Université de Sherbrooke

BIOCHEMICAL AND GENETIC ANALYSIS OF RNA PROCESSING AND DECAY

Par

Ghada Ghazal

Département de Microbiologie et d'infectiologie

Thése présentée à la Faculté de médecine et des sciences de la santé
en vue de l'obtention du garde de
philosophiae doctor (Ph.D.) en Microbiologie

Membres du Jury :

Dr. Nicholas Proudfoot, raporteur externe à l'université

Dr. François Bachand, raporteur externe au programme

Dr. Raymund Wellinger, président et directeur du département

Dr. Sherif Abou Elela, directeur de thèse

31 août 2009

Your file Votre référence
ISBN: 978-0-494-62832-4
Our file Notre référence
ISBN: 978-0-494-62832-4

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ABREVIATIONS

bp	Base pair
CPSF	Cleavage polyadenylation specific factors
CTD	Carboxy-terminal domain
CUT	Cryptic unstable transcript
dsRBD	Double stranded RNA binding domain
dsRBP	Double stranded RNA binding protein
ETS	External transcript
mRNA	Messenger RNA
nt	nucleotide
N-term	N-terminal domain
NUCD	Nuclease domain
PAZ	PIWI/Argonaute/Zwille
Pol II	Polymerase II
Pre-mRNA	Premature mRNA
Pre-rRNA	premature ribosomal RNA
rDNA	Ribosomal DNA

RNAi	RNA interference
RNP	Ribonucleoprotein
snRNP	Small nuclear ribonucleoprotein
UTR	Untranslated region

Université de Sherbrooke

L'ANALYSE BIOCHIMIQUE ET GÉNÉTIGUE DE LA MATURATION ET LA DÉGRADATION DE L'ARN

Par

Ghada Ghazal

Département de Microbiologie et d'infectiologie

Thése présentée à la Faculté de médecine et des sciences de la santé

en vue de l'obtention du garde de

philosophiae doctor (Ph.D.) en Microbiologie

L'expression des gènes est le conduit par lequel l'information génétique est traduite dans les phénotypes cellulaires. Récemment, il a été démontré que le programme de l'expression des gènes dans les cellules de mammifères est régi, au moins en partie par l'expression d'ARN double brin court (ARNdb). Ce mode de régulation des gènes est influencé par un grand groupe de protéines de liaison à l'ARN double brin qui peuvent soit stabiliser ou déclencher la dégradation de l'ARN double brin. En effet, les ribonucléases (RNases) spécifiques à l'ARN double brin jouent un rôle important dans l'expression des gènes. Dans la plupart des eucaryotes, les membres de la famille des RNase III spécifiques à l'ARNdb déclenchent la dégradation de l'ARN et initient la réponse immune de la cellule. Un défaut dans l'activité de la RNase III (DICER) inhibe l'expression des gènes et favorise le développement du cancer. D'autre part, la surexpression de la RNase III bloque l'infection virale. Cependant, très peu est connu sur la fonction de

gestion domestique des RNases III chez les eucaryotes et le mécanisme par lequel ils font la distinction entre les espèces d'ARN cellulaire et l'infection virale. Cette thèse pave la voie sur la manière dont les ARNdbs sont choisis pour être clivés et démontre leur contribution dans le mécanisme de l'ARN en utilisant la levure comme modèle d'étude. Initialement, les déterminants de réactivité de la RNase III chez la levure (Rnt1p) ont été identifiés *in vitro* et utilisés pour étudier l'impact global de Rnt1p sur la maturation des ARNs non-codants. Les résultats indiquent que Rnt1p est nécessaire pour la maturation de tous les petits ARN nucléolaires (snoRNAs) impliqués dans la méthylation de l'ARNr et ils identifient un nouveau rôle de Rnt1p dans la maturation des snoRNAs introniques. Il a été démontré que le clivage de Rnt1p contribue à coordonner l'expression de certaines protéines ribosomales et des snoRNA contenus dans leurs introns. La maturation du snoRNA à partir de l'ARN pré-messager bloque l'expression du gène hôte, alors qu'en retardant la maturation du snoRNA, celle-ci se séroule sur l'intron excisé ce qui permet l'expression des deux gènes. De cette façon, la cellule peut coordonner soigneusement la quantité de protéines ribosomales et de snoRNAs requises pour la biogénèse des ribosomes. En outre, l'analyse globale de la maturation des snoRNAs a identifié de nouveaux signaux de clivage de Rnt1p qui ne présentent pas un motif de séquence conservé.

Cette constatation a conduit à la conclusion que Rnt1p utilise une vaste combinaison de motifs structuraux pour identifier ses substrats et augmenter ainsi le nombre de cibles potentielles de dégradation *in vivo*. Pour évaluer

cette possibilité, une nouvelle recherche de motifs pouvant être clivés par Rnt1p a été effectuée. Fait intéressant, de nombreux signaux de clivage de Rnt1p ont été identifiés dans des régions intergéniques qui n'encodent aucun transcrit d'ARN connus. In *vivo*, les résultats démontrent que Rnt1p est capable de terminer la transcription des ARNms non-polyadenylés et participe à un mécanisme de surveillance contre la continuation de transcription (read-through). Cette découverte démontre un lien direct entre la Rnt1p et la machinerie de la transcription des ARN messagers, et prévoit un nouveau mécanisme de la terminaison de la transcription indépendante de la polyadénylation. Ensemble, les travaux décrits dans cette thèse présentent un exemple de la façon dont les RNase III chez les eucaryotes identifient leurs substrats et présentent un modèle dans lequel la transcription de l'ARN, sa maturation et sa stabilité sont liés.

Mots clés : dsRNA, RNases, Rnt1p, snoRNA, transcription termination, 3'end formation

BIOCHEMICAL AND GENETIC ANALYSIS OF RNA PROCESSING AND DECAY

Par

Ghada Ghazal

Département de Microbiologie et d'infectiologie

Thèse présentée à la Faculté de médecine et des sciences de la santé

en vue de l'obtention du garde de

philosophiae doctor (Ph.D.) en Microbiologie

Gene expression is the conduit by which genetic information is connected into cellular phenotypes. Recently, it was shown that gene expression in mammalian cells is governed, at least in part, by the expression of short double stranded RNA (dsRNA). This mode of gene regulation is influenced by a large group of dsRNA binding proteins that could either stabilize or trigger the degradation of dsRNA. Indeed, double stranded RNA (dsRNA) specific ribonucleases (RNases) play an important role in regulating gene expression. In most eukaryotes, members of the dsRNA specific RNase III family trigger RNA degradation and initiate cellular immune response. Disruption of human RNase III (Dicer) deregulates fetal gene expression and promotes the development of cancer. However, very little is known about the housekeeping function of eukaryotic RNase III and the mechanism by which they distinguish between exogenous and endogenous cellular RNA species. This thesis elucidates how dsRNAs are selected for cleavage and demonstrates their contribution to RNA metabolism in yeast as model eukaryote. Initially, the reactivity determinants of yeast RNase III (Rnt1p) were identified *in vitro* and

used to study the global impact of Rnt1p on the processing of non-coding RNA. The results indicate that Rnt1p is required for the processing of all small nucleolar RNAs (snoRNAs) involved in rRNA methylation and identify a new role of Rnt1p in the processing of intronic snoRNAs. It was shown that Rnt1p cleavage helps to coordinate the expression of some ribosomal protein genes hosting intronic snoRNAs. Direct snoRNA processing from the pre-mRNA blocks the expression of the host gene, while delayed snoRNA processing from the excised intron allows the expression of both genes. In this way, the cell can carefully calibrate the amount of snoRNA and ribosomal proteins required for ribosome biogenesis. In addition, a global analysis of snoRNA processing identified new forms of Rnt1p cleavage signals that do not exhibit a conserved sequence motif but instead use a new RNA fold to recruit the enzyme to the cleavage site. This finding led to the conclusion that Rnt1p may use a wide combination of structural motifs to identify its substrates and thus increases the theoretical number of potential degradation targets *in vivo*. To evaluate this possibility, a new search for snoRNA independent Rnt1p cleavage targets was performed. Interestingly, many Rnt1p cleavage signals were identified in intergenic regions devoid of known RNA transcripts. *In vivo*, it was shown that Rnt1p induce the termination of non-polyadenylated transcripts and functions as a surveillance mechanism for transcription read-through. This finding directly links Rnt1p to the transcription machinery and provides a new mechanism for polyadenylation independent transcription termination. Together the work described in this thesis presents an example

of how eukaryotic RNase III may identify its substrates and present a case study where transcription, RNA processing and stability are linked.

Key Words: dsRNA, RNases, Rnt1p, snoRNA, transcription termination, 3'end formation

INTRODUCTION

1. Regulation of gene expression

Modern cells are defined by their gene sequence but shaped by their protein make up (Herbert and Rich, 1999a). The transformation of genetic information from deoxyribonucleic acid sequence (DNA) into protein is an essential and tightly controlled process termed "gene expression" (Granneman and Baserga, 2005; Hinnebusch, 1990). In its simplest form, gene expression is a passive conduit of stored genetic information with little influence on the phenotypic outcome (Herbert and Rich, 1999b). However, even in the simplest of live forms like viruses, the process leading to protein production can greatly influence the organism's function and may even determine its chance to survive (Katze and Agy, 1990; Naryshkin et al., 1998; Stoltzfus and Madsen, 2006). Therefore, cells have developed a highly refined mechanism to control the expression time and amount of each gene and used it to fine-tune the accumulation of any particular protein at a specific time (Haile and Papadopoulou, 2007; Harrison, 1990; Izawa and Inoue, 2009). Regulation of gene expression provides cells with the flexibility they need to face changes in their environment and increase the versatility of their protein functions (Hengge-Aronis, 2002; Marles-Wright and Lewis, 2007; Wassarman, 2002). For example, yeast genes involved in glucose metabolism are expressed in the presence of glucose, while those required for gluconeogenesis are repressed (Gelade et al., 2003). In bacteria, regulation of gene expression is

mostly a response mechanism to rapidly changing environment (Klaenhammer et al., 2007). In contrast, changes in mammalian gene expression respond to the need for cell specialization and differentiation (Harrison, 1990).

The mechanism regulating gene expression varies depending on the organism and gene function. In bacteria, gene regulation is ingrained in the genome structure (Rocha, 2008). For example, genes with related function are clustered into "operons" to allow coordinated expression of proteins with interdependent functions (Rocha, 2008). In eukaryotes, the genome structure and mechanism of gene expression is drastically different than that of bacteria (Mateos-Langerak et al., 2007). Genes are normally not organized by function and transcription is physically separated from translation by the nuclear membrane. Eukaryotic genes need not only to include information about their transcriptional program but also need to embed in the RNA information that dictates its stability, export, translatability and the nature of the protein it produces (Zhai et al., 2008). Impairing any of these steps may signal RNA degradation and abort the expression process. Therefore, in eukaryotes co-regulation of proteins cannot be achieved by a simple switch, a single factor or even a single step of gene expression. For simplicity, eukaryotic gene regulation is often separated into four classes; 1) transcriptional, 2) posttranscriptional, 3) translational and 4) posttranslational gene regulation (Nolan and Cogoni, 2004). Transcriptional gene regulation influences the overall amount of the primary gene products and is often used as a master on

and off switch of gene expression. In reality, however, it is important to note that eukaryotic gene regulation is an integrated process where one level of gene expression affects the other.

2. Transcription

The most direct way to control the expression of a gene is to regulate its rate of transcription; that is, the rate at which RNA polymerases transcribe genes into messenger RNA (mRNA) (Westholm et al., 2008). The basic mechanism of transcription is the same in all organisms where DNA dependent RNA polymerases recognize a specific DNA sequence and use it to polymerize free nucleotides into ribonucleic acid chains (Lee and Young, 2000). The main difference between bacterial and eukaryotic transcription machinery is in the number of RNA polymerases and the associated transcription factors. In Bacteria, all genes are transcribed by the same RNA polymerase (Balleza et al., 2009), whereas eukaryotes use three different nuclear polymerases (RNA Pol I-III) (Chambon, 1975; Roeder and Rutter, 1970). These polymerases differ in the number and type of subunits they contain, as well as the class of RNAs they transcribe; that is, RNA Pol I transcribes ribosomal RNAs (rRNAs) (Kuhn et al., 2007), RNA Pol II (Meyer et al., 2009) transcribes RNAs that will become messenger RNAs (mRNAs) and also small regulatory RNAs, whereas RNA Pol III transcribes small RNAs such as transfer RNAs (tRNAs). Because RNA Pol II transcribes protein-encoding genes, it has been the main

target of transcriptional regulation. Transcription begins with the binding of the polymerase to the promoter region which is essential for correct positioning and assembly of Pol II and the general transcription factors in a state termed preinitiation complex (PIC). Next, a marked conformational change allowing the active Pol II to open the template strand of the promoter and starts the initiation of transcription. After synthesis of ~30 bases of RNA, the 5' end of the RNA is modified by adding a cap that consists of a modified guanine nucleotide. Pol II then releases its contacts with the core promoter and the rest of the transcription machinery and enters the transcription elongation phase (Figure 1). Factors that promote productive RNA chain synthesis, RNA processing, RNA export and chromatin modification can all be recruited to elongating Pol II (Bentley, 2002). A key step of the transition of Pol II to the elongation mode of RNA synthesis is an extensive phosphorylation of the RNA polymerase II tail, carboxy-terminal domain "CTD" (Figure 1). This C-terminal domain consists of a long tandem array of repeated seven-amino-acid sequences, containing two serines (serine 2, serine 5) per repeat that can be phosphorylated. In addition phosphorylation of a third serine in position 7 was recently identified and its impact on transcription is currentlybeing investigated (Kim et al., 2009). As transcription precedes the two major sites of phosphorylation (serine 2, serine 5), predominate, and the CTD of the Pol II undergoes conformational changes to recruit the termination factors (Bentley, 2005; Zorio and Bentley, 2004; Bentley, 1999).Transcription termination is an important process as it enhances gene expression by

facilitating polymerase recycling and thus maintains a pool of available polymerase (Dye and Proudfoot, 1999). Once transcribed, RNA is normally processed to produce the mature form and either exported to the cytoplasm for translation or assembled into functional RNP in the nucleus (RNP).

3. RNA Maturation

RNA maturation is the process by which a nascent RNA is transformed into a stable functional form. For a mRNA this means capping, polyadenylation, and removal of intronic sequence through splicing (Meyer et al., 2009). On the other hand, for non-coding RNA like small nuclear RNA (snRNA) or ribosomal RNA (rRNA) maturation means the removal of transcribed spacers, modification of the 5' and 3' ends, and assembly into an active RNA protein complex (Lafontaine and Tollervey, 1995; Nazar, 2004; Reddy and Busch, 1983). In both cases, the aim of this process is to remove non-functional sequence, ensure the quality of the transcribed RNA and increase the versatility of RNA functions. Each step of RNA maturation involves a complex machinery of RNA and protein factors capable of specifically recognize its target and modify it according to a pre-determined and precise program (Fischer et al., 1991; Maxwell and Fournier, 1995; Wahl et al., 2009)

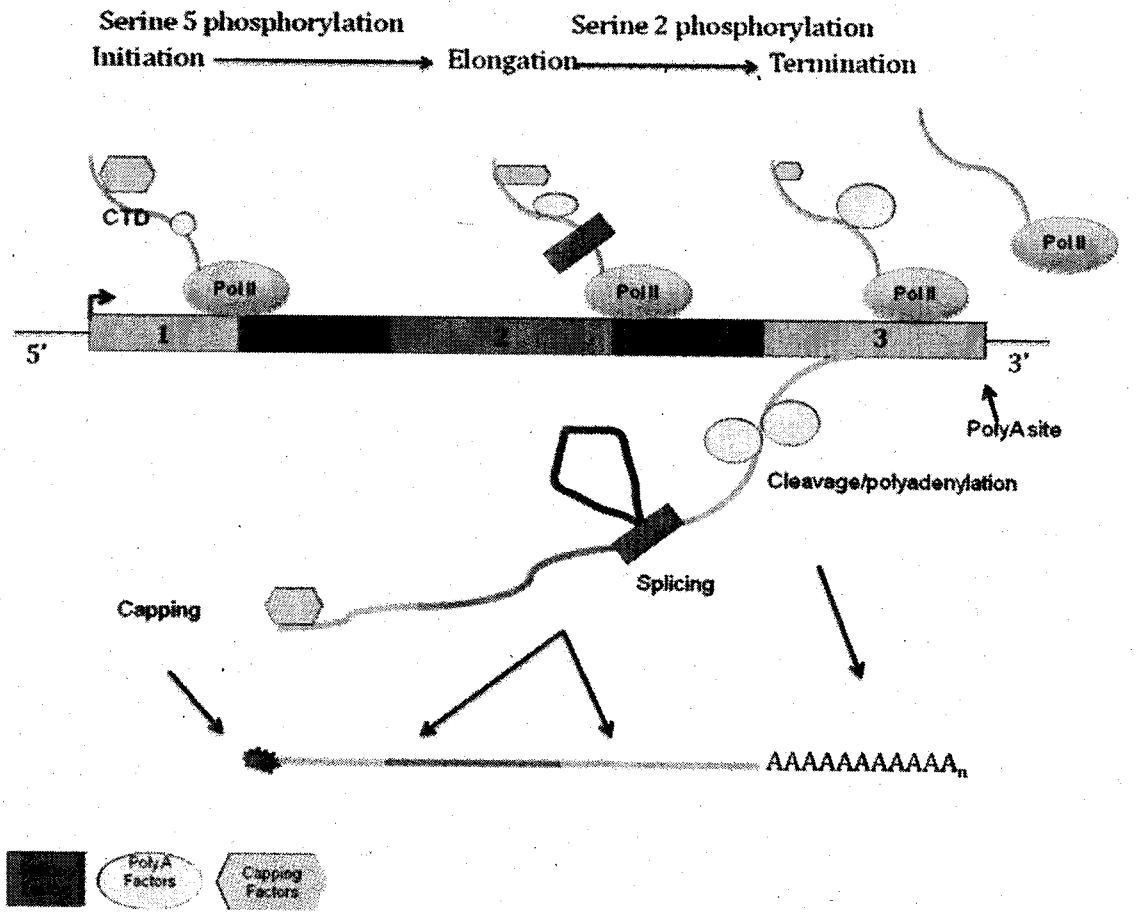


Figure 1. Schematic representation of co-transcriptional processing.

Processing factors interact with the Pol II machinery via the carboxy-terminal domain (CTD) of the largest subunit of RNA Pol II. Capping enzymes are recruited to the 5'ends of genes. As Pol II traverses the gene, splicing factors associate with the transcription complex. Phosphorylation of Ser2 and Ser5 residues in the CTD heptad repeats is indicated in yellow circles. Exon numbers are marked in colors. Introns are shown in black boxes. The red star represents the cap structure (adapted from Zorio and Bentley, 2004).

3.1 Capping

Capping is a specific form of 5' end modification that occurs during transcription by RNA polymerase II (Pol II). RNA produced by Pol I or III are not capped and these polymerases do not associate with capping enzymes (Gu and Lima, 2005). In general caps confer stability to mRNAs by protecting them from digestion by exonucleases. However the crucial role of the 5' cap of the mRNAs is to position the ribosome to initiate translation through the binding of the initiation factor CBPI. In fact, some viruses such as poliovirus prevent capped cellular mRNAs from being translated into proteins. This enables poliovirus to take over the protein synthesizing machinery in the infected cell to make new viruses (Thompson and Sarnow, 2000). The 5' cap is generated by the addition of a guanosine to the extreme 5' end of the nascent mRNA by the guanylyl transferase enzyme, this guanosine later converted into 7-methylguanosine by the guanine methyltransferase (Furuichi and Shatkin, 2000; Gu and Lima, 2005; Shuman, 2001). These dimeric capping enzymes are recruited to the phosporylated carboxy-terminal domain (CTD) of the Polymerase II at the early stages of RNA synthesis (Figure 1) (Viladevall et al., 2009). After the RNA is capped, elongation factors required for splicing and termination are recruited to the CTD of the Pol II.

3.2 Splicing

Splicing is a process by which intervening sequence (introns) are removed and the protein encoding fragments (exons) are joined together to generate mature mRNA ready for translation (Rio, 1993; Umen and Guthrie, 1995). This process if performed by a large RNA protein complex (Spliceosomal complex) that ensures the fidelity and efficiency of intron removal (Wahl et al., 2009). Splicing allows cells to swap exons during development and thus modify protein sequence and function as the cellular functions change (Irimia et al., 2009; Mattaj and Hamm, 1989; Rio, 1993). In addition, splicing is also believed to contribute to genome complexity and increase the diversity of protein functions through the process of alternative splicing (Boue et al., 2003; Kim et al., 2008; Kriventseva et al., 2003; Park and Graveley, 2007). Indeed most human genes are now believed to be alternatively spliced (Wang et al., 2008).

Although introns are often considered disposable junk DNA, it is becoming increasingly clear that information in these non-coding sequences can directly or indirectly affect gene expression (Le Hir et al., 2003). Introns can influence every level of RNA metabolism from transcription (Finkbeiner, 2001) to RNA stability and thus may have a major impact on cell function and fitness. For example, mutations in conserved intron sequences may lead to several

human diseases like the neurodegenerative disorders Friedreich ataxia, (Baralle et al., 2008; Lewandowska et al., 2005), or to spinal muscular atrophy (SMA) (Kashima et al., 2007). In addition, it is now accepted that introns carry a plethora of non-coding RNA signal sequences required for RNA modification (Fedorov et al., 2005; Lim et al., 2002; Ooi et al., 1998a), translational regulation and RNA degradation (Lin et al., 2008; Lin et al., 2006; Ying et al., 2008) Indeed, the majority of human microRNAs (miRNAs), implicated in RNA interference, and small nucleolar RNAs (snoRNAs), required for the modification of rRNA, are found in intronic sequence (Bortolin and Kiss, 1998; Lin et al., 2006; Tanaka-Fujita et al., 2007; Tanaka et al., 2000)

In budding yeast, only a minority of genes contains introns and only a handful of these may undergo alternative splicing (Parenteau et al., 2008). Splicing in yeast, however, plays an important role in regulating gene expression under specific conditions. For examples expression of the ribosomal protein *RPL32* is autoregulated through interaction between the Rpl32 protein and the sequence near the splice site of its own mRNA (Li et al., 1996; Vilardell and Warner, 1997) Splicing can also regulate the steady state level of gene expression. For example, the intronic sequence in the RNA binding protein *YRA1,* which couples transcription to export, was shown to reduce gene expression and its removal causes dramatic increase in expression that affects cell viability at high-temperature (Preker and Guthrie, 2006).

Therefore, splicing is not only important for generating mature mRNA but also for providing an additional regulatory layer that increases protein diversity and fine-tune gene expression.

3.3 Termination and 3'end formation

Traditionally, transcription termination and formation of the 3' end were considered two separate and sequential processes. However, recent studies are indicating that these two events are tightly linked and the interference with one may impair the other (Cui et al., 2008; Kaplan et al., 2005; Proudfoot, 2004). In most eukaryotes, the generation of the 3' end and transcription termination are initiated by cleavage of pre-mRNA 20-30 nucleotides upstream of the polyadenylation site (Figure 1). This endonucleolytic cleavage occurs within a consensus sequence of AAUAAA by a multisubunit cleavage / polyadenylation specificity factor (CPSF) (Zarudnaya et al., 2002). Once the CPSF generated the new mRNA 3' end, the poly (A) polymerase Pap1p uses it to catalyze the addition of up to 250 adenine residues to the cleaved 3'end of mRNA (Kuehn et al., 2009; Mandart and Parker, 1995). Normally, the addition of the canonical poly(A) tails by the CPSF / Pap1p machinery increases RNA stability and ensures RNA export to the cytoplasm (Noe et al., 1999). However, in yeast it was recently found that the addition of a short Poly(A) tail (20-40 adenines) by the non-conventional poly(A) polymerase Trf4p signals rapid RNA degradation (Figure 2) (Arigo et al., 2006; Neil et al.,

2009; Thiebaut et al., 2006). Thus, while long processive mRNA polyadenylation increases RNA stability, short disruptive poly(A) tails signal the rapid degradation of cryptic unstable transcripts "CUT" by exoribonuclease (Arigo et al., 2006; Pandey and Marzluff, 1987). However, not all mRNAs are regulated through polyadenylation. In metazoan cells, replication-dependent histone mRNAs are not poyadenylated (Pandey and Marzluff, 1987). In this case, formation of the 3' end of the mRNA occurs by endonucleolytic cleavage of pre-mRNA to release the mature form. This specific cleavage requires several trans-acting factors, including a protein, the stem-loop binding protein (SLBP), which binds to a 26-nucleotide long hairpin; and a small nuclear RNP, U7 snRNP (Davila Lopez and Samuelsson, 2008; Wagner and Marzluff, 2006). This indicates that Pol II transcription does not necessarily lead to the generation of polyadenylated RNA. Indeed, Pol II transcribes many non-coding RNAs that do not possess a poly(A) tail, like snRNAs and snoRNAs (Grzechnik and Kufel, 2008; Guffanti et al., 2006; Jacobs et al., 2004).

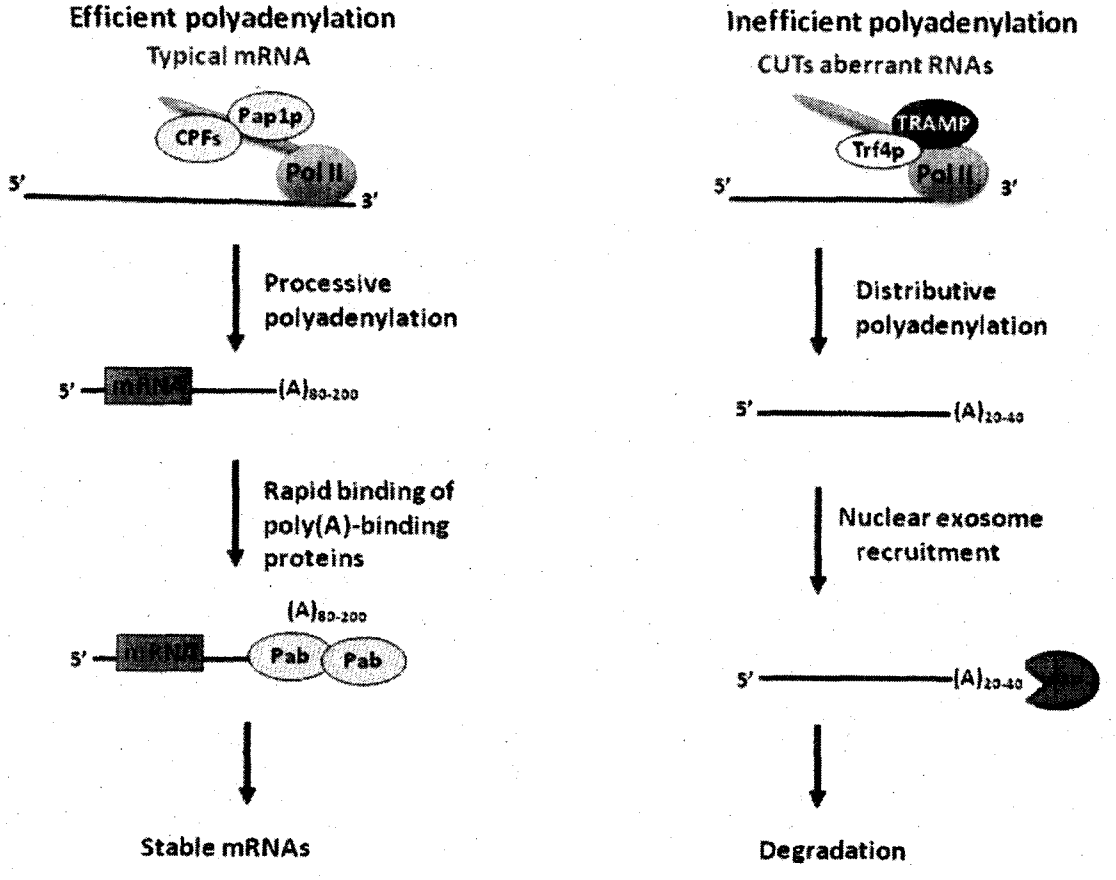


Figure 2. Models for polyadenylation dependent RNA stability.

Abbreviations: CPF, cleavage and polyadenylation factors of mRNAs; TRAMP: Trf4p/Air2p/Mtr4p polyadenylation complex. The lengths of the pol(A) tails added by each of the poly(A) polymerases are approximate (adapted from Chanfreau, 2005).

Non-coding RNAs, such as snRNAs and snoRNAs are synthesized as larger precursors by Pol II from independent transcription units, polycistronic precursors, or excised from introns (Lafontaine and Tollervey, 1995; Ooi et al., 1998b). In general, all Pol II transcribed snRNAs (U1, U2, U4 and U5) are generated as independent transcriptional units that do not require the polyadenylation machinery for the formation of their 3' ends (Forbes et al., 1983; Krol et al., 1983; Marz et al., 2008). Instead, the mature 3' end of these RNAs is generally determined by the binding sites of proteins involved in the assembly of the Spliceosomal RNP complex (Gornemann et al., 2005; Mougin et al., 2002). The assembly of these snRNAs into functional RNPs is essential for protecting the mature 3' end from exoribonuleolytic cleavage (Staley and Woolford, 2009). Like snRNA, snoRNA mature 3' ends are marked and maintained by the binding sites of protein components of the snoRNP (Ballarino et al., 2005; Morlando et al., 2004; Verheggen et al., 2002). It is widely accepted that methylation snoRNAs are stabilized by the binding of C/D box protein complex, while pseudouridylation snoRNA are maintained by the binding of H/ACA box proteins (Kiss et al., 2006; Morlando et al., 2004; Preti et al., 2006). However, the exact protein component that marks the 3' end is not clear. In general, it is believed that transcription termination and formation of the 3' end of the majority of pre- snRNA and snoRNA at least in yeast involves a complex of two RNA-binding proteins, Nrd1 and Nab3, and a putative RNA helicase, Sen1. Nrd1 interacts with the C-terminal domain (CTD) of Pol II and with the exosome to link termination

with processing (Carroll et al., 2004; Steinmetz et al., 2001). The maturation of non-coding RNA is then completed by the 3' end trimming of exonucleases and the assembly of the snoRNPs complex. Recently, it was proposed that the recruitment of the machinery required for 3' end formation to the transcription termination site is dictated by the phosphorylation state of the Pol II C-terminal domain (CTD) depends on the transcript size (Figure 3) (Gudipati et al., 2008; Vasiljeva et al., 2008). In particular this idea states that in the early state of transcription,Ser5 residues of the CTD become hypermethylated and thus recruit the Nrd1 termination complex, which has a preference for the Ser5 phosphorylated CTDs (Gudipati et al., 2008). This binding triggers the formation of the 3' end of short RNA transcripts like snRNA, snoRNA or cryptic non-polyadenylated transcripts destined for degradation (Gudipati et al., 2008; Vasiljeva et al., 2008). As the transcript length increases, the phosphorylation of Pol II Ser5 decreases and the phosphorylation of Ser2 increases favoring the binding of the polyadenylation dependent processing machinery leading to the formation of mRNA 3' ends (Gudipati et al., 2008; Morlando et al., 2004; Vasiljeva et al., 2008). It should be noted, however, that transcript size is in sufficient to indicate the nature of the 3' end. It is likely that the combination of the sequence near the transcription termination site and the length of the transcript determines the final outcome. Once the 3' end is formed, transcription is terminated and Pol II falls off its DNA template.

There are two major models proposed for transcription termination. The "anti-

terminator" or 'allosteric" model proposes that transcription through the termination signal changes the properties of the elongating Pol II complex (Figure 3) (Epshtein et al., 2007), perhaps by dissociation of positive elongation factors or recruitment of termination factors (Erie, 2002). In this model termination apparently occurs without cleavage as in the case of snoRNA or snRNA termination (Osheim et al., 1999; Steinmetz et al., 2001; Tran et al., 2001; Egloff et al., 2008; Richard and Manley, 2009). In the second so called "torpedo" model cleavage of the nascent RNA transcript at the poly (A) site transmits a signal to Pol II, leading to the destabilization of the elongation complex (Tollervey, 2004). This torpedo model is conserved between human and yeast (Luo and Bentley, 2004; West et al., 2004). Recently, it was shown that the 5' to 3' exoribonuclease Rat1p in yeast and its homologue in human Xrn2 induce termination by degrading the 3' end RNA fragment generated by the poly (A) processing machinery (West et al., 2004). It is believed the degradation of RNA near the RNAPII transcription complex destabilizes the complex and induces its dissociation from the DNA template. It remains unclear however, how these long non-polyadenylated transcripts like the spliceosomal component U2snRNA (Abou Elela and Ares, 1998) or the telomerase RNA (TLC1) (Chapon et al., 1997) are terminated. These RNAs are too long for NRD1 dependent transcription termination and do not require polyadenylation for 3' end formation.

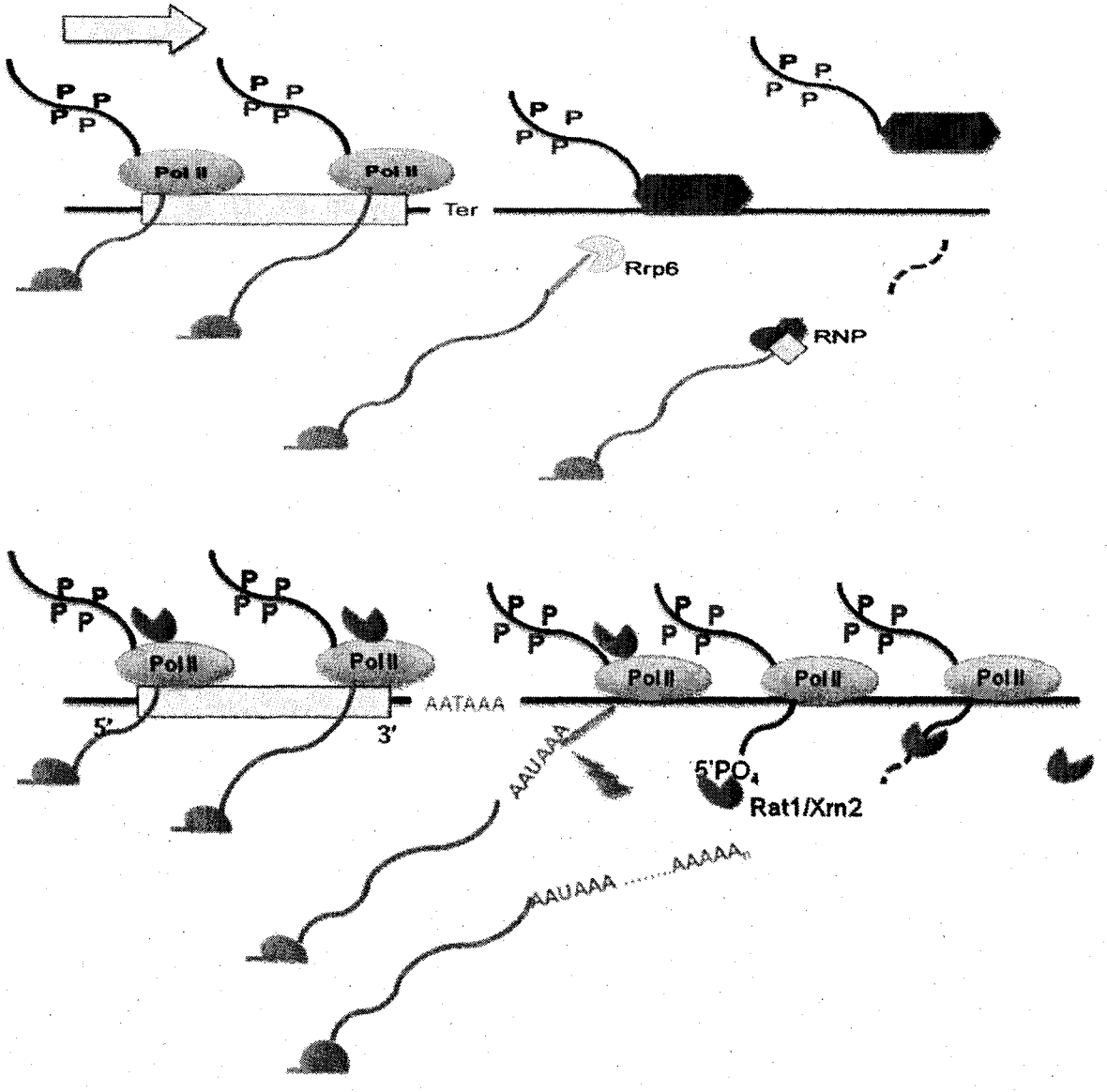


Figure 3. Transcription termination by RNA polymerase II.

(A) Allosteric model. During elongation within the gene (blue box), Pol II is in highly processive conformation (green oval), and changes in conformation (red ocatogan). Phosphorylation states of the Pol II CTD (black line) are marked in (red P). (B) Torpedo model. RNA downstream of the poly (A) cleavage site (blue line) is digested by a 5'-3' exonuclease Rat1 in yeast, and Xrn2 in humans (blue pacman) which tracks with Pol II and then leads to its dissociation (adapted from Luo and Bentley, 2004).

4. RNA turnover and degradation

RNA turnover and degradation are obligatory processes of every mRNA in cell. RNA turnover defines the natural cycle of RNA degradation in which RNA is degraded after a certain time following its transcription (Lombardo et al., 1992; Zhai et al., 2008). The time it takes any specific RNA to naturally degrade determines its relative stability, influences the number of times it is translated and the amount of proteins it can produce. Therefore, the half-life of each RNA must be programmed as a factor of its function. For example, house-keeping genes are generally transcribed into mRNAs with long half-lives like glycolytic enzyme GAPDH mRNA with half-life >24h (Lekas et al., 2000). On the other hand, proteins that are required only at particular times during the cell cycle, or during differentiation or growth have short half-lives such as *c-myc* mRNA (Loflin et al., 1999). The stability of the different mRNAs are generally determined by the length and structure of the 5' and 3' UTR as well as the length of the poly(A) tail (Bloch, 1999; Wang et al., 2005). The overall principals of mRNA degradation are conserved in both yeast and mammals. These features normally influence the degradation of mRNA in the cytoplasm, which is tightly linked to translation and often initiated by the removal of the 3' poly(A) tail followed by the removal of the 5' cap and exoribonucleolytic degradation (Santiago et al., 1987; Stripecke et al., 1994; Wang et al., 2005; Zhai et al., 2008). However, certain RNAs contain special destabilization sequence elements that signal the recruitment of *trans*-acting factors like ribonculeases (Tourriere et al., 2002). Destabilization elements

like the AU rich (ARE) decay signals (Maitra et al., 2008) or stem loop structures like the iron response elements (IRES) are usually found in the untranslated regions of mRNAs (Rothenberger et al., 1990; Constable et al., 1992; Thomson et al., 1999; Cairo et al., 2002;). The presence of these signals induces endoribonucleolytic cleavage and rapid RNA decay. Changing the activity of these elements has strong impact on cell metabolism and may lead to disease development (Palmer et al., 2008). RNA degradation can also be used as a quality control or surveillance mechanism to identify mRNAs in the cytoplasm lacking translation-termination codons (non-stop decay) (Vasudevan et al., 2002) or containing premature termination codons (nonsense decay) (Neu-Yilik and Kulozik, 2008) or that undergo translation stalling (no-go decay) (Passos et al., 2009).

RNA degradation can also occur in the nucleus (Kuai et al., 2005). In yeast, nuclear degradation routinely eliminates excised pre-rRNA spacer fragments, introns, and short cryptic RNAs (Allmang et al., 1999; van Hoof et al., 2000; Mitchell et al., 2003; Peng et al., 2003; Gonzales et al., 2005). In fact, given the large number of pre-rRNA, snRNA and snoRNA that is processed in the nucleolus or the nucleoplasm, one could imagine that the majority of RNA processing activities occur in the nucleus and not in the cytoplasm. In addition to this routine disposal of unused RNA spacers, nuclear degradation may also serve as an early quality control surveillance mechanism to eliminate transcriptional, processing and assembly errors (Skruzny et al., 2009). For example, premature or aberrant transcriptional termination may alter mRNA 3'

end formation and thus lead to the recruitment of the 3' – 5' nuclear exosome complex that degrades the unwanted product (Arigo et al., 2006). Nuclear surveillance is particularly important to eliminate the RNA components of unassembled or incorrectly assembled RNP complexes. Assembly of defective RNA would lead to generation of faulty ribosomes or spliceosomes that may compromise the survival of the entire cell. Indeed, any mutation that prevents protein binding to rRNA, snRNA or snoRNA often leads to rapid RNA degradation by the nuclear exosome complex (Lee et al., 1995; Lee and Nazar, 1997; Good et al., 1997; Hilleren et al., 2001; Morlando et al., 2004; Passos et al., 2009; Skruzny et al., 2009) In all cases, nuclear and cytoplasmic RNA decay and turnover are entirely dependent on the accuracy and efficiency of ribonucleases that quickly react to changes in inter- and intra-cellular conditions.

5. Ribonucleases

Ribonucleases (RNases) are enzymes that specifically cleave RNA phosphodiester bonds (Nicholson, 1999). Ribonucleases are divided into two classes, exoribonucleases and endoribonucleases (Nicholson, 1999). Most RNases are protein enzymes and will be discussed in details below. However, RNA based ribozyme activity has also been identified such as the tRNA processing RNase P and the mitochondrial RNA processing enzyme (MRP), which is required for pre-rRNA processing (Lindahl and Zengel, 1995; Reddy

and Shimba, 1995; Reilly and Schmitt, 1995; Tollervey, 1995). All known, exoribonucleases are protein enzymes that degrade RNA by using free 3' or 5'ends as entry sites, whereas endorinucleases recognize internal RNA sequence or structure (Deutscher, 1993; Virtanen and Astrom, 1997; Deutscher and Li, 2001; Brouwer et al., 2001; Andrade et al., 2009). These enzymes are found in all cells and each possesses special activity and specificity that suits particular cellular functions. The expression, specificity and recruitment of these ribonucleases plays an important role in shaping the RNA degradation and turnover program of all cells.

5.1 Exoribonucleases

Exoribonucleases are key components of the RNA-surveillance machinery and can be divided into 6 families based on their protein fold and activities (See table 1). These families (RNR, DEDD, RBN, PDX and RRP4) possess 3' - 5' exoribonuclease activity and one (5PX) includes RNases with 5' – 3' exoribonucleolytic activities (Zuo and Deutscher, 2001). In yeast, 3' – 5' exoribonculeases in general function as integrated complexes called exosomes that can be found in the nucleus and / or the cytoplasm (Decker, 1998; Raijmakers et al., 2004; Lykke-Andersen et al., 2009). Nuclear and cytoplasmic exosomes share ten common components from 4 different families (RNR, DEDD, PDX and RRP4) (Schneider et al., 2009). However, the RNase Ski7 are found exclusively in the cytoplasmic complex (Araki et al.,

2001) and the RNase Rrp6 and the putative nucleic-acid-binding protein Rrp47 is found only in the nuclear complex (Mitchell et al., 2003). Rrp6 is a key nuclear 3' – 5' nuclear exoribonuclease (Burkard and Butler, 2000). It directly contributes to the hydrolytic activity of the nuclear exosome and confers distributive exonucleolytic activity on unstructured and poly (A)-extended RNA (Graham et al., 2009). The nuclear exosome functions in the 3' processing of the precursors of stable RNAs, including rRNA and many snoRNAs (Kent et al., 2009; Kufel et al., 2000; Torchet et al., 2002). The nuclear exosome is also responsible for the surveillance and degradation of aberrant nuclear precursors of many types of RNA including pre-mRNAs, pre-tRNAs and pre-rRNAs.

The 5PX family currently contains two 5' – 3' exoribonucleases one is nuclear (Rat1p) (Li et al., 2006), and the other is cytoplasmic (Xrn1p). Xrn1p is the main cytoplasmic mechanism for uncapped RNA and also plays a major role in controlling the steady state level of mRNA in yeast (Brown et al., 2000; Long and McNally, 2003). Indeed, Xrn1p acts at the final step of mRNA degradation, following the normal deadenylation and decapping of mRNA. On the other hand, deletion of the nuclear exoribonuclease Rat1p impairs transcription termination and leads to the accumulation of many processing by-products (Henry et al., 1994). Both Rat1p and Xrn1p share similar substrate.

Superfamily	E.coli members	S.cerevisiae members	Distribution in other organisms	Cataytic features
RNR	RNase II RNase R	-	Most eubacteria	3' to 5' processive
	-	Rrp44 (Di3) Msu1 (Dss1) Ssd1	All eukaryotes	
DEDD	RNase D	Rrp6	All eukaryotes certain eubacteria	3' to 5' distributive Some have DNase activity
	RNaseT	-	γ-proteobacteria	
	oligoribonuclease	Ynt20 (Rex2)	All eukaryotes certain eubacteria	
	-	Pan2 Rex1,3,4	All eukaryotes	
RBN	RNase BN	-	Certain eubacteria	3' to 5' distributive
PDX	PNPase	-	Most eubacteria Certain eukaryotes	3' to 5' processive phosphate-dependent
	RNase PH	Rrp41-43,45,46 Mtr3	All kingdoms	3' to 5' distributive phosphate-depenedent
RRP4	-	Rrp4,40 Csl4	All eukaryotes Most archea	3' to 5' distributive
5PX	-	Xrn1 Rat1	All eukaryotes	5' to 3' processive

Table 1. Summary of exoribonuclease superfamilies.

(Zuo and Deutscher, 2001)

specificity and RNA that escapes Rat1p degradation in the nucleus is seamlessly degraded by Xrn1p in the cytoplasm (Poole and Stevens, 1995). In general, it seems that nuclear and cytoplasmic degradation share overlapping substrate specificity.

5.2 Endoribonucleases

Until recently, endoribonucleases were thought to play a minor role in RNA degradation. The fact that endoribonucleases recognize specific internal sequences or structural motifs make them more specialized than exoribonucleases that can recognize any RNA with free termini (MacBeth and Patterson, 1998; Kennell, 2002; Saida and Odaert, 2007). Therefore, it was thought that endoribonucleases affect a small group of transcripts with limited impact on overall gene expression. However, it is becoming increasingly clear that despite the relatively high specificity of the ribonucleases, they can still maintain broad substrate specificity. Indeed, RNA mediated endoribonucleolytic cleavage of RNA that is often called RNA interfrerence (RNAi) is rapidly becoming the largest and most studied mechanism of RNA degradation (Yamaguchi and Inouye, 2009). In yeast, there are 9 well characterized protein endoribonucleases with a broad range of activities including the debranching of excised intron lariats (e. g. Dbr1p) (Khalid et al., 2005), RNA decay activities (e.g. Dis3p) (Schaeffer et al., 2009), 3' endoribonucleolytic, ribonuclease activities (e. g. Ngl2p) (Faber et al., 2002),

and RNase III cleavage activities (Rnt1p) (Lamontagne et al., 2001). In higher eukaryotes, many of these enzymes are conserved; e.g RNase H or Dis3p, while other have evolved new functions such as yeast RNase III (Rnt1p) (Lamontagne et al., 2001; Carmell and Hannon, 2004; Ji, 2008; Lebreton et al., 2008; Schultz and Champoux, 2008). In higher eukaryotes orthologues of RNase III (Dicer and Drosha) induce and regulate the mechanism of RNA interference (RNAi) that seems to govern most conditional mRNA degradation in mammalian cells (Carmell and Hannon, 2004). This mechanism of RNA degradation is found in most eukaryotes with the notable exception of budding yeast. RNAi achieves RNA degradation through the formation of RNA duplex *in trans*. This process can be induced by the cleavage of long duplex RNA or through the processing of short structured pre-micro RNAs (miRNA) (Tijsterman and Plasterk, 2004; Filipowicz, 2005; Hutvagner, 2005; Kim et al., 2006a). In all cases, the final outcome is the formation of short RNA duplex of 21-22 base pairs that is simultaneously integrated into the RNA-induced silencing complex (RISC) that dissociates the RNA duplex and pairs one of the two RNA strands with its complementary target (Ji, 2008; Kim et al., 2006a). Perfect pairing with the target induces enodribonucleolytic cleavage of the target RNA by a 5' end-dependent endoribonuclease called argonaute (Brodersen and Voinnet, 2009). Imperfect pairing of the RISC associated RNA strand with the 3' end of the target RNA does not solicit cleavage but instead inhibits mRNA translation (Pillai et al., 2005; Pillai et al., 2007). In this

way, the cell can conditionally target endoribonucleolytic cleavage of any RNA using the same set of endoribonucleases.

6. RNase III family

Members of the RNase III family are found in all species examined with the exception of archaebacteria, where the functions of RNase III are carried-out by the bulge-helix-bulge nuclease (BHB) (Lamontagne et al., 2001). All proteins classified into this family show homology with the structural elements of the founding member, *Escherichia coli* RNase III (Nicholson, 1996, 1999). These structural elements include a nuclease domain (NUCD) that exhibits a conserved signature motif, and a dsRNA binding domain (dsRBD) that contains a motif specific to the dsRNA binding protein (dsRBP) family. The RNase III family can be divided into 4 classes, based on additional protein features and organizations (Figure 4) (Lamontagne et al., 2001). Class I includes bacterial enzymes that possess a single N-terminal NUCD and a C-terminal dsRBD (Gan et al., 2006). Class II enzymes are identified by the presence of a highly variable N-terminal extension and include fungal RNase III (Lamontagne and Abou Elela, 2001). Class III enzymes contain two NUCDs and include plant and vertebrate enzymes (Lee et al., 2003b). Class IV includes the RNAi enzyme Dicer, which possesses a N-terminal helicase domain (Lee et al., 2004). The sequence homology between orthologues varies between 84% to 20%, depending on the evolutionary distance. Most

RNase IIIs display low sequence specificity *in vitro* and usually cleave any duplex RNA with low sequence complexity (Lamontagne and Abou Elela, 2004). In contrast, RNase IIIs are highly specific and mostly target short RNA hairpins *in vivo*. This surprisingly high *in vivo* specificity prevents complementation, even between closely related species.

6.1 Protein structure of budding yeast RNase III

Budding Yeast Rnt1p is a 471 aa protein (54.5 kDa) that exhibits the features of class II RNase IIIs. The dsRBD motif is located at the C-terminus (positions 372-440) and has about 25% identity with other RNase IIIs. A unique 32 aa extension at the C-terminus is required for nucleolar localization (Lamontagne et al., 2000). The Rnt1p 162 aa NUCD contains the RNase III signature sequence implicated in catalysis. In addition, Rnt1p possesses a 199 aa N-terminal domain (N-term), that is unique to eukaryotic RNase IIIs. This N-terminal extension has no apparent functional motifs and it is not highly conserved, even among members of the *Saccharomyces* species. Deletion of the N-term renders Rnt1p salt sensitive and reduces cleavage efficiency both *in vivo* and *in vitro*. Rnt1p functions as a homodimer formed by interactions between the dsRBD and the N-term (Lamontagne and Abou Elela 2001).

