tRNA-like leader-trailer interaction promotes 3′-end maturation of MALAT1

SEYED-FAKHREDDIN TORABI,[1,2] SUZANNE J. DEGREGORIO,[1,2] and JOAN A. STEITZ[1,2]

[1]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06536, USA
[2]Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06536, USA

ABSTRACT

Human metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) is a nuclear long noncoding RNA (lncRNA) that is highly overexpressed in many cancer tissues and plays important roles in tumor progression and metastasis. The MALAT1 primary transcript contains evolutionarily conserved structural elements in its 3′-terminal region: a triple helix forming element called element for nuclear expression (ENE) and a downstream tRNA-like structure called mascRNA. Instead of being polyadenylated, mature MALAT1 is generated by recognition and processing of the mascRNA by RNase P. A genomically encoded A-rich tract at the new 3′ end of MALAT1, which is generated upon RNase P cleavage, forms a triple helical structure with the upstream ENE. Triplex formation is vital for stabilization of the mature transcript and for subsequent accumulation and oncogenic activity of MALAT1. Here, we demonstrate that efficient 3′-end maturation of MALAT1 is dependent on an interaction between the A-rich tract and the mascRNA 3′ trailer. Using mutational analyses of cell-based reporter accumulation, we show that an extended mascRNA acceptor stem and formation of a single bulged A 5′ to the RNase P cleavage site are required for efficient maturation of the nascent MALAT1 3′ end. Our results should benefit the development of therapeutic approaches to cancer through targeting MALAT1.

Keywords: MALAT1 3′-end maturation; A-rich tract; t-RNA like trailer; RNase P; RNA triple helix

INTRODUCTION

Eukaryotic genomes encode thousands of long noncoding RNAs (lncRNAs) that are emerging as important players in a wide range of cellular processes (Statello et al. 2021). It is estimated that there are ∼18,000 lncRNA human genes. A growing fraction of them is directly linked to various diseases including cancer (Ali et al. 2018; Frankish et al. 2019; Statello et al. 2021). Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), also known as nuclear-enriched abundant transcript 2 (NEAT2), is one of the most extensively studied nuclear-localized lncRNAs (Wilusz 2016; Arun et al. 2020). MALAT1 is strongly up-regulated in a broad spectrum of tumor types and is considered oncogenic in breast, lung, liver and many other cancers (Lin et al. 2007; Gutschner et al. 2013). Because of accumulating evidence that MALAT1 plays a critical role in the progression and metastasis of numerous cancers, it is being actively studied as a potential therapeutic target (Arun et al. 2016, 2018; Lin et al. 2018; Pan et al. 2020).

Although the 3′ ends of most lncRNAs are formed through canonical cleavage and polyadenylation, MALAT1 3′-end maturation is carried out by RNase P cleavage ∼0.4 kb upstream of the classical cleavage and polyadenylation site (Wilusz et al. 2008). Highly conserved RNA structural elements residing in the 3′-proximal region of MALAT1 allowed identification of MALAT1 homologs in other vertebrates (Zhang et al. 2017). These structural elements are: (i) an element for nuclear expression (ENE), which is a U-rich internal loop flanked by double helices (Conrad and Steitz 2005), and (ii) a tRNA-like structure called MALAT1-associated small cytoplasmic RNA (mascRNA) (Wilusz et al. 2008). In addition, an A-rich tract is located between the ENE and mascRNA. Upon RNase P cleavage, the released 3′-genomically encoded A-rich tract forms a blunt-ended triple helix structure with the ENE motif, which stabilizes MALAT1 by shielding its 3′ end from exonucleolytic degradation (Brown et al. 2014). Unlike mature MALAT1, which is nuclear, mascRNA localizes to the cytoplasm (Wilusz et al. 2008).

Corresponding authors: seyed.torabi@yale.edu, joan.steitz@yale.edu
Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.078810.121.

Without 3′-end maturation and release of the A-rich tract, MALAT1 cannot accumulate to high abundance inside the cell (Brown et al. 2012). Maturation of the MALAT1 3′ end and formation of the triple helix structure not only stabilize the transcript but also up-regulate translation in a GFP expression system (Wilusz et al. 2012).

