results. Tan-Wong and coworkers suggested that, in a regulated repressed state, the loop conformation transcription of *BRCA1* is disabled and maintains *BRCA1* gene expression. It is known that loss of *BRCA1* expression can be carcinogenic, so investigation into the manipulation of another layer of control on gene expression may lead to beneficial new cancer treatment targets. If *BRCA1* is regulated by a gene loop, would the synthesis and introduction of functional loop-promoting proteins (e.g. Ssu72) to tumor cells in which *BRCA1* is repressed aid in tumor suppression? Understanding the molecular mechanisms of gene loop function and consequences in HIV infection/persistence and breast tumorigenesis are clear areas of experimental priority.

References

Bataile, A.R., Jeronimo, C., Jacques, P. E., Laramée, L., Fortin, M.E., Forest, A, …& Robert, F. (2012). A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Molecular Cell. 45*, 158-170.

Castelnuovo, M., & Stutz, F. (2013). Gene loops and HDACs to promote transcription directionality. *Nucleus. 4*, 92-94.

Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2012). Capturing chromosome conformation. *Science. 295*, 1306-1311.

El Kaderi, B., Medler, S., Raghunayakula, S., & Ansari, A. (2009). Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. *The Journal of Biological Chemistry. 284*, 25015-25025.

Ganem, C., Devaux, F., Torchet, C., Jacq, C., Quevillon-Cheruel, S., Labesse, G., Facca, C., & Faye, G. (2003). Ssu72 is a phosphatase essential for transcription termination of snoRNAs and specific mRNAs in yeast. *The EMBO Journal. 22*, 1588-1598.

Grzechnik, P., Tan-Wong, S.M., & Proudfoot, N.J. (2014). Terminate and make a loop: regulation of transcriptional directionality. *Trends in Biochemical Sciences. 39*, 319-327.

Hampsey, M., Singh, B.N., Ansari, A., Lainé, J.P., & Krishnamurthy, S. (2011). Control of eukaryotic gene expression: gene loops and transcriptional memory. *Advances in Enzyme Regulation. 51*, 118-125.

He, X., Khan, A.U., Cheng, H., Pappas, D.L. Jr, Hampsey, M., & Moore, C.L. (2003). Functional interactions between the transcription and mRNA 3' end processing machineries mediated by Ssu72 and Sub1. *Genes & Development. 17*, 1030-1042.

Hobson, D.J., Wei, W., Steinmetz, L.M., & Svejstrup, J.Q. (2012). RNA polymerase II collision interrupts convergent transcription. *Molecular Cell. 48*, 365-374.

Kazerouninia, A., Ngo, B., & Martinson, H.G. (2010). Poly(A) signal-dependent degradation of unprocessed nascent transcripts accompanies poly(A) signal-dependent transcriptional pausing *in vitro*. *RNA. 16*, 197-210.

Kim, H.S., Baek, K.H., Ha, G.H., Lee, J.C., Kim, Y.N., Lee, J., …& Lee, C.W. (2010). The hsSsu72 phosphatase is a cohesin-binding protein that regulates the resolution of sister chromatid arm cohesion. *The EMBO Journal. 29*, 3544-3557

Keuhner, J.N., Pearson, E.L., & Moore, C. (2011). Unraveling the means to an end: RNA polymerase II transcription termination. *Nature Reviews Molecular Cell Biology. 12*, 283-294.

Lainé, J.P., Singh, B.N., Krishnamurthy, S., & Hampsey, M. (2009). A physiological role for gene loops in yeast. *Genes & Development. 23*, 2604-2609.

Meinhart, A., Silberzahn, T., & Cramer, P. (2003). The mRNA transcription/processing factor Ssu72 is a potential tyrosine phosphatase. *Journal of Biological Chemistry. 278*, 15917-15921.

O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., & Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. *Nature Genetics. 36*, 1014-1018.

Perkins, K.J., Lusic, M., Mitar, I., Giacca, M., & Proudfoot, N.J. (2008). Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. *Molecular Cell. 29*, 56-68.

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., …& Jensen, T.H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. *Science. 322*, 1851-1854.

Reyes-Reyes, M., & Hampsey, M. (2007). Role for the Ssu72 C-terminal domain phosphatase in RNA polymerase II transcription elongation. *Molecular and Cellular Biology. 27*, 926-936.

Rhee, H.S., & Pugh, B.F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature 483*, 295-301.

Seila, A.C., Core, L.J., Lis, J.T., & Sharp, P.A. (2009). Divergent transcription: a new feature of active promoters. *Cell Cycle. 8*, 2557-2564.

Singh, B.N., & Hampsey, M. (2007). A transcription-independent role for TFIIB in gene looping. *Molecular Cell. 27*, 806-816.

Sun, Z.W. & Hampsey, M. (1996). Synthetic enhancement of a TFIIB defect by a mutation in SSU72, an essential yeast gene encoding a novel protein that affects transcription start site selection in vivo. *Molecular and Cellular Biology. 16*, 1557-1566.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E.,…& Proudfoot, N.J. (2010). Gene loops enhance transcriptional directionality. *Science. 338*, 671-675.

Tan-Wong, S.M., Wijayatilake, H.D., & Proudfoot, N.J. (2009). Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. *Genes & Development. 23*, 2610-2624.

Tan-Wong, S.M., French, J.D., Proudfoot, N.J., & Brown, M.A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Proceedings of the National Academy of Science U S A. 105*, 5160-5165.

Wei, W., Pelechano, V., Järvelin, A.I., & Steinmetz, L.M. (2011). Functional consequences of bidirectional promoters. *Trends in Genetics. 27*, 267-276.

Werner-Allen, J.W., Lee, C.J., Liu, P., Nicely, N.I., Wang, S., Greenleaf, A.L., & Zhou, P. (2011). Cis-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. *The Journal of Biological Chemistry. 286*, 5717-5726.

Wu, W., Pinto, I., Chen, B.S., & Hampsey, M. (1999). Mutational analysis of yeast TFIIB. A functional relationship between Ssu72 and Sub1/Tsp1 defined by allele-specific interactions with TFIIB. *Genetics. 153*, 643–652.

Xiang, K., Manley, J.L., & Tong, L. (2012). An unexpected binding mode for a Pol II CTD peptidephosphorylated at Ser7 in the active site of the CTD phosphatase Ssu72. *Genes & Development. 26*, 2265-2270.

Xiang, K., Nagaike, T., Xiang, S., Kilic, T., Beh, M.M.,
Manley, J.L., & Tong, L. (2010). Crystal structure of the
human symplekin-Ssu72-CTD phosphopeptide complex.
Nature. *467*, 729-733.

Xu, Z., Wei, W., Gagneur, J., Clauder-Münster, S., Smolik,
M., Huber, W., & Steinmetz, L.M. (2011). Antisense
expression increases gene expression variability and locus
interdependency. *Molecular Systems Biology*. *7*, 468.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster,
S., Camblong, J., … & Steinmetz, L.M. (2009).
Bidirectional promoters generate pervasive transcription in
yeast. *Nature*. *457*, 1033-1037.

.Zhang, D.W., Mosley, A.L., Ramisetty, S.R., Rodríguez-
Molina, J.B., Washburn, M.P., & Ansari, A.Z. (2012).
Ssu72 phosphatase-dependent erasure of phospho-Ser7
marks on the RNA polymerase II C-terminal domain is
essential for viability and transcription termination. *The
Journal of Biological Chemistry*. *287*, 8541-8551.