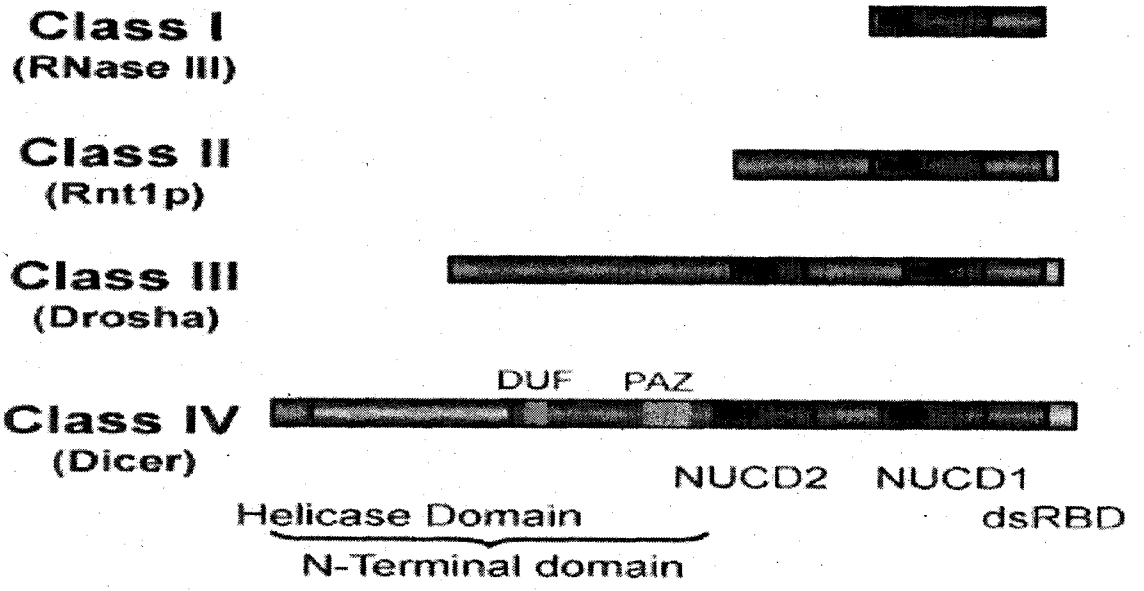


Figure 4: Schematic representation of the RNase III family.

Green boxes represent the dsRNA binding domain (dsRBD), yellow boxes represent the amino acid residues that extend beyond the dsRBD (CTE), red boxes represent the nuclease domain (NUCD). Blue boxes represent the N-terminal domain.

Based on the crystal structure of two bacterial RNase IIIs, specifically those of *Aquifex aeolicus* (Gan et al., 2006), the Rnt1p homodimer is predicted to form an antiparallel dimer (Figure 5B). If Rnt1p folds like bacterial RNase III, the two NUCDs of the homodimer would form a valley supported by a ball-and-socket junction involving highly conserved amino acids. Secondary structure prediction suggests that Rnt1p maintains the conserved $\propto\beta\beta\beta\propto$ structure of dsRBD and of several of the seven helices in the bacterial NUCD. Recently, crystal and solution structures of Rnt1p dsRBD confirmed the presence of the classical $\propto\beta\beta\beta\propto$ structure, and revealed an additional helix near the C-terminus ($\propto$ 3) unique to Rnt1p (Figure 5 A) (Leulliot et al., 2004; Wu et al., 2004). The solution structure of the dsRBD RNA complex indicates that the additional helix is not located near the RNA, but that it could influence the binding of RNA to $\propto$ 1.

Studies of E.coli RNase III suggest that the substrate selection is influenced by antideterminant nucleotides (Rudinger et al., 1996; Zhang and Nicholson, 1997). This means that the absence of a nucleotide or structure, and not its presence, triggers RNA binding and cleavage. As more RNaseIII are tested, it is becoming increasingly clear that eukaryotic RNase IIIs possess a different mechanism of substrate selectivity. Unlike other RNase III, Rnt1p recognizes substrates with conserved stem-loop structures. Most Rnt1p substrates exhibit conserved AGNN tetraloop structure (Lamontagne et al. 2001; Lebars et al., 2001; Lamontagne et al., 2003; Lamontagne et al., 2004).

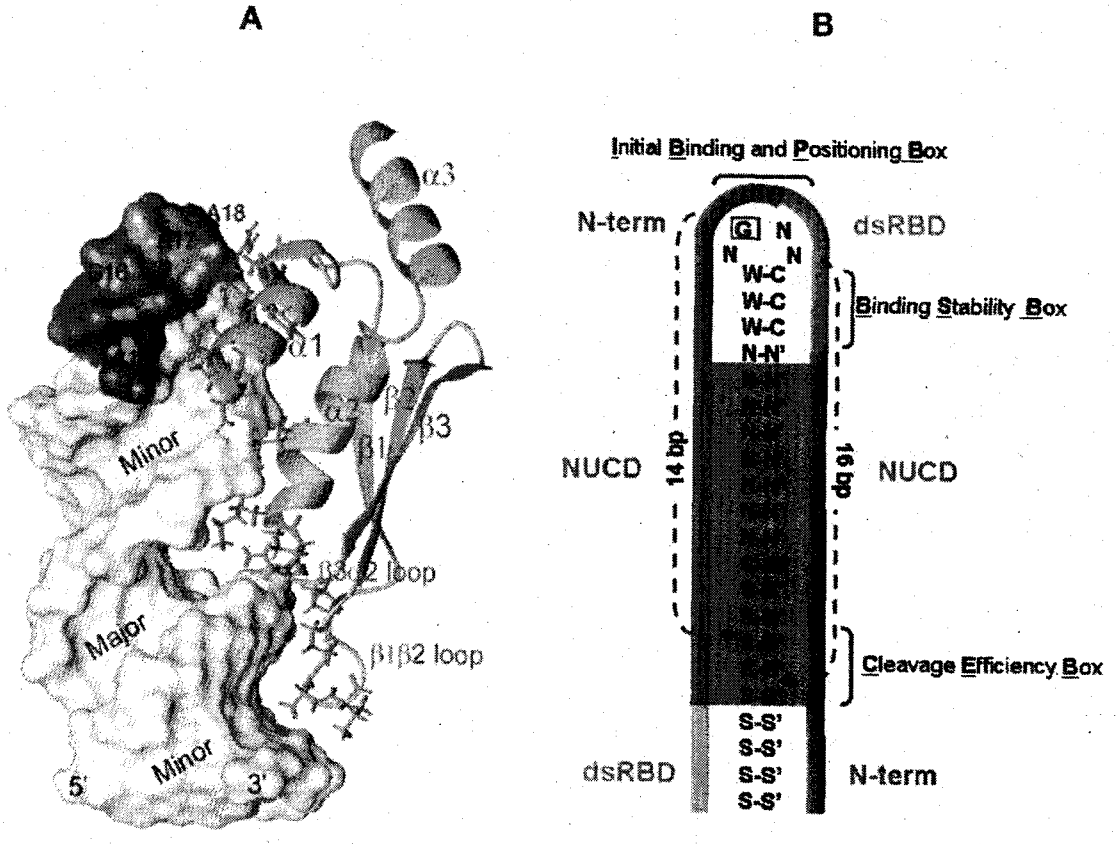


Figure 5. Rnt1p mechanism of action.

(A) Interactions between Rnt1p dsRBD and snR47 RNA. Solvent accessible surface of the RNA showing the major and minor grooves with AGAA tetraloop nucleotides colored red, blue, and orange, and the protein in yellow ribbon. Protein side chains that interact with the RNA are shown in sticks. Adapted from Wu et al 2004. (B) A model for the mode of action of Rnt1p. Rnt1p adopts a homodimer confirmation in the presence of RNA. Sequences influencing binding and cleavage are indicated. Arrows showing the positions of the cleavage sites between 14 and 16bp from the tetraloop NGNN.

Rnt1p cleaves at a fixed distance from the conserved loop, generating a product with staggered ends (Figure 5B). The solution structures of two different substrates of Rnt1p reveal a common fold for the terminal loop with the universal G in *syn* conformation and extensive base stacking (Lebars et al., 2001). The structure suggests that Rnt1p recognizes the shape of the tetraloop for the initial interaction with its substrate. This is in marked contrast with the recognition mode of most dsRNA binding proteins including RNase III, which interact primarily with the minor groove of the double helix and recognize the shape of the A-form dsRNA (Wu et al., 2004).

6.2 Cellular functions of yeast RNase III

RNT1 is not an essential gene but its deletion causes severe growth defects, temperature sensitivity, hypersensitivity to heavy metals and sporulation defects (Abou Elela et al., 1996; Abou Elela and Ares, 1998). Normal wild type cell doubling is around 2-3 hours at 30°C, however cells lacking Rnt1p will divide in around 7 hours at 26°C (Abou Elela and Ares, 1998). Deletion of Rnt1p also affects cell cycle progression and nuclear division; yet this effect has been shown to be independent of Rnt1p cleavage activity (Catala et al., 2004). Rnt1p localizes to various places depending on the phases of the cell cycle; in G1 till the end of S phase, Rnt1p is localized to the nucleolus, in G2 phase to the end of mitosis, Rnt1p is present at the nucleoplasm. However, Rnt1p is not detected in the cytoplasm, even when it is overexpressed. Rnt1p

is also important for the co-localisation of nuclear proteins implicated in the maturation of rRNA. Rnt1p is shown to bind to Gar1p and this interaction is required for the nuclear localization of Gar1p, Nhp2 and Cbp5p that are involved in the processing of H/ACA snoRNAs (Tremblay et al., 2002b).

The primary function of Rnt1p is the processing of pre-rRNA. This processing is achieved by the cleavage of a stem-loop structure at the 3' external transcript spacer (3'ETS) of the pre-rRNA (Abou Elela et al., 1996). Recenlty, it has been shown that Rnt1p is not only implicated in the maturation of rRNA but also its transcription. Rnt1p interacts with two subunits of the polymerase I (Pol I) (Catala et al 2008). Moreover deletion of Rnt1p inhibits the synthesis of rRNA and alters the conformation of the chromatin at the ribosomal DNA locus (Catala et al., 2008). Finally, Rnt1p is also required for the termination of rDNA transcription by Pol I (Prescott et al., 2004; Kawauchi et al., 2008).

Other functions of Rnt1p include the processing of all Pol II transcribed snRNAs (Chanfreau et al., 1997; Abou Elela and Ares, 1998; Seipelt et al., 1999) and a number of snoRNAs (Chanfreau et al., 1998a). Initially, Rnt1p cleaves a stem-loop structure capped with the canonical AGNN tetraloop at the 3'end of the snoRNAs. This cleavage allows the trimming by the exosome and further assembly of the Ribonucleoproteins (RNP) to complete the maturation. In addition to the role of Rnt1p in the maturation of non-coding RNAs, Rnt1p can also trigger mRNA degradation. In this case, Rnt1p cleavage occurs in the coding sequence of the mRNA to control its expression in response to cellular environmental cues (Ge et al., 2005;

Larose et al., 2007). The presence of a Rnt1p cleavage signal in an mRNA coding sequence triggers the degradation of the nuclear fraction of the RNA presumably by the nuclear exosome. It is not clear, however, how Rnt1p differentiates between mRNA scheduled for degradation from mRNA exported to the cytoplasm for translation. It is also unclear how Rnt1p cleavage is activated by environmental and cellular conditions.

7. Aim of the project

In the beginning of this study only a handful of Rnt1p cleavage signals in non-coding RNA were known and the mechanism by which the enzyme selects its substrate was restricted an apparent preference for AGNN hairpins. Therefore, we aimed at understanding the mechanism by which Rnt1p identifies its substrate and developed tools to find and characterize new Rnt1p substrates. This essential biochemical approach led us to discover new forms of Rnt1p substrates that do not require an AGNN tetraloop for cleavage and identified new and unexpected functions of Rnt1p in transcription termination. The initial work succeeded in identifying Rnt1p cleavage sites in all known non-coding RNAs and revealed a new mechanism by which long-range interaction induces excision of intron embedded snoRNA. Examination of the newly discovered Rnt1p cleavage sites suggested that the enzyme uses a flexible substrate recognition mechanism that tolerates broad variation in primary and tertiary structures. Based on this finding, we modified our

model for substrate recognition and scanned the yeast genome for new forms of cleavage targets that are not associated with snoRNAs. Surprisingly, we uncovered a large number of Rnt1p cleavage sites in intergenic sequences. Biochemical and genetic studies of the intergenic cleavage signals suggested a new role of Rnt1p in inducing transcription termination of long non-polyadenylated RNAs. In addition, this study directly linked dsRNA specific endoribonucleolytic activity to the transcription complex. The central findings of this thesis challenge the common views of distinct RNA transcription, processing and decay and provide a model in which gene expression is not simply defined by the decision to transcribe a gene but rather by the stability of the nascent transcripts.

ARTICLE 1

Ghazal, G., Ge, D. Gervais-Bird, J., Gagnon, J., and Abou Elela, S. (2005) Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. Mol Cell Biol. 2005 April; 25(8): 2981–2994.

Préambule

Dans cette étude, nous avons cherché de nouveaux substrats de Rnt1p en examinant le profil d'expression de tous les snoRNAs en présence et en absence de Rnt1p. En parallèle, nous avons développé un programme qui identifie les signaux de clivage de Rnt1p connus à proximité des séquences des snoRNAs. Une combinaison d'approches*in silico* et *in vitro* a identifié tous les substrats connus de Rnt1p et révélé 7 nouveaux substrats snoRNAs. Une enquête minutieuse du rôle de Rnt1p dans la maturation de ces snoRNAs a montré que les signaux de clivage de Rnt1p sont plus grands que ce que l'on croyait auparavant. Ce travail montre l'implication de Rnt1p dans la maturation des différentes organisations des snoRNAs. Fait intéressant, nous avons montré que la maturation des snoRNAs encodés dans les introns des gènes ribosomales représente un nouveau mécanisme qui coordonne la production des isoformes des protéines ribosomales et de leurs snoRNAs associés. J'ai effectué toutes les expériences à l'exception de l'analyse bioinformatique et un des Northern blots montré dans la Figure 3.

ARTICLE 1

Ghazal, G., Ge, D. Gervais-Bird, J., Gagnon, J., and Abou Elela, S. (2005) Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. Mol Cell Biol. 2005 April; 25(8): 2981–2994.

Summary

In this study we have searched for new Rnt1p substrates by examining the expression profile of all known snoRNAs before and after the deletion of Rnt1p. In parallel, we have developed a program that identifies Rnt1p cleavage signals near known snoRNA sequences. A combined *in silico* and *in vitro* approach identified all known substrates of Rnt1p and revealed 7 new snoRNA associated substrates. Careful investigation of the role of Rnt1p in the maturation of these snoRNAs showed that Rnt1p cleavage signals are larger than what was previously believed. This work shows the implication of Rnt1p in the processing of different snoRNAs organization. Interestingly, we have shown that processing of snoRNAs imbedded in the introns of ribosomal genes represents a new mechanism that coordinates the production of ribosomal protein isoforms and their associated snoRNAs. I conducted all experimental data with the exception of the bioinformatic analysis and one of the Nothern blots shown in Figure 3.

Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals

Ghada Ghazal, Dongling Ge, Julien Gervais-Bird, Jules Gagnon, and Sherif Abou Elela*

RNA Group / Groupe ARN

Département de Microbiologie et d'Infectiologie,

Faculté de Médecine, Université de Sherbrooke,

Sherbrooke, Québec, Canada J1H 5N4

Running Title: Genomic Approach for the Identification of Rnt1p Substrates

Keywords: RNase III, Rnt1p, snoRNA, dsRNA, Processing.

Word Count: 54.942

Abstract

In *Saccharomyces cerevisiae*, the maturation of both pre-ribosomal RNA (pre-rRNA) and pre-small nucleolar RNAs (pre-snoRNAs) involves common factors, thereby providing a potential mechanism for the co-regulation of snoRNA and rRNA synthesis. In this study we examined the global impact of the dsRNA specific ribonuclease Rnt1p, which is required for pre-rRNA processing, on the maturation of all known snoRNAs. *In silico* searches for Rnt1p cleavage signals, and genome-wide analysis of the Rnt1p-dependent expression profile, identified 7 new Rnt1p substrates. Interestingly, two of the newly identified Rnt1p-dependent snoRNAs snR39 and snR59 are located in the introns of the ribosomal proteins (r-protein) genes RPL7A and RPL7B. *In vitro* and *in vivo* experiments indicated that snR39 is normally processed from the lariat of RPL7A, suggesting that the expressions of RPL7A and snR39 are linked. In contrast, snR59 is produced by a direct cleavage of the RPL7B pre-mRNA indicating that a single pre-mRNA transcript cannot be spliced to produce a mature RPL7B mRNA and processed by Rnt1p to produce a mature snR59 simultaneously. The results presented here reveal a new role of yeast RNase III in the processing of intron-encoded snoRNAs that permits independent regulation of the host mRNA and its associated snoRNA.

Introduction

Bacterial pre-rRNA processing is carried out by a defined set of nucleases (Apirion, 1983; Apirion and Gegenheimer, 1981; Apirion and Miczak, 1993; Perry, 1976; Srivastava and Schlessinger, 1990). Key among this set is RNase III, initially isolated by its ability to bind and cleave duplex RNA (Robertson, 1967; Robertson et al., 1968). RNase III generates the immediate precursors to the mature 16S and 23S rRNA from the primary transcripts by cleaving within two extended RNA duplexes formed by long range interactions that pair the termini of each rRNA (Bram et al., 1980; Young and Steitz, 1978). These long-range interactions provide a simple method of coordinating the processing events at both ends of the transcript. In eukaryotes, pre-rRNA processing is more complex and requires many more snoRNAs and protein components with overlapping functions (Eichler and Craig, 1994; Fromont-Racine et al., 2003; Granneman and Baserga, 2004; Pederson, 1998; Reeder, 1990). For example, the removal of the 5' external transcribed spacer (ETS) requires 4 snoRNAs (U3, snR30, U14, and snR10), and about 64 snoRNAs are required for rRNA modifications (Lafontaine and Tollervey, 1995; Venema and Tollervey, 1995). snoRNAs are divided in two major subclasses: the first includes box C/D snoRNAs that mostly function as a guide for the methylation of rRNA (Bachellerie and Cavaille, 1997; Kiss-Laszlo et al., 1996; Kiss-Laszlo et al., 1998; Tycowski et al., 1996); while, the second includes H/ACA snoRNAs that guide RNA pseudouridine formation (Lafontaine et al., 1998; Ni et al., 1997; Watkins et al., 1998). Most mammalian snoRNAs are encoded within intron

sequences and are processed from either unspliced precursors or lariat species (Hirose et al., 2003; Hirose and Steitz, 2001; Zhou et al., 2004). In yeast, most snoRNAs are transcribed either as independent units, or as a part of polycistronic transcript, while only 7 of the 66 known snoRNAs are located in the introns of mRNAs (Filipowicz et al., 1999; Petfalski et al., 1998; Tollervey and Kiss, 1997). Several polycistronic snoRNAs, and few monocistronic ones are processed by Rnt1p, the orthologue of the bacterial RNase III (Lamontagne et al., 2001), which is also required for the processing of the pre-rRNA's 3' end (Abou Elela et al., 1996; Chanfreau et al., 1998a; Chanfreau et al., 1998b; Kufel et al., 1999; Lee et al., 2003a). Following processing by Rnt1p, the RNAs are trimmed by exonucleases producing the mature ends (Kufel et al., 2000; van Hoof et al., 2000).

Unlike other RNase IIIs, Rnt1p recognizes substrates with conserved stem-loop structures and has a low affinity for generic RNA duplexes (Lamontagne and Abou Elela, 2004). Most Rnt1p substrates exhibit a conserved AGNN tetraloop structure (Chanfreau et al., 2000; Lamontagne et al., 2003; Lebars et al., 2001; Wu et al., 2001). Rnt1p, cleaves at a fixed distance from the conserved loop, generating a product with staggered ends (Lamontagne et al., 2003). Mutations (Lamontagne et al., 2003), chemical protection assay (Lamontagne and Abou Elela, 2004), chemical interference (Chanfreau et al., 2000) and NMR analysis (Lamontagne et al., 2003) indicate that Rnt1p binding and cleavage are regulated by reactivity epitopes grouped into three

boxes (Fig. 1A). These are the initial binding and positioning box (IBPB), located at the tetraloop; the binding stability box (BSB), located adjacent to the tetraloop; and, the cleavage efficiency box (CEB), located near the cleavage site. Alteration of both the sequences of the IBPB and the BSB inhibits cleavage and reduces binding, while alteration of the CEB sequence inhibits cleavage without affecting the binding efficiency. Thus, despite the lack of universally conserved residues, the nucleotide composition of the reactivity epitopes contributes to substrate selectivity. The second nucleotide of the IBPB is believed to be universally conserved and changing it to any nucleotide other than G reduces binding to known substrates and blocks cleavage (Chanfreau et al., 2000; Lamontagne and Abou Elela, 2004; Lamontagne et al., 2003; Nagel and Ares, 2000). Recently, the solution structure of the Rnt1p substrate complex indicated that the enzyme interacts with the minor groove adjacent to the 3' end of the tetraloop, and it was suggested the substrate recognition depends on the shape of the groove (Wu et al., 2004). However, accurate identification of the universal features of Rnt1p substrates requires the identification of a large set of substrates that allows statistical analysis of the cleavage signals.

In this study we have searched for new Rnt1p substrates by examining the expression profile of all known snoRNAs before and after the deletion of Rnt1p. In parallel, we have developed a program that identifies potential Rnt1p cleavage signals near known snoRNA sequences. All newly identified

substrates were tested for cleavage *in vitro,* and their contribution to snoRNA processing was verified *in vivo*. Our combined *in silico* and *in vitro* approach identified all known substrates of Rnt1p and revealed 7 new snoRNA substrates. In general, monitoring the expression of snoRNAs was most effective when Rnt1p cleavage is not redundant with other processing events that could lead to the maturation of the snoRNA in question. In contrast, the *in silico* screen was most effective in identifying snoRNAs that harbor conserved processing signals, regardless of their processing pathway *in vivo*.

Materials and Methods

Strains and plasmids

Yeast cells were grown and manipulated using standard procedures (27, 52). The effect of Rnt1p depletion was studied using the strains W303-1A and Δ*rnt1* (18). The Δ*rnt1* Δ*dbr1* strain (Ooi et al., 1998b)was constructed by crossing Δ*rnt1* cells with Δ*dbr1* cells kindly provided by J. D. Boeke, Johns Hopkins University. The temperature sensitive strain was a recreation of the *rnt1-ts* strains described earlier (Nagel and Ares, 2000). The temperature sensitive strain *prp2ts* was a kind gift from Ren-Jang Lin, City of Hope.

Microarray Based Analysis of snoRNA Expression

The microarray experiment was conducted and analyzed by the Genome Quebec Innovation Center (Montreal, QC, Canada). The RNA was extracted from W303, Δ*rnt1* and *RNT1-TS* cells (Lamontagne et al., 2000; Nagel and Ares, 2000) grown at either 26°C or 37°C in either YC complete media or YC – leu media (Tremblay et al., 2002a).

In silico Screen of Rnt1p Substrates

The sequence homology score was calculated using two methods. One looks for homology to the sequences conserved in all known substrates and the other uses an algorithm that searches for the best sequence homology to any single known substrate. For both methods, a nucleotide probability matrix was generated from the alignment of known substrates using their tetraloop as an anchor point. In the second method, a score was given to each substrate based on its sequence homology to all known substrates and the sum of the probability of its nucleotides in relation to the distance to the tetraloop. Higher significance was given to the nucleotides near the tetraloop. In order to identify the best sequence homology to a known substrate, an intermediate score was weighted for each known substrate. Only the highest intermediate score was kept. The intermediate scores were calculated by comparing the nucleotides of potential substrates to those of known substrates. For each nucleotide

comparison, when the two nucleotides were identical, the intermediate score was raised by a distance-weighted factor. When the two nucleotides were not identical, the intermediate score was only raised according to the probability matrix multiplied by the weight factor. In addition to thermostability, an evaluation of the secondary structure of the potential substrates was based on the quality of its stem. In this study, this was calculated by giving a positive score to nucleotides downstream of the tetraloop that pair to upstream nucleotides and vice-versa. This score was weighted according to the nucleotide's distance to the tetraloop, with higher significance being given to the nucleotides close to the tetraloop. For any potential substrate, the sum of its nucleotide scores represented the quality of its stem.

In vitro RNA cleavage

Cleavage reactions were performed essentially as described earlier (Lamontagne et al., 2003) using either 0.2 pmol of recombinant Rnt1p (Lamontagne and Abou Elela, 2001), or total cell extracts (Lamontagne et al., 2004). For the *in vitro* cleavage assay, 2 fmol of internally labeled RNA were incubated in the presence of either 0.2 pmol of Rnt1p for 20 minutes at 30°C in 20µl reaction buffer (30 mM Tris pH 7.5, 5 mM spermidine, 10 mM MgCl2, 0,1 mM dithiothreitol (DTT), and 0,1 mM EDTA pH7.5). Yeast extracts were

prepared using three liters of yeast culture (W303 or Δ*RNT1* strain) grown to 0.8 O.D.600 at 26°C in YEPD. Cells were harvested by centrifugation, washed, and resuspended in 0.4 times the cell pellet's volumes of AGK buffer (10 mM HEPES, pH 7.9, 1.5 mM MgCl2, 200 mM KCl, 10% glycerol, 1mM PMSF, 1 mM benzamidine, 1 µg/ml leupeptin, 1 µg/ml aprotinin, 1 µg/ml pepstatin A, and 1 µg/ml antipain). Following cell lysis in liquid nitrogen, the frozen powder was transferred to a centrifuge tube and centrifuged at 18900 x g for 30 min. The supernatant was then centrifuged at 94000 x g for 30 min and dialyzed for 3 h against 2 liters of dialysis buffer (20 mM HEPES, pH 7.0, 0.2 mM EDTA, 0.5 mM DTT, 50 mM KCl, and 20% glycerol). Finally, the extract was centrifuged at 18900 x g for 20 minutes and the supernatant was stored at -80°C. The model snR55 substrate was generated by T7 RNA polymerase (Abou Elela and Ares, 1998). 5'end labeled RNA was produced as described (Lamontagne and Abou Elela, 2004). The snR55, snR56, and snR48 templates were prepared by *in vitro* transcription using PCR products as templates. Each PCR product was obtained from genomic DNA using a forward primer (F) carrying a T7 promoter located 200-300 nts upstream of the mature 5' end of snoRNA and a reverse primer (R) located 200-300 nts downstream of the mature 3' end of the snoRNA. The oligonucleotides are listed in supplementary oligonucleotides list 1. Cleavage of total RNA extracted from both wild type and Δ*rnt1* cells was conducted as described earlier (Catala et al., 2004) using total RNA (50 µg) incubated with recombinant Rnt1p (10 pmol) in the reaction buffer described above.

Northern blot analysis

Northern blot analyses were performed with total RNA (10-15 µg) run on 4-8% denaturing polyacrylamide gel as described earlier (Abou Elela and Ares, 1998). The RNA was visualized using either randomly labeled probes, or 5' end labeled oligonucleotides. The oligonucleotides used are listed in supplementary oligonucleotides list 2.

Primer extension

Primer extension reaction were performed as described earlier (Abou Elela et al., 1996). The oligonucleotides used for the primer extensions of snR50, snR52, snR54, snR56, snR57, snR58, snR59 snR60, snR62, snR64, snR67, snR68, snR69 and snR71 are listed in supplementary oligonucleotides list 2. Oligonucleotides specific to snR67, snR55 and snR39 are listed in supplementary oligonucleotides list 3.

Result

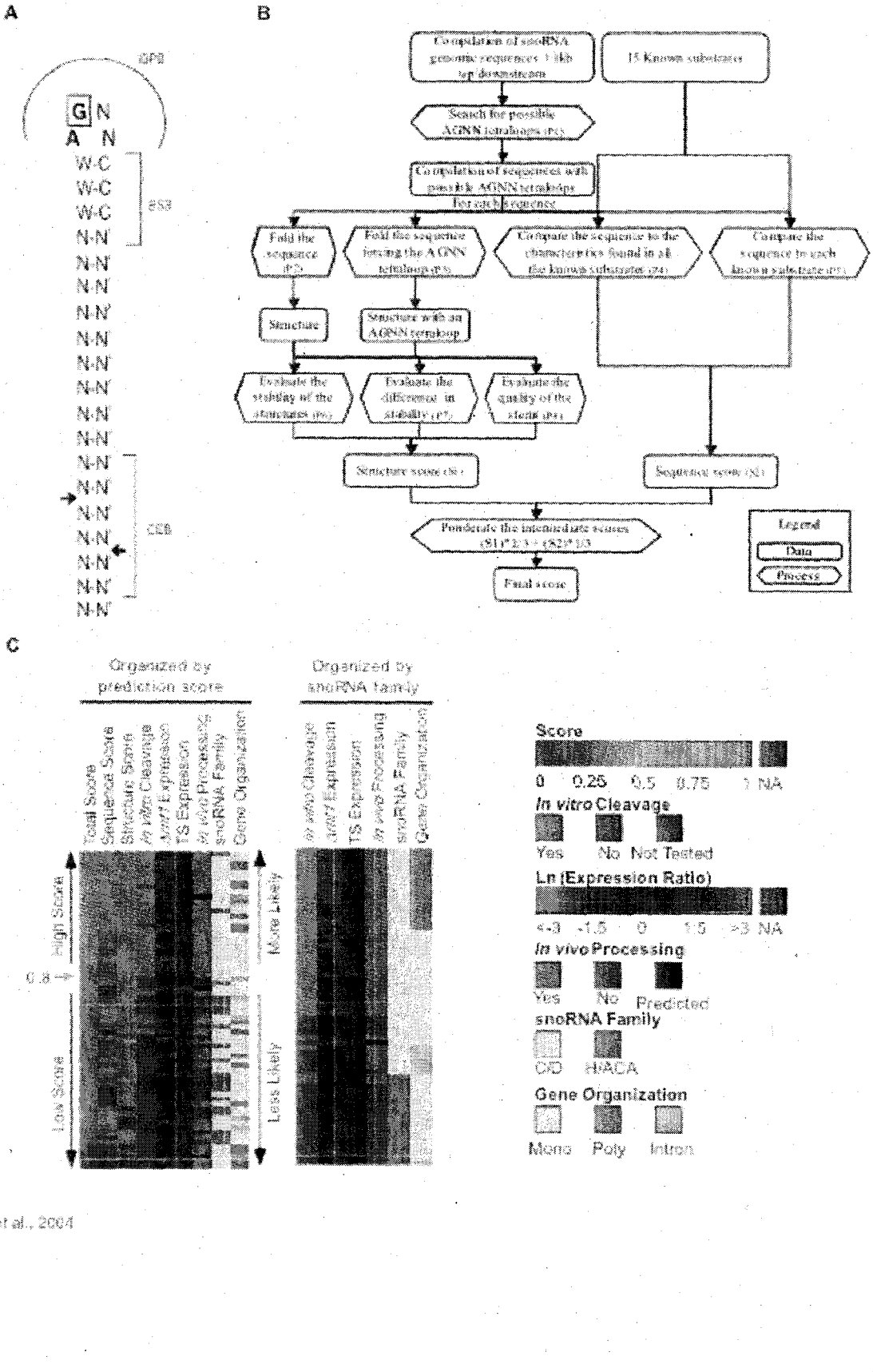
In silico and *in vivo* search for Rnt1p dependent snoRNA

In order to identify snoRNAs that require Rnt1p cleavage for their maturation we screened the sequences within 1Kb upstream and downstream of all

known yeast snoRNAs for stem-loop structures that resemble known Rnt1p substrates (Fig. 1A). As indicated in figure 1B, the sequences were scanned for the presence of an AGNN tetraloop followed by three Watson and Crick base pairs. Scores were given for each substrate based on sequence homology, secondary structure stability and secondary structure similarity to known substrates. Based on the scores of known substrates we have set the score cutoff to 0.8. Above this cutoff the identified RNA structures are expected to be cleaved by Rnt1p. A total of 64 snoRNAs with potential stem-loops were identified, but only 26 obtained a score above 0.8 and two did not associate with any predicted structure (Fig. 1C). The other 38 snoRNAs were found near structures with a score lower than 0.8. As expected, all tested structures with a score higher than 0.8 were cleaved by Rnt1p *in vitro*, validating the efficiency of the selection scheme (Supplementary table 1). Interestingly, most of Rnt1p substrates with score higher than 0.8 were found near C/D box snoRNAs and only three were found near H/ACA box snoRNAs (Fig. 1C and supplementary table 1). The 38 snoRNA associated structures with a score lower than 0.8 include 2 cleaved by Rnt1p in a loop independent but Nop1-dependant manner (16) and 5 associated with snoRNAs expressed as part of polycistronic units and processed using a stem associated with an adjacent snoRNA within the polycistronic unit. On the other hand, 22 other snoRNAs associated with structures that were not cleaved by Rnt1p and their processing were not affected by RNT1 deletion in vivo (Supplementary Table 1 and reference therein). The snoRNA snR17b (U3) carried a cleavage

signals identical to that of snR17a and thus was presumed to be processed by Rnt1p (Kufel et al., 2000) but was not directly tested. Finally, 4 snoRNAs were processed by Rnt1p using structures that differ from the canonical AGNN tetraloop motif that was used in the search for Rnt1p cleavage sites (Fig. 1). Only 4 AGNN stem-loop structures with a score between 0.675 - 0.8 were cleaved by Rnt1p *in vitro*. Therefore, by using a score cutoff of 0.8 we have missed 4 substrates with canonical AGNN tetraloop resulting in a false negative rate of 11%. *In silico* search using NGNN as the starting motif did not identify new snoRNA associated motifs that are affected by the deletion of RNT1 *in vivo* (data not shown).

In order to assess the efficiency of the AGNN tetraloop as indicator for Rnt1p dependent snoRNAs, we have examined the expression profile of snoRNAs in the presence and absence of Rnt1p. The microarray based expression profile of wild type cells was compared to that of cells carrying a complete deletion of Rnt1p. In parallel, the snoRNA expression profile of cells carrying a temperature sensitive allele of Rnt1p grown at the permissive temperature was compared to that of cells grown at the restrictive temperature. It is important to note that the microarray analysis will not differentiate between precursor and mature snoRNA and therefore the increase in the expression level of any snoRNA could reflect an accumulation of a pre-snoRNA, mature



Ohazal et al., 2004
Figure 1

Fig. 1. A combined *in silico* and *in vitro* approach identifies Rnt1p cleavage signals near known snoRNAs. (A) Schematic representation of a model RNA substrate illustrating the features used for the selection of new Rnt1p substrates. Arrows indicate the sites of cleavage. N represents any nucleotide, N' its counter part. W-C indicates a position where base pairing is predominant and is required for optimal activity (Lamontagne et al., 2003). The initial binding and positioning box (IBPB) indicates nucleotides that position the enzyme for cleavage. The binding stability box (BSB) indicates nucleotides that stabilize the binding of Rnt1p and enhance cleavage. The cleavage efficiency box (CEB) indicates nucleotides that directly contribute to the Rnt1p cleavage efficiency without affecting substrate binding. (B) Methods used for the selection of Rnt1p substrates *in silico*. The illustrated procedure takes as input two independent compilations: one contains snoRNA sequences and the other contains known Rnt1p substrates. P1 indicates the search performed using RNAMotif (Macke et al., 2001). P2 and P3 indicate RNA folding obtained with the Vienna RNA package (Schuster et al., 1994) and the minimum free energy folding algorithm (Schuster et al., 1994) and using the dynamic programming algorithm (McCaskill, 1990). P4 and P5 indicate comparisons performed using an algorithm developed during this study. P6 and P7 indicate an evaluation performed with the Vienna RNA package (Schuster et al., 1994). P8 indicates the evaluation performed using an algorithm developed in the course of this study. S1 indicates a score for which the highest value 1 is for structures with a long and stable stem having

few bulges and an internal loop. The lowest value of S1 is given to a sequence that, when unconstrained, does not fold into an AGNN tetraloop and has a high stability difference as compared to the constrained structure. S2 indicates a score for which the highest value of 1 is given to primary sequences that share the characteristics common to all known substrates and is highly similar to at least one of the known substrates. The lowest value of S2 (0) is given to a sequence that does not resemble any of the known substrates. (C) Summary of the data obtained from the prediction algorithm, the microarray data and both *in vitro* and *in vivo* validation. Details of the experimental data and the references of previously published results are indicated in supplementary table 1. A score was given to each potential Rnt1p cleavage sites within 1KB of all known snoRNAs. The score was assigned as described in 1B. *In vitro* cleavage was tested by the incubation of total RNA with recombinant Rnt1p as described in the Material and Methods section. Expression level detected upon the inactivation of Rnt1p temperature mutant (TS Expression) was recorded after 4 hours shift to the non-permissive temperature. *In vivo* processing was assessed by northern blot analysis of RNA extracted from Δ*rnt1* cells and a defect in processing was scored by the accumulation of a snoRNA precursor in the absence of Rnt1p. Information about the snoRNA families and gene organization were obtained from the snoRNA database (Samarsky and Fournier, 1999). The snoRNA were organized either according to the prediction score (left panel) or according to the snoRNA gene family (right). Notice that most C/D box snoRNA are

processed by Rnt1p while only few H/ACA box snoRNA reside near Rnt1p cleavage signals.

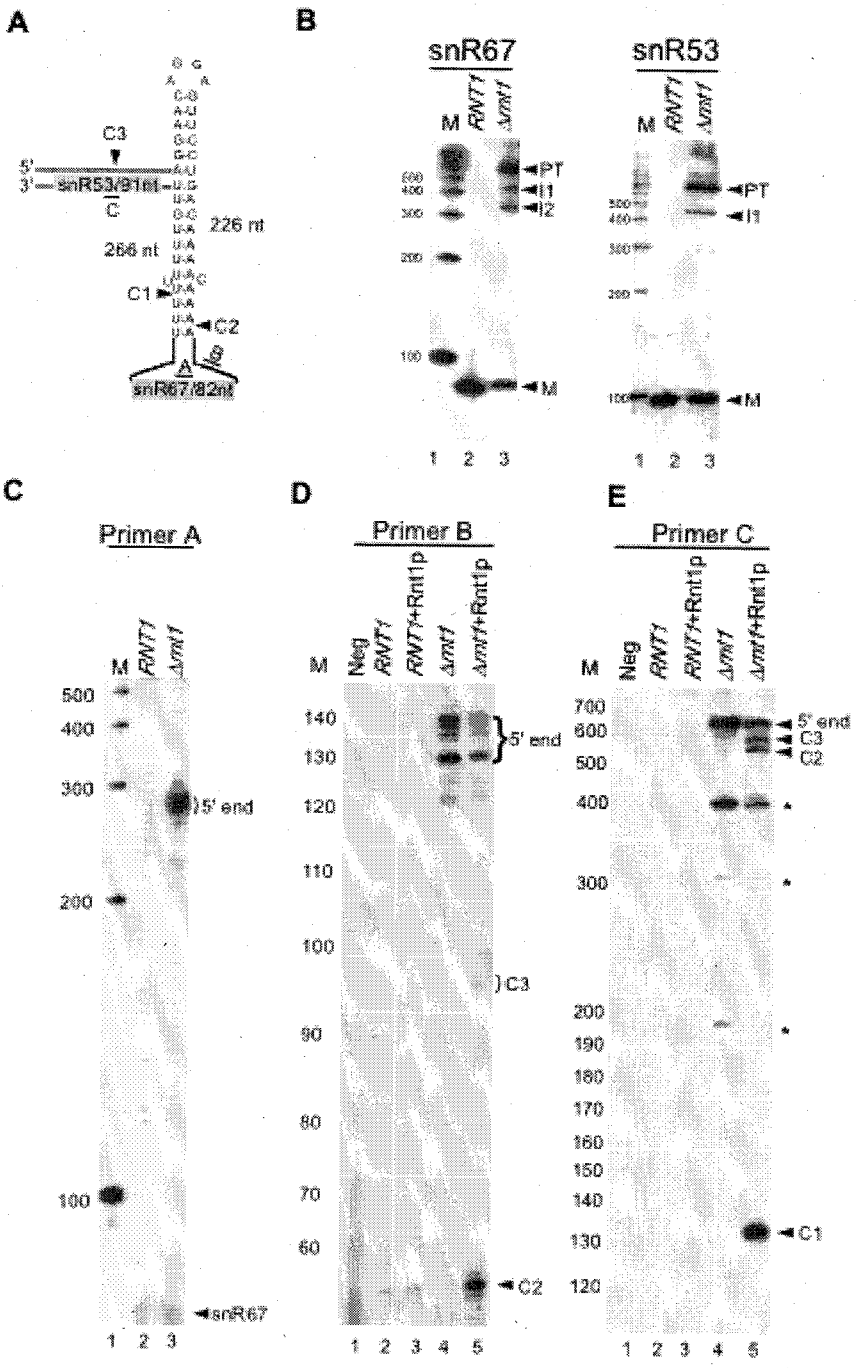
snoRNA or both. The expression level of most snoRNAs located near stem-loops that were cleaved by Rnt1p *in vitro* was increased more than 1.5 fold upon the deletion of Rnt1p (Fig. 1C). The expression level of only 5 snoRNAs near cleavable Rnt1p processing signal did not increase in the absence of Rnt1p probably due to rapid degradation of the unprocessed RNA transcript. All but 9 of the snoRNAs that were cleaved *in vitro* were affected 4 hours after a shift to the restrictive temperature. Most of the *in vitro* substrates that were not overexpressed in Δ*rnt1* cells were associated with monocistronic or intron encoded snoRNAs. The most sensitive substrates to Rnt1p deletion or inactivation were those expressed as polycistronic units. All independently transcribed snoRNAs and intron-encoded snoRNAs except snR42 with scores inferior to 0.8 did not exhibit Rnt1p-dependent expression. The result of the *in silico* and *in vivo* screens identified 7 new substrates and indicated that all but 22 snoRNAs, mostly H/ACA snoRNAs, are processed by Rnt1p.

Identification of Rnt1p substrates that form through Long-range base-pairing

The ideal Rnt1p substrate (Fig. 1A) is a perfect uninterrupted A-U rich stem capped with AGNN tetraloop, a feature which is very easy to identify by searching for a stable structural motif. However, most of Rnt1p substrates are interrupted stems that in many cases are not stable when taken out of their

RNA context. Furthermore, it has been previously shown that some of Rnt1p substrates can form through long-range interactions that are very difficult to identify using conventional motif-based searches or folding programs like mfold (Zuker, 2003; Zuker and Jacobson, 1995, 1998). In contrast, the *in silico* screen we have developed is capable of identifying AGNN tetraloops with three base-pair stems regardless of either the context or the global folding of the targeted RNA. Consequently, we were able to identify a hidden 3 bp stem capped with an AGGA tetraloop that could not form a stable local stem within the polycistronic unit of snR53/snR67 (Fig. 2A). Northern blot analysis of RNA extracted from either wild type or Δ*rnt1* cells hybridized to a probe corresponding to the mature sequence of snR67 revealed the accumulation of a large RNA precursor in Δ*rnt1* cells and a decrease in the level of the mature snR67 (Fig. 2B). Hybridization to a probe corresponding to snR53 also showed an accumulation of a precursor corresponding to the size of the unprocessed polycistronic transcript. However, the level of the mature snR53 was less affected than that of snR67, suggesting that the localization of snR53 near the 3' end of the primary transcript makes it less sensitive to Rnt1p deletion. Extension of a primer hybridized to the mature snR67 sequence confirmed the accumulation of an RNA species that extends to the predicted 5' end of the polycistronic subunit in the absence of Rnt1p (Fig. 2C). The capacity of Rnt1p to directly cleave the primary transcript of snR53/snR67 was tested *in vitro* using recombinant Rnt1p and total RNA extracted from Δ*rnt1* cells. Northern blot analysis of total RNA cleaved by

Rnt1p *in vitro* confirmed that the primary transcript of snR53/snR67 is a direct substrate of Rnt1p (data not shown). Primer extension of total RNA incubated with recombinant Rnt1p revealed three



Ghazal et al., 2004
Figure 2

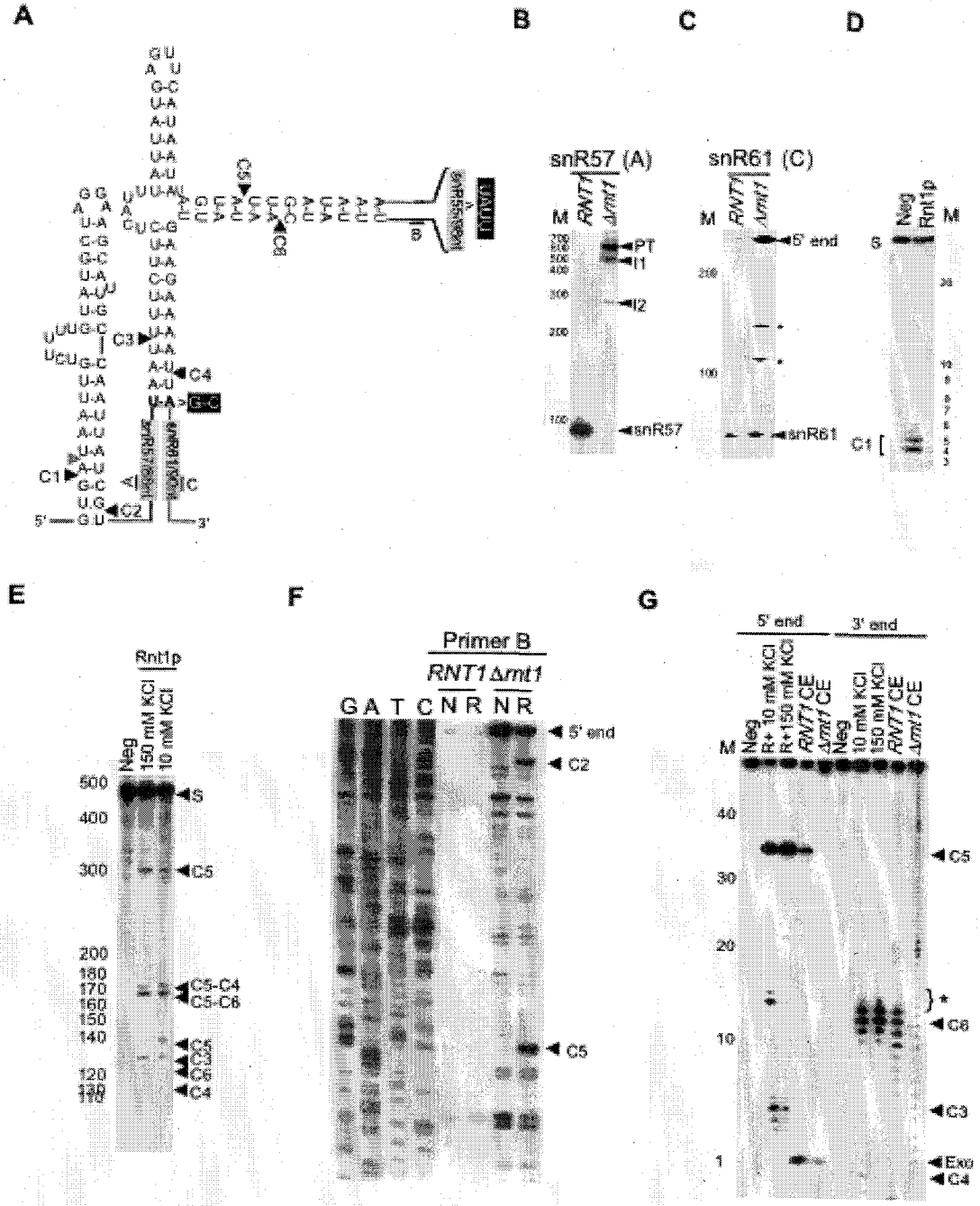
Fig.2. Rnt1p is required for the maturation of the polycistronic snR53/snR67 unit. (A) Schematic representation of the predicted stem loop structure that forms through long-range interactions between the sequences surrounding snR53 and snR67. The arrowheads indicate the position of the cleavage site identified *in vitro*. (B) Northern blot analysis of the snR67 and snR53 expression patterns both in the presence and the absence of Rnt1p. RNA was extracted from either wild type or Δ*rnt1* cells, separated on a 6% acrylamide gel, and hybridized to radioactive probes corresponding to the mature sequence of either snR67 or snR53. The positions of the mature snoRNA (M), the processing intermediates (I1 and I2) and the primary transcript (PT) are indicated on the right. (C) Primer extension mapping of the mature and extended termini of snR67. RNA extracted from both wild type and Δ*rnt1* cells was subjected to primer extension using primer A which is complementary to the coding sequence of snR67. Position of the mature RNA (snR67) and the extended forms detected in the absence of Rnt1p (5' end) are indicated on the right. (D) Mapping Rnt1p cleavages upstream of snR67. RNA extracted from either wild type or Δ*rnt1* cells was incubated with recombinant Rnt1p enzyme. The cleaved RNA was subjected to primer extension using primer B that hybridizes to sequence upstream of snR67. The positions of the cleavage sites (C2 and C3) and of the 5' ends are indicated on the right. (E) Mapping of the Rnt1p cleavage sites downstream of snR67. RNA was treated as

described in D, but the primer extension was performed using a primer (C) that hybridizes downstream of snR67.

cleavages, two near the 5' end of snR67 and one between snR53 and snR67 (Fig. 2D and E). This indicates the presence of two redundant cleavages at the 5' end of snR67. We conclude that the snR67/snR53 polycistronic transcript is a direct substrate of Rnt1p *in vitro,* and that Rnt1p is required for the efficient processing of both snR67 and snR53 *in vivo*.