Previous studies have demonstrated that enzymes involved in tRNA biogenesis, including RNase P, RNase Z, and CCA-adding enzymes, are responsible for the maturation and processing of both the MALAT1 3′ end and the mascRNA (Wilusz et al. 2008; Kuhn et al. 2015). First, RNase P cleavage occurs downstream from the A-rich tract, followed by RNase Z cleavage, which cuts off the mascRNA 3′-trailer sequence (Wilusz et al. 2008). Despite little sequence conservation in the majority of the MALAT1 sequence, the triple helix forming region and mascRNA are highly conserved across vertebrate species (Quinn and Chang 2016; Zhang et al. 2017). In addition, the 7 nt of the mascRNA 3′ trailer, which is removed by RNase Z and is absent from the processed mascRNA, are highly conserved among vertebrates (Fig. 1A). Yet, 3′-trailer sequences in tRNAs are generally not conserved, nor are they a major determinant of substrate recognition by mammalian RNase Z (Nashimoto 2000).

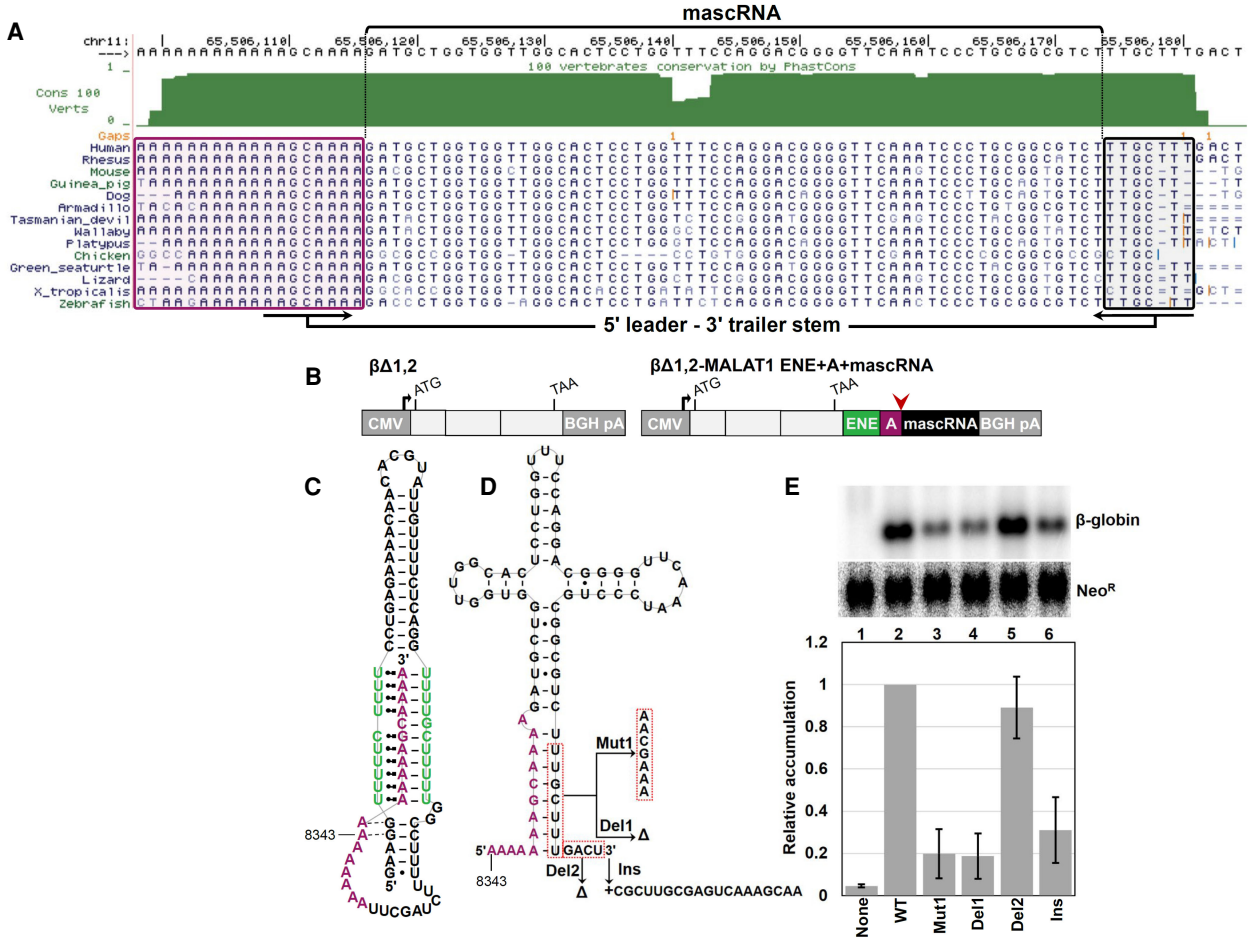

FIGURE 1. 3′-terminal region of the nascent MALAT1 contains evolutionarily conserved sequences that are predicted to interact before 3′-end maturation. (*A*) mascRNA is located immediately downstream from the MALAT1 A-rich tract (boxed in purple). A 7-nt long mascRNA 3′ trailer (boxed in black), which is highly conserved among vertebrates, is fully complementary to the A-rich tract. (*B*) Schematic of the intronless β-globin reporter construct containing a cytomegalovirus promoter, a human βΔ1,2 gene, and a bovine growth hormone poly(A) site (BGH pA). In ENE-containing reporters, the ENE (green), the A-rich tract (purple), and the mascRNA (black) are inserted upstream of the poly(A) site. The RNase P cleavage site is marked by an arrowhead. (*C*) Schematic of the MALAT1 ENE + A-rich tract shows the triple helical structure formed at the 3′ end of mature MALAT1. The U-rich internal loops are in green and the A-rich tract is in purple. (*D*) The human mascRNA is predicted to fold into a tRNA-like structure with an extended acceptor stem that is formed through base-pairing interactions between the upstream A-rich tract and the downstream mascRNA 3′ trailer. A single A residue is predicted to bulge immediately upstream of the RNase P cleavage site. Nucleotides targeted