A single stem-loop structure directs cleavage in two independent stems

The *in silico* prediction assay identified a short stem capped with AGUU tetraloop located in the middle of snR57/snR55/snR61 polycistronic unit. This single tetraloop was predicted to direct cleavages in two independent stems, one releases snR61 and the other releases both snR55 and snR57. An additional canonical stem was also identified at the 5' end of snR57 (Fig.3A). Very little signal of the mature snR57, snR61, and snR55 were detected upon deletion of *RNT1* (Figure 3B and data not shown). In contrast, an RNA species corresponding to the size of the unprocessed precursor was detected, indicating that Rnt1p is required for the processing of this transcription unit (Figure 3B). Extension of a primer corresponding to the sequence at the 5' end of snR57 (primer A) confirms that in the absence of Rnt1p the processing of the stem at the 5' end does not occur (Figure 3C). Incubation of recombinant Rnt1p with a T7 transcribed model substrate corresponding to the predicted stem-loop structure upstream of snR57 resulted in a specific cleavage at the predicted distance from the AGGA loop



Ghazal et al., 2004
Figure 3

Fig. 3. Rnt1p uses a single Tetraloop to release three different snoRNAs. (A) Illustration of the predicted cleavage sites associated with the snR57/snR55/snR61 cluster. Arrowheads indicate the cleavage sites identified *in vitro*. The positions of the primers used are indicated by letters. The positions of the mutations used in G are indicated in black boxes. (B) Northern blot analysis of snR57. Northern analysis was performed as in figure 2B. (C) Mapping of the 5' end of snR57. Primer extension was performed as described in figure 2C. The asterisk indicates truncated 5' end species that accumulate in the absence of Rnt1p. (D) Cleavage of a model substrate representing the stem-loop structure found upstream of snR57. The T7 transcribed RNA representing the 45 nts long stem-loop structure near the 5' end of snR57 was 5' end labeled and incubated with recombinant Rnt1p. The cleavage product corresponding to a cleavage in position C1 is indicated on the left. (E) An internally labeled T7 transcribed substrate corresponding to the entire snR57/snR55/snR61 cluster was incubated in the presence of Rnt1p. The cleaved RNA was separated using a 12% PAGE and directly visualized by autoradiography. The bands corresponding to the different cleavage sites are indicated on the right. (F) Primer extension mapping of total RNA cleaved by Rnt1p *in vitro*. The RNA was cleaved to completion then incubated with Primer B that corresponds to the sequence near the 3' end of snR61 (indicated in A). Unrelated DNA sequence was used as a marker. (G) *In vitro* cleavage of an artificial substrate confirming Rnt1p's capacity to direct

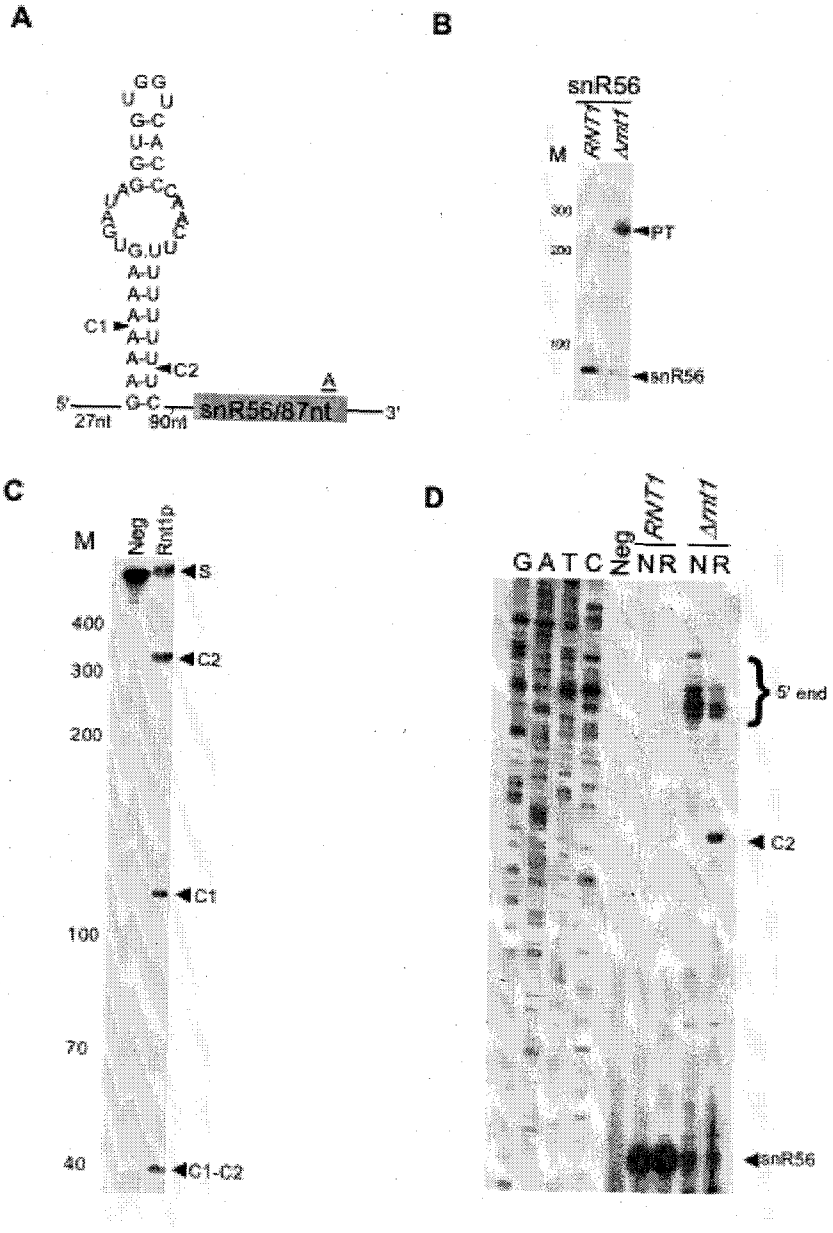
4 cleavage events using a single binding site. A T7 transcript of an RNA harboring the AGUU stem loop structure found in the snR57/snR55/snR61 cluster was tested for cleavage *in vitro*. The snR55 sequence of this RNA was replaced by a UAUU tetraloop, and the snR57 and snR61 sequence was replaced by terminating the RNA with a G-C base pair. The RNA was labeled either at the 3' or the 5' end, and was incubated either in the presence of recombinant Rnt1p, wild type cell extracts, or Δ*rnt1* cell extracts. The cleavage products are indicated on the right (C1-C6). (Exo) indicates products produced by exonucleases found in cell extracts. The asterisk indicates unspecific cleavages that occur under low salt conditions.

(Fig. 3D). This indicates that the predicted stem is a direct substrate of Rnt1p. In order to test whether or not the predicted structure located between snR55 and snR57/snR61 is cleaved by Rnt1p, and to map the cleavage site, we produced a T7 transcript that corresponded to the entire region located between snR57 and snR61 and incubated it with recombinant Rnt1p. Northern blot analysis revealed 6 different cleavage products consistent with four cleavage events in two independent stems at the right distance from the tetraloop. Two cleavages release snR55, one releases the 3' end of snR57 and one liberates the 5' end of snR61 (Fig. 3E and data not shown). Mutation of the AGUU tetraloop to GAAA blocks all cleavages in both stems (data not shown). These data clearly demonstrate that the predicted AGUU tetraloop is required for cleavage events that occur in the sequence that separate snR57 / snR55 from snR61.

The predicted cleavages at positions C4 and C6 were confirmed by reverse transcription using a primer corresponding to the mature sequence of snR55 (data not shown). The cleavage sites predicted 5' to snR55 were determined by extending a primer corresponding to the sequence at the 3' end of snR55 (primer B) after incubation of total RNA with recombinant Rnt1p (Fig. 3F). In the context of the native RNA subunit that accumulates in the absence of Rnt1p (Figure 3B and C), we detected *in vitro* cleavage at the 5' of snR55 (C5) and at the 5' end of snR57 (C2). These same cleavage sites were also detected using a primer corresponding to the mature sequence of snR61

(data not shown). Failure to detect cleavage at C3 using a primer at the 3' end of C5 (Fig. 3A and F) indicates that under this cleavage condition all substrates that were cleaved at C3 were also cleaved at C5. Cleavage at C3 is detected using specific probes at the 5' end of C3 and probes corresponding to the sequence that separates C3 and C5 cleavage sites (Fig. 3E). We were also able to detect the product arising from cleavage at C3 using substrate labeled at the 5' end of C3 (data not shown). Together, these data suggest that a single tetraloop may direct Rnt1p cleavages within two separate stems. In order to directly examine this possibility, we produced a short T7 transcript representing a model substrate that maintains the two stems linked to the AGUU tetraloop binding site. The first stem ends with the two nucleotides located below the C4 cleavage site and the other terminates with a UAUU tetraloop that replaces the naturally occurring snR61 sequence. The UAUU tetraloop cannot direct cleavage by itself and no tetraloop other the AGUU could was found within this RNA transcript. Thus any cleavage detected in the adjacent stem would be directed by the AGUU tetraloop. The model T7 substrate was labeled at either the 3' or the 5' ends and incubated with recombinant Rnt1p either at a low monovalent salt concentration or at physiological salt concentration. As shown in figure 3G, all four cleavages are detectable at both salt concentrations. The cleavage sites were mapped based on the size of the released fragments (Fig. 3G). The simultaneous detection of cleavages at C3/C5 and C4/C6 within a single end labeled RNA species suggests that the four cleavages are not produced from a single

binding event. Incubation of the model substrate in either a wild type or a Δ*rnt1* cell extract confirmed that the native Rnt1p could also cleave substrates with bifurcated stems. Consequently, kinetic analysis using similar bifurcated RNA substrate showed that each stem is cleaved by a distinct binding event (Lamontagne and Abou Elela, Unpublished results).



Ghazal et al., 2004
Figure 4

Fig. 4. Rnt1p does not require the presence of an AGNN tetraloop for the maturation of snR56. A schematic representation of the predicted cleavage sites associated with the snR56 cluster. The cleavage sites identified *in vitro* are indicated by arrowheads. (B) Northern blot analysis of snR56. RNA extraction and Northern analysis were performed as described in Figure 2B. (C) *In vitro* cleavage of a model substrate representing the cleavage signals found near snR56. Internally labeled T7 transcribed RNA possessing the stem-loop structure indicated in A was incubated with recombinant Rnt1p. The cleavage products were separated by PAGE and visualized by autoradiography. (D) Mapping of the Rnt1p cleavage site using primer extension. The primer extension was performed as described in Figure 2C. The mature RNA, extended ends and the cleavage site are indicated on the right.

The AGNN tetraloop is not conserved near all Rnt1p dependent snoRNAs

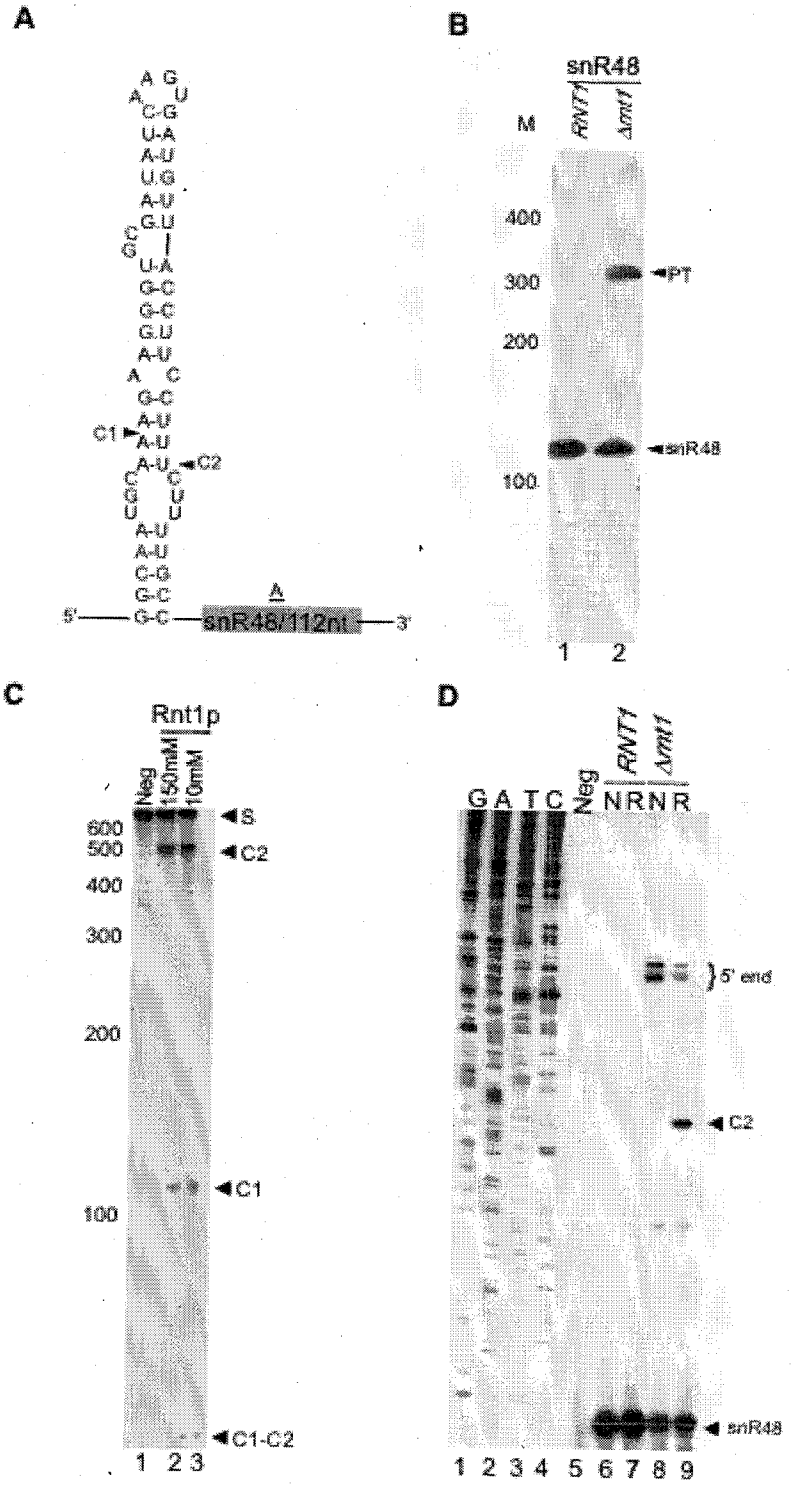
Analysis of the expression profiles of snoRNAs suggested that the monocistronic snR56 and snR48 snoRNAs are Rnt1p dependent. However, we failed to identify AGNN tetraloops near the termini of these snoRNAs. Northern blot analysis confirmed that the deletion of Rnt1p impairs the processing of these two snoRNAs (Fig. 4B and Fig. 5B). *In vitro* cleavage of a T7 transcribed RNA corresponding to the 5' end of snR56 indicated that Rnt1p could directly cleave the RNA near snR56 despite the absence of an AGNN tetraloop (Fig. 5C). Extension of a primer corresponding to a portion of the snR56 mature sequence confirmed a cleavage at a stem capped with UGGU that occurs at the predicted distance from the loop (Figure 4D). While this study was in progress it was also reported that Rnt1p could cleave a stem in the intron of RPL18A that is capped with UGGU (Danin-Kreiselman et al., 2003). Thus Rnt1p does not require A in the first position of the tetraloop for cleavage. Similarly, *in vitro* cleavage of a model substrate near the 5' end of snR48 confirmed that Rnt1p could cleave this substrate. However, no canonical stem capped with an NGNN tetraloop was found in the vicinity of the cleavage. Primer extension analysis indicated a cleavage near an AAGU terminal tetraloop. Since all known substrates to date contain a G in the second position of the tetraloop, and it has previously been shown that changing this G to any other nucleotide blocks cleavage *in vitro*(Chanfreau et al., 2000), we presume that a special feature of either this RNA stem or of the

combination of this special loop and stem sequence, permits the Rnt1p cleavage in this case. We conclude that the AGNN tetraloop is not essential for the Rnt1p dependent maturation of snoRNA.

Direct processing of intron embedded snoRNAs by Rnt1p

Two of the predicted Rnt1p cleavage signals were located in the introns of the pre-mRNA of the ribosomal protein RPL7A, and that of its nearly identical isoform RPL7B, near snR39 and snR59, respectively. Consistently, the microarray expression profile indicated that the expression of both snoRNAs was induced upon the deletion of *RNT1* (Supplementary Table 1). In order to examine the impact of Rnt1p on the processing of snR39 and snR59 we monitored the RNA profile of these two RNAs in both the absence and the presence of Rnt1p. As shown in figure 6B, the deletion of Rnt1p did not inhibit the accumulation of mature snR39, or of the mature RPL7A mRNA (Fig. 6C), but caused mild accumulation of the unspliced pre-mRNA precursor is observed. The Northern blot analysis suggests that the increase in the expression of snR39 observed by microarray analysis (Supplementary Table 1) is due to either an accumulation of the unspliced mRNA or partially degraded snoRNA and not due to the accumulation of snR39 or its immediate precursor. In contrast, deletion of Rnt1p causes the accumulation of a precursor of snR59, a significant accumulation of the pre-mRNA and a

reduction in the level of mature snR59 (Fig. 7B). The accumulation of RPL7B pre-mRNA, did not result in a decrease in the level of the mature mRNA



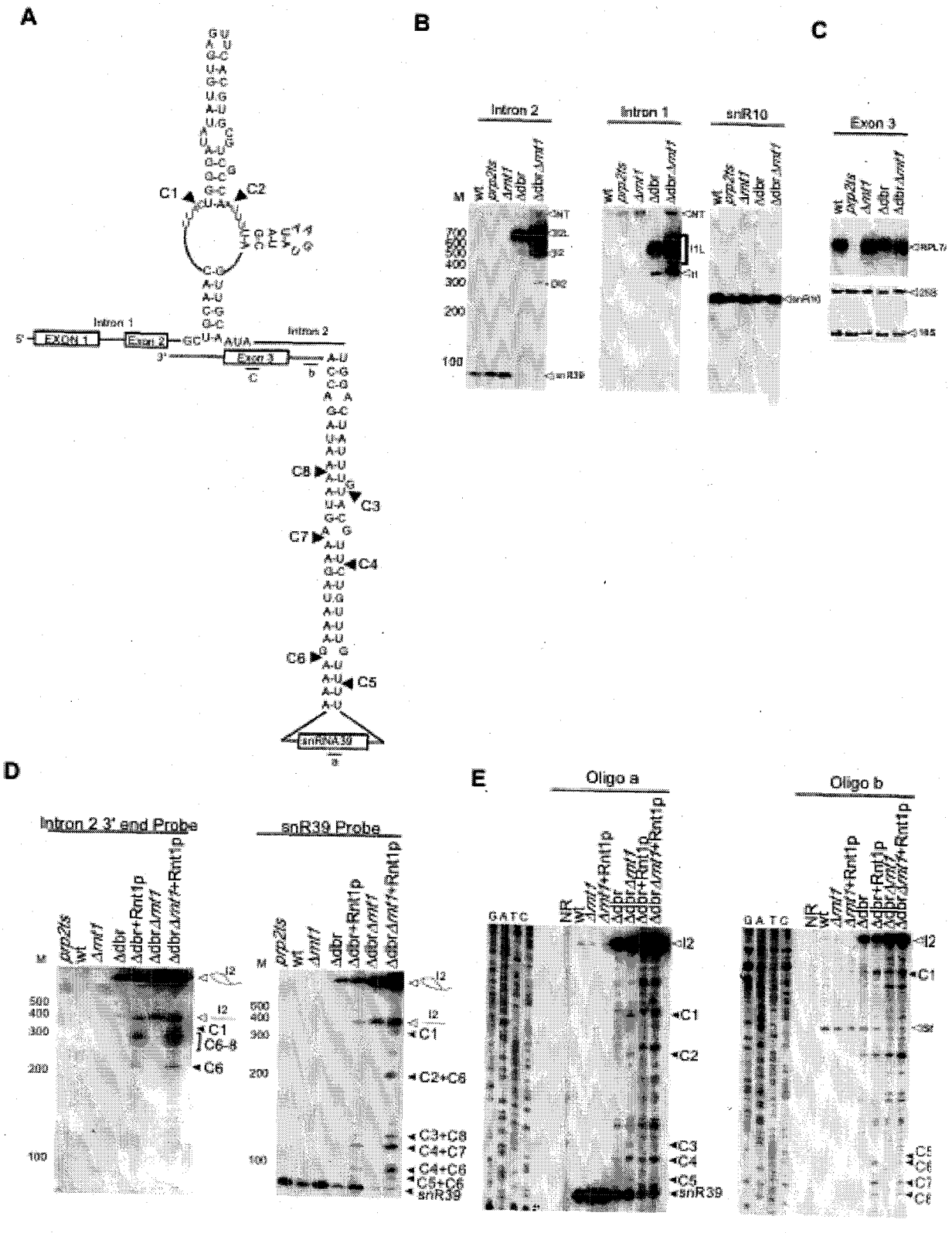
Ghazal et al., 2004
Figure 5

Fig. 5. Rnt1p is required for the processing of snR48 in the absence of any detectable NGNN tetraloop. (A) A schematic representation of the predicted stem near the 5' end of snR48. The cleavage sites identified *in vitro* are indicated by arrowheads. (B) Northern blot analysis of snR48. RNA extraction and Northern analysis were performed as described in Figure 2B. (C) *In vitro* cleavage of a model substrate representing the cleavage signals found near snR48. Internally labeled T7 transcribed RNA exhibiting the stem-loop structure indicated in A was incubated with recombinant Rnt1p and the cleavage products were separated by PAGE and visualized by autoradiography. (D) Mapping of the Rnt1p cleavage site using primer extension. The primer extension was conducted as described in 2C. The mature RNA, extended ends and the cleavage site are indicated on the right.

suggesting that splicing is not impaired, as is observed upon inactivation of the essential splicing factor Prp2p (Fig. 7B).

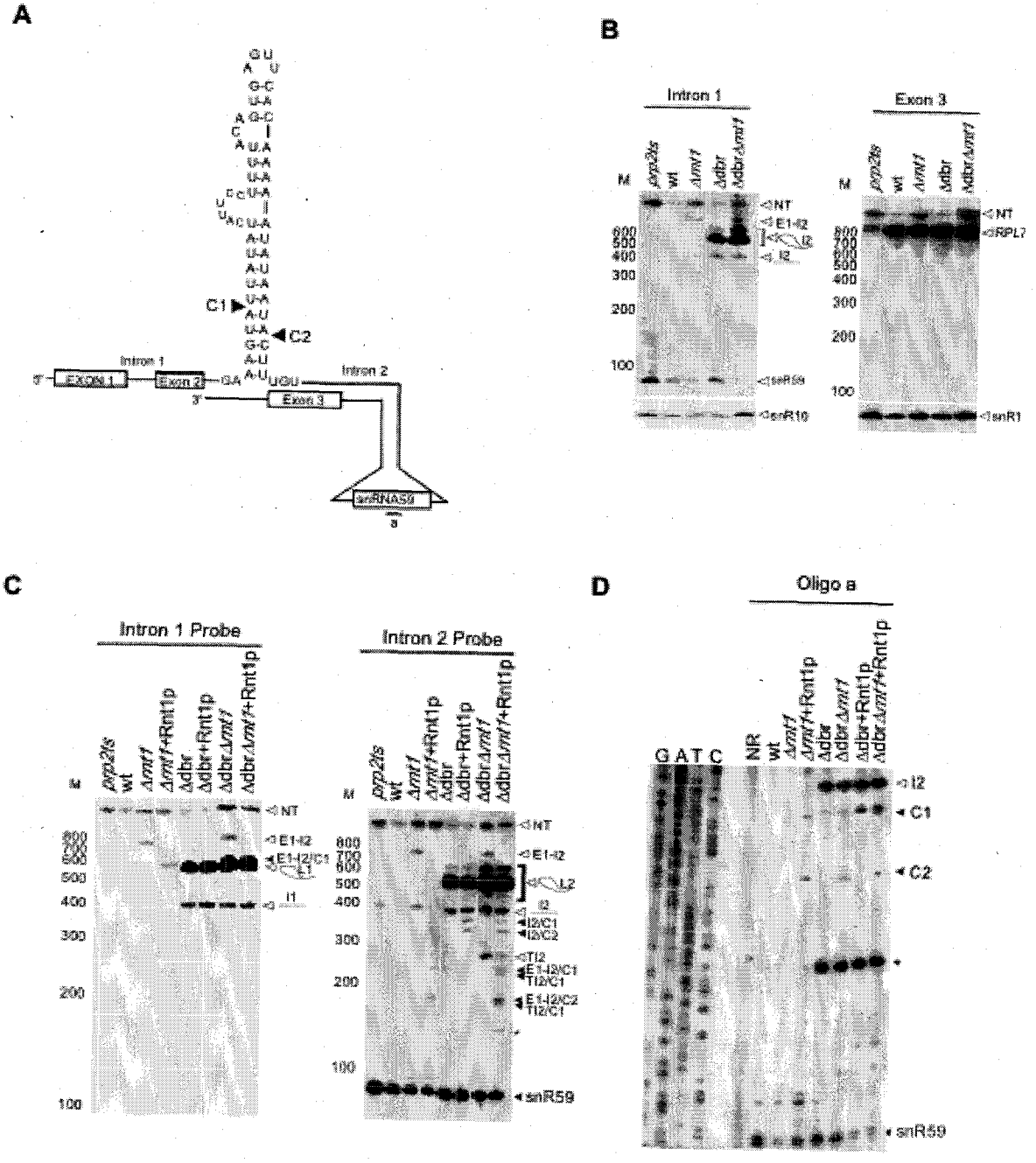
In order to examine the possibility of a lariat dependent processing pathway of snR39 and snR59 we studied the accumulation of these two snoRNAs in both cells lacking the debranching enzyme Dbr1p and in cells containing deletions in the *DBR1* and *RNT1* genes. In the case of snR39 the deletion of *DBR1* resulted in the accumulation of a lariat containing snR39 and in a severe reduction in the mature form of snR39. Deletion of both *DBR1* and *RNT1* completely abolished the production of mature snR39 (Fig. 6B). These data indicate that Rnt1p plays a minor role in releasing snR39 from the lariat, while the major processing pathway of snR39 is through the trimming of the debranched lariat. *In vitro* cleavage assays and primer extension mapping of the cleavage site revealed that Rnt1p could cleave the lariat containing the snR39, but not the primary mRNA transcript. Eight different cleavage sites were mapped near both the 5' and 3' ends of snR39 (Fig. 6D and E). Two of these are located at the predicted distance from the identified AGUU tetraloop, while the 6 others are not near any recognizable tetraloop motif. It is possible that an alternative fold brings these cleavage sites close to the identified terminal tetraloop. In contrast, both *in vitro* cleavage and cleavage site mapping of snR59 indicated that Rnt1p targets the pre-mRNA, and not the produced lariat for cleavage (Fig. 7C and D). The cleavage of Rnt1p occurred at the predicted distance from the AGUU tetraloop (Fig. 7A and D). These data indicate that Rnt1p is required for the maturation of snR59. We

conclude that while the processing of snR39 is dependent on the splicing of RPL7A pre-mRNA, the splicing of RPL7B pre-mRNA is inhibited when snR59 is processed by Rnt1p.



Ghazal et al., 2004
Figure 6

Fig. 6. Rnt1p assists in releasing the intron-encoded snR39 from the lariat of RPL7A pre-mRNA. (A) Schematic representation of the structure found near snR39 within the intron of RPL7A. The cleavage sites identified *in vitro* are indicated by arrowheads. (B) Northern blot analysis of snR39 and the associated mRNA sequence. RNA extraction and Northern analysis were performed as described in figure 2B. Probes corresponding to the snR39 mature sequence (intron 2), intron 1, or exon 3 were used. A probe corresponding to snR10 that is not affected by Rnt1p deletion is shown as a control for both loading and RNA quality. (C) Northern blot analysis using a probe specific to exon 3 of RPL7A. The RNA used was extracted and manipulated as described in B except that it was fractionated on 1% agarose gel. (D) *In vitro* cleavage of total RNA. RNA extracted from either wild type cells or Δ*rnt1* cell was incubated with recombinant Rnt1p. Northern blot analysis using probes specific to the sequence near either the 3' end of snR39 or the mature sequence were used to display the cleavage products. The positions of the different cleavages are indicated on the right. L designates the position of the lariat sequence. (E) Primer extension mapping of Rnt1p cleavage *in vitro*. The experiment was conducted as described in Figure 2C.



Ghazal et al., 2004
Figure 7

Fig. 7. Rnt1p is required for the processing of the intron encoded snR59 from the pre-mRNA of RPL7B. (A) Schematic representation of the structure found near snR59 within the intron of RPL7B. The cleavage sites identified *in vitro* are indicated by arrowheads. (B) Northern blot analysis of snR59 and the associated mRNA sequence. Probes corresponding to either intron 1 or exon 3 were used. A probe corresponding to snR10 that is not affected by Rnt1p deletion is shown as a control for both loading and RNA quality. E1 and E2 indicate exons 1 and 2 respectively, I1 indicates intron 1, I2 indicates intron two. L indicates the splice lariat. NT indicates the nascent transcript. C indicates the site of cleavage. TI2 indicates truncated fragments of intron 2. (D) *In vitro* cleavage of total RNA. RNA was extracted from either wild type or Δ*rnt1* cells was incubated with recombinant Rnt1p. Northern blot analysis using probes specific to the sequences near either the 3' end of snR59 or the mature sequence were used to display the cleavage products. The positions of the different cleavages are indicated on the right. L designates the position of the lariat sequence. (E) Primer extension mapping of Rnt1p cleavage *in vitro*. The experiment was conducted as described in Figure 2C.

Discussion

In this study we presented a combined *in silico*, *in vitro* and *in vivo* approach for the detection of Rnt1p substrates and demonstrated its utility for the identification of key processing events. Most Rnt1p processing signals were found near box C/D snoRNAs, while only threex were found near box H/ACA snoRNAs; indicating either a distinct evolutionary origin or a distinct regulatory pathway for each class of snoRNA. The location and organization of Rnt1p cleavage signals was found to vary from one snoRNA transcript to another. Rnt1p cleaved substrates formed through base-pairing between distantly located RNA sequences thereby ensuring the maturation of both ends of the targeted snoRNA (Fig. 2 and 3), as previously suggested for the bacterial RNase III (Young and Steitz, 1978). In other cases, a single NGNN tetraloop directed cleavages at two distinct cleavage sites thereby relating the processing of two adjacent snoRNAs (Fig. 3). The cleavage signals of all monocistronic snoRNAs that are processed by Rnt1 were found near the 5' end except the two isoforms of U3 snoRNA that contain introns and matures through splicing (Kufel et al., 2000). Finally, depending on the nature and context of Rnt1p cleavage signal, the processing of intron-encoded snoRNAs could be either linked to or separated from the splicing of the host pre-mRNA allowing a flexible control of the snoRNA associated redundant r-protein isoforms (Fig 6 and 7). The data presented here indicate the capacity of Rnt1p processing signals to provide a flexible tool to relate r-RNA, snoRNA, and r-protein production.

In yeast most intron containing pre-mRNAs encode r-proteins and are redundant (Planta and Raue, 1988). For years it has been suggested that this organization is important for fine-tuning the expression of the proteins and linking it to rRNA production. Indeed, in several cases the expression of one r-protein isoform regulates the splicing or the mRNA transport of another. However, it is not clear how the production of two r-protein isoforms might be regulated if they harbor a functionally distinct snoRNA as is the case for RPL7A and RPL7B. The two proteins are nearly identical, and have similarity to the *E. coli* L30 and rat L7 ribosomal proteins (Marchfelder et al., 1998; Yon et al., 1991). Deletion of RPL7A that harbors snR39 within its introns moderately impairs growth and affects budding (Mizuta et al., 1995); however, deletion of RPL7B that harbor snR59 in its intron has no effect on growth (Mizuta et al., 1995). Deletion of both genes is lethal, reflecting their housekeeping function as part of the ribosome. Like other r-proteins, the expression of these two isoforms needs to be regulated in order to achieve an equimolar production of both the protein and the rRNA it binds. For example, the expression of these two proteins is shutdown along with that of all other r-proteins when no rRNA is produced. However, controlling the transcriptional level will also affect the snoRNAs encoded within the introns of these proteins, which are required for the production of normal, mature, rRNA. Although the snoRNAs embedded in the introns of both RPL7 isoforms are not essential like most snoRNAs but they are conserved among fungi

(Samarsky and Fournier, 1999; Samarsky et al., 1998) and their expression is expected to be controlled like most component of the ribosome biogenesis machinery (Peng et al., 2003). In this study we show that in the case of RPL7B cleavage by Rnt1p could release snR59 while preventing the production of RPL7B. In contrast, snR39 production appears to be linked to RPL7A since Rnt1p could only cleave the splicing byproduct of RPL7A, and not the mature RNA. Therefore, the cells could control the overall level by slowing the splicing of RPL7B, which will lead to an increase in the cleavage of Rnt1p, producing snR59 and reducing the amount of RPL7B. This is consistent with the fact that deletion of RPL7A has more effects on growth than the deletion of RPL7B. Introns of r-protein mRNA were previously searched for Rnt1p cleavage sites and no cleavage was detected other than those identified within the introns of RPS22B and RPL18 pre-mRNA (Danin-Kreiselman et al., 2003). We have searched all other mRNAs containing introns in yeast (Grate and Ares, 2002) for potential Rnt1p cleavage sites and we did not find mRNA with intronic stem-loops above 0.8 that are significantly overexpressed upon Rnt1p deletion other than those previously identified (Gagnon and Abou Elela unpublished result). Indicating that Rnt1p cleavage within pre-mRNA introns might be restricted to r-proteins possibly to help regulates ribosome biogenesis. However, we cannot exclude a more general but redundant role of Rnt1p in the regulation of intron containing mRNAs that cannot be easily detected by the deletion of *RNT1*.

Acknowledgements

The authors would like to thank J. D. Boeke for the Δ*dbr1* strain and Ren-Jang Lin for the *prp2* temperature sensitive strain. We also thank Stéphanie Larose **Erreur ! Aucun nom n'a été donné au signet.**for preparing the yeast cell extract used in this study. We are indebted to all members of the Abou Elela lab for stimulating discussion and critical reading of the manuscript. This work was supported by grant # 216854 from the Natural Sciences and Engineering Research Council of Canada (NSERC). S. A. is a Chercheur-Boursier Junior II of the Fonds de la Recherche en Santé du Québec (FRSQ).

Supplementary Table 1. Yeast snoRNA processing signals and Rnt1p dependent expression

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
snR4	186	N	C/D	Mono	N	N	N	NT	-1.04/5.08	N	NR	(Chanfreau et al., 1998a)
snR13	124	N	C/D	Mono	N	N	N	NT	-3.2/-1.10	N	25S	TS, (Chanfreau et al., 1998a)
snR17a	333	+41	C/D	Mono	AGGU	L	Y	Y	2.32/2.02	Y	C18S	(Chanfreau et al., 1998a), (Kufel et al., 2000)
snR17b	332	+44	C/D	Mono	AGGU	L	Y	NT	ND	P	C18S	(Chanfreau et al., 1998a), (Kufel et al., 2000)

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
snR39B	95	-79	C/D	Mono	AGUU	L	Y	Y	5.91/3.84	Y	25S	TS, (Chanfreau et al., 1998a)
snR40	97	-86	C/D	Mono	AGUU	LRI	Y	Y	11.3/1.7	Y	18S, 25S	(Chanfreau et al., 2000), (Chanfreau et al., 1998a)
snR45	172	N	C/D	Mono	N	N	N	NT	-1.73/2.89	N	NR	(Chanfreau et al., 1998a)
snR47	99	-54	C/D	Mono	AGAA	L	Y	Y	2.77/1.12	Y	18S, 25S	(Chanfreau et al., 2000), (Chanfreau et al., 1998a)
snR48	112	-95	C/D	Mono	AAGU	L	Y	Y	18.08/3.19	Y	25S	TS
snR50	89	-94	C/D	Mono	AGUC	LRI	Y	Y	18.66/1.96	Y	25S	TS, (Lee et al., 2003a)

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
snR52	92	-82	C/D	Mono	AGAU	L	Y	Y	6.08/-1.29	Y	18S, 25S	TS, (Lee et al., 2003a)
snR56	87	-90	C/D	Mono	UGGU	L	Y	Y	12.17/1.56	Y	18S	TS
snR58	96	-57	C/D	Mono	AGAU	L	Y	Y	1.38/2.31	Y	25S	TS, (Lee et al., 2003a)
snR60	103	-89	C/D	Mono	AGGU	LRI	Y	Y	6.43/1.32	Y	25S	TS, (Lee et al., 2003a)
snR62	100	-40	C/D	Mono	AGUG	L	Y	Y	1.65/1.59	Y	25S	TS, (Lee et al., 2003a)
snR63	255	-79	C/D	Mono	AGUU	L	Y	Y	8.06/3.05	Y	25S	TS, (Lee et al., 2003a)
snR64	101	-81	C/D	Mono	AGCA	L	Y	Y	56.35/15.77	Y	25S	TS, (Lee et al., 2003a)
snR65	100	-72	C/D	Mono	AGAA	L	Y	Y	10.77/1.07	Y	25S	TS, (Lee et al., 2003a)
snR66	85	-55	C/D	Mono	AGAU	L	Y	Y	-2.95/1.01	Y	25S	(Lee et al., 2003a)
snR68	136	-112	C/D	Mono	AGGA	L	Y	Y	1.95/1.93	Y	25S	TS, (Lee et al., 2003a)

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
snR69	101	-113	C/D	Mono	AGGA	LRI	Y	Y	13.59/2.71	Y	25S	TS, (Lee et al., 2003a)
snR71	89	-71	C/D	Mono	GGUU	L	Y	Y	33/2.40	Y	25S	TS, (Lee et al., 2003a)
snR79	85	-72	C/D	Mono	AGGA	L	Y	Y	-1.19/1.10	Y	18S	(Chanfreau et al., 2000), (Chanfreau et al., 1998a)
snR41[l]	110	-63	C/D	Poly	AGUA	LRI	Y	Y	41.97/4.78	Y	18S	(Chanfreau et al., 2000), (Chanfreau et al., 1998a)
snR70	165	PA	C/D	Poly	PA	PA	PA	PA	21.49/2.63	Y	18S	(Chanfreau et al., 1998a)
snR51	107	-36	C/D	Poly	AGUU	L	Y	Y	17.06/1.92	Y	18S, 25S	(Chanfreau et al., 2000), (Chanfreau

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
												et al., 1998a)
snR53	91	PA	C/D	Poly	PA	PA	PA	PA	12.93/7.24	Y	18S	TS
snR67	82	-188	C/D	Poly	AGGA	LRI	Y	Y	5.31/1.36	Y	25S	TS
snR61	89	PA	C/D	Poly	PA	PA	PA	PA	10.66/3.86	Y	25S	TS
snR55	98	-41	C/D	Poly	AGUU	LRI	Y	Y	52.39/7.44	Y	18S	TS
snR57	88	-97	C/D	Poly	AGGA	L	Y	Y	38.09/2.30	Y	18S	TS
snR190	190	-59	C/D	Poly	AGUU	LRI	Y	Y	25.31/2.48	Y	25S	(Lee et al., 2003a), (Chanfreau et al., 1998a), (Chanfreau et al., 1998b)
snR14	126	PA	C/D	Poly	PA	PA	PA	PA	1.81/7.73	Y	18S	(Lee et al., 2003a), (Chanfreau et al., 1998a), (Chanfreau et al.,

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
												1998b)
snR72	91	PA	C/D	Poly	PA	PA	PA	PA	16.46/2.69	Y	25S	TS, (Lee et al., 2003a), (Chanfreau et al., 1998a), (Qu et al., 1999)
snR73	103	+93	C/D	Poly	AGUU	LRI	Y	Y	30.95/2.79	Y	25S	(Chanfreau et al., 1998a), (Qu et al., 1999)
snR74	80	PA	C/D	Poly	PA	PA	PA	PA	86.76/5.70	Y	18S	(Chanfreau et al., 1998a), (Lee et al., 2003a), (Qu et al., 1999)
snR75	85	+73	C/D	Poly	AGUU	LRI	Y	Y	8.47/3.08	Y	25S	(Chanfreau et al., 1998a), (Qu et al., 1999)
snR76	104	-54	C/D	Poly	AGUA	L	Y	Y	17.31/4.70	Y	25S	(Chanfreau et al.,

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
												1998a), (Qu et al., 1999)
snR77	84	PA	C/D	Poly	PA	PA	PA	PA	8.82/2.08	Y	18S	(Chanfreau et al., 1998a), (Lee et al., 2003a), (Qu et al., 1999)
snR78	82	+59	C/D	Poly	AGUA	LRI	Y	Y	50.33/2.55	Y	25S	(Chanfreau et al., 1998a), (Qu et al., 1999)
snR18	102	N	C/D	Intr	N	N	N	NOD	-1.43/-1.37	NOD	25S	(Chanfreau et al., 1998a), (Giorgi et al., 2001)
snR24	89	N	C/D	Intr	N	N	N	N	-1.4/-1.3	N	25S	TS, (Chanfreau et al., 1998a)
snR38	95	N	C/D	Intr	N	N	N	NOD	-1.01/1.03	NOD	25S	(Chanfreau et al.,

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
												1998a), (Giorgi et al., 2001)
snR39	89	-111	C/D	Intr	AGUU	L	Y	Y	2.47/-1.13	DBD	25S	(Chanfreau et al., 1998a), TS
snR54	86	N	C/D	Intr	N	N	N	W	1.49/-1.41	N	18S	TS
snR59	78	-107	C/D	Intr	AGUU	L	Y	Y	3.47/1.54	Y	25S	TS
snR3	194	N	H/ACA	Mono	N	N	N	NT	1.12/1.72	N	25S	(Chanfreau et al., 1998a)
snR5	197	N	H/ACA	Mono	N	N	N	NT	-1.35/2.02	N	25S	(Chanfreau et al., 1998a)
snR8	189	N	H/ACA	Mono	N	N	N	NT	-2.06/1.07	N	25S	(Chanfreau et al., 1998a)
snR9	187	N	H/ACA	Mono	N	N	N	NT	-1.15/1.59	N	NR	TS, (Chanfreau et al., 1998a)

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
snR10	245	N	H/ACA	Mono	N	N	N	NT	-1.45/1.35	N	25S	(Chanfreau et al., 1998a)
snR11	258	N	H/ACA	Mono	N	N	N	NT	-1.89/1.16	N	25S	TS, (Chanfreau et al., 1998a)
snR30	608	N	H/ACA	Mono	N	N	N	NT	-1.4/1.8	N	C18S	(Chanfreau et al., 1998a)
snR31	222	N	H/ACA	Mono	N	N	N	NT	1.22/1.96	N	18S	(Chanfreau et al., 1998a)
snR32	188	N	H/ACA	Mono	N	N	N	NT	-2.45/-1.66	N	25S	(Chanfreau et al., 1998a)
snR33	183	N	H/ACA	Mono	N	N	N	NT	-1.29/2.22	N	25S	(Chanfreau et al., 1998a)
snR34	203	N	H/ACA	Mono	N	N	N	NT	-1.36/-1.11	N	25S	(Chanfreau et al., 1998a)

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
snR35	204	N	H/ACA	Mono	N	N	N	NT	-1.74/1.19	N	18S	(Chanfreau et al., 1998a)
snR36	182	-91	H/ACA	Mono	AGUA	L	Y	Y	1.36/1.50	Y	18S	(Chanfreau et al., 2000), (Chanfreau et al., 1998a)
snR37	386	N	H/ACA	Mono	N	N	N	NT	-1.67/1.04	N	25S	(Chanfreau et al., 1998a)
snR42	351	N	H/ACA	Mono	N	N	N	NT	2.45/1.55	N	25S	(Chanfreau et al., 1998a)
snR43	209	-59	H/ACA	Mono	AGUG	L	Y	Y	1.12/1.64	Y	NR	(Chanfreau et al., 2000), (Chanfreau et al., 1998a)
snR46	197	-114	C/D	Mono	AGGA	L	Y	Y	1.54/-1.61	Y	25S	(Chanfreau et al.,

Name	Size	Position[a]	Family	Gene	Loop[b]	Stem[c]	Fold[d]	Cleavage[e]	Expression[f]	Processing[g]	Target[h]	References[k]
												2000), (Chanfreau et al., 1998a)
snR49	170	N	H/ACA	Mono	N	N	N	NT	ND	N	25S	(Chanfreau et al., 1998a)
snR189	192	N	H/ACA	Mono	N	N	N	NT	-1.11/1.11	N	18S, 25S	(Chanfreau et al., 1998a)
snR44	211	N	H/ACA	Intr	N	N	N	NT	-1.5/-3.27	N	18S, 25S	(Chanfreau et al., 1998a)

L : Local Interaction.

LRI : Long Range Interaction.

NT : Not tested.

PA : Processed from an adjacent cleavage assigned to a neighboring snoRNA in a polycistronic cluster.

NOD : Nop1p and Rnt1p dependent.

DBD : Dbr1p and Rnt1p dependent.

ND : No data in microarray.

NR: Not reported.

W: Very weak cleavage far from any recognizable cleavage signal, thus could be unspecific.

Y: Indicates the presence of a stem, correct folding, substrates that are affected by the deletion of Rnt1p, or substrate cleaved by Rnt1p *in vitro.*

N: Indicates no stem found, no stable structure, not affected Rnt1p deletion, or not cleaved by Rnt1p.

TS: Indicates snoRNAs examined during the course of this study.

[a] The position of the predicted stem loop structure from the mature ends of snoRNA.

[b] The sequence of the predicted tetraloop.

[c] The origin and the nature of folding of the predicted stem.

[d] The ability of the predicted stem to fold in the context of the entire snoRNA precursor sequence.

[e] The result of the cleavage assay using recombinant Rnt1p and either a model substrate representing the stem in question, or using total RNA extracted from Δ*rnt1* cells.

[f] Effect of Rnt1p deletion or inactivation on snoRNA expression as determined using microarray analysis. The first number indicates the fold increase or decrease in the snoRNA in Δ*rnt1p* cells when compared with wild type. The second number indicates the fold increase or decrease upon shifting cells carrying a temperature sensitive allele of RNT1 to the restrictive temperature as compared to wild type cells grown at the same temperature.

[g] Effect on processing was determined by northern blot analysis, and to identify the snoRNA precursor that accumulate in the absence of Rnt1p.

[h] The rRNA targets that is modified or cleaved by the snoRNA in question.

[k] Reference was given to all reports that examined the snoRNA for cleavage by Rnt1p. The general information about the snoRNA were retrieved from the snoRNA table (Samarsky and Fournier, 1999).

[l] snR41 the identified tetraloop does not correspond to published data.

Supplementary Oligonucleotides List 1:

snR48F:

TAATACGACTCACTATAGCGCGACATCATATACCTTTGTCCGCTG

snR48R: AACGACATCTGGGCGTAAACGACACC

snR50F: TAATACGACTCACTATAGCAAAATGACCTTTCCCTCC

snR50R: ATAAGATCAGCTATGCCAGCTTG

snR53F: TAATACGACTCACTATAGTTGTCGCCCCAAGCGGATC

snR53R: ACATAAGATCAAAGATGAAACTTGC

snR60F: TAATACGACTCACTATAGACTTCGTCGCTTTCTCCTCC

snR60R: CTTCCACGAATGCACAGGGAG

snR71F: TAATACGACTCACTATAGCAAAGTATATAAGAGACCATAAACAG

snR71R: CACAGCGAACCAGATCGAGATG

snR56F: TAATACGACTCACTATAGGGCGCCGCAAACCCCTCC;

snR56R: CCATCTTTTTCACAGGCGGTGTC

snR55: TAATACGACTCACTATAGAGGAAGTATATGCAGGACATATTGTG;

snR55R : TTGGTTCAGAAGCAGAACTGAATAG

Supplementary Oligonucleotides List 2:

snR48: CTTCACATCCTAACATTAGAGATGCC

snR53: GAACACGTTCATGATTAGCATGGAG

snR56: GCAATATTGTTATCTGCAAAACTTCG

snR57: CCTAATTCACAATATGTCCTGCATATAC

snR58: GGAGGGTCTAATCTCCTTCAGAAG

snR61: TTGGTTCAGAAGCAGAACTGAATAG

snR67: GATCCGCTTGGGGCGACAACTTAG

snR39: GGTGATAGTTACGACAGCATCGTCAATG

snR59: GCCGAAAGATGGTGATTAAACGACAG

Exon3 RPl7A F: CAAAAGGAATACGAAACTGCT

Exon3 Rpl7A R : GTTACCGAAAGAACCACCTT

Intron1 Rpl7A F : AGAAGTATACTAGTTTCCGC

Intron1 Rpl7AR : ACATCGAATTGAAAACACCA

Intron2 Rpl7AF: GTATGTTAAACTTTTGCTTAC

Intron2 Rpl7A R: AATACCAGATGGAAAACACAG

Intron 1 Rpl7B F: TACGACCTTATTTGGTAACTAGTTTGTTGT

Intron 1Rpl7BR: GCAGAGACATGCTGGAAATATCTATCAATG

Intron2 Rpl7BF: GTTCATTTACCATGTTTGAAAGA

Intron2 Rpl7BR: GATTCTTGGCATATTCTCACTC

Exon3 Rpl7B F: GGAATACGAAACTGCTGAAAG

Exon3 Rpl7B R: CAGATGGGTTGGACAACTTG

Supplementary Oligonucleotides List 3:

snR67 TTTATAAGCATACGCAAACAG

snR55 GGAAATAATGAAAATCCAGGTAAT

snR39 AATACCAGATGGAAAACACAG

References

1. **Abou Elela, S., and M. Ares, Jr.** 1998. Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA. EMBO J **17:**3738-46.

2. **Abou Elela, S., H. Igel, and M. Ares, Jr.** 1996. RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. Cell **85:**115-24.

3. **Apirion, D.** 1983. RNA processing in a unicellular microorganism: implications for eukaryotic cells. Prog Nucleic Acid Res Mol Biol **30:**1-40.

4. **Apirion, D., and P. Gegenheimer.** 1981. Processing of bacterial RNA. FEBS Lett **125:**1-9.

5. **Apirion, D., and A. Miczak.** 1993. RNA processing in prokaryotic cells. Bioessays **15:**113-20.

6. **Bachellerie, J. P., and J. Cavaille.** 1997. Guiding ribose methylation of rRNA. Trends Biochem Sci **22:**257-61.

7. **Bram, R. J., R. A. Young, and J. A. Steitz.** 1980. The ribonuclease III site flanking 23S sequences in the 30S ribosomal precursor RNA of E. coli. Cell **19:**393-401.

8. **Catala, M., B. Lamontagne, S. Larose, G. Ghazal, and S. A. Elela.** 2004. Cell cycle-dependent nuclear localization of yeast RNase III is required for efficient cell division. Mol Biol Cell **15:**3015-30.

9. **Chanfreau, G., M. Buckle, and A. Jacquier.** 2000. Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. Proc Natl Acad Sci U S A **97:**3142-7.

10. **Chanfreau, G., P. Legrain, and A. Jacquier.** 1998b. Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. J Mol Biol **284:**975-88.

11. **Chanfreau, G., G. Rotondo, P. Legrain, and A. Jacquier.** 1998a. Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. EMBO J **17:**3726-37.

12. **Danin-Kreiselman, M., C. Y. Lee, and G. Chanfreau.** 2003. RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. Mol Cell **11:**1279-89.

13. **Eichler, D. C., and N. Craig.** 1994. Processing of eukaryotic ribosomal RNA. Prog Nucleic Acid Res Mol Biol **49:**197-239.

14. **Filipowicz, W., P. Pelczar, V. Pogacic, and F. Dragon.** 1999. Structure and biogenesis of small nucleolar RNAs acting as guides for ribosomal RNA modification. Acta Biochim Pol **46:**377-89.

15. **Fromont-Racine, M., B. Senger, C. Saveanu, and F. Fasiolo.** 2003. Ribosome assembly in eukaryotes. Gene **313:**17-42.

16. **Giorgi, C., A. Fatica, R. Nagel, and I. Bozzoni.** 2001. Release of U18 snoRNA from its host intron requires interaction of Nop1p with the Rnt1p endonuclease. EMBO J **20:**6856-6865.

17. **Granneman, S., and S. J. Baserga.** 2004. Ribosome biogenesis: of knobs and RNA processing. Exp Cell Res **296:**43-50.

18. **Grate, L., and M. Ares, Jr.** 2002. Searching yeast intron data at Ares lab Web site. Methods Enzymol **350:**380-92.

19. **Hirose, T., M. D. Shu, and J. A. Steitz.** 2003. Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. Mol Cell **12:**113-23.

20. **Hirose, T., and J. A. Steitz.** 2001. Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. Proc Natl Acad Sci U S A **98:**12914-9.

21. **Kiss-Laszlo, Z., Y. Henry, J. P. Bachellerie, M. Caizergues-Ferrer, and T. Kiss.** 1996. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell **85:**1077-88.

22. **Kiss-Laszlo, Z., Y. Henry, and T. Kiss.** 1998. Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. EMBO J **17:**797-807.

23. **Kufel, J., C. Allmang, G. Chanfreau, E. Petfalski, D. L. Lafontaine, and D. Tollervey.** 2000. Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. Mol Cell Biol **20:**5415-24.

24. **Kufel, J., B. Dichtl, and D. Tollervey.** 1999. Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3' ETS but not the 5' ETS. RNA **5:**909-17.

25. **Lafontaine, D., and D. Tollervey.** 1995. Trans-acting factors in yeast pre-rRNA and pre-snoRNA processing. Biochem Cell Biol **73:**803-12.

26. **Lafontaine, D. L. J., C. Bousquet-Antonelli, Y. Henry, M. Caizergues-Ferrer, and D. Tollervey.** 1998. The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. Genes Dev **12:**527-37.

27. **Lamontagne, B., and S. Abou Elela.** 2001. Purification and characterization of Saccharomyces cerevisiae Rnt1p nuclease, vol. 342. Academic Press, San diego, CA.

28. **Lamontagne, B., and S. A. Elela.** 2004. Evaluation of the RNA determinants for bacterial and yeast RNase III binding and cleavage. J Biol Chem **279:**2231-41.

29. **Lamontagne, B., G. Ghazal, I. Lebars, S. Yoshizawa, D. Fourmy, and S. A. Elela.** 2003. Sequence dependence of substrate recognition and cleavage by yeast RNase III. J Mol Biol **327:**985-1000.

30. **Lamontagne, B., R. N. Hannoush, M. J. Damha, and S. Abou Elela.** 2004. Molecular requirements for duplex recognition and cleavage by eukaryotic RNase III: discovery of an RNA-dependent DNA cleavage activity of yeast Rnt1p. J Mol Biol **338:**401-18.

31. **Lamontagne, B., S. Larose, J. Boulanger, and S. A. Elela.** 2001. The RNase III family: a conserved structure and expanding functions in eukaryotic dsRNA metabolism. Curr Issues Mol Biol **3:**71-8.

32. **Lamontagne, B., A. Tremblay, and S. Abou Elela.** 2000. The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. Mol Cell Biol **20:**1104-15.

33. **Lebars, I., B. Lamontagne, S. Yoshizawa, S. Aboul-Elela, and D. Fourmy.** 2001. Solution structure of conserved AGNN tetraloops: insights into Rnt1p RNA processing. EMBO J **20:**7250-8.

34. **Lee, C. Y., A. Lee, and G. Chanfreau.** 2003. The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of S. cerevisiae box C/D snoRNAs. RNA **9:**1362-70.

35. **Macke, T. J., D. J. Ecker, R. R. Gutell, D. Gautheret, D. A. Case, and R. Sampath.** 2001. RNAMotif, an RNA secondary structure definition and search algorithm. Nucleic Acids Res **29:**4724-35.

36. **Marchfelder, A., D. A. Clayton, and A. Brennicke.** 1998. The gene for ribosomal protein L7a-1 in Schizosaccharomyces pombe contains an intron after the initiation codon. Biochim Biophys Acta **1397:**146-50.

37. **McCaskill, J. S.** 1990. The equilibrium partition function and base pair binding probabilities for RNA secondary structure. Biopolymers **29:**1105-19.

38. **Mizuta, K., T. Hashimoto, and E. Otaka.** 1995. The evolutionary relationships between homologs of ribosomal YL8 protein and YL8-like proteins. Curr Genet **28:**19-25.

39. **Nagel, R., and M. Ares, Jr.** 2000. Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5'-AGNN-3' tetraloop. RNA **6:**1142-56.

40. **Ni, J., A. L. Tien, and M. J. Fournier.** 1997. Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. Cell **89:**565-73.

41. **Ooi, S. L., D. A. Samarsky, M. J. Fournier, and J. D. Boeke.** 1998. Intronic snoRNA biosynthesis in Saccharomyces cerevisiae depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA [In Process Citation]. RNA **4:**1096-110.

42. **Pederson, T.** 1998. The plurifunctional nucleolus. Nucleic Acids Res **26:**3871-6.

43. **Peng, W. T., M. D. Robinson, S. Mnaimneh, N. J. Krogan, G. Cagney, Q. Morris, A. P. Davierwala, J. Grigull, X. Yang, W. Zhang, N. Mitsakakis, O. W. Ryan, N. Datta, V. Jojic, C. Pal, V. Canadien, D. Richards, B. Beattie, L. F. Wu, S. J. Altschuler, S. Roweis, B. J. Frey, A. Emili, J. F. Greenblatt, and T. R. Hughes.** 2003. A panoramic view of yeast noncoding RNA processing. Cell **113:**919-33.

44. **Perry, R. P.** 1976. Processing of RNA. Annu Rev Biochem **45:**605-29.

45. **Petfalski, E., T. Dandekar, Y. Henry, and D. Tollervey.** 1998. Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. Mol Cell Biol **18:**1181-9.

46. **Planta, R. J., and H. A. Raue.** 1988. Control of ribosome biogenesis in yeast. Trends Genet **4:**64-8.

47. **Qu, L. H., A. Henras, Y. J. Lu, H. Zhou, W. X. Zhou, Y. Q. Zhu, J. Zhao, Y. Henry, M. Caizergues-Ferrer, and J. P. Bachellerie.** 1999. Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. Mol Cell Biol **19:**1144-58.

48. **Reeder, R. H.** 1990. rRNA synthesis in the nucleolus. Trends Genet **6:**390-5.

49. **Robertson, H. D.** 1967. A nuclease specific for double-stranded RNA. Virology **12:**718.

50. **Robertson, H. D., R. E. Webster, and N. D. Zinder.** 1968. Purification and properties of ribonuclease III from Escherichia coli. J Biol Chem **243:**82-91.

51. **Samarsky, D. A., and M. J. Fournier.** 1999. A comprehensive database for the small nucleolar RNAs from Saccharomyces cerevisiae. Nucleic Acids Res **27:**161-4.

52. **Samarsky, D. A., M. J. Fournier, R. H. Singer, and E. Bertrand.** 1998. The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. EMBO J **17:**3747-57.

53. **Schuster, P., W. Fontana, P. F. Stadler, and I. L. Hofacker.** 1994. From sequences to shapes and back: a case study in RNA secondary structures. Proc R Soc Lond B Biol Sci **255:**279-84.

54. **Srivastava, A. K., and D. Schlessinger.** 1990. Mechanism and regulation of bacterial ribosomal RNA processing. Annu Rev Microbiol **44:**105-29.

55. **Tollervey, D., and T. Kiss.** 1997. Function and synthesis of small nucleolar RNAs. Curr Opin Cell Biol **9:**337-42.

56. **Tremblay, A., B. Lamontagne, M. Catala, Y. Yam, S. Larose, L. Good, and S. A. Elela.** 2002. A physical interaction between Gar1p and Rnt1p is required for the nuclear import of H/ACA small nucleolar RNA-associated proteins. Mol Cell Biol **22:**4792-802.

57. **Tycowski, K. T., C. M. Smith, M. D. Shu, and J. A. Steitz.** 1996. A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in Xenopus. Proc Natl Acad Sci U S A **93:**14480-5.

58. **van Hoof, A., P. Lennertz, and R. Parker.** 2000. Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Mol Cell Biol **20:**441-52.

59. **Venema, J., and D. Tollervey.** 1995. Processing of pre-ribosomal RNA in Saccharomyces cerevisiae. Yeast **11:**1629-50.

60. **Watkins, N. J., A. Gottschalk, G. Neubauer, B. Kastner, P. Fabrizio, M. Mann, and R. Luhrmann.** 1998. Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA- binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. RNA **4:**1549-68.

61. **Wu, H., A. Henras, G. Chanfreau, and J. Feigon.** 2004. Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. Proc Natl Acad Sci U S A **101:**8307-12.

62. **Wu, H., P. K. Yang, S. E. Butcher, S. Kang, G. Chanfreau, and J. Feigon.** 2001. A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. EMBO J **20:**7240-9.

63. **Yon, J., A. Giallongo, and M. Fried.** 1991. The organization and expression of the Saccharomyces cerevisiae L4 ribosomal protein genes and their identification as the homologues of the mammalian ribosomal protein gene L7a. Mol Gen Genet **227:**72-80.

64. **Young, R. A., and J. A. Steitz.** 1978. Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in Escherichia coli ribosomal precursor RNA. Proc Natl Acad Sci U S A **75:**3593-7.

65. **Zhou, H., J. Zhao, C. H. Yu, Q. J. Luo, Y. Q. Chen, Y. Xiao, and L. H. Qu.** 2004. Identification of a novel box C/D snoRNA from mouse nucleolar cDNA library. Gene **327:**99-105.

66. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res **31:**3406-15.

67. **Zuker, M., and A. B. Jacobson.** 1998. Using reliability information to annotate RNA secondary structures. RNA **4:**669-79.

68. **Zuker, M., and A. B. Jacobson.** 1995. "Well-determined" regions in RNA secondary structure prediction: analysis of small subunit ribosomal RNA. Nucleic Acids Res **23:**2791-8.

ARTICLE 2

Erreur ! Source du renvoi introuvable.and Abou Elela, S. (2006) Characterization of the Reactivity Determinants of a Novel Hairpin Substrate of yeast RNase III. J Mol Biol. 2006 Oct 20;363(2):332-44

Préambule

Le manuscrit décrit la découverte d'une nouvelle classe de substrats de Rnt1p qui n'exige pas la conservation de la tetraboucle NGNN, mais utilise plutôt des combinaisons différentes des séquencesde la tige et de la boucle. Les résultats présentés dans le manuscrit révèlentun mécanisme alternatif pour la sélection du substrat par la RNase III et proposent un plus large spectre de substrat que ce quel'on pensait auparavant.

J'ai effectué tous les travaux expérimentaux de ce document.

ARTICLE 2

Erreur ! Source du renvoi introuvable.and Abou Elela, S. (2006) Characterization of the Reactivity Determinants of a Novel Hairpin Substrate of yeast RNase III. J Mol Biol. 2006 Oct 20;363(2):332-44

Summary

The manuscript describes the discovery of a new class of Rnt1p substrates that does not require the conserved NGNN tetraloop for cleavage, but instead uses alternative combinations of loop and stem sequence. The results presented in the manuscript reveal an alternative mechanism for substrate selection by RNase III and suggest a broader substrate spectrum than previously believed. I have conducted all the experimental work in this paper.

Characterization of the Reactivity Determinants of a Novel Hairpin Substrate of yeast RNase III

Ghada Ghazal and Sherif Abou Elela*

Groupe ARN / RNA Group

Département de Microbiologie et d'Infectiologie, Faculté de Médecine,

Université de Sherbrooke

Sherbrooke, Québec, Canada J1H 5N4

Summary

RNase III enzymes form a conserved family of proteins thatspecifically cleave double-stranded (dsRNA). These proteins are involved in a variety of cellular functions, including the processing of many non-coding RNAs, mRNA decay, and RNA interference. Yeast RNase III (Rnt1p) selects its substrate by recognizing the structure generated by a conserved NGNN tetraloop (G2-loop). Mutations of the invariant guanosine stringently inhibit binding and cleavage of all known Rnt1p substrates. Surprisingly, we have found that the 5' end of snoRNA 48 is processed by Rnt1p in the absence a G2-loop. Instead, biochemical and structural analyses revealed that cleavage in this case is directed by a hairpin capped with an AAGU tetraloop, with a preferred adenosine in the first position (A1-loop). Chemical probing indicated that A1-loops adopt a distinct structure that varies at the 3' end where Rnt1p interacts with G2-loops. Consistently, chemical footprinting and chemical interference

assays indicate that Rnt1p binds to G2- and A1-loops using different sets of nucleotides. Also, cleavage and binding assays showed that the N-terminal (N-term) domain of Rnt1p aids selection of A1-capped hairpins. Together, the results suggest that Rnt1p recognizes at least two distinct classes of tetraloops using flexible protein RNA interactions. This underscores the capacity of double stranded RNA binding proteins (dsRBPs) to use several recognition motifs for substrate identification.

Keywords: dsRNA, RNA recognition, Ribonucleases, Rnt1p, Tetraloop, snoRNA

Introduction

Members of the RNase III family(Conrad and Rauhut, 2002; Lamontagne et al., 2001; Nicholson, 1999) are found in all species examined with the exception of archaebacteria, where the functions of RNase III are carried-out by the bulge-helix-bulge nuclease (BHB).(Lykke-Andersen et al., 1997) Membership requires homology with the structural elements of the founding member, *Escherichia coli* RNase III.(Nicholson, 1999) Most RNase IIIs display low sequence specificity *in vitro* and usually cleave any duplex RNA with no obvious structural or sequence motifs.(Lamontagne and Abou Elela, 2004) In contrast, RNases III are highly specific *in vivo*, and mostly target short RNA hairpins.(Ghazal et al., 2005; Nicholson, 1996) This surprisingly high specificity *in vivo* prevents complementation, even between closely related species.(Mitra and Bechhofer, 1994; Rotondo et al., 1997)

Yeast Rnt1p(Abou Elela et al., 1996) exhibits the main features of class II RNase IIIs, which include a dsRBD, a nuclease domain (NUCD), and an N-terminal extension (N-term).(Lamontagne and Abou Elela, 2001) The dsRBD motif is located at the C-terminus and has approximately 25% identity with other RNase IIIs.(Lamontagne et al., 2001) Rnt1p dsRBD is distinguished by a unique C-terminal extension required for nucleolar localization.(Catala et al., 2004) The enzyme NUCD contains the RNase III signature sequence implicated in catalysis and possesses an N-term extension that is unique to

eukaryotic RNase III.(Lamontagne et al., 2000) The N-term contributes to the formation of Rnt1p homodimer and is required for efficient cleavage at high salt concentrations(Lamontagne et al., 2000). Recently, crystal (Leulliot et al., 2004) and solution(Wu et al., 2004) structures of Rnt1p dsRBD confirmed the classical $\propto\beta\beta\beta\propto$ structure, and revealed an additional helix near the C-terminus ($\propto$ 3) that is unique to Rnt1p. The solution structure of the dsRBD / RNA complex(Wu et al., 2004) indicates that the additional helix is not located near the substrate RNA, but that it could influence the binding of RNA to $\propto$ 1.(Leulliot et al., 2004)

Studies of *E. coli* RNase III suggest that substrate selection is regulated by antideterminant nucleotides.(Zhang and Nicholson, 1997) This means that the absence of a nucleotide or structure, and not its presence, triggers RNA binding and cleavage. As more RNase IIIs are tested, it is becoming increasingly clear that eukaryotic RNase IIIs possess a different mechanism of substrate selectivity in which antideterminants play a minor role.(Lamontagne and Abou Elela, 2004) For example, Rnt1p prefers substrates that exhibit an NGNN tetraloop structure(Chanfreau et al., 2000) (Figure 1A). This unique affinity to NGNN tetraloops (G2-loops) is likely due to changes in the protein structure, such as the addition of helix ($\propto$ 3) at the end of Rnt1p dsRBD. The structure of the Rnt1p dsRBD/RNA complex (Wu et al., 2004) shows that the protein monomer contacts both the RNA major and

minor grooves, but, surprisingly, not the conserved G in the second position of the terminal tetraloop.(Lamontagne and Abou Elela, 2004)Deletion of this tetraloop, or substitution of the universally conserved G, blocks cleavage and reduces binding under physiological conditions.(Lamontagne et al., 2003) Because all G2-loops adopt a similar tertiary structure, it was suggested that Rnt1p recognizes the overall shape and not the sequence of the tetraloop.(Lebars et al., 2001; Wu et al., 2001) However, Rnt1p fails to cleave RNA hairpins capped with ACAA tetraloops, which have an overall conformation that is similar to G2-loops.(Staple and Butcher, 2003) This implies that Rnt1p substrate selection involves much more than tetraloop conformation recognition and suggest that Rnt1p, a member of dsRNA binding protein family, can efficiently distinguish between hairpins having closely related tetraloops.

Recently, a genome-wide search for Rnt1p substrates identified the sequence adjacent to the snoRNA 48 (snR48) as a substrate for Rnt1p (Fig. 1B).(Ghazal et al., 2005) However, analysis of the secondary structure near snR48 identified an AAGU tetraloop structure (A1-loop) instead of the canonical G2-loop near the cleavage site, indicating that Rnt1p may use different motifs for RNA cleavage. In this study we examined the features of this new class of substrates and compared it to that of the classical G2-loop family in order to understand how Rnt1p discriminates between its substrates and other RNA hairpins. The results indicate that in the absence of guanosine

in the second position of the tetraloop, Rnt1p uses an alternative set of interactions involving the first and the third nucleotides of the terminal loop as well as nucleotides in the stem for substrate selection. This new set of interactions allows Rnt1p to identify A1-loops and distinguish them from others loops associated with uncleavable RNA hairpins.

Results

Rnt1p recognizes substrates with different tetraloop structure

Genome-wide analysis of the expression level of all known snoRNAs upon the deletion of Rnt1p identified the nascent transcript of snR48 as a potential substrate for Rnt1p.(Ghazal et al., 2005)*In silico*, folding of the RNA sequence upstream and downstream of the mature sequence of snR48 did not reveal a stable hairpin with a terminal G2-loop (Fig. 1A and data not shown). Instead, a thermodynamically stable stem capped with A1-loop was identified near the 5' end of snR48 (Figure 1B). Phylogenetic analysis of sequences adjacent to snR48 in four closely related *Saccharomyces* species (*sensu stricto*) supports the existence of a conserved AAGU capped stem near the 5' end of the predicted snR48 (Fig. 1B). The most conserved nucleotides of the loops are A and G in the first and third position of the loop respectively. Incubation of a T7 transcribed RNA representing an A1-substrate (R48) with recombinant Rnt1p produced cleavages 14 and 16 nt downstream from the AAGU tetraloop, (Fig. 1C), as in the case of NGNN substrates(Lamontagne et al., 2003).

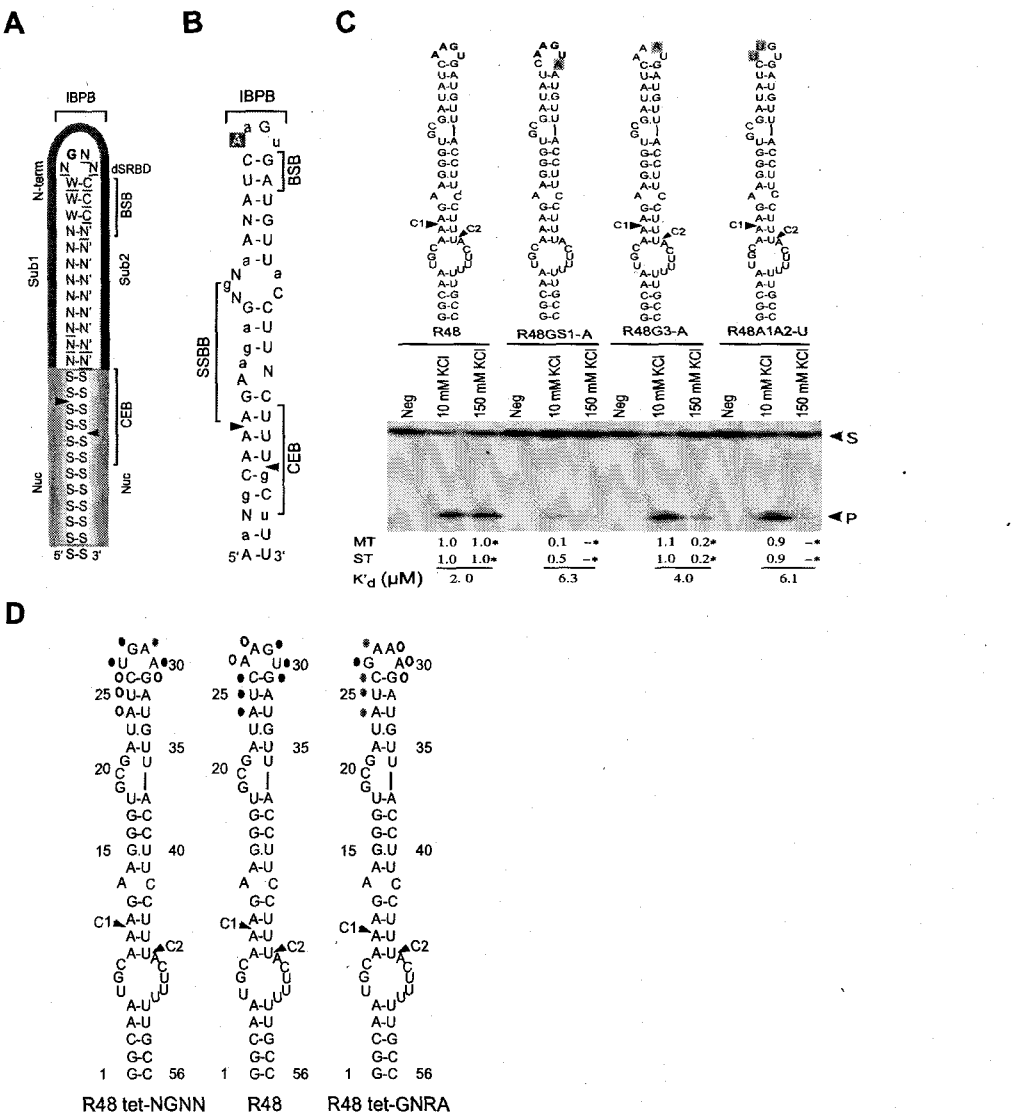
To examine the contribution of the A1-loop to substrate recognition and cleavage we introduced mutations that disrupt pairing of the nucleotides adjacent to the loop(R48GS1-A) or that change the phylogenetically

conserved loop sequence (R48G3-A and R48A1A2-U) and tested them for cleavage. All alterations of the tetraloop blocked cleavage at physiological conditions indicating that the A1-loop directs cleavage by Rnt1p. The substrates were incubated with Rnt1p, either under low salt (10 mM KCl) or physiological salt concentration (150 mM KCl). Like most members of the RNase III family, the affinity of Rnt1p to its substrate is enhanced at low salt concentration allowing the cleavage of poor substrates and this often results in cleavage at secondary cleavage sites.(Lamontagne et al., 2000; Li et al., 1993) Therefore, testing RNase III at two different salt conditions allowed us to differentiate between features of the RNA substrates that are essential for recognition by Rnt1p from those required for enhancing the association with the enzyme. To monitor the cleavage efficiency under both single and multiple turnover conditions, relative cleavage rates were calculated either in the presence of excess substrate (10:1) or in the presence of excess protein (1:400). As shown in Figure 1C, a mutation disrupting the closing base-pair of an A1-loop (R48GS1-A) blocked cleavage under physiological salt conditions, and significantly reduced cleavage efficiency under low salt conditions. In order to determine whether or not the reduction in the cleavage observed with R48GS1-A is a result of reduced affinity to Rnt1p, we measured the ability of Rnt1p to bind this substrate in the absence of Mg^{2+} using gel mobility shift assay and compared it to R48 (Figure 1C). As expected, the dissociation constant of R48GS1-A was 3 times higher than that of the wild type substrate

R48 suggesting that at least in part the inefficient cleavage of R48GS1-A reflects reduced affinity to Rnt1p.

We noticed the presence of a phylogenetically conserved G in the third position of the A1-loop (Fig. 1B), which usually contains a variable nucleotide in G2-loops. To test the contribution of this nucleotide to Rnt1p selection of A1-substrates, we mutated this G to A (R48G3-A) and measured the impact on Rnt1p binding and cleavage. As shown in Figure 1C, the cleavage efficiency of R48G3-A was similar to that of R48 under low salt condition, but was greatly reduced under physiological salt conditions. Rnt1p binding to R48G3-A was reduced twofold when compared to R48. These data suggest that the presence of a guanosine in the third position of the A1-loops enhances cleavage by Rnt1p. A mutation that changed the AAGU to UUGU (R48A1A2-U) reduced binding and blocked cleavage under physiological salt condition, while allowing efficient cleavage under low salt conditions. To ensure that the effect of these mutations on Rnt1p cleavage is not due to global rearrangement of the RNA structure we probed the secondary structure of the different hairpins using RNases V_1,(Lowman and Draper, 1986) T1,(Nazar and Wildeman, 1981; Wildeman and Nazar, 1980, 1981) and T2(Douthwaite et al., 1983; Garrett et al., 1984; Krol et al., 1981; Vigne and Jordan, 1977). No major effect on the secondary structure was detected except in the case of R48GS1-A, which carries mutations that disrupt the closing base-pair (data not shown). As predicted, disruption of the closing

base-pair destroyed the tetraloop structure of R48GS1-A. Thus, the A1-loop promotes binding to Rnt1p and the conserved nucleotides are required for the cleavage under physiological conditions.



Ghazal and Abou Elela, 2006
Figure 1

Figure 1. Rnt1p cleaves RNA hairpins capped with AAGU tetraloops. (A) Model representing the canonical NGNN (G2) dependent substrates of Rnt1p and the nucleotides known to interact with the dsRBD. IBPB, BSB, and CEB indicate the nucleotide groups that affect Rnt1p initial binding and positioning, binding stability, and cleavage efficiency respectively. The double stranded RNA binding domain (dsRBD) is illustrated in black, the nuclease domain (Nuc) in light gray, and the N-terminal domain (N-term) in dark gray. N represents any nucleotide and N' its complementary nucleotide including wobbles. W-C indicates positions where base-pairing is required. S indicates regions where base-pairing is not required. The underlined nucleotides indicate positions where the dsRBD was shown to interact in the solution structure of dsRBD / AGNN hairpin.(Wu et al., 2004) The arrowheads indicate the sites of cleavage. (B) Consensus structure of Rnt1p cleavage signal at the 5' end of snR48. Nucleotides shown in uppercase are conserved in all species. Lowercase indicates nucleotides conserved in four species. Note the presence of a new group of nucleotides representing a stem specific binding box (SSBB) important for tetraloop independent binding and cleavage. (C) Rnt1p cleaves the conserved stem at the 5' end of snR48 in AAGU (A1) dependent manner. RNA hairpins representing either the conserved stem (R48) or versions carrying different mutations altering the loop nucleotides or structure were T7 transcribed, 5' end labeled and tested for binding and cleavage by Rnt1p. The percentage of cleavage of each substrate in single turnover conditions (ST) or multiple turnover conditions (MT) was normalized

with that obtained with the wild-type control. Average values obtained from three independent experiments carried in either 10 mM or 150 mM KCl (identified by asterisks) are indicated. The gel shown is an example of cleavage carried out in RNA excess (MT). The apparent K'_d deduced from three independent gel mobility shift assays is shown at bottom. Arrowheads indicate the position of the substrate (S) and the 5' end cleavage products (P). (D) Comparison between the hydroxyl radicals generated cleavage patterns of A1 (R48), G2 (R48 tet-NGNN), and GNRA (R48 tet-GNRA) hairpins. Circles indicate the cleavages observed near the tetraloops. Strong cleavages are indicated by black circles, moderate cleavages are indicated by gray circles, and weak cleavages are indicated by open circles. The result is an average of the 3 different experiments.

The A1-loop forms a distinct tertiary structure

The A1-loop may solicit cleavage by Rnt1p either by mimicking the structure of G2-loops or by forming an alternative structure that is independently recognized by Rnt1p. To differentiate between these two possibilities we probed the tertiary structure of A1-loop (R48) and compared it to that of G2-loops (R48 tet-NGNN) and the highly structured GNRA tetraloops (R48 tet-GNRA), which is not recognized by Rnt1p.(Lebars et al., 2001) The results suggest that the A1-loop 3' end exhibits a different structure than that formed by either A1-loop or GRNA tetraloop (Fig. 1D). The tertiary structures of each hairpin were probed using hydroxyl radicals generated by the Fe (II) – EDTA complex with hydrogen peroxide.(Tullius and Dombroski, 1985) The free radicals generated cleave the nucleic acid backbone in a manner that is insensitive to secondary structure and dependent on the surface accessibility of each ribose C4'.(Han and Dervan, 1994) Therefore, variations in cleavage intensity indicate differences in the tertiary structure of RNA molecules.(Nazar, 1991) Comparison of the cleavage patterns of R48 and R48 tet-UGAA reveals strikingly different features of the region near the tetraloops. Nucleotides within or adjacent to the tetraloop of R48 were more accessible than those of R48 tet-NGNN, with the exception of the first two adenosines (A27 and 28) of the AAGU tetraloop. A similar increase in accessibility was observed when comparing the R48 upper stem with that of either R48 tet-GNRA or R48 tet-NGNN. Nucleotides in positions A24, U25 and C26 were distinctly more accessible in R48 than in the case of both R48

tet-NGNN and R48 tet-GNRA. A1-loop exhibited specific protection in position A27 and A28 that cannot be detected in other loop sequences. We conclude that A1-loops adopt a structure different than that of both GNRA and G2-loops and suggest that the affinity of Rnt1p to A1-substrates is not due to its close similarity to the G2-structure.

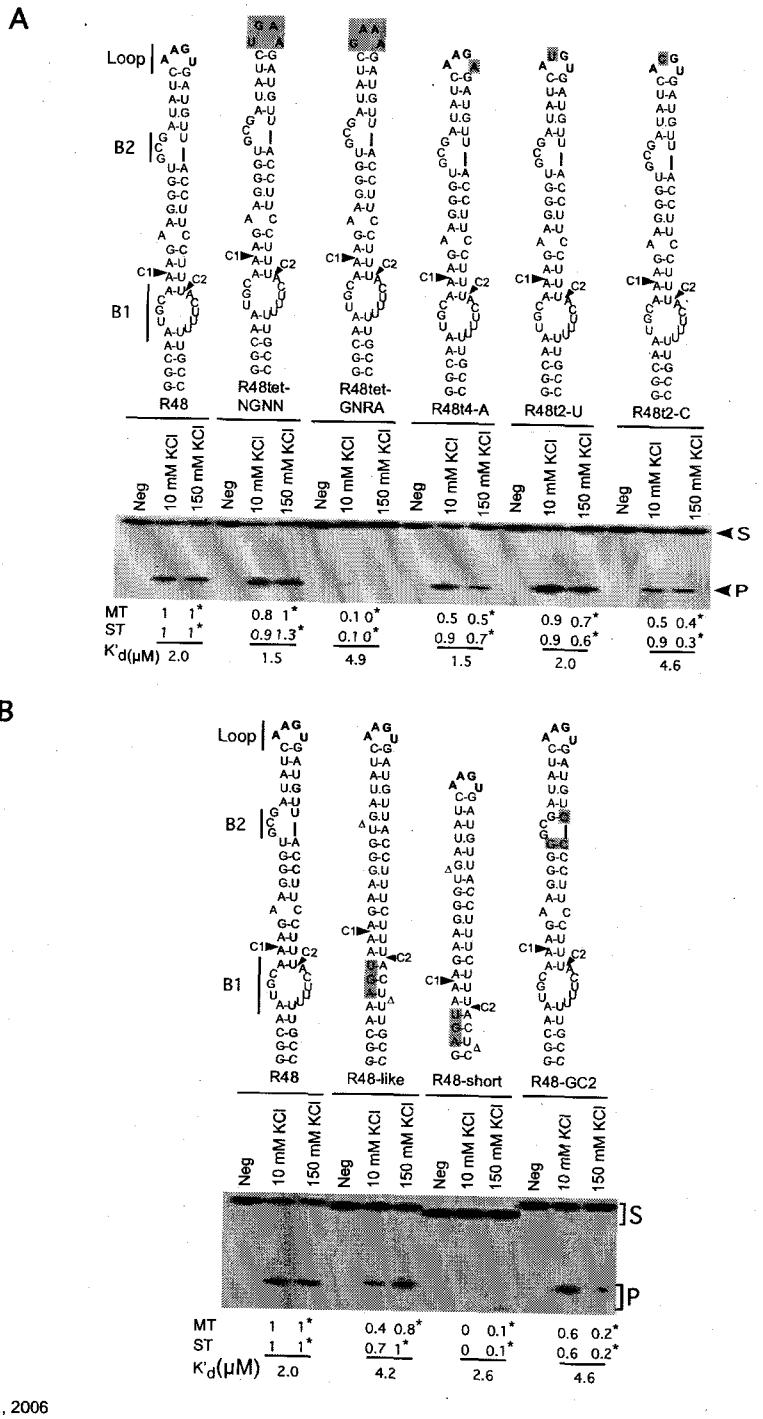
Rnt1p recognizes the stem structure of A1 substrates

Alteration of the G2-loop completely blocks Rnt1p binding and cleavage.(Lamontagne et al., 2003) In contrast, changes of the A1-loop sequence permits weak but accurate cleavage by Rnt1p (Fig. 1C), suggesting that the enzyme recognizes the A1-stem through a different binding mode. To investigate this possibility, we examined the effects of different mutations that alter the A1 loop and stem structure on Rnt1p binding and cleavage. The results indicate that while the A1-loop is required for efficient cleavage the internal bulge in the upper stem can accurately direct Rnt1p cleavage regardless of the loop sequence. As shown in Figure 2A, Rnt1p efficiently cleaved R48 under both low and physiological salt conditions. Also, replacement of the A1-loop with a G2-loop in the context of the R48 stem (R48tet-NGNN) did not significantly affect binding or cleavage, indicating that the A1-stem is functionally compatible with the G2-structure. Both R48 and R48 tet-NGNN are cleaved with a k_M value of about 1.1 μM. However, the k_{cat} of R48 tet-NGNN was 21.5 min^{-1} while that of R48 was about 9.25 min^{-1},

suggesting that the G2-loop supports a higher turnover rate than the A1-loop. As expected, replacement of the A1-loop with a GNRA tetraloop (R48tet-GAAA) that blocks cleavage and binding of G2-stems (Lamontagne et al., 2003) allowed weak, but detectable cleavage at low salt concentration when associated with the R48 stem. Unlike G2-hairpins, where single mutations that change the conserved G blocks binding to Rnt1p (Chanfreau et al., 2000; Nagel and Ares, 2000), changing the R48 tetraloop to GNRA (R48tet-GAAA) reduced binding by 2.5 fold. This suggests that A1-stems have higher affinity to Rnt1p than those found near G2-loops. Changing the sequence of the fourth nucleotide of the A1-loops to A (R48t4-A) moderately reduced cleavage under turnover conditions without affecting the binding affinity, suggesting that the sequence of the loop is required for productive interaction with the enzyme. Changing the second nucleotide of the tetraloop A28 to U (R48t2-U) or C (R48t2-C) also reduced Rnt1p cleavage, but only the change from A to C significantly reduced the binding affinity to Rnt1p. This result demonstrates the capacity of Rnt1p to use different features within the dsRNA stem for substrate recognition.

The unique ability of the R48 stem to support residual cleavage by Rnt1p (Fig. 2A) necessitated a search for the features that distinguishes it from other short RNA helices. Elimination of the unpaired nucleotides forming the two internal loops characteristic of the R48 substrate (R48-like) modestly reduced Rnt1p cleavage in low salt, but not at physiological salt concentrations (Fig.

2B). Rnt1p bound R48-like with an apparent K'_d of 4.2 μM, about two times more than that of R48. This result indicates that the internal loops of R48 are



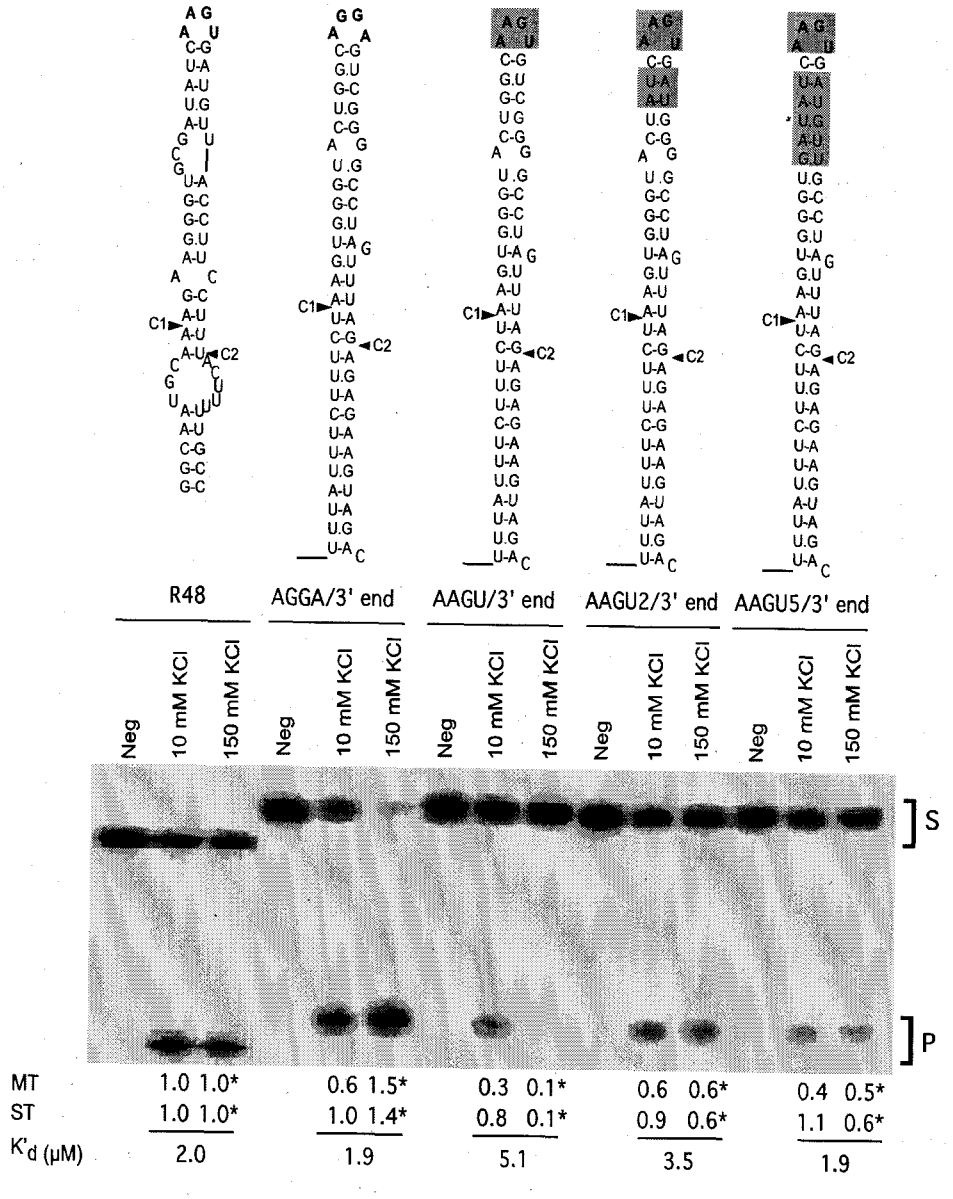
Ghazal and Abou Elela., 2006
Figure 2

Figure 2. Rnt1p recognizes both stem and loop structures of A1-dependent substrates. (A) Rnt1p binding and cleavage of RNA substrates with mutations that change the tetraloop nucleotides. The arrowheads indicate the positions of Rnt1p cleavages. The gray boxes indicate the site of mutations. The position of nucleotides deletions indicated by Δ. The different RNAs were either 5' end (shown) or 3' end labeled (not shown) and incubated either in the absence (Neg) or the presence of Rnt1p under either low (10 mM KCl) or physiological (150 mM KCl) salt concentrations. The cleavage was performed in multiple turnover (MT) and single turnover (ST) conditions. The positions of the substrates (S) and products (P) are indicated on the right. The relative cleavage rates, shown at the bottom, were determined by quantifying the products using Instant Imager. The percentage of cleavage of each substrate was normalized with that obtained with the wild-type control. The K'_d value was obtained by gel mobility shift assay and represents the average of three different experiments. (B) Binding and cleavage assay of R48 variants carrying mutations that alters the stem structure. The experiments were conducted and illustrated as described in A.

not essential for cleavage, but are required for optimal binding to Rnt1p. Surprisingly, the reduction of the R48-like stem from 24 to 20 bp (R48-short), which does not affect the cleavage of G2-substrates(Lamontagne and Abou Elela, 2004; Lamontagne et al., 2003) blocked cleavage of R48-short without significantly affecting its binding to Rnt1p (Fig. 2B). Mutations that alter the sequence context of the internal bulge in the upper stem of A1-substrate reduced cleavage, especially under physiological salt conditions. We suggest that Rnt1p independently recognizes the structure generated by the A1-stem.

To examine the capacity of A1-loops to direct Rnt1p cleavage in a canonical G2-stem we generated and tested a chimeric substrate composed of a G2-stem found near the 3' end of the 25S rRNA and an AAGU tetraloop. We chose the 3' end of the 25S rRNA(Abou Elela et al., 1996; Kufel et al., 2000) because it is the most efficiently cleaved substrate known to date.(Lamontagne and Abou Elela, 2001) As shown in Figure 3, Rnt1p efficiently bound and cleaved the natural AGGA/3'end substrate. Interestingly, replacing the G2-loop with an A1-loop (AAGU/3'end) greatly reduced binding to Rnt1p and inhibited cleavage under physiological conditions. The observed inhibition is not due to perturbations in the stem structure as confirmed by the *in silico* and *in vitro* analyses of AAGU/3'end structures (data not shown). Previous studies indicated that the reactivity of G2-substrates is influenced by the first 4 base-pairs of the stem. (Lamontagne et al., 2003) Accordingly, we changed the stem sequence near the tetraloop of AAGU/3'end to resemble

that of R48 (Fig. 3). The substrate carrying the first 3 base pairs of R48 (AAGU2/3'end) was bound and cleaved



Ghazal and Abou Elela, 2006
Figure 3

Figure 3. Rnt1p requires different stem sequence for the cleavage of A1- and G2-substrates. RNA hairpins representing the R48 structure, the G2 cleavage signal found at the 25S rRNA 3' end (AGGA/3'), AGGA/3' capped with A1-loop (AAGU/3' end) and other variants where the upper stem sequence of AAGU/3' end was substituted with that of R48 (AAGU2/3' end and AAGU5/3' end) were assayed for cleavage and binding. Gray boxes indicate the sites were sequence was substituted with that of R48. Cleavage rates and binding affinities were calculated as described in Figure 1.

more efficiently than the AAGU/3'end. Substitution of the 6 base-pairs adjacent to the tetraloop of AAGU/3'end with those found in R48 (AAGU5/3'end) did not enhance cleavage by Rnt1p. We conclude that A1- and G2-loops require different stem sequences for optimal binding and cleavage.

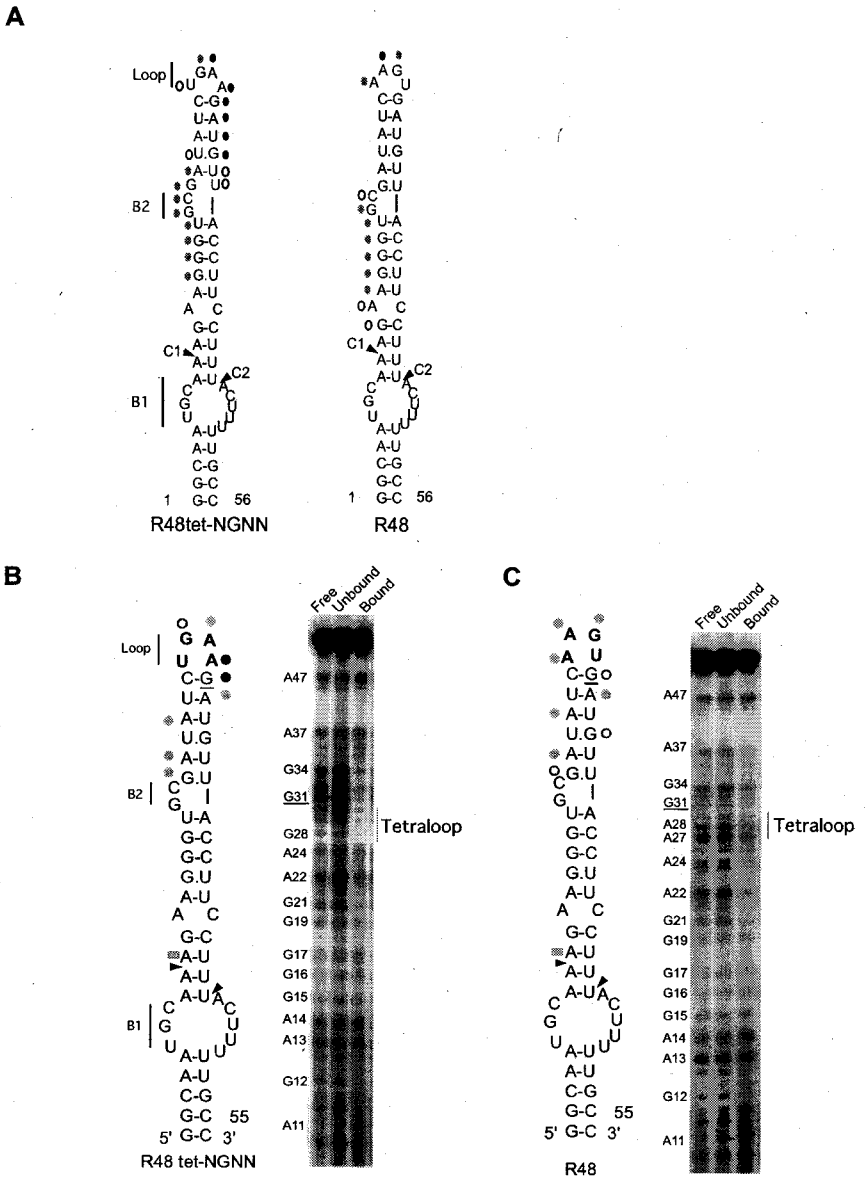
Rnt1p binds to G2- and A1-Loops using different sets of nucleotides

Rnt1p dsRBD interacts with G2-lsubstrates by forming hydrogen bonds with the nucleotides near the 3' end of the loop.(Lamontagne et al., 2003; Lamontagne et al., 2004) In order to understand how A1-substrates are selected for cleavage, we compared the A1-substrates pattern of Rnt1p-dependent protection from hydroxyl radicals generated by Fe (II)-EDTA(Han and Dervan, 1994) to that exhibited by G2-substrates. As shown in Figures 4A, R48 tet-NGNN exhibited a strong protection in the third and fourth nucleotide of the loop and the first 4 nucleotides of the stem near the loop 3' end. This protection pattern is consistent with previous chemical footprinting analysis of the G2-hairpins(Lamontagne and Abou Elela, 2004) as well as the interactions identified in the solution structure of Rnt1p dsRBD-substrate complex(Wu et al., 2004). Interestingly, a new set of footprints was observed near the second bulge of the R48 stem, but not in the canonical RNA stems associated with G2-loops.(Lamontagne and Abou Elela, 2004) The protection pattern of R48 was noticeably different from that observed for R48 tet-NGNN.

The strong G2-loop specific protection near the 3' end of the loop was not observed with R48. Instead, enhanced protection of the second nucleotide of the A1-loop and a new site of protection in the first position were observed. In addition, the stem specific footprints near the upper internal bulge were shifted downward towards the Rnt1p cleavage site in the case of the A1-substrates. The result indicates that Rnt1p selects A1-substrates by interacting with the loop 5' end, while identifying G2-substrates by binding near the loop 3' end.