for mutagenesis are outlined in red boxes. In *C* and *D*, nucleotide 8343 of the human MALAT1 is indicated. (*E*) Northern blots probed for β-globin and control neomycin resistance (Neo^R) transcripts (*top*) were quantitated by normalizing the β-globin signals to those of Neo^R (*bottom*). Reporter accumulation with the wild-type MALAT1 ENE + A-rich tract + mascRNA was set at 1. Relative accumulation was the average of at least three independent experiments ± SD.

The possible involvement of the mascRNA 3′ trailer in MALAT1 maturation has not been carefully explored. Here, we show that the conserved portion of the mascRNA 3′ trailer interacts with sequences 3′ to the A-rich tract, explaining coevolution of these regions. We investigated a series of mutations using a cellular intronless β-globin (βΔ1,2) reporter assay to demonstrate that base-pairing interactions between the A-rich tract and the mascRNA 3′ trailer are required for efficient MALAT1 3′-end maturation, subsequently enabling cellular accumulation of the reporter RNA. Our results suggest that therapeutic interventions could target MALAT1 3′-end maturation through preventing the interaction between the A-rich tract and the mascRNA 3′ trailer.

RESULTS AND DISCUSSION

We first investigated the contribution of the mascRNA 3′-trailer sequence to enhancing the cellular accumulation of βΔ1,2 reporters containing MALAT1 ENE + A-rich tract + mascRNA sequence upstream of the poly(A) site (Fig. 1B–D). The wild-type (WT) construct contains an 11-nucleotide (UUGCUUUGACU) mascRNA 3′-trailer sequence. Inclusion of the MALAT1 ENE + A-rich tract + mascRNA in the 3′ untranslated region of the reporter mRNA enabled its accumulation in HEK293T cells (Fig. 1E), which has been shown to result from formation of an RNA with a 3′-blunt ended triplex structure (Fig. 1C; Brown et al. 2014). When only the first 7 nt of the mascRNA 3′ trailer, which are highly conserved among vertebrates and are fully complementary to part of the A-rich tract (Fig. 1D), were mutated or deleted, reporter transcript accumulation dropped more than fivefold (Fig. 1E, lanes 3–4). In contrast, we observed no significant change in reporter accumulation upon deletion of the noncomplementary region of the 3′-trailer sequence (Fig. 1E, lane 5).