In order to demonstrate that Rnt1p indeed requires different sets of nucleotides to binds the A1- and G2-substrates, we examined the chemical interferences pattern of these two classes of substrates. Chemical interference analysis defines the nucleotides that reduce a substrates capacity to interact with proteins when chemically modified.(Karaoglu and Thurlow, 1991) R48 tet-NGNN and R48 hairpins were chemically modified using DEPC and incubated with Rnt1p. Free and bound RNA fractions were separated by native PAGE, eluted from gels and cleaved with aniline. The cleavage products were separated by denaturing PAGE and visualized by autoradiography (Fig. 4B and C). Nucleotides that are under-represented in the bound RNA population and at the same time over-represented in the free RNA populations are considered to be required for binding to Rnt1p. As shown in Figure 4B, R48 tet-NGNN exhibited a modification pattern consistent with an important role for the 3' end of the loop. The strongest inhibitory effect

was observed upon the modification of the nucleotides A30 and G31. Moderate interference was observed upon the modification of the third nucleotide of the loop (A29) and the second nucleotide of the stem (A32). Weak interference was also



Ghazal and Abou Elela, 2006
Figure 4

Figure 4. Rnt1p uses different nucleotides for binding A1- and G2-substrates. (A) Summary of Rnt1p dependent protection against cleavage by hydroxyl radical damages. Sites of strong, moderate and weak protections are indicated by black, gray and open circles respectively. The hydroxyl radical cleavage patterns was generated by cleavage of 5' end labeled RNA incubated with recombinant Rnt1p in the presence of hydroxyl radical generated by Fe (II) – EDTA complex with hydrogen peroxide. The cleavages were quantified using Gel Doc (Bio-Rad, Mississauga, ON, Canada). (B and C) Identification of the nucleotides required by Rnt1p for the selection of A1-hairpins. Modification interference analysis was carried out using T7 transcribed and 5' end labeled RNA representing the structure of R48 tet-NGNN and R48. The different RNAs were modified using DEPC and incubated with recombinant Rnt1p. Bound and free RNA fractions were gel purified using standard gel mobility shift assay. RNA was eluted from gel fragments and treated with aniline to generate a cleavage ladder of the modified nucleotides. The cleaved RNA was separated using 12% denaturing gels. Black, gray and open circles indicate strong, moderate and weak inhibitory effects of the modifications respectively. Gray rectangles represent modification dependent stimulatory effects.

generated by modification of the conserved G at position 28. This result is consistent with the chemical footprinting analysis (Figure 4A), previous chemical interference assay,(Chanfreau et al., 2000) and the interaction identified in the solution structure of the dsRBD/NGNN RNA complex.(Wu et al., 2004) Surprisingly, we also observed a new set of critical nucleotides (G21, A22, and A24) near the phylogenetically conserved bulge in upper part of A1-stem. Once again these data indicated that Rnt1p independently recognized the loop and the stem of the A1-substrates. The strong interference that is normally observed upon modification of the fourth nucleotides of the G2-loop and the first nucleotide of the stem (Fig. 4B)(Chanfreau et al., 2000) was dramatically reduced in the case of A1-hairpins (Fig. 4C). Strikingly, a new site of interference was detected at the first position of the A1- but not the G2-loop. Therefore, while the G2-hairpin displayed a cluster of strongly modified nucleotides near the 3' end the loop, A1-hairpins exhibited fewer and less pronounced interfering modifications within the loop. We conclude that Rnt1p recognizes A1- and G2-substrates employing alternative modes of binding that involve a different set of determinants.

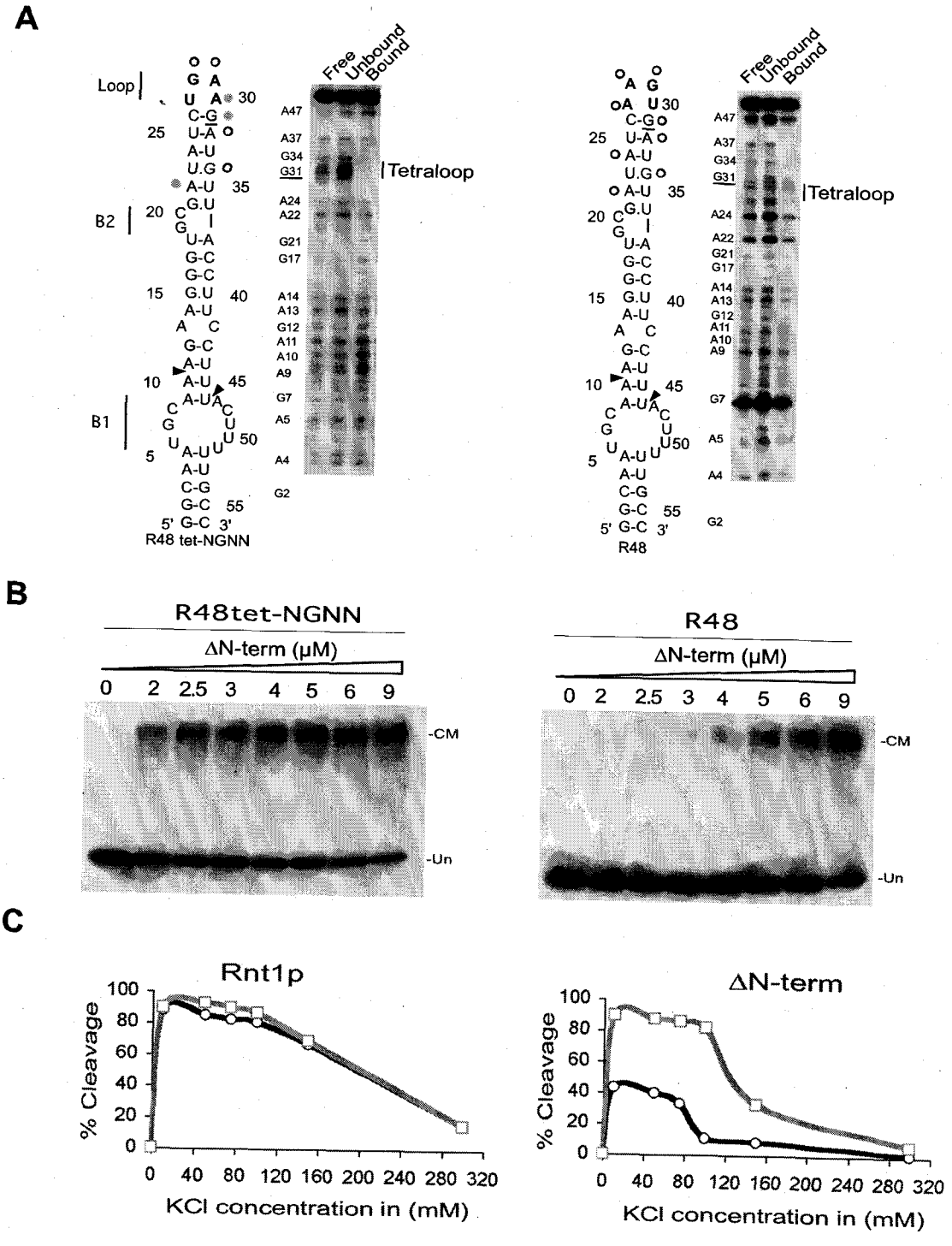
Rnt1p dsRBD selectively binds to RNA capped with A1-loops

Biochemical and structural analyses suggest that the selection of G2-substrates is performed by the dsRBD of Rnt1p. Thus it is possible that Rnt1p

also uses the dsRBD to identify the A1-substrates. To test whether the Rnt1p dsRBD preferentially binds A1-substrats, we examined the binding of protein fragments representing Rnt1p dsRBD to hairpins terminating with GNRA tetraloop (R48tet-GAAA), G2-loop (R48tet-NGNN), or A1-loop (R48). The bound and free RNAs were separated using mobility shift assay (data not shown). Surprisingly, the R48tet-GAAA substrate, which binds to the full Rnt1p enzyme with an apparent K'_d of 4.9 μM, did not bind the dsRBD even at concentrations higher than 6 μM (data not shown). As expected, R48 tet-NGNN and R48 bound to Rnt1p dsRBD efficiently with an apparent K'_d of 1.5 and 2 μM respectively. However, these two substrates exhibited different migration patterns in gel-shift assays indicating differing types of interaction with the dsRBD (data not shown). Comparison of the chemical modification patterns that interfere with the binding of the dsRBD to R48 tet-NGNN and R48 suggest that similar to the full-length Rnt1p the dsRBD binds these two substrates using different sets of nucleotides (Fig. 5A). These data show that the dsRBD is capable of discriminating between NGNN, AAGU and GAAA capped hairpins, but unlike Rnt1p it cannot independently recognize the A1-stem. We conclude that the A1-stem specific interaction with Rnt1p is not mediated by the dsRBD, but is instead mediated by either the N-terminal or nuclease domains of Rnt1p.

The N-terminal domain of Rnt1p is required for the efficient cleavage of A1-substrates

Since the dsRBD failed to bind the A1-stem we evaluated the contribution of the N-terminal domain to the binding of A1-hairpins. Mobility shift assays were performed using a truncated version of Rnt1p that extends from position 172 to 471 (ΔN-term) and contains both the dsRBD and nuclease domain. As shown in Figure 5B, the ΔN-term bound R48tet-NGNN with an apparent K'_d of 3.7 μM under physiological salt conditions, which is about 2.5 times less efficient than Rnt1p. Surprisingly, the ΔN-term bound to R48 under the same conditions with an apparent K'_d of 6.1 μM. This high K'_d suggests that the ΔN-term cleavage of A1-hairpins is much less efficient than that of G2-hairpins. To test this possibility, we compared the cleavage of R48tet-NGNN and R48 with ΔN-term and Rnt1p as control under different monovalent salt concentrations. As shown in Figure 5C, deletion of Rnt1p N-terminal domain selectively inhibited the cleavage of R48 at salt concentrations ranging from 0 to 100 mM. In contrast, cleavage of G2-hairpins (R48 tet-NGNN) was inhibited at salt concentrations above 100 mM as previously shown (Lamontagne et al., 2000). Indeed, at 100 mM salt concentration the cleavage of R48 is completely blocked while that of R48 tet-NGNN is not affected. We conclude that the N-term of Rnt1p is required for the cleavage of G2-substrate *in vivo* and suggest that Rnt1p uses different sets of amino acid residues for the selection of A1- and G2-substrates.



Ghazal and Abou Elela, 2006
Figure 5

Figure 5. Rnt1p dsRBD and N-terminal domain contributes to binding and cleavage of A1-dependent substrate. (A) Identification of the nucleotides required for interaction with dsRBD. Modification interference analysis using a truncated protein fragment representing Rnt1p dsRBD was carried out and presented as described in Figure 4. (B) Gel mobility shift assay using the Rnt1p C-terminal domain. A truncated version of Rnt1p lacking the N-terminal 191 amino acid residues (ΔN-term) was assayed for binding to both R48 tet-NGNN and R48. The positions of free (Un) and bound (CM) RNA are shown on the right. (C) The capacity of Rnt1p and the ΔN-term to cleave R48 tet-NGNN (□) and R48 (○) was assayed under increasing concentrations of monovalent salts. The cleavage reactions were carried as described and the cleavage products were quantified as in Figure 1. The percent of Rnt1p and ΔN-term cleavages were plotted versus the concentration of KCl. The data shown are the average of two experiments.

Discussion

In this study we reveal a new class of Rnt1p substrates that does not require the conserved NGNN (G2) tetraloop structure for binding or cleavage. The recognition of this new substrate depends on an AAGU (A1) tetraloop (Fig. 1) and is influenced by the stem structure (Fig. 2). Mutations in the A1-loop and stem sequence revealed a residual cleavage activity promoted by the stem structure alone (Fig. 2 and 3). Chemical probing indicated that the A1- hairpin structure is not identical to that generated by G2-loops, but instead represents a structural variant that is still recognized by Rnt1p (Fig 1). This finding is in agreement with the recently determined solution structure of the A1-hairpin that shows a different conformation from that exhibited by G2-hairpins.(Gaudin et al., 2006) Consistent with this, chemical footprinting and chemical interference assays indicated that Rnt1p binds A1- and G2-substrates using different sets of nucleotides (Fig 4). Cleavage and binding assays showed that the N-term of Rnt1p plays an important role in the selection of A1-hairpins. These results demonstrate the capacity of a member of the dsRBP family to use multiple structural motifs for substrate selection and suggest that yeast RNase III uses an adaptable network of protein RNA interactions to differentiate between closely related tetraloop structures.

The data presented here suggest that Rnt1p recognizes more than one form or spacing of the minor groove. Structural probing using hydroxyl radicals clearly showed that the local structure adjacent to the A1-loop is different than

that formed by G2-loops. Consistently, comparison between the solution structure of the A1- and G2-hairpins show that the two loops differ in the conformation of the second and fourth nucleotides of the loop.(Gaudin et al., 2006) This clearly shows that Rnt1p recognition of the A1-substrates is not due to similarity between the two structures. Indeed, hairpins capped with ACAA tetraloops assumes a similar conformation to that of G2- and A1-hairpins but is not cleaved by Rnt1p.(Gaudin et al., 2006; Staple and Butcher, 2003; Wu et al., 2001) The difference between ACAA and AGNN tetraloops is that the cytosine in the second position of the ACAA tetraloop is in *anti* glycosidic torsion angle instead of the *syn* conformation adopted by the AGNN guanosine.(Staple and Butcher, 2003) These changes in conformation by itself are sufficient to block Rnt1p cleavage of the canonical G2- but not A1-substrates. Rnt1p cleaves RNA hairpins capped with A1-loops despite the fact that the second nucleotide of the loop is in an *anti* conformation.(Gaudin et al., 2006) It is also unlikely that the sequence of the nucleotide in the second position itself blocks cleavage because the dsRBD of Rnt1p does not directly interact with this position(Wu et al., 2004) and here we show that changing the sequence of the entire tetraloop does not inhibit cleavage (Fig. 1B). Indeed, A1-hairpins were moderately cleaved by Rnt1p (Fig 2A), whereas ACAA was not (Lamontagne and Abou Elela, unpublished observation). This means that AAGU and not ACAA tetraloops contain features that allow it to be cleaved even in the absence of guanosine in the second position. We propose that the structure at the 3' end of A1-loops,

which is very different from that found in G2 (Figure 1) and ACAA hairpins, (Gaudin et al., 2006) allows it to be recognized by Rnt1p despite the structural changes in the second position of the tetraloop.

The data presented here also point to the possibility of a substrate recognition mechanism that is independent of the terminal loop. Earlier work has shown that Rnt1p may bind, but does not cleave, long RNA duplexes. (Lamontagne and Abou Elela, 2004) In addition, the *S. pombe* RNase III orthologue Pac1 may recognize an internal loop for cleavage.(Lamontagne and Abou Elela, 2004) Here we show that Rnt1p may also use, albeit poorly, an internal loop for cleavage and we have identified a stem that could be cleaved regardless of the loop sequence (Fig. 2 and data not shown). Consistently, protein footprinting assays indicate that Rnt1p interacts near the internal loop of R48 (Fig. 4A), and similar footprints were not detected in substrates lacking the internal loop.(Lamontagne and Abou Elela, 2004) This result clearly indicates that Rnt1p may bind to substrates using an alternative and tetraloop-independent set of nucleotides. The unprecedented capacity of the R48 stem to direct cleavage by Rnt1p is not coincidental because R48 must compensate for the suboptimal reactivity of its A1-loop (Fig. 2B). As shown in figure 1A, we propose that the region near the internal bulge serves as a secondary stem specific binding box (SSBB) that contributes to the binding and cleavage of R48. Given the size and locations of the different RNase III protein domains(Leulliot et al., 2004; Wu et al., 2004), it is likely that the

SSBB binds to either to the nuclease domain or to the dsRBD of second subunit of the Rnt1p homodimer, which does not interact with the tetraloop region. The capacity of Rnt1p to cleave stems with internal loops or other specific structural motifs would allow Rnt1p to recognize a subset of long RNA duplexes and thus allow this enzyme to cleave yet a greater number of RNA targets.

Flexible and adaptable substrate selection by Rnt1p appears to be encoded in the basic structure of RNase III. The crystal structure of bacterial RNase III indicates that the enzyme uses four RNA binding motifs to interact with three regions within the RNA duplex.(Gan et al., 2006) Two of the RNA binding motifs are located in the dsRBD while the other two are found in the NUCD. The two dsRBD binding sites are required for stable interaction with the RNA while the two NUCD binding sites are required for the formation of catalytically active complexes.(Gan et al., 2005, 2006) The RNase III binding configuration reduces the dependency on a specific structural feature and permits a flexible mode of RNA recognition that allows substrate identification using combinations of different structural features. This mode of substrate recognition would explain why many mutations in Rnt1p substrates that modestly reduce RNA binding greatly inhibit cleavage, while others that inhibit cleavage do not affect binding (Figs. 1C and 2). In addition, multiple interactions with the substrate may allow Rnt1p to recognize different RNA

hairpins as long as the overall network of interaction correctly position the RNA within the nuclease domain for cleavage.

Yeast RNase III is distinguished from other members of the ribonuclease III family by the presence of a long N-term that does not contain a recognizable motif.(Lamontagne et al., 2000) The contribution of the N-term to the cleavage of A1-substrate is different than that of the G2-substrates (Fig. 5C). Deletion of the N-term inhibits cleavage, but not binding to G2-hairpins(Lamontagne et al., 2000) suggesting that the N-term does not contribute to the selection of this class of substrates. In contrast, deletion of the N-term drastically inhibits both binding and cleavage of A1-hairpins (Fig. 5C), indicating that the N-term contributes to the selection of A1-substrates. It is possible that the N-term influences the selection of the substrates by interacting with, and thus stabilizing, a specific contact between the dsRBD and the RNA (Fig. 1A). The increased dependence of A1-hairpins on the N-term for cleavage suggests that Rnt1p modifies the mode of interaction to better fit suboptimal RNA structures. Together the results presented here reveal an adaptable mechanism of substrate selection that allows yeast RNase III and perhaps other members of the RNase III family to recognize a broad spectrum of structural motifs.

Methods

Phylogenetic Analysis

A search for a conserved structure that may represent Rnt1p cleavage site was conducted using BLAST (Altschul et al., 1990) in four *Saccharomycessensu stricto* species: *S. paradoxus*, *S. mikatae*, *S. bayanus*(Kellis et al., 2003) and *S. kudriavzevii.*(Notredame et al., 2000) The genome of these species codes for RNase III orthologues with high similarity to *S. cerevisiae* and thus assumed to have similar substrate specificities. The homologous sequences were searched for stable tetraloop structures upstream of the predicted snR48 5' end. In *S. cerevisiae*, the tetraloop is located at 94 nucleotides from the snR48 5' end. In the four other *Saccharomyces* species similar tetraloop were found at distances ranging from 78 to 96 nt of the predicted snoRNA 5' end. The five stem-loop sequences were aligned using T-Coffee 2.66.(Notredame et al., 2000)

Enzymatic assay

RNA transcripts used for *in vitro* cleavage were generated by T7 RNA polymerase and were 5' end labeled as described earlier.(Lamontagne et al., 2003)*In vitro* cleavage was performed using 0.8 pmol of recombinant enzyme

in a 20 µl reaction buffer (30 mM Tris-HCl pH 7.5, 150 mM KCl, 5 mM spermidine, 0.1 mM DTT. 0.1 mM EDTA pH 7.5). Cleavage reactions were performed using 3 fmol of 5' end labeled RNA, for single turnover conditions, and 8 pmol substrate for multiple turnover conditions. The cleavage products were quantified using an Instant Imager (Packard, Meriden, CT). Calculations and curves were performed using the GraphPad Prism 3.0 program (GraphPad Software, CA). All experiments were performed a minimum of three times.

Gel shift assay

RNA binding assays were performed using 2 fmol of T7 transcribed internally labeled RNA (α-CTP^{32}) prepared as previously described.(Lamontagne and Abou Elela, 2001) Gel mobility shift assays were performed in 20 µl binding buffer (20% (V/V) glycerol, 30 mM Tris pH 7.5, 150 mM KCl, 5 mM spermidine, 0.1 mM DTT, and 0.1 mM EDTA pH 7.5). Protein concentrations used in the assays ranged from 2 to 9 µM. Reactions were loaded on 4% non-denaturing PAGE and processed as previously described.(Lamontagne and Abou Elela, 2001)

Hydroxyl radical footprinting

5' end labeled T7 transcribed RNA (28 fmol) was incubated with 1 - 6 µM recombinant Rnt1p in 17 µl of 1X MOPS buffer (20 mM MOPS pH 7.5, 100 mM KCl, 0.1 mM DTT and 0.1 mM EDTA pH 7.5) and incubated on ice for 15 min.(Lamontagne and Abou Elela, 2004) Cleavage with hydroxyl radical was started by the addition of a freshly prepared mix containing 1 µl of 40 mM ferrous ammonium sulfate, 80 mM Na_3-EDTA pH 8.0, 1 µl of 40 mM sodium ascorbate, and 1 µl of 2.4% hydrogen peroxide. The reactions were incubated on ice for 30 sec and stopped by adding 4 µl of 100 mM thiourea. The RNA was ethanol precipitated and resuspended in RNA loading dye (0.05% xylene cyanol and 0.05% bromophenol blue in 100% formamide). The samples were loaded on 12% PAGE, and the dried gels were exposed to film. Films were scanned and analyzed using Molecular Analyst software (Bio-Rad, Mississauga, ON, Canada). All experiments were performed at least three times.

Modification interference

A 211µl reaction mix containing about 20 nmol of 5' end labeled RNA, 1 µg of tRNA, 50 mM sodium acetate pH 4.5, and 1 mM EDTA was incubated on ice for few seconds. The RNA modifications were achieved by adding 1µl fresh diethylpyrocarbonate (DEPC) and incubated for 10 minutes at 90°C as described earlier.(Conway and Wickens, 1989; Peattie, 1979) The modified RNA was incubated with different concentrations of full length or truncated

versions of Rnt1p. Free and bound RNA populations were separated using gel mobility shift assay as described before.(Lamontagne and Abou Elela, 2001) The different RNA bands were excised from gels and the RNA was eluted overnight in 500 µl LETS buffer (0.01M Tris-HCl pH7.5, 0.1M LiCl, 0.01M Na_2EDTA pH8, and 0.2% SDS) and 500 µl phenol at 4 °C. The extracted RNAs were precipitated with ethanol and treated with aniline as described.(Peattie, 1979)

Acknowledgments

We would like to thank Jules Gagnon for help in preparing Figure 1B and bioinformatics support. We also thank Roscoe Klinck, Dominique Fourmy, Jean-Pierre Perreault, Liam Good and Bruno Lamontagne for critical reading of the manuscript. This work was supported by grant # MOP-67162 from the Canadian Institute for Health Research (CIHR). S. A. is a Chercheur-Boursier Senior of the Fonds de la Recherche en Santé du Québec.

References

1. Nicholson, A. W. (1999). Function, mechanism and regulation of bacterial ribonucleases. *FEMS Microbiol Rev***23**, 371-90.

2. Conrad, C. & Rauhut, R. (2002). Ribonuclease III: new sense from nuisance. *Int J Biochem Cell Biol***34**, 116-29.

3. Lamontagne, B., Larose, S., Boulanger, J. & Elela, S. A. (2001). The RNase III family: a conserved structure and expanding functions in eukaryotic dsRNA metabolism. *Curr Issues Mol Biol***3**, 71-8.

4. Lykke-Andersen, J., Aagaard, C., Semionenkov, M. & Garrett, R. A. (1997). Archaeal introns: splicing, intercellular mobility and evolution. *Trends Biochem Sci***22**, 326-31.

5. Lamontagne, B. & Elela, S. A. (2004). Evaluation of the RNA determinants for bacterial and yeast RNase III binding and cleavage. *J Biol Chem***279**, 2231-41.

6. Ghazal, G., Ge, D., Gervais-Bird, J., Gagnon, J. & Abou Elela, S. (2005). Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. *Mol Cell Biol***25**, 2981-94.

7. Nicholson, A. W. (1996). Structure, reactivity, and biology of double-stranded RNA. *Prog Nucleic Acid Res Mol Biol***52**, 1-65.

8. Rotondo, G., Huang, J. Y. & Frendewey, D. (1997). Substrate structure requirements of the Pac1 ribonuclease from Schizosaccharmyces pombe. *RNA***3**, 1182-93.

9. Mitra, S. & Bechhofer, D. H. (1994). Substrate specificity of an RNase III-like activity from Bacillus subtilis. *J Biol Chem***269**, 31450-6.

10. Abou Elela, S., Igel, H. & Ares, M., Jr. (1996). RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. *Cell***85**, 115-24.

11. Lamontagne, B. & Abou Elela, S. (2001). *Purification and characterization of Saccharomyces cerevisiae Rnt1p nuclease.* Methods Enzymol (Nicholllson, A. W., Ed.), 342, Academic Press, San diego, CA.

12. Catala, M., Lamontagne, B., Larose, S., Ghazal, G. & Elela, S. A. (2004). Cell cycle-dependent nuclear localization of yeast RNase III is required for efficient cell division. *Mol Biol Cell***15**, 3015-30.

13. Lamontagne, B., Tremblay, A. & Abou Elela, S. (2000). The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. *Mol Cell Biol***20**, 1104-15.

14. Leulliot, N., Quevillon-Cheruel, S., Graille, M., Van Tilbeurgh, H., Leeper, T. C., Godin, K. S., Edwards, T. E., Sigurdsson, S. T., Rozenkrants, N., Nagel, R. J., Ares, M. & Varani, G. (2004). A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. *EMBO J***23**, 2468-2477.

15. Wu, H., Henras, A., Chanfreau, G. & Feigon, J. (2004). Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc Natl Acad Sci U S A***101**, 8307-12.

16. Zhang, K. & Nicholson, A. W. (1997). Regulation of ribonuclease III processing by double-helical sequence antideterminants. *Proc Natl Acad Sci U S A***94**, 13437-41.

17. Chanfreau, G., Buckle, M. & Jacquier, A. (2000). Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. *Proc Natl Acad Sci U S A***97**, 3142-7.

18. Lamontagne, B., Ghazal, G., Lebars, I., Yoshizawa, S., Fourmy, D. & Elela, S. A. (2003). Sequence dependence of substrate recognition and cleavage by yeast RNase III. *J Mol Biol***327**, 985-1000.

19. Lebars, I., Lamontagne, B., Yoshizawa, S., Abou-Elela, S. & Fourmy, D. (2001). Solution structure of conserved AGNN tetraloops: insights into Rnt1p RNA processing. *EMBO J***20**, 7250-8.

20. Wu, H., Yang, P. K., Butcher, S. E., Kang, S., Chanfreau, G. & Feigon, J. (2001). A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. *EMBO J***20**, 7240-9.

21. Staple, D. W. & Butcher, S. E. (2003). Solution structure of the HIV-1 frameshift inducing stem-loop RNA. *Nucleic Acids Res***31**, 4326-31.

22. Li, H. L., Chelladurai, B. S., Zhang, K. & Nicholson, A. W. (1993). Ribonuclease III cleavage of a bacteriophage T7 processing signal. Divalent cation specificity, and specific anion effects. *Nucleic Acids Res***21**, 1919-25.

23. Lowman, H. B. & Draper, D. E. (1986). On the recognition of helical RNA by cobra venom V1 nuclease. *J Biol Chem***261**, 5396-403.

24. Nazar, R. N. & Wildeman, A. G. (1981). Altered features in the secondary structure of Vicia faba 5.8s rRNA. *Nucleic Acids Res***9**, 5345-58.

25. Wildeman, A. G. & Nazar, R. N. (1981). Studies on the secondary structure of 5.8 S rRNA from a thermophile, Thermomyces lanuginosus. *J Biol Chem***256**, 5675-82.

26. Wildeman, A. G. & Nazar, R. N. (1980). Nucleotide sequence of wheat chloroplastid 4.5 S ribonucleic acid. Sequence homologies in 4.5 S RNA species. *J Biol Chem***255**, 11896-900.

27. Garrett, R. A., Christensen, A. & Douthwaite, S. (1984). Higher-order structure in the 3'-terminal domain VI of the 23 S ribosomal RNAs from Escherichia coli and Bacillus stearothermophilus. *J Mol Biol***179**, 689-712.

28. Douthwaite, S., Christensen, A. & Garrett, R. A. (1983). Higher order structure in the 3'-minor domain of small subunit ribosomal RNAs from a gram negative bacterium, a gram positive bacterium and a eukaryote. *J Mol Biol***169**, 249-79.

29. Krol, A., Gallinaro, H., Lazar, E., Jacob, M. & Branlant, C. (1981). The nuclear 5S RNAs from chicken, rat and man. U5 RNAs are encoded by multiple genes. *Nucleic Acids Res***9**, 769-87.

30. Vigne, R. & Jordan, B. R. (1977). Partial enzyme digestion studies on Escherichia coli, Pseudomonas, Chlorella, Drosophila, HeLa and yeast 5S RNAs support a general class of 5S RNA models. *J Mol Evol***10**, 77-86.

31. Tullius, T. D. & Dombroski, B. A. (1985). Iron(II) EDTA used to measure the helical twist along any DNA molecule. *Science***230**, 679-81.

32. Han, H. & Dervan, P. B. (1994). Visualization of RNA tertiary structure by RNA-EDTA.Fe(II) autocleavage: analysis of tRNA(Phe) with uridine-EDTA.Fe(II) at position 47. *Proc Natl Acad Sci U S A***91**, 4955-9.

33. Nazar, R. N. (1991). Higher order structure of the ribosomal 5 S RNA. *J Biol Chem***266**, 4562-7.

34. Nagel, R. & Ares, M., Jr. (2000). Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5'-AGNN-3' tetraloop. *RNA***6**, 1142-56.

35. Kufel, J., Allmang, C., Chanfreau, G., Petfalski, E., Lafontaine, D. L. & Tollervey, D. (2000). Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol Cell Biol***20**, 5415-24.

36. Lamontagne, B., Hannoush, R. N., Damha, M. J. & Abou Elela, S. (2004). Molecular requirements for duplex recognition and cleavage by eukaryotic RNase III: discovery of an RNA-dependent DNA cleavage activity of yeast Rnt1p. *J Mol Biol***338**, 401-18.

37. Karaoglu, D. & Thurlow, D. L. (1991). A chemical interference study on the interaction of ribosomal protein L11 from Escherichia coli with RNA molecules containing its binding site from 23S rRNA. *Nucleic Acids Res***19**, 5293-300.

38. Gaudin, C., Ghazal, G., Yoshizawa, S., Abou Elela, S. & Fourmy, D. (2006). Structure of an AAGU tetraloop and it implication in Rnt1p RNA processing. *J Mol Biol.*

39. Gan, J., Tropea, J. E., Austin, B. P., Court, D. L., Waugh, D. S. & Ji, X. (2006). Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell***124**, 355-66.

40. Gan, J., Tropea, J. E., Austin, B. P., Court, D. L., Waugh, D. S. & Ji, X. (2005). Intermediate states of ribonuclease III in complex with double-stranded RNA. *Structure***13**, 1435-42.

41. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol***215**, 403-10.

42. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature***423**, 241-54.

43. Notredame, C., Higgins, D. G. & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol***302**, 205-17.

44. Conway, L. & Wickens, M. (1989). Modification interference analysis of reactions using RNA substrates. *Methods Enzymol***180**, 369-79.

45. Peattie, D. A. (1979). Direct chemical method for sequencing RNA. *Proc Natl Acad Sci U S A***76**, 1760-4.

ARTICLE 3

Ghazal, G., Gagnon, J., Jacques, PE., Landry, JR., Roberts, F. and Abou Elela, S.(2009) Yeast RNase III triggers polyadenylation independent transcription termination. Molecular Cell (accepted July 14, 2009).

Préambule

Ce manuscrit décrit la découverte d'un nouveau mode de terminaison de la transcription indépendant de la polyadenylation. Ce document montre comment la transcription des gènes exprimant un long ARN peut se terminer sans l'ajout de la queue Poly A. Dans cette étude, nous montrons que le clivage par l'endoribonucléase Rnt1p mène à la terminaison de la transcription selon le modèle "Torpedo". En outre, nous montrons que ce nouveau mode de terminaison de la transcription joue un rôle important dans la prévention de l'interférence entre les gènes à proximité et permet l'utilisation de la terminaison de la transcription en tant que mécanisme de régulation des gènes. En effet, nous fournissons une étude de cas démontrant comment la terminaison par Rnt1p rendle mécanisme d'autorégulationcapable de contrôler l'expression de protéines liant l'ARN. J'ai effectué tous les travaux expérimentaux dans ce document, sauf pour l'analyse bioinformatique et les données de la puce del'immunoprécipition de la chromatine (CHIP-Chip) présentées dans le tableau 2.

ARTICLE 3

Ghazal, G., Gagnon, J., Jacques, PE., Landry, JR., Roberts, F. and Abou Elela, S.(2009) Yeast RNase III triggers polyadenylation independent transcription termination. Molecular Cell (accepted July 14, 2009).

Summary

This manuscript describes the discovery of new polyadenylation- independent mode of transcription termination. This paper shows how the transcription of genes expressing long RNA transcripts may terminate without the addition of Poly A tail. In this study, we show that transcription may terminate by an independent endonuclease (i. e. Rnt1p) to trigger cleavage dependent or "Torpedo" termination. In addition, we demonstrate that this new mode of transcription termination plays an important role in preventing interference between genes in close proximity and allows the use of alternative transcription termination as a mechanism for gene regulation. Indeed, we provide a complete case study demonstrating how Rnt1p dependent termination makes the autoregulatory mechanism controlling the expression of the RNA binding protein. I have conducted all experimental work in this paper except for the bioinformatic analysis and the ChIP on chip data shown in Table 2.

Yeast RNase III triggers polyadenylation independent transcription termination

Ghada Ghazal[1], Jules Gagnon[1], Pierre-Étienne Jacques[2], Josette-Renée Landry[2], François Roberts[2,3] and Sherif Abou Elela[1*].

[1] RNA Group, *Département de microbiologie et d'infectiologie, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1H 5N4.

[2] Institut de recherches cliniques de Montréal, 110 Avenue des Pins Ouest, Montréal, Québec, Canada, H2W 1R7.

[3] Département de Médecine, Faculté de Médecine, Université de Montréal.

*Corresponding author

Sherif Abou Elela, Ph. D.

Département de microbiologie et d'infectiologie,

Faculté de médecine et des sciences de la santé,

Université de Sherbrooke,

3001 12e Ave Nord,

Summary

Transcription termination of messenger RNA (mRNA) is normally achieved by polyadenylation followed by Rat1p dependent 5'-3' exoribonuleolytic degradation of the downstream transcript. Here we show that the yeast orthologue of the dsRNA-specific ribonuclease III (Rnt1p) may trigger Rat1p dependent termination of RNA transcripts that fail to terminate near polyadenylation signals. Rnt1p cleavage sites were found downstream of several genes and the deletion of *RNT1* resulted in transcription read-through.

Inactivation of Rat1p impaired Rnt1p dependent termination and resulted in the accumulation of 3' end cleavage products. These results support a model for transcription termination in which co-transcriptional cleavage by Rnt1p provides access for exoribonucleases in the absence of polyadenylation signals.

Key words: budding yeast / transcription termination / mRNA expression / RNase III / Rnt1p

Introduction

Transcription termination plays an important role in determining the fate and function of the RNA. For example, formation of polyadenylated RNA could signal protein translation, while aberrant termination may trigger RNA degradation (Zhao et al., 1999). There are currently two models for transcription termination in eukaryotes; the first is called the "torpedo" model, which is the predominant mode of termination in protein coding genes (Luo et al., 2006; Tollervey, 2004), and the other is the "allosteric" model, which appears to be favoured in genes producingshort non-coding RNA (Kim et al., 2006b; Vasiljeva et al., 2008). In the case of "torpedo" termination the polyadenylation signal that is often found near the end of protein coding genes triggers an endonucleolytic RNA cleavage generating an entry site for the 5'-3' exoribonuclease Rat1p that in turn destabilizes theRNAP II elongation complex (Kim et al., 2004). On the other hand, the "allosteric" mode of termination does not require cleavage or the presence of a polyadenylation signal but depends on the binding of a termination complex in close proximity to the promoter (Carroll et al., 2004; Vasiljeva et al., 2008).

In yeast, mutations that change the 3' end sequence of mRNAs or inactivate the exoribonuclease Rat1p result in transcription read-through that often terminates before the promoter of the downstream genes (Kim et al., 2004;

Luo et al., 2006). Indeed, the intergenic regions in yeast are normally littered with non-canonical polyadenylation sites that become active upon the disruption of the primary site of transcription termination (Grec et al., 2000; Milligan et al., 2005). Therefore, simple defects in a canonical termination site will not automatically lead to the production of polycistronic RNA transcripts. Consistently, most of the confirmed RNAP II-transcribed polycistronic transcripts in yeast are processed by the yeast dsRNA specific RNase III (Rnt1p) and does not require polyadenylation signals for termination (Ghazal et al., 2005). Rnt1p cleaves specific RNA stems terminating with NGNN or AAGU tetraloops (Ghazal and Elela, 2006) found near pre-rRNA (Abou Elela et al., 1996; Kufel et al., 1999), snRNAs (Abou Elela and Ares, 1998), snoRNAs (Ghazal et al., 2005) or occasionally within mRNA coding sequence (Ge et al., 2005; Larose et al., 2007). Recently, Rnt1p was also shown to promote termination of RNAP I by giving access to Rat1p in a mechanism analogous to that of the polyadenylation-dependent "torpedo" mode of termination (El Hage et al., 2008; Kawauchi et al., 2008). However, the impact of Rnt1p on RNAP II transcription remains unexplored.

Since Rnt1p processes clusters ofRNAP II transcribed snoRNAs in yeast (Ghazal et al., 2005) we explored the possibility that the enzyme also influences the expression of neighbouring protein coding genes. Accordingly, we have searched for clusters of open reading frames (ORF) separated by a canonical Rnt1p cleavage signal and identified several conserved regions

with the potential to produce dicistronic transcripts. One of these transcripts was expressed *in vivo* in an Rnt1p dependent manner. In depth mutational analysis and chromatin immunoprecipitation of this transcription unit indicated, that Rnt1p is required for terminating mRNA transcripts that fail to terminate near polyadenylation signals. Genome-wide search for Rnt1p dependent transcription termination sites identified additional genes that require Rnt1p for alternative transcription termination. Together, the results presented here reveal a mechanism for gene regulation in which Rnt1p triggers mRNA degradation by inducing polyadenylation independent "torpedo" like transcription termination.

Results

Rnt1p represses the expression of dicistronic mRNA in yeast

Rnt1p cleaves RNA stems terminating with NGNN tetraloops to initiate the processing of polycistronic snoRNA (Ghazal et al., 2005). Accordingly, we reasoned that the presence of Rnt1p cleavage signal within a cluster of genes could be indicative of dicistronic mRNA expression. To examine this possibility, we searched a group of 5 *sensu strictoSaccharomyces* species (Herrero, 2005; Liti et al., 2006) for the presence of conserved NGNN stem-loops located between gene-pairs transcribed in the same direction (Figure 1A). Three conserved tetraloops were found but only one in a dicistronic transcript containing the sequence of the *NPL3* and *GPI17* genes was detected in the absence of Rnt1p (Figure 1B). Using gene specific probes, we monitored the expression of *NPL3* and *GPI17* in the presence (*RNT1)* or the absence *(rnt1*Δ) of *RNT1.* As shown in figure 1C, the probe specific to the *NPL3* coding sequence (Probe I) detected the mature Npl3 mRNA (Russell and Tollervey, 1995; Russell and Tollervey, 1992) in RNA extracted from wild type cells (Lane 1). In contrast, *rnt1*Δ RNA (Lane 3) exhibited two additional large RNA species. Hybridizing *RNT1* RNA with *GPI17* specific probes (Probe IV) highlighted a band (Lane 13) corresponding to the predicted size of the

mature *GPI17* mRNA (Zhu et al., 2005). Surprisingly, the expression of the mature *GPI17* mRNA was reduced in *rnt1*Δ RNA (Lane 15) and a large transcript migrating with the same speed as that observed with the *NPL3* specific probe (Lane 15) was detected. This indicates that the deletion of *RNT1* inhibits the expression of *GPI17* and leads to the accumulation of a large transcript containing both *Npl3* and *Gpi17* sequences. This was further confirmed by a probe (Probe III) hybridizing to the intergenic region (Lane 9). Consistently, western blot using antibodies against Npl3p and Gpi17p (Figure 1D) revealed that while *RNT1* deletion does not affect Npl3 mRNA translation it inhibits the production of Gpi17p. Since *GPI17* is essential (Zhu et al., 2005), we presume that a small amount of proteins, below the detection level of the western blot, is expressed in *rnt1*Δ cells. In any case, the results clearly show that Rnt1p is required for the normal expression of Gpi17p.

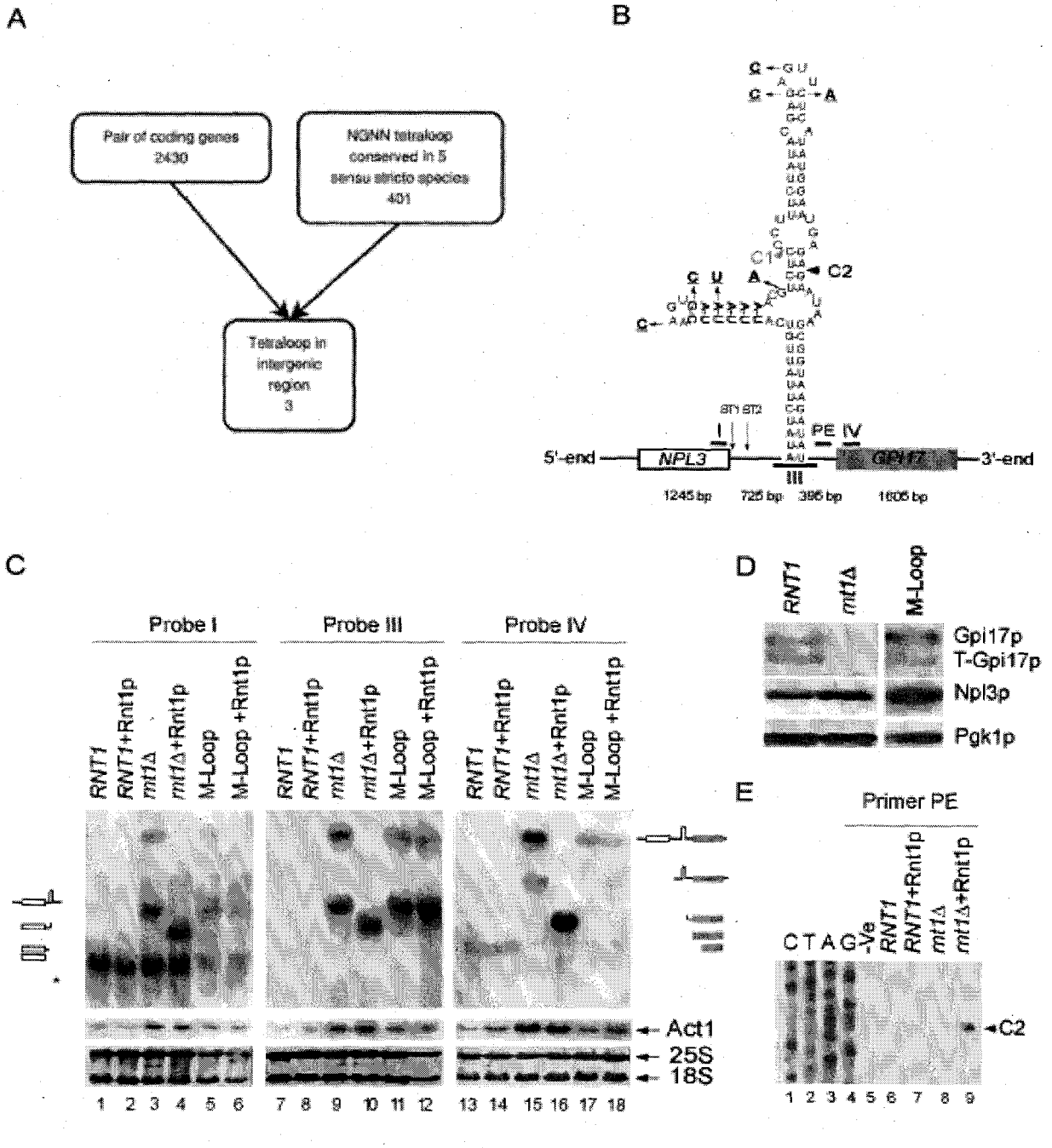


Figure 1
Ghazal et al., 2009

Figure 1. Identification of Rnt1p dependent dicistronic RNA. (A) *In silico* search of conserved gene clusters separated by NGNN stem-loop structures using *Saccharomycescerevisiae* Genome Database (SGD) annotations. (B) Illustration of *NPL3-GPI17* gene cluster showing the size of each gene fragment at bottom. Positions of Northern blot probes (I and IV) and reverse transcription primers are shown on top. C1 and C2 indicate the position of the predicted cleavage sites. Black arrowhead indicates cleavage observed *in vitro* using total RNA. Grey arrowhead indicates cleavage observed with a 5' end labeled model substrate *in vitro* (data not shown). Point mutations disrupting Rnt1p cleavage site are shown in bold. ST1 and ST2 indicate the position of previously reported polyadenylation sites near the *NPL3* 3' end (Chinnusamy et al., 2008; Steinmetz et al., 2006). (C) Rnt1p is required for the cleavage of the extended Npl3-Gpi17 RNA *in vivo* and *in vitro*. RNA was extracted from wild type (*RNT1*), *rnt1*Δ, and from cells carrying mutations in Rnt1p cleavage site (M-Loop) and incubated either alone or in the presence of recombinant Rnt1p. Schematics of the different RNA transcripts are indicated beside each gel. Open and grey boxes represent *NPL3* and *GPI7* ORFs respectively. (D) Western blot analysis of the Npl3p and Gpi17p. Proteins were extracted from *RNT1*, *rnt1*Δ and M-Loop cells separated on 12% SDS gel and visualized using antibodies specific to Gpi17p, Npl3p or the control Pgk1p. Note that Gpi17p exists in two forms, a full-length membrane-bound version (Gpi17p) and a truncated free form (T-Gpi17p) (Zhu et al., 2005). (E) Rnt1p cleaves Npl3-Gpi17 extended RNA *in vitro.* Reverse

transcription using a primer downstream of the predicted cleavage site was performed using RNA extracted from wild type or *rnt1*Δ cells incubated with or without recombinant Rnt1p. Sequencing of DNA corresponding to the same region is indicated on the left as a marker. The position of the cleavage (C2) is indicated on the right.

Direct cleavage of the extended *Npl3-Gpi17* mRNA by Rnt1p was confirmed *in vitro* using recombinant Rnt1p. Total RNA extracted from *RNT1* or *rnt1*Δ cells was incubated with recombinant Rnt1p and the impact on Npl3 and Gpi17 mRNA was monitored by Northern blotting (Figure 1C). As expected, Rnt1p did not affect the mature *Npl3* (Lanes 2) or *Gpi17* (Lanes 14) mRNAs. On the other hand, Rnt1p converted the large extended RNA transcripts observed in *rnt1*Δ cells to smaller fragments (Lanes 4, 10, and 16) corresponding to the predicted cleavage products (Figure 1B). The exact location of Rnt1p cleavage was determined by reverse transcription using a primer (PE) complementary to the sequence downstream of the loop predicted to bind Rnt1p (Figure 1B). Again, no cleavage product was detected in wild type RNA (Figure 1E lanes 6 and 7), while a band corresponding to a cleavage 16 nucleotides downstream of the conserved NGNN tetraloop (C2) was observed in *rnt1*Δ RNA in the presence (Lane 9) and not in the absence (Lane 8) of recombinant Rnt1p. To directly determine the impact of Rnt1p cleavage on the expression of *NPL3-GPI17,* we mutated Rnt1p cleavage signal and monitored the effect on mRNA synthesis. Six point-mutations were introduced in the two stem-loop structures (Ghazal and Elela, 2006) predicted to be cleaved by Rnt1p (Figure 1B) and the impact was monitored by Northern blot. A large *Npl3- Gpi17* transcript similar to that detected in *rnt1*Δ RNA (Lanes, 3, 9 and 15) was observed in cells harbouring the stem-loop mutations (M-Loop) (Lanes 5, 11, and 17). The extended RNA produced from the gene carrying mutations in the loops was not cleaved by recombinant

Rnt1p *in vitro* confirming that the cleavage site was indeed disrupted (Lanes 6, 12, and 18). However, the disruption of the cleavage signal reduced the expression of *GPI17* to a lesser extend than *RNT1* deletion (Figure 1C). Indeed, western blot analysis indicated that the stem loop mutation does not inhibit the expression of Gpi17p (Figure 1D). Therefore, while direct cleavage by Rnt1p is required for inhibiting the accumulation of the Npl3-Gpi17 transcript the presence of Rnt1p itself may play additional role in modulating the expression of *GPI17*.

NPL3 termination-sequence induces the accumulation of Rnt1p dependent read-through transcripts

To determine the elements regulating the expression of GPI17, we deleted the predicted promoter regions of either the *NPL3* or *GPI17* gene (Figure 2A) and monitored the impact on RNA expression. The deletions were achieved by inserting a *URA3* gene fused to *ADH1* termination sequence (Akada et al., 2006; Noble and Johnson, 2005) upstream of the translation start codon of the chromosomal copies of either *NPL3* or *GPI17*. Deletion of the *NPL3* promoter (*npl3prΔ*) blocked the expression of Npl3 mRNA (Figure 3B) in both the presence (Lane 4) and the absence of *RNT1* (Lane 5). The deletion of the *NPL3* promoter increased the expression of Gpi17 mRNA in *rnt1Δ* cells (Lane 14) but not in *RNT1* cells (Lane 13). Notably, deletion of the *NPL3* promoter

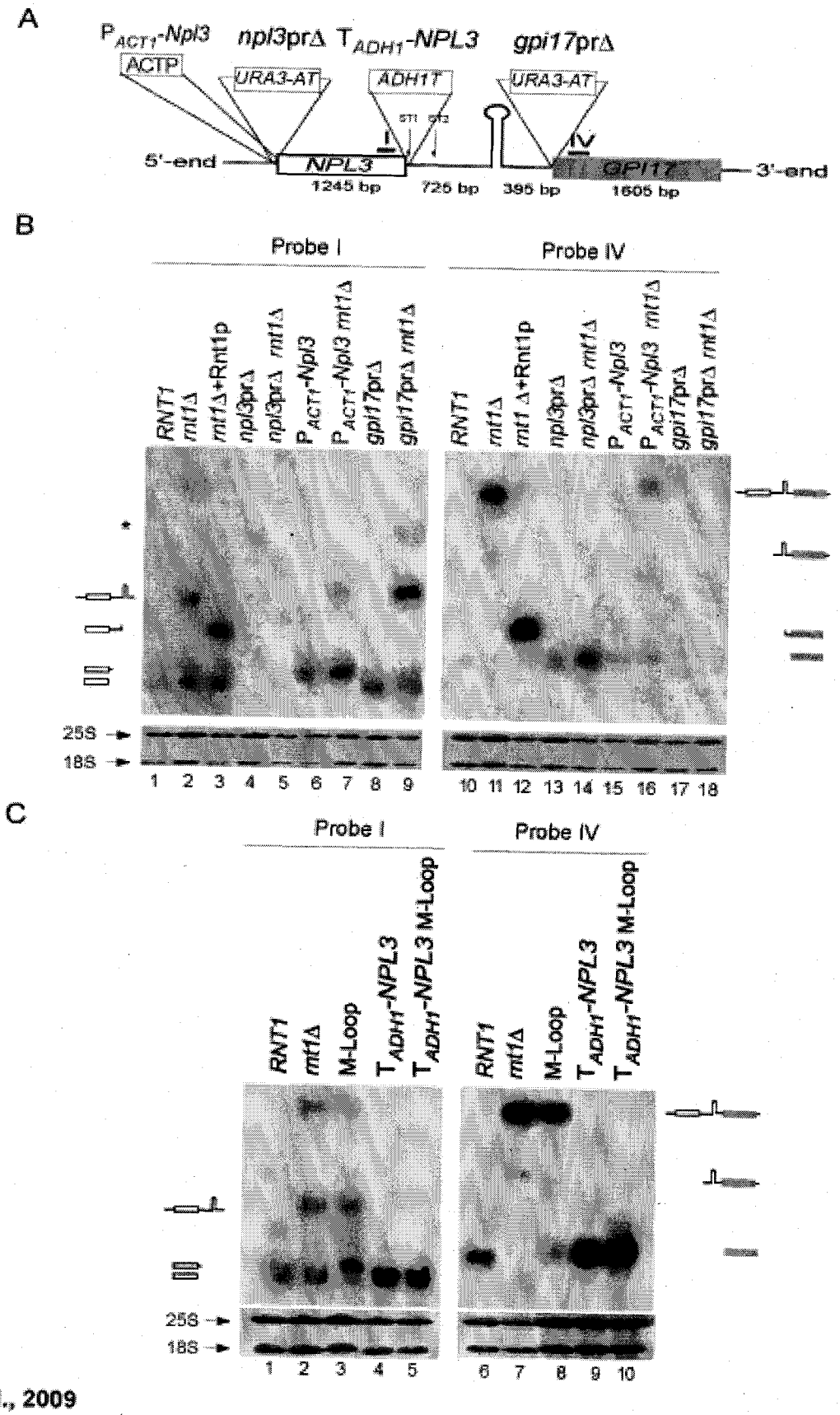


Figure 2
Ghazal et al., 2009

Figure 2. Identification of *cis*-acting elements controlling the expression of the Npl3-Gpi17 RNA. (A) Schematics representation of the different mutations introduced in the promoter and termination regions of *NPL3* and *GPI17.* The promoter of *NPL3* was either replaced by *ACT1* promoter (ACT1P) or deleted by inserting a *URA3* gene linked to a strong *ADH1* terminator (*URA3*-AT). The 3' end sequence containing the two reported *NPL3* polyadenylation signals (Steinmetz et al., 2006) was replaced by a strong *ADH1* terminator (ADH1T) (Blancafort et al., 1997). The promoter region of *GPI17* was replaced by *URA3*-AT. All replacements and deletions were carried in the chromosomal copies of the genes and the names of the resulting yeast strains are shown on top. (B) *NPL3* promoter is not required for the production of mature Gpi17 mRNA. Northern blot analysis was performed using RNA extracted from cells carrying the different mutations and visualized by probes specific to either *NPL3* (I) or Gpi17 (IV) mRNA. The stained rRNA is shown as a loading control. (C) The *NPL3* 3' end is required for transcriptional read-through. Northern blot analysis was performed using RNA extracted from the different mutations as described in B.

abolished the expression of the read-through Npl3-Gpi17 RNA even in the absence of *RNT1*. This indicates that the *NPL3* promoter is required for the expression of the Npl3-Gpi17 RNA but not for the synthesis of normal Gpi17 mRNA. In addition, these results indicate that Rnt1p is not required for transcription initiated from the *GPI17* promoter. Changes in the *GPI17* expression upon the deletion of Rnt1p and *NPL3* promoter may stem from the general effect of *RNT1* deletion on stress and membrane related proteins (Ge et al., 2005; Lee et al., 2005; Tremblay, 2002). Deletion of the *GPI17* promoter (*gpi17prΔ*) abolished the expression of Gpi17 mRNA (Lane 17) without affecting the expression of Npl3 mRNA in wild type cells (Lane 8). Replacement of *GPI17* promoter by *URA3* gene reduced the length of the Npl3 read-through transcripts accumulating in *rnt1Δ* cells (Lane 9). These new extended transcripts did not hybridize to probes specific to *GPI17* (Lane 18) indicating that *NPL3* transcription terminated upstream of the *URA3* gene (Figure 2A). We conclude that transcription read-through of *NPL3* is dependent at least in part on the sequence near the gene 3' end.

The transcriptional read-through of *NPL3* may be influenced by the promoter or termination sequence. To differentiate between these two possibilities, we first replaced the *NPL3* promoter with that of *ACT1* and monitored the impact on RNA expression (Figure 2A and B). As expected, transcripts produced from the *ACT1* promoter (P_{ACT1}-*NPL3*) were slightly larger than those driven from the endogenous *NPL3* promoter due to changes in the transcription start

site, irrespective of *RNT1* expression (Lanes 6 and 7). On the other hand, P_{ACT1}-*NPL3* did not significantly change the expression level of either mature Npl3 (Lane 6) or Gpi17 mRNA (Lane 15) when expressed in wild type cells. Changing *NPL3* promoter in *rnt1*Δ cells reduced transcriptional read-through (Lane 7 and 16) and permitted the expression of mature Gpi17 mRNA (Lane 16). This suggests that either *ACT1* promoter enhances termination near the canonical *NPL3* termination site or that the changes in the site of transcription initiation influence the termination efficiency.

To evaluate the impact of *NPL3* termination on the accumulation of the extended Npl3-Gpi17 transcripts, we replaced the sequence between the translation stop codon of *NPL3* and the polyadenylation signal with the *ADH1* termination signal (Figure 2A). The mutations were introduced either in *RNT1* cells or in cells carrying mutations in Rnt1p cleavage signals. As shown in figure 2C, the introduction of the *ADH1* terminator abolished the expression of read-through transcripts regardless of Rnt1p cleavage (Lanes 4, 5, 9 and 10). This demonstrates that the production of extended Npl3-Gpi17 RNA is largely due to the leaky termination of the *NPL3* gene. Therefore, Rnt1p seems to function as a fail-safe terminator of *NPL3*.

Rat1p is required for transcription termination downstream of Rnt1p cleavage site

To determine the fate of the RNA cleaved by Rnt1p, we examined the impact of known exoribonucleases on the expression of *NPL3* and *GPI17* transcripts. As shown in figure 3, deletion of the nuclear 3'-5' exoribonuclease Rrp6p induced the expression of 3' extended Npl3 RNA that migrates with the same speed as Rnt1p cleavage products (lanes 4, 13, 22, and 31). This suggests that Rnt1p 5' end cleavage products are degraded by Rrp6p under normal growth conditions. Surprisingly, inactivation of a temperature-sensitive allele of the 5'-3' nuclear exonuclease Rat1p (*rat1-1*) reduced the expression of both mature Npl3 and Gpi17 mRNAs (Lanes 6, 15, 24 and 33). The general reduction in mRNA could be explained by poor termination and subsequent degradation of the aberrant RNA since Rat1p was previously shown to be required for the termination of polyadenylated mRNA (Kim et al., 2004). Indeed, Rat1p may affect the 3' end formation by influencing the recruitment of 3' end processing factor and the choice of polyadenylation site (Luo et al., 2006). At permissive temperature the deletion of the cytoplasmic 5'-3' exoribonuclease *XRN1* in *rat1-1* cells resulted in little perturbation of the *Npl3 mRNA,* while causing an accumulation of a 5' end-extended Gpi17 RNA species consistent with the 3' end cleavage product of Rnt1p (Lanes 7, 16, 25, and 34). At restrictive temperature, the *rat1-1xrn1Δ* and *rat1-1* cells exhibited the same profile of *NPL3* expression (Lanes 6, 8, 15, and 17).

However, *rat1-1xrn1*Δ RNA exhibited a new transcript corresponding to the size of the stem-loop structure cleaved by Rnt1p (Lane 26). In addition, other bands corresponding to Rnt1p 3' end cleavage products containing the *GPI17* sequence were detected (Lane 35). Deletion of the nonsense mediated decay exoribonuclease Upf1p had little effect on RNA expression (Lanes 9, 18, 27, 36). This result indicates that Rnt1p cleavage leads to the degradation of the 3' end cleavage product by Rat1p in the nucleus or, surviving this, by Xrn1p in the cytoplasm. We conclude that Rnt1p cleavage generates an entry site for the 5'-3' exoribonuclease Rat1p, which causes "torpedo" like transcription termination.

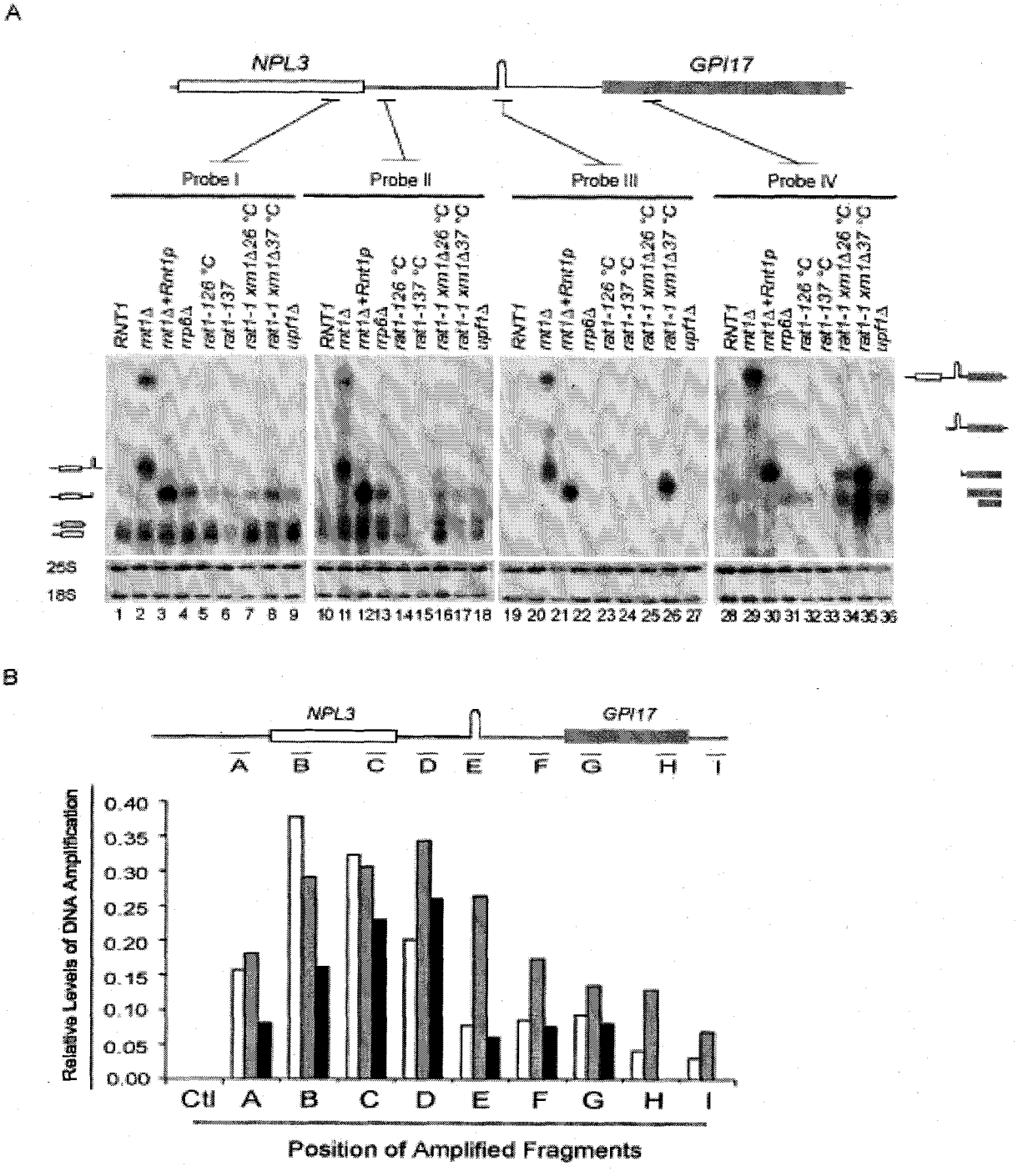


Figure 3
Ghazal et al., 2009

Figure 3. Rnt1p cleavage triggers transcription termination. (A) Degradation of Rnt1p cleavage product by 5'-3' exoribonuclease is required for the expression of *GPI17*. RNA was extracted from cells lacking the 3'-5' nuclear exoribonuclease Rrp6p (*rrp6*Δ), cells expressing a temperature sensitive allele of the 5'-3' nuclear exoribonuclease *RAT1* grown at permissive (*rat1-1* 26°C)or restrictive conditions (*rat1-1* 37°C), rat1-1 cells lacking 5'-3' cytoplasmic exoribonuclease *XRN1* grown at the permissive (*rat1-1 xrn1*Δ 26°C) or restrictive (*rat1-1 xrn1*Δ 37°C) temperature, and cells lacking the non-sense mediated decay ribonuclease Upf1p (*upf1*Δ). The different RNAs were visualized by probes complementary to different regions of the *NPL3-GPI17* cluster. The rRNA is included as loading control. (B) Rnt1p enhances the transcription termination of *NPL3*. Chromatin Immunoprecipitations were performed using antibodies against the RNAP II protein subunit Rpb1p in the presence (white column) or the absence (grey column) of *RNT1*. The association of Rnt1p with transcribed genes was examined using antibodies against Rnt1p (black column). The precipitated DNA was amplified by real-time PCR using primers specific to different regions within the *NPL3-GPI17* clusters (indicated on the top). A total of two biological and three technical replicates were used to calculate the relative levels of DNA precipitated and the average values are indicated. A primer-pair amplifying a known untranscribed region of chromosome V was used as negative control (Ctl). Standard deviations between replicate experiments was ± 0.05.

To link Rnt1p directly to transcription termination, the pattern of RNAP II association with the *NPL3* and *GPI17* genes was examined by chromatin immunoprecipitation (ChIP) using antibodies against the *RPB1* subunit (Malagon et al., 2006) in the presence or the absence of Rnt1p. As expected, in *RNT1* cells (Figure 3B) the RNAP II co-immunoprecipitated with DNA fragments corresponding to the promoter (A) and coding sequence of *NPL3* (B and C) but not with known untranscribed regions of chromosome V (Ctl). DNA corresponding to the intergenic region between *NPL3* and *GPI17* (D, E and F, white columns) co-immunoprecipitated as or more efficiently than the DNA corresponding to the coding region or the sequence downstream of *GPI17* (G, H and I, white columns). Strikingly, the deletion of *RNT1* significantly increased the association of RNAP II with the intergenic region between *NPL3* and *GPI17* (D, E, and F, grey columns) and the *GPI17* ORF (G, H and I, grey columns). This suggests that Rnt1p is required for the efficient termination of *NPL3*.

In order to understand how Rnt1p influences transcription termination, we immunoprecipitated Rnt1p and monitored its association with the actively transcribed *NPL3* and *GPI17* genes (Figure 3B, black columns). Interestingly, Rnt1p co-precipitated with fragments corresponding to *NPL3* promoter region (A), *NPL3* coding sequence (B and C), and the *NPL3* transcription termination site (D). Weak associations with the intergenic region (E and F) and the 5' end (G) of *GPI17* were also detected. Rnt1p did not co-precipitate with *GPI17*

3' end fragments (H and I). The strongest association was found between Rnt1p and fragment immediately adjacent to *NPL3* 3' end. Inactivation of a temperature sensitive allele of RNAP II also inhibited the association of Rnt1p with all DNA fragments (data not shown). These data indicate that Rnt1p associates with actively transcribed DNA and is required for transcription termination downstream of *NPL3* 3' end.

RNT1 deletion perturbs transcription termination of several RNAP II transcribed genes

To examine the possibility that Rnt1p mediated the transcription termination of genes other than *NPL3* we initiated a new search *in silico* looking for all conserved Rnt1p cleavage signals downstream of known polyadenylation sites. As indicated in Supplemental Table 1, five stem-loop structures other than that near *NPL3* 3' end were found with score above 0.8, which was previously established as reasonable cut-off (Ghazal et al., 2005). Three of the newly identified genes were either not expressed under vegetative growth or the identified Rnt1p stem-loop overlapped with the sequence of downstream tRNA (data not shown). Interestingly two additional genes known to code for RNA binding proteins (*NAB2* and *RPL8A*) were identified. Northern blot analysis of the mRNA produced by these two genes indicated that *RNT1* deletion causes transcriptional read-through downstream of the

canonical termination site (Supplemental Table 1 and Figure 4A and B). In addition, deletion of *RNT1* increased the amount of the RNA generated from the *NAB2* (Figure 4A lanes 1 and 2) and *RPL8A* genes (Figure 4B lane 1 and 2). This increase in expression could be due either to an increase in transcription rate as evident in the case of *RPL8*A (Figure 4D), increased RNA stability in the absence of *RNT1* or combination of both factors. Deletion of the exoribonuclease *XRN1* and the inactivation of the *RAT1* resulted in the accumulation of an extended Rat1 RNA and fragments corresponding to the sequence downstream of Rnt1p cleavage site located near the 3' end of Rpl8A and Nab2 (Figure 4A and B lane 9). Chromatin immunoprecipitation indicated that Rnt1p interact with the transcriptional units of *NAB2* and *RPL8A* and confirmed transcription read-through of these genes upon the deletion of *RNT1* (Figure 4C and D). These data further confirm the role of Rnt1p in suppressing the accumulation of transcription read-through products by generating an entry site for the exoribonculeases Rat1p and Xrn1p.

In order to examine the global impact of Rnt1p on transcription termination we analyzed the overall pattern of RNAP II occupancy in the presence and the absence of *RNT1*. RNAP II specific chromatin immunoprecipitation was performed as described above from *RNT1* or *rnt1*Δ cells and the extracted DNA hybridized to a DNA microarray containing an average of 4 probes per

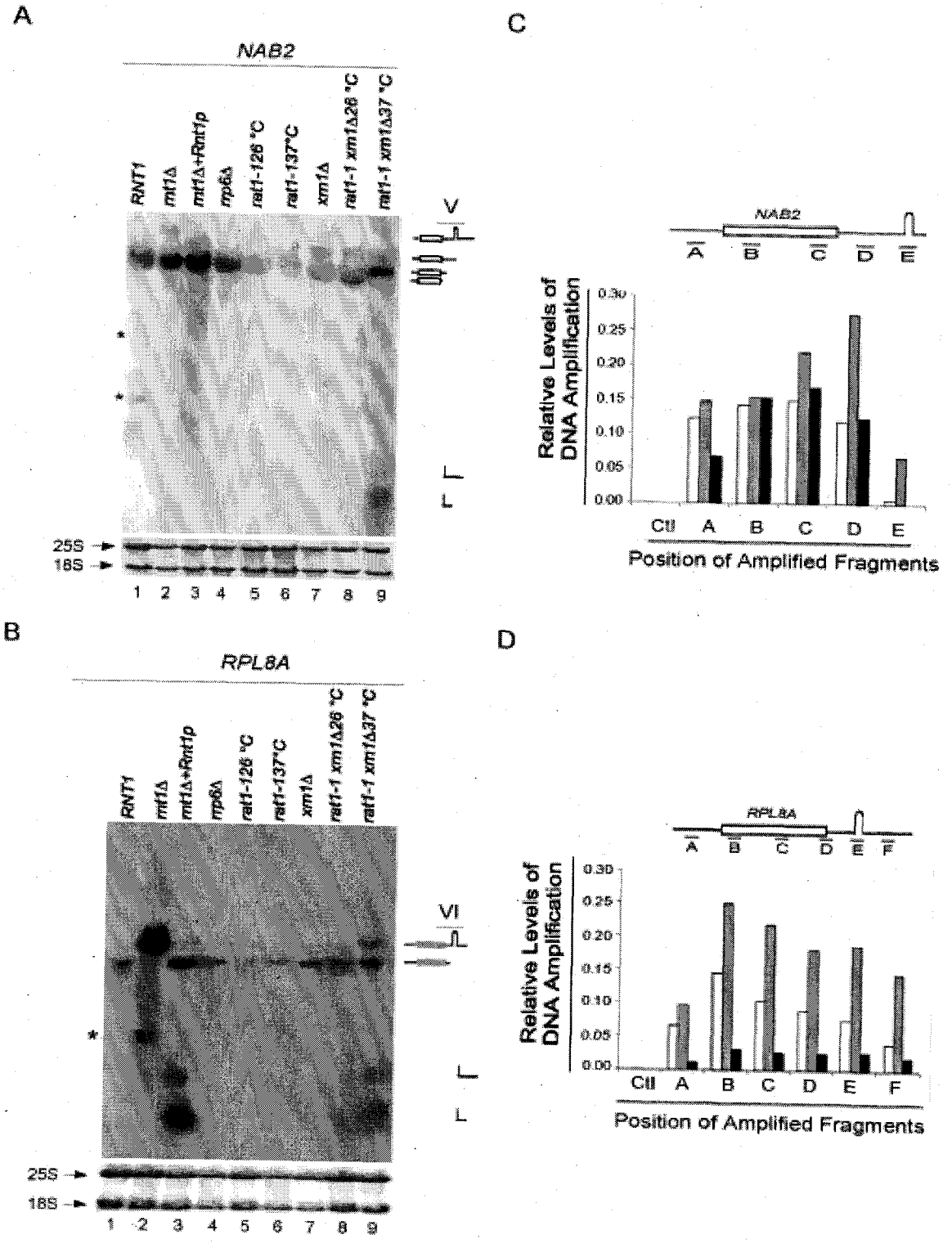


Figure 4
Ghazal et al., 2009

Figure 4. Rnt1p cleavage signal identifies sites of alternative transcription termination in genes coding for RNA binding proteins. Extended Nab2 (A) and Rpl8A (B) RNA are cleaved by Rnt1p *in vivo* and *in vitro*. RNA was extracted from *RNT1*, *rnt1*Δ, *rrp6*Δ, *xrn1*Δ cells or cells expressing a temperature sensitive allele of *RAT1* grown at permissive (*rat1-1* 26°C)or restrictive conditions (*rat1-1* 37°C) or rat1-1 cells lacking 5'-3' cytoplasmic exoribonuclease *XRN1* grown at the permissive (*rat1-1 xrn1*Δ 26°C) or restrictive (*rat1-1 xrn1*Δ 37°C) temperature. *In vitro* cleavage assay of RNA extracted from *rnt1*Δ cells was carried by incubation with recombinant Rnt1p (*rnt1*Δ + Rnt1p). RNAs were visualized by probes complementary to either a sequence near *NAB2* 3' end (V) or unique sequence that is found in the 3'UTR of *RPL8A* and not the *RPL8B* isoform (VI). The rRNA is included as loading control. The asterisk indicates 3' end RNA degradation products that were observed occasionally. The 3' end cleavage project of NAB2 generated by Rnt1p *in vitro* was too faint to detect in the exposure shown. *NAB2* (C) and *RPL8A* (D) Chromatin immunoprecipitations were performed using antibodies against the RNAP II protein subunit Rpb1p in the presence (white column) or absence (grey column) of *RNT1* or using antibodies against Rnt1p (Black column) in wild type cells as described in Figure 3B. Immunoprecipitations of *RPL8A* chromatin was performed either in wild type strain (data not shown) or strains lacking the *RPL8B* to avoid cross-amplification of homologous sequence (D). The data were obtained and calculated as mentioned in Figure 3B.

kilobase across the whole yeast genome. Extended association of RNAP II at the 3' end of 39 genes that could not be attributed to overlapping or neighbouring genes were identified (Table 1). The length of these extensions varied from 230 to 2440 nts with an average extension length of 815 nts. As expected, the transcription read-through near the *NPL3* and *RPL8A* genes was identified. However, the *NAB2* gene was not detected because it is located next to a very highly transcribed gene that generates an RNAP II signal overlapping with the intergenic region located downstream from the *NAB2* gene. The accuracy of the systematic analysis of RNAP II read-through was tested by quantitative PCR on 10 genes identified by ChIP-chip. In all cases, the increased association of RNAP II was confirmed by quantitative PCR (data not shown). Interestingly, the ChIP-chip approach identified a transcriptional read-through in two non-coding RNA (U2 snRNA (Abou Elela and Ares, 1998) and snR190 (Chanfreau et al., 1998b)) that were shown to be processed by Rnt1p. In both cases the deletion of *RNT1* leads to the transcription of extended RNA species (Abou Elela and Ares, 1998; Chanfreau et al., 1998b). To link the RNAP II profile to RNA expression we examined the RNA transcripts produced by 29 genes displaying transcriptional read-through by ChIP-chip. Northern blot analysis indicated that 24 of those 29 genes indeed produce a larger RNA species in *rnt1*Δ cells (Table 1). Interestingly, most of these genes (18/24) were also overexpressed in the mutant cells. In about a third of the cases (8/24) some transcriptional read-through could be detected in wild type cells, suggesting that some RNAP

II molecules can escape from Rnt1p surveillance. Only four of the genes we have tested were downregulated in *rnt1*Δ cells, while three showed no difference in expression. Most of these genes (6/7), however, produced extended transcript in *rnt1*Δ cells, suggesting that the termination function of Rnt1p is not linked to its role in the regulation of transcription level. As shown in figure 5, three of Rnt1p dependent transcriptional read-through resulted in the accumulation of discitronic transcripts that include the sequence of two neighbouring genes while, in other cases, the extension terminated in the intergenic region. Together these data suggest that Rnt1p impact on transcription is not limited to NPL3-GPI17 cluster and may extend to genes with different functions.

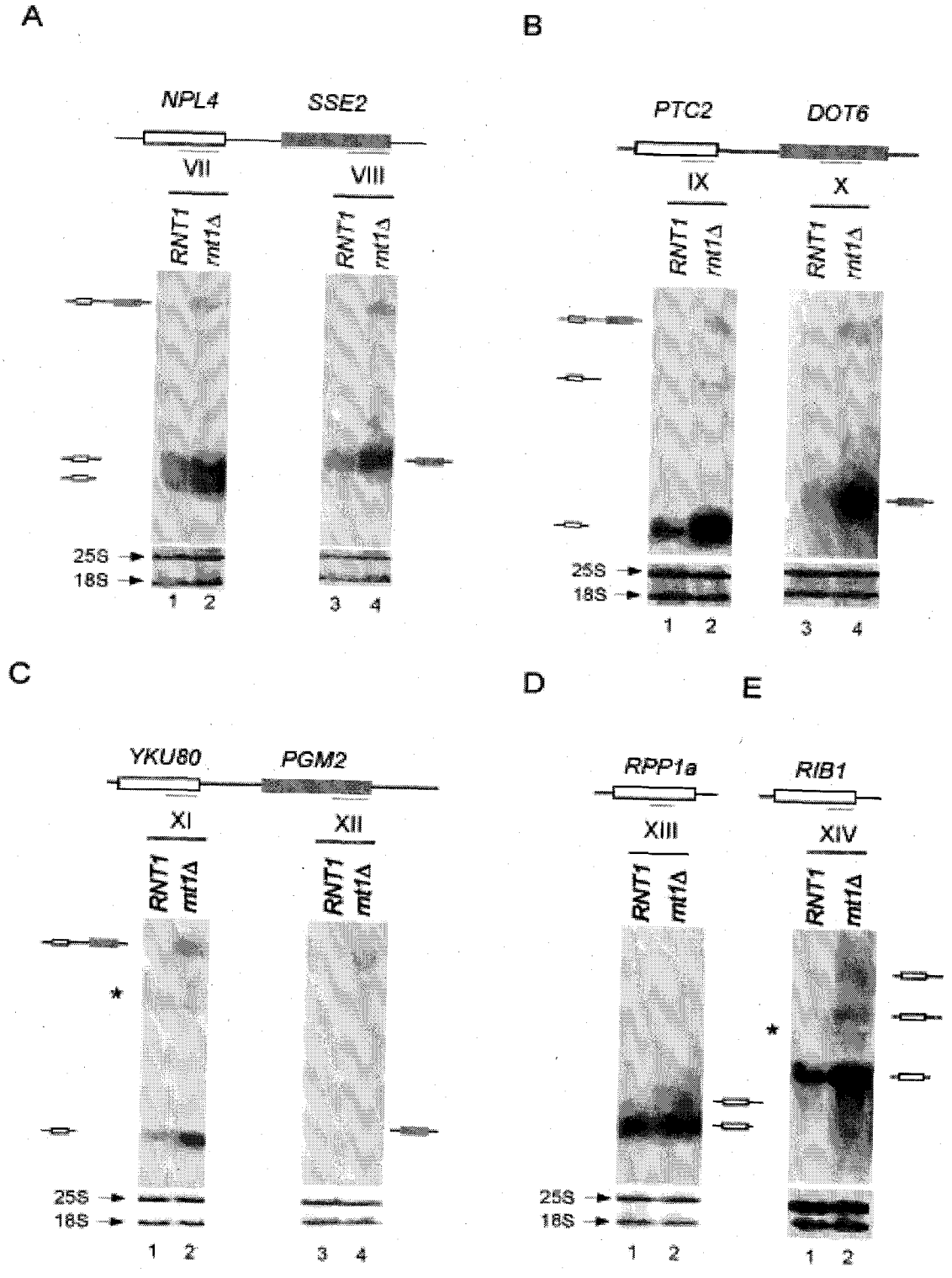


Figure 5
Ghazal et al., 2009

Figure 5. RNAP II chromatin immunoprecipitation in the absence of *RNT1* identifies genes with alternative transcription termination.Northern blot analysis was carried out as described in Figure 1. The probe position relative to the gene structure is shown on top. The nature of the different transcripts is schematically represented on the side. The asterisk indicate cross hybridization with rRNA. In all cases extensions were not detected in cells expressing *RNT1* even after prolonged exposure.

Discussion

In this study, we have shown that Rnt1p cleavage signal may function as polyadenylation independent fail-safe terminator. Deletion of the dsRNA specific ribonuclease Rnt1p induced the expression of a long read-through transcript containing the sequence of the *NPL3* and *GPI17* genes (Figure 1). *In vitro*, recombinant Rnt1p cleaved the extended RNA species at the predicted NGNN stem-loop structure in the absence of any other factors (Figure 1). *In vivo* Rnt1p was found associated with actively transcribed *NPL3* and its deletion resulted in transcriptional read-through interfering with the transcription of the downstream gene coding for Gpi17p (Figure 3). Rnt1p dependent transcriptional read-through was also detected in several genes with a variety of functions indicating that the impact of Rnt1p on transcription is not limited to a single gene (Table 1). Together the results reveal a mode of gene regulation were polyadenylation independent transcription termination triggers degradation of nascent RNA transcripts.

Npl3p is an RNA-binding protein implicated in the export of mRNA and it functions as an antagonist of transcription termination (Burkard and Butler, 2000; Krebber et al., 1999; Lund et al., 2008). In addition, it was recently proposed that phosphorylated Npl3p inhibits efficient recognition of the canonical polyadenylation signal of its own transcript (Lund et al., 2008).

However, the mechanism by which Npl3p influences mRNA processing remained unclear. Npl3p binds its actively transcribed gene and overexpressing Npl3p causes transcription read-through that is normally inhibited by Rnt1p (data not shown). The endonucleolytic cleavage prevents read-through and preserves the transcriptional activity of the downstream genes. Indeed, the biological advantage conferred by this mechanism is evident from the conservation of the Rnt1p cleavage signal in five closely related *Saccharomyces* species (Figure 1A). This phenomenon is not an isolated event since we have found other RNA binding proteins (Table 1 and Figure 4) that could benefit from Rnt1p triggered termination. For examples, Nab2p (Roth et al., 2005) and Rpl8ap (Cusick, 1994) are known to bind RNA and Nab2 was shown to be autoregulated via the induction of transcription read-through (Roth et al., 2005).

The two current models of RNAP II transcription termination (torpedo and allosteric) do not explain how the transcription of long non-polyadenylated RNA terminates. The "torpedo" model requires cleavage near the polyadenylation signals and the "allosteric" model functions only with short non-coding RNA (Kim et al., 2006b; Luo et al., 2006). We propose a modified "torpedo" mode of termination (Figure 6) in which Rnt1p circumvent the need for polyadenylation signals by generating an entry site for the Rat1p exonuclease leading to termination of transcription. In this model, polyadenylation is not required and therefore long RNA can be produced

without being obligatorily transported and translated. Recruitment of Rnt1p to the termination site is likely to be signalled by phosphorylation dependent interaction of the RNA II CTD. Interestingly, Rnt1p was shown to interact with the RNAP II CTD in a two hybrid system when the phosphorylation site required for either polyadenylation dependent termination (Serine 2) or that required for Nrd1p complex dependent non-polyadenylated RNA termination (Serine 5) is mutated (Ghazal and Abou Elela unpublished data). This suggests that indeed Rnt1p represent transcription termination alternative in situations where neither conventional "torpedo" nor "allosteric" modes of termination are possible (e. g. long non-polyadenylated RNA). Indeed, it was recently shown mechanistically in a model system that Rnt1p elicit RNAP II termination by a torpedo mechanism (see Rondon et al., this issue). Rnt1p cleaves the 3' end of the non-polyadenylated U2 snRNA and in the absence of Rnt1p a longer polyadenylated transcript is produced (Abou Elela and Ares, 1998). Indeed, deletion of Rnt1p leads to transcriptional read-through in the U2 gene (Table 1). This mode of transcription termination is not unique to RNAP II. Rnt1p cleavage at the 25S pre-rRNA gives access to Rat1p allowing it to terminate transcription by RNAP I in a "torpedo" like fashion (El Hage et al., 2008).

The impact of Rnt1p on transcription is not limited to transcription termination. In many cases we have observed an overall increase in RNAP II occupancy associated with an increase in gene expression in the absence of *RNT1*

(Table 1). The effect of Rnt1p on the level of transcription and termination are not necessarily linked. In the case of *NPL3*, the disruption of Rnt1p cleavage site lead to transcriptional read-through but the level of expression is lower than that observed upon the deletion of *RNT1* (Figure 1C). On the other hand, increased gene expression in *rnt1*Δ do not necessarily lead to transcription read-through (Larose et al., 2007). Indeed, genome-wide analysis of gene expression in the absence of Rnt1p identified many RNA transcripts that are over-expressed upon the deletion of *RNT1* and the vast majority did not exhibit changes in the site of transcription termination (Ge et al., 2005). Therefore, in certain cases the recruitment of Rnt1p to the active transcription complex may directly modulate transcription independent of the cleavage at the 3' end of the nascent RNA.

Discovering that Rnt1p cleavage induces Rat1p dependent transcription termination mandates re-examination of Rnt1p function in RNA processing. It is currently accepted that Rnt1p processes most non-coding RNA in yeast including pre-rRNA, snRNA and snoRNA (Abou Elela and Ares, 1998; Abou Elela et al., 1996; Ghazal et al., 2005). However, it is not clear why this processing step is necessary and why in certain cases the lack of this processing leads to the generation of polyadenylated RNA (Abou Elela and Ares, 1998). The results presented here suggest that in many cases Rnt1p cleavage is not introduced as an obligatory processing step but rather as a transcription terminator required in order to avoid the polyadenylation of

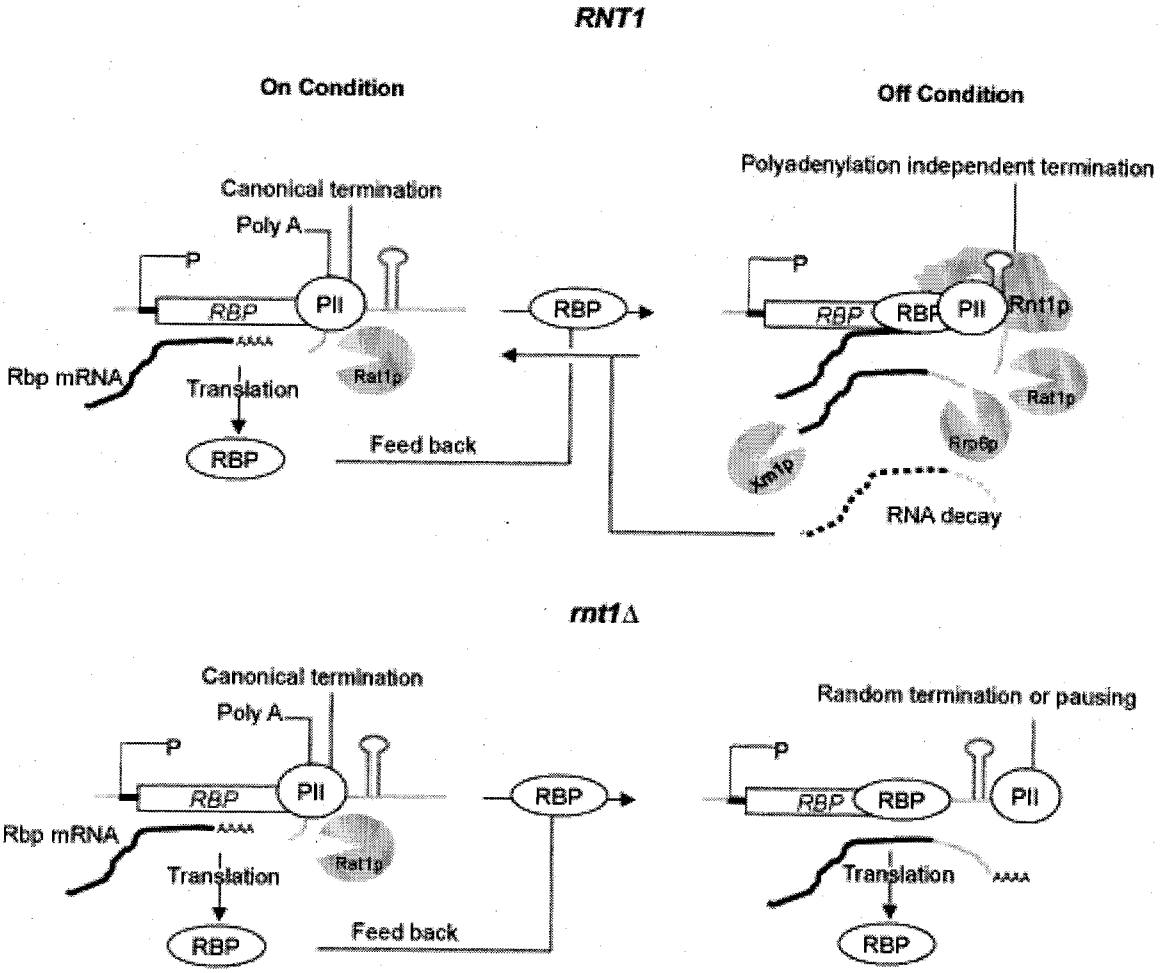


Figure 6
Ghazal et al., 2009

Figure 6. Model describing the impact of Rnt1p on transcription termination and mRNA stability. Under normal conditions,(*RNT1*) transcription of RNA binding proteins (RBP) genes like *NPL3*, *NAB2*, or *RPL8A* is autoregulated. When the amount of RBPs is low (On condition), transcription terminates at the canonical site via Rat1p dependent "torpedo" mechanism leading to the production of mature RNA and protein synthesis. When the RBPs accumulate in the cell (Off condition), they bind near the termination site of their gene leading to transcription read-through up to Rnt1p cleavage signal downstream. Cotranscriptional cleavage by Rnt1p gives access to Rat1p leading to "torpedo" like termination. However, in this case the resulting RNA is rapidly degraded by the exoribonuclease Rrp6p, Rat1p and Xrn1p. In the absence of *RNT1,* transcription continues to the polyadenylation signal of the downstream gene or until it meats a cryptic polyadenylation site in the intergenic region. In both cases, the polyadenylated RNA is transported and translated disrupting the auto-regulatory circuit of the RBP.

aberrant RNA. This interpretation is also compatible with the role of Rnt1p in mRNA degradation (Ge et al., 2005; Larose et al., 2007). In this case, Rnt1p will not simply degrade the newly transcribed mRNA but will also terminate transcription. This indeed explains why an enzyme localized in the nucleus plays a role in the regulation of mostly cytoplasmic RNA species like mRNA. However, the discovery that Rnt1p cleavage elicits transcription termination raises questions about the mechanism of polycistronic snoRNA processing. In this scenario, Rnt1p cleaves between snoRNAs that are transcribed as a single transcript leading to the maturation of these different RNAs (Ghazal et al., 2005). Therefore, if Rnt1p cleavage leads to transcription termination, the downstream snoRNA will not be produced. This apparent paradox could be explained by the presence of specific sequence elements or transcription factors that specifically prevent the 5' end generated by Rnt1p cleavage of these gene clusters from being digested by Rat1p. Case by case studies of Rnt1p cleavage and its link to termination will reveal the existence of these elements. Meanwhile, the data presented here reveal a model of polyadenylation independent transcription termination and provide a mechanism by which transcription termination may regulate gene expression.

Experimental Procedures

Strains and plasmids

Yeast strains were grown and manipulated using standard procedures (Guthrie and Fink, 1991). Yeast strains used in this study are listed in supplemental Table 1. For details, see Supplemental material file 1.

Search for Rnt1p cleavage signals

All uninterrupted pairs of ORFs transcribed in the same orientation were identified in the April 9th 2008 version of *Saccharomyces* Genome Database (SGD). Independently, all conserved NGNN-capped stem-loops in five *sensu strictoSaccharomyces* species (*S. cerevisiae, S. paradoxus, S. mikatae, S. kudriavzevii and S. bayanus*) were identified using the genome multiple alignment from UCSC (http://hgdownload.cse.ucsc.edu/goldenPath/ sacCer1/bigZips/multizYeast.zip). To assess conservation, three criteria where considered: the conserved G in position two of the tetraloop, the capacity of the two closing base pairs to form canonical Watson-Crick base pairs and the formation of 23 nt NGNN-capped stem-loop as predicted by Vienna RNA 1.6.5. When these three criteria were validated at the same position in the alignment for the five species, the stem-loop was considered as conserved. We found that three of these conserved NGNN-capped stem-loops are located in the intergenic region between consecutive coding transcripts.

RNA analysis

RNA extractions, Northern hybridization and primer extension were performed as previously described(Ghazal et al., 2005). Primer extension was performed using (20 ng/µl) of reverse primer CAAATTCTTTGAAATTAGCCTGACCCAAAC, and 10 µg of RNA. The primers used to generate the randomly labeled probes (Perbal, 1988) used for Northern blots are listed in Supplemental Table 1. Cleavage of total RNA was conducted as previously described (Ghazal et al., 2005) using 50 µg of total RNA and 8 pmol of purified Rnt1p (Lamontagne and Abou Elela, 2001). Standard 1.2 % agarose or 6% polyacrylamide gels were used to separate low and high molecular weight RNAs respectively.

Western Blotting

Total protein extracts and western blot analysis were performed as described before (Catala et al., 2008). Cells were grown in 50 ml culture. Proteins (10-20 µg) were loaded on 12% SDS gel, transferred and incubated with 1:3000 dilution of antibodies against Npl3p (Russell and Tollervey, 1992) and 1:500 Gpi17p (Zhu et al., 2005). Anti-rabbit HRP was used as secondary antibody at a dilution of 1:80 000. Pgk1p was detected using anti-mouse HRP antibody as a secondary at a dilution of 1:16000 (Sigma-Aldrich Canada Ltd., Oakville, Ont).

Chromatin immunoprecipitation and microarray analysis

Chromatin extracts were prepared as described (Strahl-Bolsinger et al., 1997; Taggart et al., 2002). Immunoprecipitations were performed with monoclonal anti-Rpb1 8WG16 (Covance, Berkeley, CA) and polyclonal anti-Rnt1p (Lamontagne et al., 2000) as described earlier (Catala et al., 2008). The method used for quantitative PCR amplification is outlined in supplemental material. ChIP material was labeled and hybridized on DNA microarrays (Agilent Technologies) containing 44,290 Tm-adjusted 60-mer probes covering the entire yeast genome for an average density of one probe every 287 bp (±100 bp) as described before (Rufiange et al., 2007). The data were normalized and replicates were combined using a weighted average method as described previously (Rufiange et al., 2007). The combined datasets are available supplemental file 2. Comparing RNAP II density beyond the 3' end of ORFs identified genes with termination defects in *rnt1*Δ cells. In addition, those exhibiting previously noted changes in expression after the deletion of *RNT1* were closely inspected to ensure that no obvious candidates are missed through the automated selection process.

Accession Numbers

The ChIP-chip data in this paper have been deposited in NCBIs Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO series accession number GSE16784.

Acknowledgements

We thank Momchil Vodenicharov for help with chromatin immunoprecipitation and Julie Parenteau with yeast genetics. We also thank Jean-François Lucier for the design of PCR primers. We are indebted for Mark Schmitt, Julie Parenteau, Mathieu Catala and Bruno Lamontagne for critical reading of the manuscript. This work was supported by a grant from the Canadian Institute of Health Research. S. A. is a Chercheur Boursier National of the Fonds de la Recherche en Santé du Québec. F.R. holds a New Investigator Award from the Canadian Institute of Health Research. P-É.J. holds a post-doctoral award from the IRCM training program in cancer research funded by the CIHR. J.-R.L is a research fellow of the Terry Fox Foundation through an award from the National Cancer Institute of Canada.

References:

Abou Elela, S., and Ares, M., Jr. (1998). Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA. EMBO Journal *17*, 3738-3746.

Abou Elela, S., Igel, H., and Ares, M., Jr. (1996). RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. Cell *85*, 115-124.

Akada, R., Kitagawa, T., Kaneko, S., Toyonaga, D., Ito, S., Kakihara, Y., Hoshida, H., Morimura, S., Kondo, A., and Kida, K. (2006). PCR-mediated seamless gene deletion and marker recycling in Saccharomyces cerevisiae. Yeast *23*, 399-405.

Blancafort, P., Ferbeyre, G., Sariol, C., and Cedergren, R. (1997). PolI-driven integrative expression vectors for yeast. J Biotechnol *56*, 41-47.

Burkard, K.T., and Butler, J.S. (2000). A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. Mol Cell Biol *20*, 604-616.

Carroll, K.L., Pradhan, D.A., Granek, J.A., Clarke, N.D., and Corden, J.L. (2004). Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol Cell Biol *24*, 6241-6252.

Catala, M., Tremblay, M., Samson, E., Conconi, A., and Abou Elela, S. (2008). Deletion of Rnt1p alters the proportion of open versus closed rRNA gene repeats in yeast. Mol Cell Biol *28*, 619-629.

Chanfreau, G., Rotondo, G., Legrain, P., and Jacquier, A. (1998). Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. Embo J *17*, 3726-3737.

Chinnusamy, V., Gong, Z., and Zhu, J.K. (2008). Nuclear RNA export and its importance in abiotic stress responses of plants. Curr Top Microbiol Immunol *326*, 235-255.

Cusick, M.E. (1994). Purification and identification of two major single-stranded binding proteins of yeast Saccharomyces cerevisiae as ribosomal protein L4 and histone H2B. Biochim Biophys Acta *1217*, 31-40.

El Hage, A., Koper, M., Kufel, J., and Tollervey, D. (2008). Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. Genes Dev *22*, 1069-1081.

Ge, D., Lamontagne, B., and Abou Elela, S. (2005). RNase III-Mediated Silencing of a Glucose-Dependent Repressor in Yeast. Curr Biol *15*, 140-145.

Ghazal, G., and Elela, S.A. (2006). Characterization of the reactivity determinants of a novel hairpin substrate of yeast RNase III. J. Mol. Biol. *363*, 332-344.

Ghazal, G., Ge, D., Gervais-Bird, J., Gagnon, J., and Abou Elela, S. (2005). Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. Mol Cell Biol *25*, 2981-2994.

Grec, S., Wang, Y., Le Guen, L., Negrouk, V., and Boutry, M. (2000). Cryptic polyadenylation sites within the coding sequence of three yeast genes expressed in tobacco. Gene *242*, 87-95.

Guthrie, C., and Fink, G.R. (1991). Guide to Yeast Genetics and Molecular Biology (San Diego, CA: Academic Press).

Herrero, E. (2005). Evolutionary relationships between Saccharomyces cerevisiae and other fungal species as determined from genome comparisons. Rev Iberoam Micol *22*, 217-222.

Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., and Proudfoot, N.J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. Genes Dev *22*, 1082-1092.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517-522.

Kim, M., Vasiljeva, L., Rando, O.J., Zhelkovsky, A., Moore, C., and Buratowski, S. (2006). Distinct pathways for snoRNA and mRNA termination. Mol Cell *24*, 723-734.

Krebber, H., Taura, T., Lee, M.S., and Silver, P.A. (1999). Uncoupling of the hnRNP Npl3p from mRNAs during the stress-induced block in mRNA export. Genes Dev *13*, 1994-2004.

Kufel, J., Dichtl, B., and Tollervey, D. (1999). Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3' ETS but not the 5' ETS. RNA *5*, 909-917.

Lamontagne, B., and Abou Elela, S. (2001). Purification and characterization of Saccharomyces cerevisiae Rnt1p nuclease, Vol 342 (San diego, CA: Academic Press).

Lamontagne, B., Tremblay, A., and Abou Elela, S. (2000). The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. Mol Cell Biol *20*, 1104-1115.

Larose, S., Laterreur, N., Ghazal, G., Gagnon, J., Wellinger, R.J., and Elela, S.A. (2007). RNase III-dependent regulation of yeast telomerase. J Biol Chem *282*, 4373-4381.

Lee, A., Henras, A.K., and Chanfreau, G. (2005). Multiple RNA surveillance pathways limit aberrant expression of iron uptake mRNAs and prevent iron toxicity in S. cerevisiae. Mol Cell *19*, 39-51.

Liti, G., Barton, D.B., and Louis, E.J. (2006). Sequence diversity, reproductive isolation and species concepts in Saccharomyces. Genetics *174*, 839-850.

Lund, M.K., Kress, T.L., and Guthrie, C. (2008). Autoregulation of Npl3, a yeast SR protein, requires a novel downstream region and serine phosphorylation. Mol Cell Biol *28*, 3873-3881.

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

Malagon, F., Kireeva, M.L., Shafer, B.K., Lubkowska, L., Kashlev, M., and Strathern, J.N. (2006). Mutations in the Saccharomyces cerevisiae RPB1 gene conferring hypersensitivity to 6-azauracil. Genetics *172*, 2201-2209.

Milligan, L., Torchet, C., Allmang, C., Shipman, T., and Tollervey, D. (2005). A nuclear surveillance pathway for mRNAs with defective polyadenylation. Mol Cell Biol *25*, 9996-10004.

Noble, S.M., and Johnson, A.D. (2005). Strains and strategies for large-scale gene deletion studies of the diploid human fungal pathogen Candida albicans. Eukaryot Cell *4*, 298-309.

Perbal, B. (1988). A Practical Guide to Molecular Cloning, Second edn (New York: John Wiley and Sons, Inc.).

Roth, K.M., Wolf, M.K., Rossi, M., and Butler, J.S. (2005). The nuclear exosome contributes to autogenous control of NAB2 mRNA levels. Mol Cell Biol *25*, 1577-1585.

Rufiange, A., Jacques, P.E., Bhat, W., Robert, F., and Nourani, A. (2007). Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. Mol Cell *27*, 393-405.

Russell, I., and Tollervey, D. (1995). Yeast Nop3p has structural and functional similarities to mammalian pre- mRNA binding proteins. Eur J Cell Biol *66*, 293-301.

Russell, I.D., and Tollervey, D. (1992). NOP3 is an essential yeast protein which is required for pre-rRNA processing. J Cell Biol *119*, 737-747.

Steinmetz, E.J., Warren, C.L., Kuehner, J.N., Panbehi, B., Ansari, A.Z., and Brow, D.A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell *24*, 735-746.

Strahl-Bolsinger, S., Hecht, A., Luo, K., and Grunstein, M. (1997). SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. Genes Dev. *11*, 83-93.

Taggart, A.K., Teng, S.C., and Zakian, V.A. (2002). Est1p as a cell cycle-regulated activator of telomere-bound telomerase. Science *297*, 1023-1026.

Tollervey, D. (2004). Molecular biology: termination by torpedo. Nature *432*, 456-457.

Tremblay, A. (2002). Etude de la fonction de la RNase III eucaryote et identification de ses partenaires cellulaires dans un criblage double-hybrides. In Department of Microbiology (Sherbrooke, Université de Sherbrooke).

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

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

Zhu, Y., Fraering, P., Vionnet, C., and Conzelmann, A. (2005). Gpi17p does not stably interact with other subunits of glycosylphosphatidylinositol transamidase in Saccharomyces cerevisiae. Biochim Biophys Acta *1735*, 79-88.

Table 1: Genome-wide screen for *RNT1* dependent transcription termination

Gene Name	ORF	Function	Extension length[a]	Fold change[b]	Expression[c]	Extension[d]
NPL3	YDR432W	RNA Metabolism	2440	5.3	No change	NE
U2	LSR1	RNA Metabolism	750	2.0	Decreased	NE
snR190	snR190	Ribosome Biogenesis	660	8.5	Decreased	NE
RRP1	YDR033W	Ribosome Biogenesis	510	3.9	Increased	E
NOP1	YDL014W	Ribosome Biogenesis	910	7.8	Decreased	No
RPL40A	YIL148W	Ribosomal Protein	740	3.5	ND	ND
RPL27B	YDR471W	Ribosomal Protein	560	1.6	ND	ND
RPL27A	YHR010W	Ribosomal Protein	290	2.5	Increased	NE[e]
RPL8A[f]	YHL033C	Ribosomal Protein	650	5.8	Decreased	NE
RPP1a[f]	YDL081C	Translation	400	1.1	No change	NE
RPS14B	YJL191W	Ribosomal Protein	750	2.9	ND	ND
TEF2	YBR118W	Translation	1000	4.8	ND	ND
TOS1	YBR162C	Cell Wall	1070	2.4	Increased	No
YPS3	YLR121C	Membrane	1010	2.3	Increased	E
MRH1	YDR087C	Membrane	820	2.5	Increased	No
ZEO1	YOL109W	Membrane	230	1.5	Increased	E

MIC17	YMR002W	Membrane	440	2.7	Increased	E
LYP1	YNL268W	Membrane	1650	3.5	Increased	NE
FTR1	YER145C	Membrane	2430	5.6	Increased	No
NPL4[f]	YBR170C	Membrane	1270	1.8	Increased	NE
TPI1	YDR050C	Glycolysis	470	2.8	Increased	E
CDC19	YAL038W	Glycolysis	450	4.1	ND	ND
PRE6	YOL038W	Glycolysis	2430	2.2	Increased	E
RPN12	YFR052W	Proteosome	750	2.2	ND	ND
OTU1	YFL044C	Protein Degradation	2150	3.3	ND	ND
SBA1	YKL117W	Protein Folding	470	2.7	Increased	No
LGE1	YPL055C	Histone Modifications	690	2.6	Increased	E
MED7	YOL135C	Transcription	820	3.0	ND	ND
SUT1	YGL162W	Transcription	430	2.3	ND	ND
NCB2	YDR397C	Transcription	970	2.6	ND	ND
PTC2[f]	YER089C	DNA Damage	770	4.4	Increased	NE
YKU80[f]	YMR106C	DNA Damage	2400	1.6	Increased	NE
SIM1	YIL123W	DNA Replication	580	3.1	No change	NE[e]
GIC2	YDR309C	Budding	1160	2.7	Increased	NE
OPI6	YDL096C	Dubious	460	3.2	Increased	E
RIB1	YBL033C	Response to drug	680	2.9	Increased	NE
-	YDR524W-C	Unknown	320	1.5	Increased	NE

[a] The length of the extension (in nucleotides) predicted by the RNAP II ChIP-chip upon the deletion of *RNT1*. [b] The *rnt1*Δ / *RNT1* fold change of RNAP II occupancy in the identified extension adjusted by the difference of RNAP on the complete gene. [c-d] Variation in the level of expression of mature RNA fragment [c] or size [d] as detected by Northern blot (ND, not determined; E, RNA extension detected in wild type strains that is increased in the absence of *RNT1*; NE, new extension detected only in the absence of *RNT1* and; No, no extension). [e] Extension due to intron retention. [f] Manually selected genes. The changes in the expression of U2 and snR190 in the absence of *RNT1* were previously reported (Abou Elela and Ares, 1998; Chanfreau et al., 1998b).

DISCUSSION

1. A flexible mechanism of substrate selectivity increases the spectrum of Rnt1p cleavage targets

Members of the RNase III family are defined by their ability to identify and cleave simple RNA duplexes. Indeed, all RNase III orthologues can cleave long A form RNA helices with very low sequence complexity *in vitro*, albeit with variable efficiencies. *In vivo*, these enzymes are highly species specific and usually cleave short structured RNA. This conditional substrate selectivity and apparent promiscuity *in vitro* makes substrate prediction, which is essential for the understanding of the enzyme *in vivo* function, very difficult. Therefore, in the initial part of this study, we aimed at probing the substrate specificity of a model RNase III (Rnt1p) under defined *in vitro* conditions and validating the finding *in vivo*. By using a combined *in vivo* and *in vitro* approach, we managed to develop a substrate definition capable of predicting potential Rnt1p processing signals and produce a more accurate model for the mechanism of substrate selection. Rnt1p substrate selectivity was viewed as a fixed process in which the enzyme cleaves at a fixed nucleotide-distance from the AGNN tetraloop structure (Lamontagne et al., 2003). The guanosine in the second position was presumed to be universally conserved and thus believed to be essential for cleavage (Lebars et al., 2001). In this study we demonstrate that the enzyme may recognize different forms of the tetraloop sequence and the substrate is not defined by a single structure or sequence

element but as an integrated unit of reactivity determinants (Lamontagne et al., 2003). Furthermore, we have found that the cleavage site is not determined through a strict counting of the number of base pairs from the tetraloop structure. Instead, cleavage sites are selected as a function of the three-dimensional helical distance that separates the enzyme primary binding site and the position of scissile bond. Together these findings explain how a multifunctional enzyme can achieve precise cleavage while acting on a broad base of reactive substrates. This work was later supported by the solution structure of the substrate bound Rnt1p dsRBD, which indicated that the enzyme does not directly interact with the conserved G in the second position but rather used the minor groove at the 3' end for binding (Figure 6A). Furthermore, comprehensive nucleotides substitution in the upper stem loop structure confirmed that the enzyme does not use a fixed set of hydrogen bonds to bind its substrate, but rather uses a flexible network that could adapt to a variety of structures (Lavoie and Abou Elela, 2008). This enzyme flexibility is also applicable to other members of the RNase III family. For example, S. pombe RNase III (Pac1) may cleave long generic RNA duplex with little specificity, while accurately cleaving a selected group of short RNA hairpins at a fixed distance from a conserved internal loop structure (Rotondo et al., 1997; Lamontagne and Abou Elela, 2004). In human cells , the nuclear RNase III Drosha identifies a large terminal loop or uses chaperon proteins to identify the stem base. However, in the case of the RNA interference enzyme Dicer, substrate definition seems to be drastically modified by the evolution of

a new single stranded RNA specific domain called PAZ (Yan et al., 2003). This domain was shown to be important for substrate selectivity and the contribution of the dsRBD to the substrate selectivity of this enzyme remains unclear.

1.1 The AGNN tetraloop sequence is not essential for RNA cleavage

The first known substrates of Rnt1p were RNA processing signals found in the transcribed spacers near rRNA, snRNAs and snoRNAs. Examination of this limited group of related substrates revealed the presence of a common AGNN tetraloop near the cleavage site. Mutation of the first two nucleotides of the loop blocked cleavage suggesting that the AGNN tetraloop is a universal Rnt1p cleavage signal essential for substrate recognition. Solution structure of the AGNN tetraloop revealed a fold in which the guanosine at the 2^{nd} position is in *syn* orientation and the last two nucleotides are stacked (Lebars et al., 2001). Therefore, it was hypothesized that Rnt1p recognizes the RNA fold and not the sequence of AGNN tetraloop. However, the enzyme failed to cleave RNAs with similar tetraloop structures. Indeed, we found that Rnt1p can bind but not cleave the HIV-1 frameshift inducing ACAA tetraloop structure (data not shown), which forms a nearly identical AGNN fold (Stapler et Butcher 2003). This indicated that binding does not necessarily lead to cleavage and that Rnt1p can differentiate between closely related RNA folds. In addition, this observation necessitated a new look at the substrate

specificity of Rnt1p and the substrate reactivity epitopes. The discovery and analysis of a large group of substrates reported in this thesis suggestes that Rnt1p cleavage may occur in sequences devoid of AGNN tetraloop. Instead the cleavage may be solicited be NGNN or AAGU tetraloops. The AAGU does not solicit Rnt1p cleavage by mimicking the AGNN fold. Solution structure of AAGU tetraloop indicates that unlike the AGNN structure, where G is in *syn* position, the 2nd nucleotide of the AAGU tetraloop exists in *anti* confirmation similar to that of the HIV stem-loop structure. However, the major difference between AAGU and both the AGNN and ACAA tetraloops is at the 3' end of the loop where the last nucleotide uracil was found to be extruded from the tetraloop (Gaudin et al 2006). The changes in the tetraloop 3' end appears to modify the structure of the minor groove found in the AAGU tetraloop while allowing better access to the 5' end nucleotide of the loops than in the case of the AGNN structure (Figure 6B). Indeed, we found that unlike the AGNN teraloops where the Rnt1p mostly interacts with the 3' sequences, the AAGU tetraloop interaction was more evenly distributed across the tetraloop nucleotides. This study allowed the reclassification of Rnt1p substrates into two classes. The first is the G2 class, which contains an NGNN tetraloop and the other is the A1 class, which include RNA with AAGU tetraloop. Each of these two classes uses different sets of hydrogen bonds to bind the enzyme and displays varying dependency on the adjacent stem sequence and structure.

1.2 Contribution of the stem sequence to cleavage efficiency

Bacterial RNase III uses negative selection to distinguish between its substrate and other RNA duplexes. This antideterminant mode of substrate selection means that it is the absence of specific sequence for A form RNA helix and not its presence that drives binding and cleavage (Zhang and Nicholson, 1997). In yeast, it was thought the fold of the terminal tetraloop is sufficient for substrate selection with no or little influence of the RNA stem. During my study, I was able to demonstrate that in addition to the AGNN tetraloop, stem sequences can also play a role in determining the binding affinity and cleavage efficiency (Lamontagne et al 2003). By using mutational analysis we were able to dissect three regions in the substrate that affect binding and cleavage (Figure 5B). The initial binding and recognition box represented by tetraloop drives initial contact and positioning of the enzyme, while binding efficiency is ensured by the base pairs downstream of the loop. Finally, catalysis is influenced by the sequence immediately adjacent to the cleavage site. Through a combination of these three elements the enzyme activity could be finely tuned to achieve correct cleavage at the correct time. For example, substrates that required maximum cleavage activity like pre-rRNA exhibit the most effective combination of these elements while substrates found in the middle of mRNA may use less reactive combinations.

The contribution of the stem sequence to Rnt1p binding and cleavage becomes more evident as the tetraloop reactivity decreases. Indeed, the

presence of an internal loop at specific positions enhances cleavage of stems capped with A1 tetraloops, which is known to be less reactive than G2 tetraloops. This internal tetraloop sequence in combination with specific stem sequences may direct weak cleavage even in presence of GNRA tetraloop, a form of loop that deters binding to Rnt1p. This mode of action resembles that identified in fission yeast where the *Sshizosaccharomyces pombe* RNase III orthologue Pac1 recognizes an internal loop for cleavage (Rotondo et al., 1997). Together these findings indicate that while Rnt1p has developed high affinity to the structure generated by G2 and A1, it may cleave RNA duplex with a specific structure as within other members of the RNase III family.

1.3 The helical rulers cleavage mechanism re-defined

Members of dsRNA binding proteins recognize the A form of a dsRNA. The conserved dsRNA binding domain "dsRBD", identifies an A-form dsRNA of 11bp by interacting with a narrow deep major groove formed between 2 shallow minor grooves. Thus, the enzyme distinguishes between dsRNA and dsDNA by the distance between the two minor grooves (Abou Elela et al., 1996; Wu et al., 2004; Gan et al., 2006). In yeast, Rnt1p dsRBD has a similar structrue as that of other dsRBPs, and consists of 2 helices and 3 beta sheets. These results suggest a similar role of dsRNA recognition where the enzyme measures the distance between the 2 minor grooves. However, Rnt1p possesses an additional helix $\propto 3$, which was proposed to play an

important role in increasing the enzyme affinity to RNA hairpins (Figure 6A) (Abou Elela et al., 1996; Leulliot et al., 2004; Wu et al., 2004; Gan et al., 2006). The additional helix modifies the position of amino acid sequences that normally interact with the A form helix minor groove and make them fit better with the minor groove formed by the tetraloop 3' end. Based on this mechanism and experimental observation of the cleavage site, it was proposed that the enzyme functions as a helical ruler by counting the number of base pairs from the tetraloop. Indeed, normally the enzyme cleaves at each side of the stem 14 and 16 bp pairs below of the tetraloop. In this study, we showed that *in vivo*, Rnt1p processes polycistronic and intron-encoded snoRNAs by cleaving a forked RNA stem that forms through long-range RNA interactions using a single tetraloop. In this case, the distance from the tetraloop to the cleavage site is interrupted with stems that may extend hundreds of nucleotides and thus strict base pair counting is not possible. Instead, the cleavage signal seems to be selected by default based on the alignment with the enzyme nuclease domain, which forms as a consequence of an interaction between tetraloop minor groove and the dsRBD. This theory is supported by an independent study showing that Rnt1p could cleave coaxialy stacked structures formed *in trans* (Lamontagne et al 2007). Crystal structure of the bacterial RNase III suggests that the enzyme first binds its substrate through interactions with the dsRBD leading to conformational changes that form the catalytic core. The catalytic core is created through intermolecular homodimerization of the nuclease domains of each of the two

enzyme subunits. The distance within the catalytic core is a perfect fit for the size and shape of an A form RNA helix. Based on this, one can hypothesize that Rnt1p recognizes its substrate by first binding to the tetraloop as the anchor point using its dsRBD and that this binding leads to an automatic fitting in the nuclease domain and cleavage site selection. In this case, any additional stems or structures that interrupt the continuity of the tetraloop stem sequence could be simply accommodated in the open space between the two subunits with minimum impact on the RNP complex.

1.4 Flexible protein conformation adopts to changes in the substrate structure

In addition to the conserved nuclease and dsRNA binding domains found in bacterial RNase III, eukaryotic RNase IIIs including Rnt1p possess an additional N-terminal domain. In higher eukaryotes, the N-terminal domain may harbor a single stranded RNA recognition domain (PAZ), helicase domain, or even an extra nuclease domain. However, in budding yeast Rnt1p and fission yeast Pac1, the N-terminal domains do not share any of these structural motifs. *In vivo*, deletion of Rnt1p inhibits growth and *in vitro* deletion of Rnt1p N-terminal reduces RNA cleavage by 30% in physiological salt concentrations. Increasing the salt condition completely inhibits cleavage of the ΔN-term enzyme suggesting that the N-terminal domain contributes to the stability of the enzyme substrate structure. The role of the N-terminal domain appears to be much more pronounced in the case of A1 class of substrates.

Deletion of the N-terminal domain does not affect the binding of Rnt1p to AGNN tetraloop while binding of the AAGU tetraloops under the same conditions is strongly inhibited (Ghazal and Elela, 2006). Indeed, chemical modifications and footprinting indicated that deletion of the N-terminal changes the binding pattern near the AAGU and not the AGNN loop (Ghazal and Elela, 2006). Earlier studies, showed that the N-terminal domain by itself couldn't bind RNA (Lamontagne et al., 2000). However, the N-terminal domain can influence the dsRBD binding and hence changes the substrate affinity. Rnt1p functions as a head-to-tail homodimer where the dsRBD forms intermolecular interaction with the N-terminal domain. Therefore, we propose that the N-terminal domain interacts with the dsRBD to stabilize its interaction with the N-terminal domain (Lamontagne et al., 2001). Normally, this interaction is strong in the case of G2 substrate and thus the N-terminal impact could be observed only on cleavage and in high salt conditions (Figure 6 C, D). However, in the case of an A1 loop the interaction with the dsRBD and substrate appears to be weaker and thus requires the additional stability provide by the N-terminal domain for binding and cleavage even at low salt concentration. According to this hypothesis, dsRBD / N-terminal domain interaction forms a "cap" capable of identifying non-canonical loops by providing added stability that cannot be achieved by the dsRBD itself. In this way the enzyme could compensate for the substrates short comings and increases its range of potential substrates.

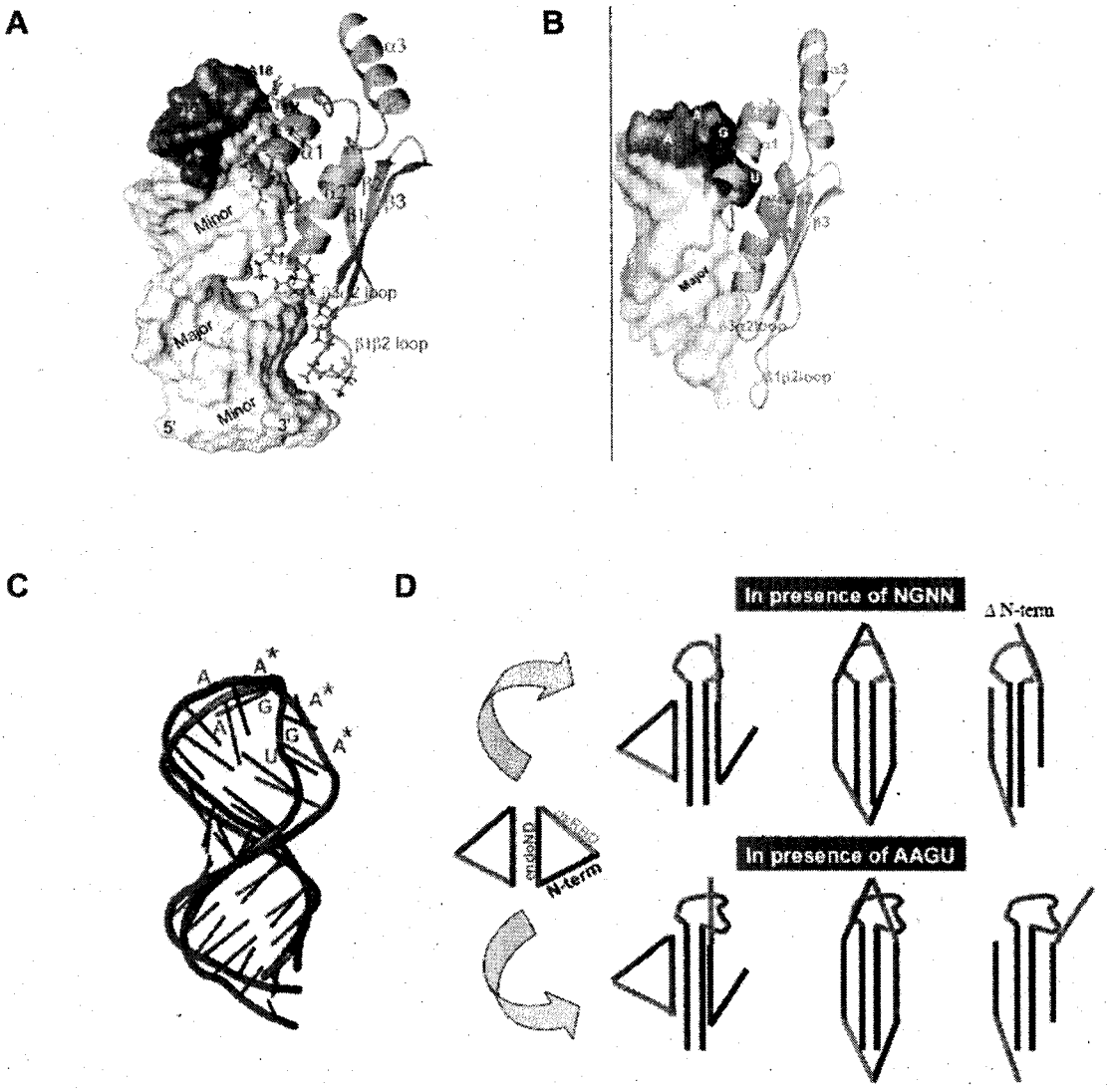


Figure 6. Rnt1p Interaction with stem-loop capped with either NGNN or AAGU tetraloop.

(A) Solution structure of the dsRBD of Rnt1p in complex with snR47 capped with AGAA tetraloop. dsRBD recognizes the shape of the minor groove formed by the tetraloop (B) Hypothetical model showing the interaction of the dsRBD of Rnt1p with snR48 capped with AAGU tetraloop. (C) Superposition of dsRNA with AGAA tertaloop shown in orange and AAGU tetraloop shown in purple. (D) In presence of NGNN tetraloop the binding of Rnt1p is initiated at the 3'end of the loop through the dsRBD then a homodimer will form through an intermolecular interaction between both subunits. In presence of AAGU tetraloop, the extrusion of the last nucleotide U, allows the dsRBD to bind to the 5' end of the loop.

2. The architecture of the pre-snoRNA transcript defines the processing mechanism

In eukaryotic cells, production of mature ribosomes includes rRNA pseudouridylation and methylation by a large set of sequence specific snoRNA protein complexes. Pseudouridylation is guided by a group of snoRNAs exhibiting a conserved H/ACA structural motif, while methylaion is performed by C/D box containing snoRNAs. The two snoRNA families appear to have evolved recently since they are not found in bacteria where RNA modification is carried out by protein enzymes. Indeed, these RNA families in budding yeast appear to be organized differently. The transcription of H/ACA box genes is mostly produced from independent promoters and less than a handful are produced from intronic mRNA. On the other hand, nearly half of the C/D snoRNAs are derived either from large polycistornic RNA or as part of mRNA introns. In this study, we showed that this different gene organization also has a major impact on the RNA maturation pathway. While only 2 H/ACA snoRNAs are processed by Rnt1p (snR43 and snR36), almost all C/D box snoRNAs require Rnt1p for maturation (Ghazal et al., 2005). It is curious however, to note that even monocistronic C/D snoRNAs are processed by Rnt1p at their 5' end and not at the 3' end as in the case of snRNAs. This mode of processing may be a relic of older polycistronic snoRNA clusters that lost snoRNA members during the evolution or evolved due to differences in the regulation and function of C/D and H/ACA box snoRNAs.

2.1 Rnt1p cleavage is essential for the processing of polycistronic snoRNA

In general, polycistronic clusters are common features of bacterial and not eukatyotic genomes. However, transcripts that carry more than one gene sequence exist in certain eukaryotes e.g. in worms (Evans and Blumenthal, 2000). In these cases, genes of related function are expressed from the same promoter.However,unlike bacterial mRNAs they are not translated as a single unit (McCarthy, 1990). Instead, they are processed to discrete messages by *trans*-splicing before translation (Evans and Blumenthal, 2000). In yeast, the only well studied clusters are those formed by snoRNAs. This mode of transcription ensures coordinated expression of several snoRNAs with similar functions. However, these snoRNAs cannot mature by simple exoribonucleolytic cleavage. Here we showed that the release of polycistronic snoRNAs requires endonucleolytic cleavage by Rnt1p (Ghazal et al., 2005). The release of polycistronic snRNAs is normally achieved by long-range interactions between the termini of each snoRNA. In this way a single processing step could release both snoRNA 5' and 3' ends. It is interesting to note that this exact architecture of the processing signals was observed in bacterial rRNA processing by RNase III. In this case, the ends of the 16S and 23S rRNAs are paired and a single cleavage event releases the rRNA to be trimmed to the mature ends by exoribonucleases. Therefore, while eukaryotic rRNA appears to have lost this feature and instead uses separate 5' and 3'

end processing signals, snoRNA at least in yeast stands as evidence for the evolutionary origin of these processing events.

2.2 Is the cleavage of the monocistornic RNA an evolutionary relic of gene clusters or a mechanism for quality control?

The rational for implicating endoribonucleases in the maturation of polycistronic snoRNA is obvious since there is no readily clear alternative for the excision of internal RNA fragments without destroying the flanking sequence. However, the involvement of Rnt1p in the processing of monocistronic RNA units is more complicated to explain. Cleavage at RNA ends could easily be achieved by exonulceases unless the sequence is highly structured or recessed. However, in reality endoribonucleases may contribute to the maturation of monocistronic RNA by inducing or accelerating the process of degradation. For example, cleavage of pre-snoRNAs by Rnt1p initiates the degradation of this otherwise stable RNA species. In this way, only correctly assembled snoRNAs are protected from degadation leading to the formation of mature RNP complex.

Processing of snoRNAs at the 5' end may also signify a required step for the modification of the 5' end. The default Pol II transcripts are capped with 7-methyl guanosine mRNA while many C/D and not H/ACA box snoRNAs exhibit 2,2,7-trimethyl guanosine (TMG) cap. Indeed, most 5' end processed

snoRNA forms a TMG cap (Chanfreau et al., 1998a). Therefore, it is possible to suggest that Rnt1p processing is required for the generation of TMG capped snoRNA. In this case, one would hypothesize that the snoRNA would be produced with normal mRNA like cap that protect the RNA from degradation and that Rnt1p cleaves the RNA once the TMG capping machine is assembled to ensure maximum RNA stability.

2.3 Processing of intronic snoRNA and the regulation of pre-mRNA splicing

In higher eukaryotes, almost all RNAs contain introns and therefore, it is not surprising that one finds snoRNA takes advantage of this and uses introns as an expression vehicle. Indeed, most snoRNAs in human do not share the function of their host genes. In contrast, the yeast genome is very compact with only about 5% of genes containing introns. However, almost all ribosomal protein genes contain introns and many harbour snoRNAs that modify the same rRNA and bind the protein encoded by the host pre-mRNA. Therefore, in yeast, embedding snoRNA in introns seems to help coordinating gene function. In this study we demonstrated that Rnt1p can release two different snoRNAs encoded in the introns of a pair of duplicated ribosomal genes RPL7A and RPL7B. In these cases, Rnt1p provides two alternative modes of intronic snoRNA processing. Excision of snR39 occurs in the lariat after processing occurs and as such the role of Rnt1p cleavage is to accelerate

processing and prevent trapping, in case debranching does not occur (Ghazal et al., 2005). This also means that snoRNA processing is dependent on the splicing of the host genes. However, snR59 processing occurs in the pre-mRNA stage before splicing occurs and thus processing of snRN59 is mutually exclusive with the production of the host gene. The differences in the processing of these two snoRNAs may reflect the need to coordinate the expression of their host protein coding genes. Expression of duplicated ribosomal genes needs to be carefully regulated to produce the exact amount required for ribosome assembly. Rnt1p cleavage provides the opportunity to inhibit the expression of RPL7B without affecting the expression of the associated snoRNA. One can imagine that overexpression of RPL7A would inhibit the splicing of RPL7B leading to the accumulation of pre-mRNA that could be cleaved by Rnt1p to ensure steady state expression of snR59. In addition, Rnt1p appears to regulate the accumulation of unspliced mRNA independently of snoRNA processing, suggesting that this mode of regulation might be a more general mechanism of gene regulation. Recently, it has been shown in mammalian cells that Drosha (Rnt1p-orthologue) cleaves intron-encoded miRNA co-transcriptionally before the host intron is spliced out (Morlando M. et al 2009). Indeed, some of the newly discovered miRNAs have snoRNA precursors (Kim and Kim, 2007). This suggests that certain aspects of Rnt1p are maintained in human cells.

3. New model for transcription termination

In eukaryotes, it is proposed that transcription termination may occur either through allosteric modifications of the RNA polymerase complex (allosteric model) or through exoribonucleolytic cleavage of the RNA 5' end generated by the polaydenylation machinery (torpedo model). The allosteric mode of termination is achieved through conformational changes of Pol II upon binding of Nrd1 complex in close proximity to the promoter (Carroll et al., 2004; Vasiljeva et al., 2008). On the other hand, torpedo termination demands the recruitment of the exoribonuclease Rat1p to degrade the remaining 3' product attached to the transcription complex. This degradation activity releases through a yet to be identified mechanism Pol II for recycling. Neither of these models explains how long RNA transcripts devoid of polaydenylation signals are created. Here we showed that Rnt1p may mediate a "torpedo" like transcription termination in the absence of a polyadenylation signal by giving an access point to Rat1p (Figure 7). This new model of transcription may promote termination of stable non-polyadenylated RNA like U2 snRNA or degrade transcriptional readthrough transcripts of protein coding genes.

3.1 Transcription termination of non-polyadenylated RNA

Transcription termination of non-coding RNA and the formation of the 3' end are currently seen as two independent processes. Transcription termination of non-coding RNAs is achieved through the binding of the Nrd1/Nab3 complex

to loosely conserved sequence motif near the termination site (Steinmetz et al., 2001). Once termination is achieved and the RNP assembly process started, the final 3' end is formed by exoribonucleolytic trimming to the mature site near the binding site of the protein components of the RNP complex. It was also suggested that Rnt1p cleaves Pol II transcripts like U2 snRNA to remove a pre-formed poly(A) tail and trigger exoribonucleolytic maturation (Abou Elela and Ares 1998). However, in this study, it was shown that Rnt1p not only helps forming the mature 3' end of U2 but it also participates in transcription termination. Deletion of *RNT1* increased the association of Pol II with the 3' end of U2 snRNA, suggesting increased transcription read-through and inhibited termination at the canonical site (Ghazal et al., 2009). Similarly, the processing signal of Rnt1p at the 3' end of the 25S pre-rRNA has recently been shown to play a role in transcription termination by Pol I. As in the case of Pol II, Rnt1p cleavage at the 25S pre-rRNA gives access to Rat1p allowing it to terminate transcription by Pol I in a "Torpedo" like fashion (Kawauchi et al 2008). Transcription termination by Rnt1p is not exclusive to non-coding RNA. Protein coding genes also use Rnt1p for transcription termination (Ge at al 2005). However, in this case, the termination product is not polyadenylated and is rapidly degraded. Thus, a single cleavage event by Rnt1p could induce transcription termination leading to the formation of a stable snRNA complex or the degradation of an mRNA transcript. The question remains as to how Rnt1p identifies targeted genes and how it competes with polyadenylation signals that are often found near Rnt1p dependent termination sites?

It is now clear that Rnt1p cleavage near a transcription termination site occurs co-transcriptionally. We have shown that Rnt1p co-immunoprecipitate with actively transcribed DNA and is released near the cleavage signals (Ghazal et al 2009). This means that Rnt1p is recruited very soon after transcription starts and is released once it found the target RNA. However, it is not clear how Rnt1p identifies the right transcription complex to bind. We propose that the recruitment of Rnt1p to the transcription complex depends on the phosphorylation state of the Pol II CTD. Recent studies, suggest that phosphorylation of the CTD Ser5 recruits the Nrd1 termination complex, while phosphorylation of Ser2 recruits the cleavage and polyadenylation factors. The choice of termination mode depends on sequences in the transcribed RNA, as well as on the distance of the termination signal to the promoter. Rnt1p is known to co-immunoprecipitate with the Nrd1/Nab3 complex and physically interacts with the helicase Sen1 a component of this complex. This interaction would ensure the recruitment of Rnt1p to Ser5 and its presence when non-coding 3' end forms. Surprisingly, our preliminary data using two-hybrid assay suggest that Rnt1p binds poorly to Ser2 / Ser 5 phosphorylated CTD (data not shown). However, mutation of either serine 2 or serine 5 dramatically increases interaction with Rnt1p. This result is consistent with an idea that Rnt1p is recruited to the transcriptional unit in a phosphorylation dependent manner specifically when at least one of the two canonical termination machineries cannot bind to the transcriptional complex. This would indeed confirm that Rnt1p provides an alternative transcription

termination mechanism in situations where neither conventional "torpedo" nor "allosteric" mode of termination is possible (long non-polyadenylated RNA) (Figure 7).

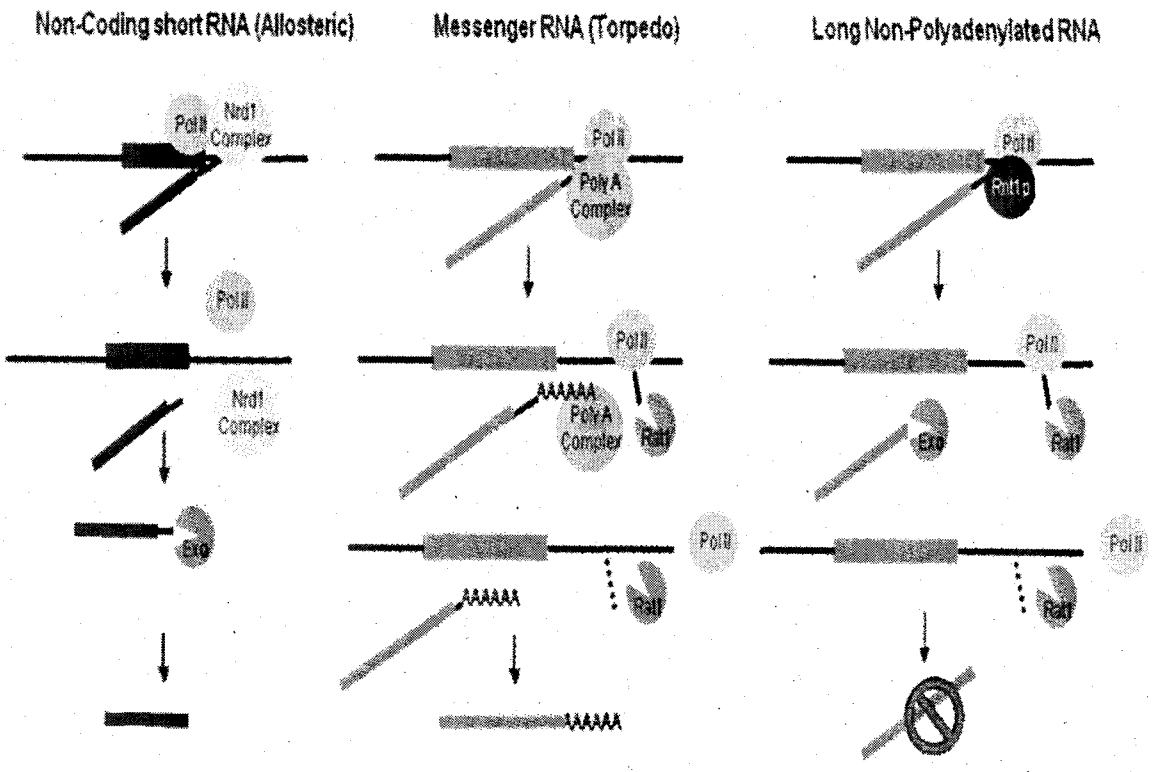


Figure 7. Model representing three modes of transcription termination.

(A) Allosteric model requires the recruitment of Nrd1 complex at the 3'end of the non-coding RNA. 3'end formation is achieved through trimming by exonucleases. (B) Torpedo model requires the recruitment of the cleavage/polyaenylation complex followed by the addition of the poly (A) at the 3'end of the mRNA. The exonuclease Rat1p is recruited co-transcriptionally to degrade allowing dissociation of the Pol II. (C) Termination by Rnt1p. Long non-polyadenylated RNA is cleaved at its 3'end through endonucleolytic cleavage of Rnt1p co-transcriptionally. The cleavage product of Rnt1p is then degraded by the torpedo exonuclease Rat1p.

3.2 Transcription termination dependent RNA decay

In addition to their roles in regulating gene expression, ribonucleases play an important role in inhibiting the expression of non-functional or aberrant RNAs. For example, the 3' – 5' exoribonuclease Rrp6p was recently shown to play an important role in repressing the expression of short cryptic unstable transcripts "CUT" (Arigo et al., 2006). In this case, inefficient transcription termination leads to the synthesis of aberrant polyadenylated RNA, or CUT, that is rapidly degraded by the exonuclease Rrp6p and associated exosome complex in the nucleus. Similarly, it has been shown that defects in Nrd1/Nab3 termination of non-coding RNAs located upstream of protein-coding genes causes transcription read-through that gives rise to an unstable polycistronic transcript that is normally polyadenylated and degraded. In this study, we have shown that Rnt1p provides an additional layer for genome surveillance to prevent transcriptional readthrough of neighboring genes. Deletion of *RNT1* causes transcription readthrough resulting in the accumulation of RNA transcripts with extended 3' ends or the production of long ploycistronic RNA units (Ghazal et al 2009). In the absence of Rnt1p, these aberrant transcripts are polyadenylated and exported to the cytoplasm, where they associate with polyribosomes and produce proteins (data not shown). This observation underscores the importance of Rnt1p surveillance since in the absence of the enzyme; transcription readthrough of polyadenylation signals will not destroy the aberrant transcript. Instead, a new

and perhaps mutated transcript will be translated leading to the production of unwanted products.

A large group of Rnt1p dependent RNA transcripts encode RNA binding proteins, a few of which modulate transcription termination. Indeed, Rnt1p influences the termination of known modifiers of polyadenylation and termination such as Npl3p and Nab2p. The expression of NPL3 and NAB2 is autoregulated (Lund et al., 2008) by the binding of Npl3p and Nab2p to the transcription termination site of its own pre-mRNA. We have found that the overexpression of Npl3p does not simply alter local termination, as was previously suggested, but instead leads to transcriptional read-through that is normally inhibited co-transcriptionally by Rnt1p cleavage. Furthermore, chromatin immnuoprecipitation shows that Npl3p directly associates with its actively transcribed gene at its 3'end (data not shown) where Rnt1p was found to bind and therefore could directly influence termination. These results suggest that the cell has developed a regulatory mechanism that under normal conditions, where antiterminator concentration like Npl3 is adequate, endonucleolytic cleavage of Rnt1p 'terminator' will occur to control the endogenous level of the mRNA. However, under certain environmental cues where antiterminator concentration increases, the activity of the terminator is inhibited. The results in this thesis present an example of how a simple endoribonucleolytic cleavage could have a strong and varying impact on gene expression based on the timing, position and strength of the cleavage signals.

4. Towards an integrated mode of gene expression

Regulation of gene expression takes place at different steps of gene expression, i.e. transcription, RNA maturation and/ or translation. Factors involved in each step were believed to act independently of the other and often this regulation of gene expression is studied at one level or another. However, It is now becoming clear that each level of gene expression can affect the other. For example, factors affecting RNA degradation may be recruited very early to the transcriptional complex and their absence may modulate transcription efficiency. Indeed the exoribonuclease Rat1p is recruited very early to the transcription complex to cleave the transcription product and induce termination. Inactivation of Rat1p not only influences termination and RNA stability, but also influences the choice of polyadenylation site (Kim et al., 2004; West et al., 2004). Similarly, the non-coding RNA termination Nrd1 complex recruits ribonucleases to the actively transcribed genes to ensure either the proper maturation of the product or its degradation, if not properly assembled into a stable RNP (Vasiljeva and Buratowski, 2006). In this case, transcription, termination, processing, assembly and even degradation machineries are all recruited simultaneously to the site of RNA synthesis. Indeed, we have shown that Rnt1p acts co-transcriptionally as a terminator, processing complex, or as a trigger for RNA decay. Together these observations support the view of an integrated RNA synthesis where the different levels of RNA production are physically linked.

Interestingly, our study also suggests that the role of Rnt1p may even extend to the regulation of gene silencing. It has been believed that dsRNA dependent gene silencing is referred to the post-transcriptional RNA degradation or RNA interference (RNAi), that involved eukaryotic RNase III (Drosha and Dicer) through production of short non-coding RNA (Moazed et al., 2006). Data from a variety of organisms have shown that RNA interference can cause chromatin modifications leading to transcriptional gene silencing. In this case, non-coding RNAs assemble on silent chromatin. Recent data from fission yeast suggest that chromatin-dependent gene silencing is achieved, at least in part; through RNA turnover mechanisms that use components of the RNAi pathway as well as polyadenylation-dependent RNA decay (Buhler and Moazed, 2007; Buhler et al., 2008). In budding yeast, where the RNAi machinery is absent, yeast RNase III Rnt1p was found to terminate transcription of pre-rRNA co-transcriptionally (El Hage et al 2008; Kawauchi et al 2008), while affecting rRNA chromatin structure of gene arrays demonstrating a link among rRNA gene chromatin, transcription and processing (Catala et al 2007). This suggests that Rnt1p can influence silencing of rDNA genes. However, it is still unclear whether or not Rnt1p contributes to silencing of Pol II transcription. To our surprise, genome-wide profiling of Pol II association with DNA indicated that the deletion of *RNT1* results in net increase in the transcription of about 200 genes. This increase in transcription in some cases associates with changes in Rnt1p dependent transcription termination, while in other cases it is independent. These results

suggest that Rnt1p may antagonize transcription activity of protein coding genes. However, The mechanism by which Rnt1p affects transcription remains unclear. It will be very interesting in the future to determine whether these changes in gene expression requires Rnt1p cleavage activity, modulated by binding of Rnt1p to the transcription complex or mediated through changes in the chromatin conformation as in the case of the rDNA repeats.

ACKNOWLEDGMENTS

I would like to thank Drs Nicholas Proudfoot, François Bachand and Raymund Wellinger for accepting to evaluate my thesis and for their critical and helpful comments.

I am greatly indebted to Sherif for his excellent supervision, constant guidance, and for believing in me. I want to extend my appreciation for his steady support, for his great understanding, and for the lessons he gave me on the importance of details to approach a research problem.

I would like to thank all the lab members for interesting and critical discussions and for their help. I would like to acknowledge Bruno Lamontagne for helping me at the beginning of my studies.

Special thanks to Dany Namroud for being an honest and best friend, who always listens to me, and helps me at any time to solve any unsolvable problem.

My grateful thanks to my parents for their unconditional encouragement. In few words I can not express how I am deeply grateful to my mother, for her patience, for her confidence in me, and for teaching me how to be persistent to accomplish any goal. I dedicate this work to her for her unwavering support to me all these years.

REFERENCES

Abou Elela, S., and Ares, M., Jr. (1998). Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA. EMBO Journal *17*, 3738-3746.

Abou Elela, S., Igel, H., and Ares, M., Jr. (1996). RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. Cell *85*, 115-124.

Akada, R., Kitagawa, T., Kaneko, S., Toyonaga, D., Ito, S., Kakihara, Y., Hoshida, H., Morimura, S., Kondo, A., and Kida, K. (2006). PCR-mediated seamless gene deletion and marker recycling in Saccharomyces cerevisiae. Yeast *23*, 399-405.

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J *18*, 5399-5410.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403-410.

Andrade, J.M., Pobre, V., Silva, I.J., Domingues, S., and Arraiano, C.M. (2009). The role of 3'-5' exoribonucleases in RNA degradation. Prog Mol Biol Transl Sci *85*, 187-229.

Apirion, D. (1983). RNA processing in a unicellular microorganism: implications for eukaryotic cells. Prog Nucleic Acid Res Mol Biol *30*, 1-40.

Apirion, D., and Gegenheimer, P. (1981). Processing of bacterial RNA. FEBS Lett *125*, 1-9.

Apirion, D., and Miczak, A. (1993). RNA processing in prokaryotic cells. Bioessays *15*, 113-120.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. EMBO J *20*, 4684-4693.

Arigo, J.T., Eyler, D.E., Carroll, K.L., and Corden, J.L. (2006). Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol Cell *23*, 841-851.

Bachellerie, J.P., and Cavaille, J. (1997). Guiding ribose methylation of rRNA. Trends Biochem Sci *22*, 257-261.

Ballarino, M., Morlando, M., Pagano, F., Fatica, A., and Bozzoni, I. (2005). The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in Saccharomyces cerevisiae. Mol Cell Biol *25*, 5396-5403.

Balleza, E., Lopez-Bojorquez, L.N., Martinez-Antonio, A., Resendis-Antonio, O., Lozada-Chavez, I., Balderas-Martinez, Y.I., Encarnacion, S., and Collado-Vides, J. (2009). Regulation by transcription factors in bacteria: beyond description. FEMS Microbiol Rev *33*, 133-151.

Baralle, M., Pastor, T., Bussani, E., and Pagani, F. (2008). Influence of Friedreich ataxia GAA noncoding repeat expansions on pre-mRNA processing. Am J Hum Genet *83*, 77-88.

Bentley, D. (2002). The mRNA assembly line: transcription and processing machines in the same factory. Curr Opin Cell Biol *14*, 336-342.

Blancafort, P., Ferbeyre, G., Sariol, C., and Cedergren, R. (1997). PolI-driven integrative expression vectors for yeast. J Biotechnol *56*, 41-47.

Bloch, K.D. (1999). Regulation of endothelial NO synthase mRNA stability: RNA-binding proteins crowd on the 3'-untranslated region. Circ Res *85*, 653-655.

Bortolin, M.L., and Kiss, T. (1998). Human U19 intron-encoded snoRNA is processed from a long primary transcript that possesses little potential for protein coding. RNA *4*, 445-454.

Boue, S., Letunic, I., and Bork, P. (2003). Alternative splicing and evolution. Bioessays *25*, 1031-1034.

Bram, R.J., Young, R.A., and Steitz, J.A. (1980). The ribonuclease III site flanking 23S sequences in the 30S ribosomal precursor RNA of E. coli. Cell *19*, 393-401.

Brodersen, P., and Voinnet, O. (2009). Revisiting the principles of microRNA target recognition and mode of action. Nat Rev Mol Cell Biol *10*, 141-148.

Brouwer, R., Pruijn, G.J., and van Venrooij, W.J. (2001). The human exosome: an autoantigenic complex of exoribonucleases in myositis and scleroderma. Arthritis Res *3*, 102-106.

Brown, J.T., Yang, X., and Johnson, A.W. (2000). Inhibition of mRNA turnover in yeast by an xrn1 mutation enhances the requirement for eIF4E binding to eIF4G and for proper capping of transcripts by Ceg1p. Genetics *155*, 31-42.

Buhler, M., and Moazed, D. (2007). Transcription and RNAi in heterochromatic gene silencing. Nat Struct Mol Biol *14*, 1041-1048.

Buhler, M., Spies, N., Bartel, D.P., and Moazed, D. (2008). TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the Schizosaccharomyces pombe siRNA pathway. Nat Struct Mol Biol *15*, 1015-1023.

Burkard, K.T., and Butler, J.S. (2000). A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. Mol Cell Biol *20*, 604-616.

Cairo, G., Recalcati, S., Pietrangelo, A., and Minotti, G. (2002). The iron regulatory proteins: targets and modulators of free radical reactions and oxidative damage. Free Radic Biol Med *32*, 1237-1243.

Carmell, M.A., and Hannon, G.J. (2004). RNase III enzymes and the initiation of gene silencing. Nat Struct Mol Biol *11*, 214-218.

Carroll, K.L., Pradhan, D.A., Granek, J.A., Clarke, N.D., and Corden, J.L. (2004). Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol Cell Biol *24*, 6241-6252.

Catala, M., Lamontagne, B., Larose, S., Ghazal, G., and Abou Elela, S. (2004). Cell cycle-dependent nuclear localization of yeast RNase III is required for efficient cell division. Mol Biol Cell *15*, 3015-3030.

Catala, M., Tremblay, M., Samson, E., Conconi, A., and Abou Elela, S. (2008). Deletion of Rnt1p alters the proportion of open versus closed rRNA gene repeats in yeast. Mol Cell Biol *28*, 619-629.

Chambon, P. (1975). Eukaryotic nuclear RNA polymerases. Annu Rev Biochem *44*, 613-638.

Chanfreau, G., Abou Elela, S., Ares, M., Jr., and Guthrie, C. (1997). Alternative 3'-end processing of U5 snRNA by RNase III. Genes Dev *11*, 2741-2751.

Chanfreau, G., Buckle, M., and Jacquier, A. (2000). Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. Proc Natl Acad Sci U S A *97*, 3142-3147.

Chanfreau, G., Legrain, P., and Jacquier, A. (1998a). Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. J Mol Biol *284*, 975-988.

Chanfreau, G., Rotondo, G., Legrain, P., and Jacquier, A. (1998b). Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. Embo J *17*, 3726-3737.

Chanfreau, G.F. (2005). CUTting genetic noise by polyadenylation-induced RNA degradation. Trends Cell Biol *15*, 635-637.

Chapon, C., Cech, T.R., and Zaug, A.J. (1997). Polyadenylation of telomerase RNA in budding yeast. RNA *3*, 1337-1351.

Chinnusamy, V., Gong, Z., and Zhu, J.K. (2008). Nuclear RNA export and its importance in abiotic stress responses of plants. Curr Top Microbiol Immunol *326*, 235-255.

Conrad, C., and Rauhut, R. (2002). Ribonuclease III: new sense from nuisance. Int J Biochem Cell Biol *34*, 116-129.

Constable, A., Quick, S., Gray, N.K., and Hentze, M.W. (1992). Modulation of the RNA-binding activity of a regulatory protein by iron in vitro: switching between enzymatic and genetic function? Proc Natl Acad Sci U S A *89*, 4554-4558.

Conway, L., and Wickens, M. (1989). Modification interference analysis of reactions using RNA substrates. Methods Enzymol *180*, 369-379.

Cui, M., Allen, M.A., Larsen, A., Macmorris, M., Han, M., and Blumenthal, T. (2008). Genes involved in pre-mRNA 3'-end formation and transcription termination revealed by a lin-15 operon Muv suppressor screen. Proc Natl Acad Sci U S A *105*, 16665-16670.

Cusick, M.E. (1994). Purification and identification of two major single-stranded binding proteins of yeast Saccharomyces cerevisiae as ribosomal protein L4 and histone H2B. Biochim Biophys Acta *1217*, 31-40.

Danin-Kreiselman, M., Lee, C.Y., and Chanfreau, G. (2003). RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. Mol Cell *11*, 1279-1289.

Davila Lopez, M., and Samuelsson, T. (2008). Early evolution of histone mRNA 3' end processing. RNA *14*, 1-10.

Decker, C.J. (1998). The exosome: a versatile RNA processing machine. Curr Biol *8*, R238-240.

Deutscher, M.P. (1993). Promiscuous exoribonucleases of Escherichia coli. J Bacteriol *175*, 4577-4583.

Deutscher, M.P., and Li, Z. (2001). Exoribonucleases and their multiple roles in RNA metabolism. Prog Nucleic Acid Res Mol Biol *66*, 67-105.

Douthwaite, S., Christensen, A., and Garrett, R.A. (1983). Higher order structure in the 3'-minor domain of small subunit ribosomal RNAs from a gram negative bacterium, a gram positive bacterium and a eukaryote. J Mol Biol *169*, 249-279.

Dye, M.J., and Proudfoot, N.J. (1999). Terminal exon definition occurs cotranscriptionally and promotes termination of RNA polymerase II. Mol Cell *3*, 371-378.

Egloff, S., O'Reilly, D., and Murphy, S. (2008). Expression of human snRNA genes from beginning to end. Biochem Soc Trans *36*, 590-594.

Eichler, D.C., and Craig, N. (1994). Processing of eukaryotic ribosomal RNA. Prog Nucleic Acid Res Mol Biol *49*, 197-239.

El Hage, A., Koper, M., Kufel, J., and Tollervey, D. (2008). Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. Genes Dev *22*, 1069-1081.

Epshtein, V., Cardinale, C.J., Ruckenstein, A.E., Borukhov, S., and Nudler, E. (2007). An allosteric path to transcription termination. Mol Cell *28*, 991-1001.

Erie, D.A. (2002). The many conformational states of RNA polymerase elongation complexes and their roles in the regulation of transcription. Biochim Biophys Acta *1577*, 224-239.

Evans, D., and Blumenthal, T. (2000). trans splicing of polycistronic Caenorhabditis elegans pre-mRNAs: analysis of the SL2 RNA. Mol Cell Biol *20*, 6659-6667.

Faber, A.W., Van Dijk, M., Raue, H.A., and Vos, J.C. (2002). Ngl2p is a Ccr4p-like RNA nuclease essential for the final step in 3'-end processing of 5.8S rRNA in Saccharomyces cerevisiae. RNA *8*, 1095-1101.

Fedorov, A., Stombaugh, J., Harr, M.W., Yu, S., Nasalean, L., and Shepelev, V. (2005). Computer identification of snoRNA genes using a Mammalian Orthologous Intron Database. Nucleic Acids Res *33*, 4578-4583.

Filipowicz, W. (2005). RNAi: the nuts and bolts of the RISC machine. Cell *122*, 17-20.

Filipowicz, W., Pelczar, P., Pogacic, V., and Dragon, F. (1999). Structure and biogenesis of small nucleolar RNAs acting as guides for ribosomal RNA modification. Acta Biochim Pol *46*, 377-389.

Finkbeiner, S. (2001). New roles for introns: sites of combinatorial regulation of Ca2+- and cyclic AMP-dependent gene transcription. Sci STKE *2001*, PE1.

Fischer, D., Weisenberger, D., and Scheer, U. (1991). Assigning functions to nucleolar structures. Chromosoma *101*, 133-140.

Forbes, D.J., Kornberg, T.B., and Kirschner, M.W. (1983). Small nuclear RNA transcription and ribonucleoprotein assembly in early Xenopus development. J Cell Biol *97*, 62-72.

Fromont-Racine, M., Senger, B., Saveanu, C., and Fasiolo, F. (2003). Ribosome assembly in eukaryotes. Gene *313*, 17-42.

Furuichi, Y., and Shatkin, A.J. (2000). Viral and cellular mRNA capping: past and prospects. Adv Virus Res *55*, 135-184.

Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2005). Intermediate states of ribonuclease III in complex with double-stranded RNA. Structure *13*, 1435-1442.

Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2006). Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. Cell *124*, 355-366.

Garrett, R.A., Christensen, A., and Douthwaite, S. (1984). Higher-order structure in the 3'-terminal domain VI of the 23 S ribosomal RNAs from Escherichia coli and Bacillus stearothermophilus. J Mol Biol *179*, 689-712.

Gaudin, C., Ghazal, G., Yoshizawa, S., Abou Elela, S., and Fourmy, D. (2006). Structure of an AAGU tetraloop and it implication in Rnt1p RNA processing. J Mol Biol.

Ge, D., Lamontagne, B., and Abou Elela, S. (2005). RNase III-Mediated Silencing of a Glucose-Dependent Repressor in Yeast. Curr Biol *15*, 140-145.

Gelade, R., Van de Velde, S., Van Dijck, P., and Thevelein, J.M. (2003). Multi-level response of the yeast genome to glucose. Genome Biol *4*, 233.

Ghazal, G., and Elela, S.A. (2006). Characterization of the reactivity determinants of a novel hairpin substrate of yeast RNase III. J Mol Biol *363*, 332-344.

Ghazal, G., Gagnon, J., Jacques, P.E., Landry, J.R., Robert, F., and Elela, S.A. (2009). Yeast RNase III triggers polyadenylation-independent transcription termination. Mol Cell *36*, 99-109.

Ghazal, G., Ge, D., Gervais-Bird, J., Gagnon, J., and Abou Elela, S. (2005). Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. Mol Cell Biol *25*, 2981-2994.

Giorgi, C., Fatica, A., Nagel, R., and Bozzoni, I. (2001). Release of U18 snoRNA from its host intron requires interaction of Nop1p with the Rnt1p endonuclease. Embo J *20*, 6856-6865.

Gonzales, F.A., Zanchin, N.I., Luz, J.S., and Oliveira, C.C. (2005). Characterization of Saccharomyces cerevisiae Nop17p, a novel Nop58p-interacting protein that is involved in Pre-rRNA processing. J Mol Biol *346*, 437-455.

Good, L., Intine, R.V., and Nazar, R.N. (1997). Interdependence in the processing of ribosomal RNAs in Schizosaccharomyces pombe. J Mol Biol *273*, 782-788.

Gornemann, J., Kotovic, K.M., Hujer, K., and Neugebauer, K.M. (2005). Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. Mol Cell *19*, 53-63.

Graham, A.C., Kiss, D.L., and Andrulis, E.D. (2009). Core exosome-independent roles for Rrp6 in cell cycle progression. Mol Biol Cell *20*, 2242-2253.

Granneman, S., and Baserga, S.J. (2004). Ribosome biogenesis: of knobs and RNA processing. Exp Cell Res *296*, 43-50.

Granneman, S., and Baserga, S.J. (2005). Crosstalk in gene expression: coupling and co-regulation of rDNA transcription, pre-ribosome assembly and pre-rRNA processing. Curr Opin Cell Biol *17*, 281-286.

Grate, L., and Ares, M., Jr. (2002). Searching yeast intron data at Ares lab Web site. Methods Enzymol *350*, 380-392.

Grec, S., Wang, Y., Le Guen, L., Negrouk, V., and Boutry, M. (2000). Cryptic polyadenylation sites within the coding sequence of three yeast genes expressed in tobacco. Gene *242*, 87-95.

Grzechnik, P., and Kufel, J. (2008). Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. Mol Cell *32*, 247-258.

Gu, M., and Lima, C.D. (2005). Processing the message: structural insights into capping and decapping mRNA. Curr Opin Struct Biol *15*, 99-106.

Gudipati, R.K., Villa, T., Boulay, J., and Libri, D. (2008). Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. Nat Struct Mol Biol *15*, 786-794.

Guffanti, E., Ferrari, R., Preti, M., Forloni, M., Harismendy, O., Lefebvre, O., and Dieci, G. (2006). A minimal promoter for TFIIIC-dependent in vitro transcription of snoRNA and tRNA genes by RNA polymerase III. J Biol Chem *281*, 23945-23957.

Guthrie, C. (1991). Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. Science *253*, 157-163.

Guthrie, C., and Fink, G.R. (1991). Guide to Yeast Genetics and Molecular Biology (San Diego, CA: Academic Press).

Haile, S., and Papadopoulou, B. (2007). Developmental regulation of gene expression in trypanosomatid parasitic protozoa. Curr Opin Microbiol *10*, 569-577.

Han, H., and Dervan, P.B. (1994). Visualization of RNA tertiary structure by RNA-EDTA.Fe(II) autocleavage: analysis of tRNA(Phe) with uridine-EDTA.Fe(II) at position 47. Proc Natl Acad Sci U S A *91*, 4955-4959.

Harrison, P.R. (1990). Molecular mechanisms involved in the regulation of gene expression during cell differentiation and development. Immunol Ser *49*, 411-464.

Hengge-Aronis, R. (2002). Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. Microbiol Mol Biol Rev *66*, 373-395, table of contents.

Henry, Y., Wood, H., Morrissey, J.P., Petfalski, E., Kearsey, S., and Tollervey, D. (1994). The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. EMBO J *13*, 2452-2463.

Herbert, A., and Rich, A. (1999a). RNA processing and the evolution of eukaryotes. Nat Genet *21*, 265-269.

Herbert, A., and Rich, A. (1999b). RNA processing in evolution. The logic of soft-wired genomes. Ann N Y Acad Sci *870*, 119-132.

Herrero, E. (2005). Evolutionary relationships between Saccharomyces cerevisiae and other fungal species as determined from genome comparisons. Rev Iberoam Micol *22*, 217-222.

Hilleren, P., McCarthy, T., Rosbash, M., Parker, R., and Jensen, T.H. (2001). Quality control of mRNA 3'-end processing is linked to the nuclear exosome. Nature *413*, 538-542.

Hinnebusch, A.G. (1990). Transcriptional and translational regulation of gene expression in the general control of amino-acid biosynthesis in Saccharomyces cerevisiae. Prog Nucleic Acid Res Mol Biol *38*, 195-240.

Hirose, T., Shu, M.D., and Steitz, J.A. (2003). Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. Mol Cell *12*, 113-123.

Hirose, T., and Steitz, J.A. (2001). Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. Proc Natl Acad Sci U S A *98*, 12914-12919.

Hutvagner, G. (2005). Small RNA asymmetry in RNAi: function in RISC assembly and gene regulation. FEBS Lett *579*, 5850-5857.

Irimia, M., Rukov, J.L., Roy, S.W., Vinther, J., and Garcia-Fernandez, J. (2009). Quantitative regulation of alternative splicing in evolution and development. Bioessays *31*, 40-50.

Izawa, S., and Inoue, Y. (2009). Post-transcriptional regulation of gene expression in yeast under ethanol stress. Biotechnol Appl Biochem *53*, 93-99.

Jacobs, E.Y., Ogiwara, I., and Weiner, A.M. (2004). Role of the C-terminal domain of RNA polymerase II in U2 snRNA transcription and 3' processing. Mol Cell Biol *24*, 846-855.

Ji, X. (2008). The mechanism of RNase III action: how dicer dices. Curr Top Microbiol Immunol *320*, 99-116.

Kaplan, C.D., Holland, M.J., and Winston, F. (2005). Interaction between transcription elongation factors and mRNA 3'-end formation at the Saccharomyces cerevisiae GAL10-GAL7 locus. J Biol Chem *280*, 913-922.

Karaoglu, D., and Thurlow, D.L. (1991). A chemical interference study on the interaction of ribosomal protein L11 from Escherichia coli with RNA molecules containing its binding site from 23S rRNA. Nucleic Acids Res *19*, 5293-5300.

Kashima, T., Rao, N., and Manley, J.L. (2007). An intronic element contributes to splicing repression in spinal muscular atrophy. Proc Natl Acad Sci U S A *104*, 3426-3431.

Katze, M.G., and Agy, M.B. (1990). Regulation of viral and cellular RNA turnover in cells infected by eukaryotic viruses including HIV-1. Enzyme *44*, 332-346.

Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., and Proudfoot, N.J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. Genes Dev *22*, 1082-1092.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E.S. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature *423*, 241-254.

Kennell, D. (2002). Processing endoribonucleases and mRNA degradation in bacteria. J Bacteriol *184*, 4645-4657; discussion 4665.

Kent, T., Lapik, Y.R., and Pestov, D.G. (2009). The 5' external transcribed spacer in mouse ribosomal RNA contains two cleavage sites. RNA *15*, 14-20.

Khalid, M.F., Damha, M.J., Shuman, S., and Schwer, B. (2005). Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. Nucleic Acids Res *33*, 6349-6360.

Kim, E., Goren, A., and Ast, G. (2008). Alternative splicing: current perspectives. Bioessays *30*, 38-47.

Kim, K., Lee, Y.S., Harris, D., Nakahara, K., and Carthew, R.W. (2006a). The RNAi pathway initiated by Dicer-2 in Drosophila. Cold Spring Harb Symp Quant Biol *71*, 39-44.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517-522.

Kim, M., Suh, H., Cho, E.J., and Buratowski, S. (2009). Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7. J Biol Chem *284*, 26421-26426.

Kim, M., Vasiljeva, L., Rando, O.J., Zhelkovsky, A., Moore, C., and Buratowski, S. (2006b). Distinct pathways for snoRNA and mRNA termination. Mol Cell *24*, 723-734.

Kim, Y.K., and Kim, V.N. (2007). Processing of intronic microRNAs. EMBO J *26*, 775-783.

Kiss-Laszlo, Z., Henry, Y., Bachellerie, J.P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell *85*, 1077-1088.

Kiss-Laszlo, Z., Henry, Y., and Kiss, T. (1998). Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA. Embo J *17*, 797-807.

Kiss, T., Fayet, E., Jady, B.E., Richard, P., and Weber, M. (2006). Biogenesis and intranuclear trafficking of human box C/D and H/ACA RNPs. Cold Spring Harb Symp Quant Biol *71*, 407-417.

Klaenhammer, T.R., Azcarate-Peril, M.A., Altermann, E., and Barrangou, R. (2007). Influence of the dairy environment on gene expression and substrate utilization in lactic acid bacteria. J Nutr *137*, 748S-750S.

Krebber, H., Taura, T., Lee, M.S., and Silver, P.A. (1999). Uncoupling of the hnRNP Npl3p from mRNAs during the stress-induced block in mRNA export. Genes Dev *13*, 1994-2004.

Kriventseva, E.V., Koch, I., Apweiler, R., Vingron, M., Bork, P., Gelfand, M.S., and Sunyaev, S. (2003). Increase of functional diversity by alternative splicing. Trends Genet *19*, 124-128.

Krol, A., Ebel, J.P., Rinke, J., and Luhrmann, R. (1983). U1, U2 and U5 small nuclear RNAs are found in plants cells. Complete nucleotide sequence of the U5 RNA family from pea nuclei. Nucleic Acids Res *11*, 8583-8594.

Krol, A., Gallinaro, H., Lazar, E., Jacob, M., and Branlant, C. (1981). The nuclear 5S RNAs from chicken, rat and man. U5 RNAs are encoded by multiple genes. Nucleic Acids Res *9*, 769-787.

Kuai, L., Das, B., and Sherman, F. (2005). A nuclear degradation pathway controls the abundance of normal mRNAs in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *102*, 13962-13967.

Kuehn, U., Guendel, M., Knoth, A., Kerwitz, Y., Ruedel, S., and Wahle, E. (2009). Poly(A) tail length is controlled by the nuclear poly(A) binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. J Biol Chem.

Kufel, J., Allmang, C., Chanfreau, G., Petfalski, E., Lafontaine, D.L., and Tollervey, D. (2000). Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. Mol Cell Biol *20*, 5415-5424.

Kufel, J., Dichtl, B., and Tollervey, D. (1999). Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3' ETS but not the 5' ETS. RNA *5*, 909-917.

Kuhn, C.D., Geiger, S.R., Baumli, S., Gartmann, M., Gerber, J., Jennebach, S., Mielke, T., Tschochner, H., Beckmann, R., and Cramer, P. (2007). Functional architecture of RNA polymerase I. Cell *131*, 1260-1272.

Lafontaine, D., and Tollervey, D. (1995). Trans-acting factors in yeast pre-rRNA and pre-snoRNA processing. Biochem Cell Biol *73*, 803-812.

Lafontaine, D.L.J., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M., and Tollervey, D. (1998). The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. Genes Dev *12*, 527-537.

Lamontagne, B., and Abou Elela, S. (2001). Purification and characterization of Saccharomyces cerevisiae Rnt1p nuclease, Vol 342 (San diego, CA: Academic Press).

Lamontagne, B., and Abou Elela, S. (2004). Evaluation of the RNA determinants for bacterial and yeast RNase III binding and cleavage. J Biol Chem *279*, 2231-2241.

Lamontagne, B., Ghazal, G., Lebars, I., Yoshizawa, S., Fourmy, D., and Abou Elela, S. (2003). Sequence dependence of substrate recognition and cleavage by yeast RNase III. J Mol Biol *327*, 985-1000.

Lamontagne, B., Hannoush, R.N., Damha, M.J., and Abou Elela, S. (2004). Molecular requirements for duplex recognition and cleavage by eukaryotic RNase III: discovery of an RNA-dependent DNA cleavage activity of yeast Rnt1p. J Mol Biol *338*, 401-418.

Lamontagne, B., Larose, S., Boulanger, J., and Abou Elela, S. (2001). The RNase III family: a conserved structure and expanding functions in eukaryotic dsRNA metabolism. Curr Issues Mol Biol *3*, 71-78.

Lamontagne, B., Tremblay, A., and Abou Elela, S. (2000). The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. Mol Cell Biol *20*, 1104-1115.

Larose, S., Laterreur, N., Ghazal, G., Gagnon, J., Wellinger, R.J., and Elela, S.A. (2007). RNase III-dependent regulation of yeast telomerase. J Biol Chem *282*, 4373-4381.

Lavoie, M., and Abou Elela, S. (2008). Yeast ribonuclease III uses a network of multiple hydrogen bonds for RNA binding and cleavage. Biochemistry *47*, 8514-8526.

Le Hir, H., Nott, A., and Moore, M.J. (2003). How introns influence and enhance eukaryotic gene expression. Trends Biochem Sci *28*, 215-220.

Lebars, I., Lamontagne, B., Yoshizawa, S., Aboul-Elela, S., and Fourmy, D. (2001). Solution structure of conserved AGNN tetraloops: insights into Rnt1p RNA processing. Embo J *20*, 7250-7258.

Lebreton, A., Tomecki, R., Dziembowski, A., and Seraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature *456*, 993-996.

Lee, A., Henras, A.K., and Chanfreau, G. (2005). Multiple RNA surveillance pathways limit aberrant expression of iron uptake mRNAs and prevent iron toxicity in S. cerevisiae. Mol Cell *19*, 39-51.

Lee, C.Y., Lee, A., and Chanfreau, G. (2003a). The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of S. cerevisiae box C/D snoRNAs. Rna *9*, 1362-1370.

Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. Annu Rev Genet *34*, 77-137.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. (2003b). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.

Lee, Y., Melekhovets, Y.F., and Nazar, R.N. (1995). Termination as a factor in "quality control" during ribosome biogenesis. J Biol Chem *270*, 28003-28005.

Lee, Y., and Nazar, R.N. (1997). Ribosomal 5 S rRNA maturation in Saccharomyces cerevisiae. J Biol Chem *272*, 15206-15212.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.

Lekas, P., Tin, K.L., Lee, C., and Prokipcak, R.D. (2000). The human cytochrome P450 1A1 mRNA is rapidly degraded in HepG2 cells. Arch Biochem Biophys *384*, 311-318.

Leulliot, N., Quevillon-Cheruel, S., Graille, M., Van Tilbeurgh, H., Leeper, T.C., Godin, K.S., Edwards, T.E., Sigurdsson, S.T., Rozenkrants, N., Nagel, R.J., *et al.* (2004). A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. Embo J *23*, 2468-2477.

Lewandowska, M.A., Stuani, C., Parvizpur, A., Baralle, F.E., and Pagani, F. (2005). Functional studies on the ATM intronic splicing processing element. Nucleic Acids Res *33*, 4007-4015.

Li, B., Vilardell, J., and Warner, J.R. (1996). An RNA structure involved in feedback regulation of splicing and of translation is critical for biological fitness. Proc Natl Acad Sci U S A *93*, 1596-1600.

Li, C.H., Irmer, H., Gudjonsdottir-Planck, D., Freese, S., Salm, H., Haile, S., Estevez, A.M., and Clayton, C. (2006). Roles of a Trypanosoma brucei 5'->3' exoribonuclease homolog in mRNA degradation. RNA *12*, 2171-2186.

Li, H.L., Chelladurai, B.S., Zhang, K., and Nicholson, A.W. (1993). Ribonuclease III cleavage of a bacteriophage T7 processing signal. Divalent cation specificity, and specific anion effects. Nucleic Acids Res *21*, 1919-1925.

Lim, Y., Lee, S.M., Kim, M., Lee, J.Y., Moon, E.P., Lee, B.J., and Kim, J. (2002). Complete genomic structure of human rpS3: identification of functional U15b snoRNA in the fifth intron. Gene *286*, 291-297.

Lin, S.L., Kim, H., and Ying, S.Y. (2008). Intron-mediated RNA interference and microRNA (miRNA). Front Biosci *13*, 2216-2230.

Lin, S.L., Miller, J.D., and Ying, S.Y. (2006). Intronic MicroRNA (miRNA). J Biomed Biotechnol *2006*, 26818.

Lindahl, L., and Zengel, J.M. (1995). RNase MRP and rRNA processing. Mol Biol Rep *22*, 69-73.

Liti, G., Barton, D.B., and Louis, E.J. (2006). Sequence diversity, reproductive isolation and species concepts in Saccharomyces. Genetics *174*, 839-850.

Loflin, P.T., Chen, C.Y., Xu, N., and Shyu, A.B. (1999). Transcriptional pulsing approaches for analysis of mRNA turnover in mammalian cells. Methods *17*, 11-20.

Lombardo, A., Cereghino, G.P., and Scheffler, I.E. (1992). Control of mRNA turnover as a mechanism of glucose repression in Saccharomyces cerevisiae. Mol Cell Biol *12*, 2941-2948.

Long, R.M., and McNally, M.T. (2003). mRNA decay: x (XRN1) marks the spot. Mol Cell *11*, 1126-1128.

Lowman, H.B., and Draper, D.E. (1986). On the recognition of helical RNA by cobra venom V1 nuclease. J Biol Chem *261*, 5396-5403.

Lund, M.K., Kress, T.L., and Guthrie, C. (2008). Autoregulation of Npl3, a yeast SR protein, requires a novel downstream region and serine phosphorylation. Mol Cell Biol *28*, 3873-3881.

Luo, W., and Bentley, D. (2004). A ribonucleolytic rat torpedoes RNA polymerase II. Cell *119*, 911-914.

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

Lykke-Andersen, J., Aagaard, C., Semionenkov, M., and Garrett, R.A. (1997). Archaeal introns: splicing, intercellular mobility and evolution. Trends Biochem Sci *22*, 326-331.

Lykke-Andersen, S., Brodersen, D.E., and Jensen, T.H. (2009). Origins and activities of the eukaryotic exosome. J Cell Sci *122*, 1487-1494.

MacBeth, K.J., and Patterson, J.L. (1998). Overview of the Leishmaniavirus endoribonuclease and functions of other endoribonucleases affecting viral gene expression. J Exp Zool *282*, 254-260.

Macke, T.J., Ecker, D.J., Gutell, R.R., Gautheret, D., Case, D.A., and Sampath, R. (2001). RNAMotif, an RNA secondary structure definition and search algorithm. Nucleic Acids Res *29*, 4724-4735.

Maitra, S., Chou, C.F., Luber, C.A., Lee, K.Y., Mann, M., and Chen, C.Y. (2008). The AU-rich element mRNA decay-promoting activity of BRF1 is regulated by mitogen-activated protein kinase-activated protein kinase 2. RNA *14*, 950-959.

Malagon, F., Kireeva, M.L., Shafer, B.K., Lubkowska, L., Kashlev, M., and Strathern, J.N. (2006). Mutations in the Saccharomyces cerevisiae RPB1 gene conferring hypersensitivity to 6-azauracil. Genetics *172*, 2201-2209.

Mandart, E., and Parker, R. (1995). Effects of mutations in the Saccharomyces cerevisiae RNA14, RNA15, and PAP1 genes on polyadenylation in vivo. Mol Cell Biol *15*, 6979-6986.

Marchfelder, A., Clayton, D.A., and Brennicke, A. (1998). The gene for ribosomal protein L7a-1 in Schizosaccharomyces pombe contains an intron after the initiation codon. Biochim Biophys Acta *1397*, 146-150.

Marles-Wright, J., and Lewis, R.J. (2007). Stress responses of bacteria. Curr Opin Struct Biol *17*, 755-760.

Marz, M., Kirsten, T., and Stadler, P.F. (2008). Evolution of spliceosomal snRNA genes in metazoan animals. J Mol Evol *67*, 594-607.

Mateos-Langerak, J., Goetze, S., Leonhardt, H., Cremer, T., van Driel, R., and Lanctot, C. (2007). Nuclear architecture: Is it important for genome function and can we prove it? J Cell Biochem *102*, 1067-1075.

Mattaj, I.W., and Hamm, J. (1989). Regulated splicing in early development and stage-specific U snRNPs. Development *105*, 183-189.

Maxwell, E.S., and Fournier, M.J. (1995). The small nucleolar RNAs. Annu Rev Biochem *64*, 897-934.

McCarthy, J.E. (1990). Post-transcriptional control in the polycistronic operon environment: studies of the atp operon of Escherichia coli. Mol Microbiol *4*, 1233-1240.

McCaskill, J.S. (1990). The equilibrium partition function and base pair binding probabilities for RNA secondary structure. Biopolymers *29*, 1105-1119.

Meyer, P.A., Ye, P., Suh, M.H., Zhang, M., and Fu, J. (2009). Structure of the 12-subunit RNA polymerase II refined with the aid of anomalous diffraction data. J Biol Chem *284*, 12933-12939.

Milligan, L., Torchet, C., Allmang, C., Shipman, T., and Tollervey, D. (2005). A nuclear surveillance pathway for mRNAs with defective polyadenylation. Mol Cell Biol *25*, 9996-10004.

Mitchell, P., Petfalski, E., Houalla, R., Podtelejnikov, A., Mann, M., and Tollervey, D. (2003). Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs. Mol Cell Biol *23*, 6982-6992.

Mitra, S., and Bechhofer, D.H. (1994). Substrate specificity of an RNase III-like activity from Bacillus subtilis. J Biol Chem *269*, 31450-31456.

Mizuta, K., Hashimoto, T., and Otaka, E. (1995). The evolutionary relationships between homologs of ribosomal YL8 protein and YL8-like proteins. Curr Genet *28*, 19-25.

Moazed, D., Buhler, M., Buker, S.M., Colmenares, S.U., Gerace, E.L., Gerber, S.A., Hong, E.J., Motamedi, M.R., Verdel, A., Villen, J., and Gygi, S.P. (2006). Studies on

the mechanism of RNAi-dependent heterochromatin assembly. Cold Spring Harb Symp Quant Biol *71*, 461-471.

Morlando, M., Ballarino, M., Greco, P., Caffarelli, E., Dichtl, B., and Bozzoni, I. (2004). Coupling between snoRNP assembly and 3' processing controls box C/D snoRNA biosynthesis in yeast. EMBO J *23*, 2392-2401.

Mougin, A., Torterotot, F., Branlant, C., Jacobson, M.R., Huang, Q., and Pederson, T. (2002). A 3'-terminal minihelix in the precursor of human spliceosomal U2 small nuclear RNA. J Biol Chem *277*, 23137-23142.

Nagel, R., and Ares, M., Jr. (2000). Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5'-AGNN-3' tetraloop. Rna *6*, 1142-1156.

Naryshkin, N.A., Gait, M.J., and Ivanovskaya, M.G. (1998). RNA recognition and regulation of HIV-1 gene expression by viral factor Tat. Biochemistry (Mosc) *63*, 489-503.

Nazar, R.N. (1991). Higher order structure of the ribosomal 5 S RNA. J Biol Chem *266*, 4562-4567.

Nazar, R.N. (2004). Ribosomal RNA processing and ribosome biogenesis in eukaryotes. IUBMB Life *56*, 457-465.

Nazar, R.N., and Wildeman, A.G. (1981). Altered features in the secondary structure of Vicia faba 5.8s rRNA. Nucleic Acids Res *9*, 5345-5358.

Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L.M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature *457*, 1038-1042.

Neu-Yilik, G., and Kulozik, A.E. (2008). NMD: multitasking between mRNA surveillance and modulation of gene expression. Adv Genet *62*, 185-243.

Ni, J., Tien, A.L., and Fournier, M.J. (1997). Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA. Cell *89*, 565-573.

Nicholson, A.W. (1996). Structure, reactivity, and biology of double-stranded RNA. Prog Nucleic Acid Res Mol Biol *52*, 1-65.

Nicholson, A.W. (1999). Function, mechanism and regulation of bacterial ribonucleases. FEMS Microbiol Rev *23*, 371-390.

Noble, S.M., and Johnson, A.D. (2005). Strains and strategies for large-scale gene deletion studies of the diploid human fungal pathogen Candida albicans. Eukaryot Cell *4*, 298-309.

Noe, V., Ciudad, C.J., and Chasin, L.A. (1999). Effect of differential polyadenylation and cell growth phase on dihydrofolate reductase mRNA stability. J Biol Chem *274*, 27807-27814.

Nolan, T., and Cogoni, C. (2004). The long hand of the small RNAs reaches into several levels of gene regulation. Biochem Cell Biol *82*, 472-481.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol *302*, 205-217.

Ooi, S.L., Samarsky, D.A., Fournier, M.J., and Boeke, J.D. (1998a). Intronic snoRNA biosynthesis in Saccharomyces cerevisiae depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. RNA *4*, 1096-1110.

Ooi, S.L., Samarsky, D.A., Fournier, M.J., and Boeke, J.D. (1998b). Intronic snoRNA biosynthesis in Saccharomyces cerevisiae depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA [In Process Citation]. Rna *4*, 1096-1110.

Osheim, Y.N., Proudfoot, N.J., and Beyer, A.L. (1999). EM visualization of transcription by RNA polymerase II: downstream termination requires a poly(A) signal but not transcript cleavage. Mol Cell *3*, 379-387.

Palmer, C.D., Mutch, B.E., Page, T.H., Horwood, N.J., and Foxwell, B.M. (2008). Bmx regulates LPS-induced IL-6 and VEGF production via mRNA stability in rheumatoid synovial fibroblasts. Biochem Biophys Res Commun *370*, 599-602.

Pandey, N.B., and Marzluff, W.F. (1987). The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. Mol Cell Biol *7*, 4557-4559.

Parenteau, J., Durand, M., Veronneau, S., Lacombe, A.A., Morin, G., Guerin, V., Cecez, B., Gervais-Bird, J., Koh, C.S., Brunelle, D., *et al.* (2008). Deletion of many yeast introns reveals a minority of genes that require splicing for function. Mol Biol Cell *19*, 1932-1941.

Park, J.W., and Graveley, B.R. (2007). Complex alternative splicing. Adv Exp Med Biol *623*, 50-63.

Passos, D.O., Doma, M.K., Shoemaker, C.J., Muhlrad, D., Green, R., Weissman, J., Hollien, J., and Parker, R. (2009). Analysis of Dom34 and its function in no-go decay. Mol Biol Cell *20*, 3025-3032.

Peattie, D.A. (1979). Direct chemical method for sequencing RNA. Proc Natl Acad Sci U S A *76*, 1760-1764.

Pederson, T. (1998). The plurifunctional nucleolus. Nucleic Acids Res *26*, 3871-3876.

Peng, W.T., Robinson, M.D., Mnaimneh, S., Krogan, N.J., Cagney, G., Morris, Q., Davierwala, A.P., Grigull, J., Yang, X., Zhang, W., *et al.* (2003). A panoramic view of yeast noncoding RNA processing. Cell *113*, 919-933.

Perbal, B. (1988). A Practical Guide to Molecular Cloning, Second edn (New York: John Wiley and Sons, Inc.).

Perry, R.P. (1976). Processing of RNA. Annu Rev Biochem *45*, 605-629.

Petfalski, E., Dandekar, T., Henry, Y., and Tollervey, D. (1998). Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. Mol Cell Biol *18*, 1181-1189.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science *309*, 1573-1576.

Pillai, R.S., Bhattacharyya, S.N., and Filipowicz, W. (2007). Repression of protein synthesis by miRNAs: how many mechanisms? Trends Cell Biol *17*, 118-126.

Planta, R.J., and Raue, H.A. (1988). Control of ribosome biogenesis in yeast. Trends Genet *4*, 64-68.

Poole, T.L., and Stevens, A. (1995). Comparison of features of the RNase activity of 5'-exonuclease-1 and 5'-exonuclease-2 of Saccharomyces cerevisiae. Nucleic Acids Symp Ser, 79-81.

Preker, P.J., and Guthrie, C. (2006). Autoregulation of the mRNA export factor Yra1p requires inefficient splicing of its pre-mRNA. RNA *12*, 994-1006.

Prescott, E.M., Osheim, Y.N., Jones, H.S., Alen, C.M., Roan, J.G., Reeder, R.H., Beyer, A.L., and Proudfoot, N.J. (2004). Transcriptional termination by RNA polymerase I requires the small subunit Rpa12p. Proc Natl Acad Sci U S A *101*, 6068-6073.

Preti, M., Guffanti, E., Valitutto, E., and Dieci, G. (2006). Assembly into snoRNP controls 5'-end maturation of a box C/D snoRNA in Saccharomyces cerevisiae. Biochem Biophys Res Commun *351*, 468-473.

Proudfoot, N. (2004). New perspectives on connecting messenger RNA 3' end formation to transcription. Curr Opin Cell Biol *16*, 272-278.

Qu, L.H., Henras, A., Lu, Y.J., Zhou, H., Zhou, W.X., Zhu, Y.Q., Zhao, J., Henry, Y., Caizergues-Ferrer, M., and Bachellerie, J.P. (1999). Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. Mol Cell Biol *19*, 1144-1158.

Raijmakers, R., Schilders, G., and Pruijn, G.J. (2004). The exosome, a molecular machine for controlled RNA degradation in both nucleus and cytoplasm. Eur J Cell Biol *83*, 175-183.

Reddy, R., and Busch, H. (1983). Small nuclear RNAs and RNA processing. Prog Nucleic Acid Res Mol Biol *30*, 127-162.

Reddy, R., and Shimba, S. (1995). Structural and functional similarities between MRP and RNase P. Mol Biol Rep *22*, 81-85.

Reeder, R.H. (1990). rRNA synthesis in the nucleolus. Trends Genet *6*, 390-395.

Reilly, T.H., and Schmitt, M.E. (1995). The yeast, Saccharomyces cerevisiae, RNase P/MRP ribonucleoprotein endoribonuclease family. Mol Biol Rep *22*, 87-93.

Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. Genes Dev *23*, 1247-1269.

Rio, D.C. (1993). Splicing of pre-mRNA: mechanism, regulation and role in development. Curr Opin Genet Dev *3*, 574-584.

Robertson, H.D. (1967). A nuclease specific for double-stranded RNA. Virology *12*, 718.

Robertson, H.D., Webster, R.E., and Zinder, N.D. (1968). Purification and properties of ribonuclease III from Escherichia coli. J Biol Chem *243*, 82-91.

Rocha, E.P. (2008). The organization of the bacterial genome. Annu Rev Genet *42*, 211-233.

Roeder, R.G., and Rutter, W.J. (1970). Specific nucleolar and nucleoplasmic RNA polymerases. Proc Natl Acad Sci U S A *65*, 675-682.

Roth, K.M., Wolf, M.K., Rossi, M., and Butler, J.S. (2005). The nuclear exosome contributes to autogenous control of NAB2 mRNA levels. Mol Cell Biol *25*, 1577-1585.

Rothenberger, S., Mullner, E.W., and Kuhn, L.C. (1990). The mRNA-binding protein which controls ferritin and transferrin receptor expression is conserved during evolution. Nucleic Acids Res *18*, 1175-1179.

Rotondo, G., Huang, J.Y., and Frendewey, D. (1997). Substrate structure requirements of the Pac1 ribonuclease from Schizosaccharmyces pombe. RNA *3*, 1182-1193.

Rudinger, J., Hillenbrandt, R., Sprinzl, M., and Giege, R. (1996). Antideterminants present in minihelix(Sec) hinder its recognition by prokaryotic elongation factor Tu. Embo J *15*, 650-657.

Rufiange, A., Jacques, P.E., Bhat, W., Robert, F., and Nourani, A. (2007). Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. Mol Cell *27*, 393-405.

Russell, I., and Tollervey, D. (1995). Yeast Nop3p has structural and functional similarities to mammalian pre- mRNA binding proteins. Eur J Cell Biol *66*, 293-301.

Russell, I.D., and Tollervey, D. (1992). NOP3 is an essential yeast protein which is required for pre-rRNA processing. J Cell Biol *119*, 737-747.

Saida, F., and Odaert, B. (2007). RNA recognition and cleavage by sequence-specific endoribonucleases. Protein Pept Lett *14*, 103-111.

Samarsky, D.A., and Fournier, M.J. (1999). A comprehensive database for the small nucleolar RNAs from Saccharomyces cerevisiae. Nucleic Acids Res *27*, 161-164.

Samarsky, D.A., Fournier, M.J., Singer, R.H., and Bertrand, E. (1998). The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. Embo J *17*, 3747-3757.

Santiago, T.C., Bettany, A.J., Purvis, I.J., and Brown, A.J. (1987). Messenger RNA stability in Saccharomyces cerevisiae: the influence of translation and poly(A) tail length. Nucleic Acids Res *15*, 2417-2429.

Schaeffer, D., Tsanova, B., Barbas, A., Reis, F.P., Dastidar, E.G., Sanchez-Rotunno, M., Arraiano, C.M., and van Hoof, A. (2009). The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat Struct Mol Biol *16*, 56-62.

Schneider, C., Leung, E., Brown, J., and Tollervey, D. (2009). The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res *37*, 1127-1140.

Schultz, S.J., and Champoux, J.J. (2008). RNase H activity: structure, specificity, and function in reverse transcription. Virus Res *134*, 86-103.

Schuster, P., Fontana, W., Stadler, P.F., and Hofacker, I.L. (1994). From sequences to shapes and back: a case study in RNA secondary structures. Proc R Soc Lond B Biol Sci *255*, 279-284.

Seipelt, R.L., Zheng, B., Asuru, A., and Rymond, B.C. (1999). U1 snRNA is cleaved by RNase III and processed through an Sm site- dependent pathway. Nucleic Acids Res *27*, 587-595.

Shuman, S. (2001). Structure, mechanism, and evolution of the mRNA capping apparatus. Prog Nucleic Acid Res Mol Biol *66*, 1-40.

Skruzny, M., Schneider, C., Racz, A., Weng, J., Tollervey, D., and Hurt, E. (2009). An endoribonuclease functionally linked to perinuclear mRNP quality control associates with the nuclear pore complexes. PLoS Biol *7*, e8.

Srivastava, A.K., and Schlessinger, D. (1990). Mechanism and regulation of bacterial ribosomal RNA processing. Annu Rev Microbiol *44*, 105-129.

Staley, J.P., and Woolford, J.L., Jr. (2009). Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. Curr Opin Cell Biol *21*, 109-118.

Staple, D.W., and Butcher, S.E. (2003). Solution structure of the HIV-1 frameshift inducing stem-loop RNA. Nucleic Acids Res *31*, 4326-4331.

Steinmetz, E.J., Conrad, N.K., Brow, D.A., and Corden, J.L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. Nature *413*, 327-331.

Steinmetz, E.J., Warren, C.L., Kuehner, J.N., Panbehi, B., Ansari, A.Z., and Brow, D.A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell *24*, 735-746.

Stoltzfus, C.M., and Madsen, J.M. (2006). Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. Curr HIV Res *4*, 43-55.

Strahl-Bolsinger, S., Hecht, A., Luo, K., and Grunstein, M. (1997). SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. Genes Dev. *11*, 83-93.

Stripecke, R., Oliveira, C.C., McCarthy, J.E., and Hentze, M.W. (1994). Proteins binding to 5' untranslated region sites: a general mechanism for translational regulation of mRNAs in human and yeast cells. Mol Cell Biol *14*, 5898-5909.

Taggart, A.K., Teng, S.C., and Zakian, V.A. (2002). Est1p as a cell cycle-regulated activator of telomere-bound telomerase. Science *297*, 1023-1026.

Tanaka-Fujita, R., Soeno, Y., Satoh, H., Nakamura, Y., and Mori, S. (2007). Human and mouse protein-noncoding snoRNA host genes with dissimilar nucleotide sequences show chromosomal synteny. RNA *13*, 811-816.

Tanaka, R., Satoh, H., Moriyama, M., Satoh, K., Morishita, Y., Yoshida, S., Watanabe, T., Nakamura, Y., and Mori, S. (2000). Intronic U50 small-nucleolar-RNA (snoRNA) host gene of no protein-coding potential is mapped at the chromosome breakpoint t(3;6)(q27;q15) of human B-cell lymphoma. Genes Cells *5*, 277-287.

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

Thompson, S.R., and Sarnow, P. (2000). Regulation of host cell translation by viruses and effects on cell function. Curr Opin Microbiol *3*, 366-370.

Thomson, A.M., Rogers, J.T., and Leedman, P.J. (1999). Iron-regulatory proteins, iron-responsive elements and ferritin mRNA translation. Int J Biochem Cell Biol *31*, 1139-1152.

Tijsterman, M., and Plasterk, R.H. (2004). Dicers at RISC; the mechanism of RNAi. Cell *117*, 1-3.

Tollervey, D. (1995). Genetic and biochemical analyses of yeast RNase MRP. Mol Biol Rep *22*, 75-79.

Tollervey, D. (2004). Molecular biology: termination by torpedo. Nature *432*, 456-457.

Tollervey, D., and Kiss, T. (1997). Function and synthesis of small nucleolar RNAs. Curr Opin Cell Biol *9*, 337-342.

Torchet, C., Bousquet-Antonelli, C., Milligan, L., Thompson, E., Kufel, J., and Tollervey, D. (2002). Processing of 3'-extended read-through transcripts by the exosome can generate functional mRNAs. Mol Cell *9*, 1285-1296.

Tourriere, H., Chebli, K., and Tazi, J. (2002). mRNA degradation machines in eukaryotic cells. Biochimie *84*, 821-837.

Tran, D.P., Kim, S.J., Park, N.J., Jew, T.M., and Martinson, H.G. (2001). Mechanism of poly(A) signal transduction to RNA polymerase II in vitro. Mol Cell Biol *21*, 7495-7508.

Tremblay, A. (2002). Etude de la fonction de la RNase III eucaryote et identification de ses partenaires cellulaires dans un criblage double-hybrides. In Department of Microbiology (Sherbrooke, Université de Sherbrooke).

Tremblay, A., Lamontagne, B., Catala, M., Yam, Y., Larose, S., Good, L., and Abou Elela, S. (2002a). A physical interaction between Gar1p and Rnt1p is required for the nuclear import of H/ACA small nucleolar RNA-associated proteins. Mol Cell Biol *22*, 4792-4802.

Tremblay, A., Lamontagne, B., Catala, M., Yam, Y., Larose, S., Good, L., and Elela, S.A. (2002b). A physical interaction between Gar1p and Rnt1pi is required for the nuclear import of H/ACA small nucleolar RNA-associated proteins. Mol Cell Biol *22*, 4792-4802.

Tullius, T.D., and Dombroski, B.A. (1985). Iron(II) EDTA used to measure the helical twist along any DNA molecule. Science *230*, 679-681.

Tycowski, K.T., Smith, C.M., Shu, M.D., and Steitz, J.A. (1996). A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in Xenopus. Proc Natl Acad Sci U S A *93*, 14480-14485.

Umen, J.G., and Guthrie, C. (1995). The second catalytic step of pre-mRNA splicing. Rna *1*, 869-885.

van Hoof, A., Lennertz, P., and Parker, R. (2000). Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. Mol Cell Biol *20*, 441-452.

Vasiljeva, L., and Buratowski, S. (2006). Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. Mol Cell *21*, 239-248.

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

Vasudevan, S., Peltz, S.W., and Wilusz, C.J. (2002). Non-stop decay--a new mRNA surveillance pathway. Bioessays *24*, 785-788.

Venema, J., and Tollervey, D. (1995). Processing of pre-ribosomal RNA in Saccharomyces cerevisiae. Yeast *11*, 1629-1650.

Verheggen, C., Lafontaine, D.L., Samarsky, D., Mouaikel, J., Blanchard, J.M., Bordonne, R., and Bertrand, E. (2002). Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. EMBO J *21*, 2736-2745.

Vigne, R., and Jordan, B.R. (1977). Partial enzyme digestion studies on Escherichia coli, Pseudomonas, Chlorella, Drosophila, HeLa and yeast 5S RNAs support a general class of 5S RNA models. J Mol Evol *10*, 77-86.

Viladevall, L., St Amour, C.V., Rosebrock, A., Schneider, S., Zhang, C., Allen, J.J., Shokat, K.M., Schwer, B., Leatherwood, J.K., and Fisher, R.P. (2009). TFIIH and P-TEFb coordinate transcription with capping enzyme recruitment at specific genes in fission yeast. Mol Cell *33*, 738-751.

Vilardell, J., and Warner, J.R. (1997). Ribosomal protein L32 of Saccharomyces cerevisiae influences both the splicing of its own transcript and the processing of rRNA. Mol Cell Biol *17*, 1959-1965.

Virtanen, A., and Astrom, J. (1997). Function and characterization of poly(A)-specific 3' exoribonucleases. Prog Mol Subcell Biol *18*, 199-220.

Wagner, E.J., and Marzluff, W.F. (2006). ZFP100, a component of the active U7 snRNP limiting for histone pre-mRNA processing, is required for entry into S phase. Mol Cell Biol *26*, 6702-6712.

Wahl, M.C., Will, C.L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701-718.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470-476.

Wang, G., Guo, X., and Floros, J. (2005). Differences in the translation efficiency and mRNA stability mediated by 5'-UTR splice variants of human SP-A1 and SP-A2 genes. Am J Physiol Lung Cell Mol Physiol *289*, L497-508.

Wassarman, K.M. (2002). Small RNAs in bacteria: diverse regulators of gene expression in response to environmental changes. Cell *109*, 141-144.

Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M., and Luhrmann, R. (1998). Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA- binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. Rna *4*, 1549-1568.

West, S., Gromak, N., and Proudfoot, N.J. (2004). Human 5' --> 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. Nature *432*, 522-525.

Westholm, J.O., Xu, F., Ronne, H., and Komorowski, J. (2008). Genome-scale study of the importance of binding site context for transcription factor binding and gene regulation. BMC Bioinformatics *9*, 484.

Wildeman, A.G., and Nazar, R.N. (1980). Nucleotide sequence of wheat chloroplastid 4.5 S ribonucleic acid. Sequence homologies in 4.5 S RNA species. J Biol Chem *255*, 11896-11900.

Wildeman, A.G., and Nazar, R.N. (1981). Studies on the secondary structure of 5.8 S rRNA from a thermophile, Thermomyces lanuginosus. J Biol Chem *256*, 5675-5682.

Wu, H., Henras, A., Chanfreau, G., and Feigon, J. (2004). Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. Proc Natl Acad Sci U S A *101*, 8307-8312.

Wu, H., Yang, P.K., Butcher, S.E., Kang, S., Chanfreau, G., and Feigon, J. (2001). A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. Embo J *20*, 7240-7249.

Yamaguchi, Y., and Inouye, M. (2009). mRNA interferases, sequence-specific endoribonucleases from the toxin-antitoxin systems. Prog Mol Biol Transl Sci *85*, 467-500.

Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468-474.

Ying, S.Y., Chang, D.C., and Lin, S.L. (2008). The microRNA (miRNA): overview of the RNA genes that modulate gene function. Mol Biotechnol *38*, 257-268.

Yon, J., Giallongo, A., and Fried, M. (1991). The organization and expression of the Saccharomyces cerevisiae L4 ribosomal protein genes and their identification as the homologues of the mammalian ribosomal protein gene L7a. Mol Gen Genet *227*, 72-80.

Young, R.A., and Steitz, J.A. (1978). Complementary sequences 1700 nucleotides apart form a ribonuclease III cleavage site in Escherichia coli ribosomal precursor RNA. Proc Natl Acad Sci U S A *75*, 3593-3597.

Zarudnaya, M.I., Kolomiets, I.M., and Hovorun, D.M. (2002). What nuclease cleaves pre-mRNA in the process of polyadenylation? IUBMB Life *54*, 27-31.

Zhai, Y., Zhong, Z., Chen, C.Y., Xia, Z., Song, L., Blackburn, M.R., and Shyu, A.B. (2008). Coordinated changes in mRNA turnover, translation, and RNA processing bodies in bronchial epithelial cells following inflammatory stimulation. Mol Cell Biol *28*, 7414-7426.

Zhang, K., and Nicholson, A.W. (1997). Regulation of ribonuclease III processing by double-helical sequence antideterminants. Proc Natl Acad Sci U S A *94*, 13437-13441.

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

Zhou, H., Zhao, J., Yu, C.H., Luo, Q.J., Chen, Y.Q., Xiao, Y., and Qu, L.H. (2004). Identification of a novel box C/D snoRNA from mouse nucleolar cDNA library. Gene *327*, 99-105.

Zhu, Y., Fraering, P., Vionnet, C., and Conzelmann, A. (2005). Gpi17p does not stably interact with other subunits of glycosylphosphatidylinositol transamidase in Saccharomyces cerevisiae. Biochim Biophys Acta *1735*, 79-88.

Zorio, D.A., and Bentley, D.L. (2004). The link between mRNA processing and transcription: communication works both ways. Exp Cell Res *296*, 91-97.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res *31*, 3406-3415.

Zuker, M., and Jacobson, A.B. (1995). "Well-determined" regions in RNA secondary structure prediction: analysis of small subunit ribosomal RNA. Nucleic Acids Res *23*, 2791-2798.

Zuker, M., and Jacobson, A.B. (1998). Using reliability information to annotate RNA secondary structures. Rna *4*, 669-679.

Zuo, Y., and Deutscher, M.P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res *29*, 1017-1026.

ANNEX