Since all tested mascRNA mutants contain the same MALAT1 ENE + A-rich tract sequence, observed differences in the accumulation of the reporter transcript could be due to differing efficiencies of processing mascRNA by RNase P, which liberates the 3′ end of the A-rich tract. Predicted secondary structures of different mascRNA mutants are shown in Supplemental Figure S1. We further examined the contribution of the 3′ trailer to reporter accumulation by inserting a sequence (see Ins in Fig. 1D) downstream from the 3′ trailer that should sequester the trailer strand by forming an exceptionally stable tetraloop with a long stem (Supplemental Fig. S1; Chen and Garcia 2013). Formation of a stable tetraloop should prevent the 3′ trailer from interacting with the A-rich tract. Despite the presence of the wild-type 3′ trailer in Ins, recruitment of the 3′ trailer into a stable tetraloop resulted in a reduction in the accumulation of the reporter transcript comparable to that of the Mut1 and Del1 mutants (Fig. 1E, lanes 3, 4 and 6). It was previously demonstrated that formation of hairpin-like structures involving a pre-tRNA 3′ trailer and its downstream sequence has no negative effect on pre-tRNA processing reactions (Levinger et al. 1998). Moreover, it was observed that the sequence of the 3′ trailer of pre-tRNA has no effect on the substrate specificity or cleavage kinetics of RNase P (Hsieh et al. 2009). Therefore, our results suggest that although mascRNA is processed by the same set of enzymes as pre-tRNAs (Wilusz et al. 2008), the mascRNA 3′ trailer plays an important role in efficient maturation of the MALAT1 3′ end. Unlike tRNAs, which lack sequence conservation in their leader and trailer sequences (Nashimoto et al. 1999), the 7 nt of the 3′ trailer proximal to the 3′ end of mascRNA exhibit high sequence conservation among vertebrates (Fig. 1A), supporting biological significance.

Next, we asked whether the accumulation of reporter transcripts containing a mutated MALAT1 triplex can be regulated by formation of a predicted mascRNA extended stem through interaction of the A-rich tract with the 3′ trailer (Fig. 2A,B). Previous studies demonstrated that deletion of a single A nucleotide at the 3′ end of the MALAT1 A-rich tract (ΔA mutant) resulted in an ∼10-fold decrease in the level of the βΔ1,2 reporter transcript (Fig. 2C, lane 3; Brown et al. 2012). Such a dramatic effect was suggested to be due to the lack of a paired 3′ terminus at the triplex–duplex junction and subsequent destabilization of the triplex motif (Brown et al. 2012). As shown in Figure 2B, the extended stem is predicted to form a single bulged A immediately upstream of the RNase P cleavage site. The βΔ1,2 reporters used in this study and earlier ones (Brown et al. 2012, 2014) all contained the same mascRNA carrying the 3′-trailer sequence shown in Figure 1D. The ΔA mutant mascRNA is predicted to form a long acceptor stem with no bulged nucleotide (Supplemental Fig. S2, structure 3). Therefore, deletion of a single U nucleotide from the 3′ trailer of the ΔA mutant (called the ΔA–ΔU mutant) should restore formation of a bulged A upstream of the RNase P cleavage site (Supplemental Fig. S2, structure 4). Indeed, the ΔA–ΔU mutant showed significantly higher reporter accumulation than that of the ΔA mutant (Fig. 2C, lane 4). To further investigate the importance of the formation of an extended stem with a bulged nucleotide between the MALAT1 A-rich tract and the mascRNA 3′ trailer, we added a single U to the mascRNA 3′ trailer, which is expected to prevent formation of the bulged A upstream of the RNase P cleavage site (Supplemental Fig. S2, structure 5). Formation of a continuous long acceptor stem resulted in a substantial reduction in the accumulation of this reporter, despite its potential to form the wild-type MALAT1 triplex motif (Fig. 2C, lane 5). These results strongly suggest that the A-rich tract forms base-pairing interactions with the 3′ trailer, creating a bulged A nucleotide upstream of the RNase P cleavage site that greatly enhances production of the mature MALAT1.

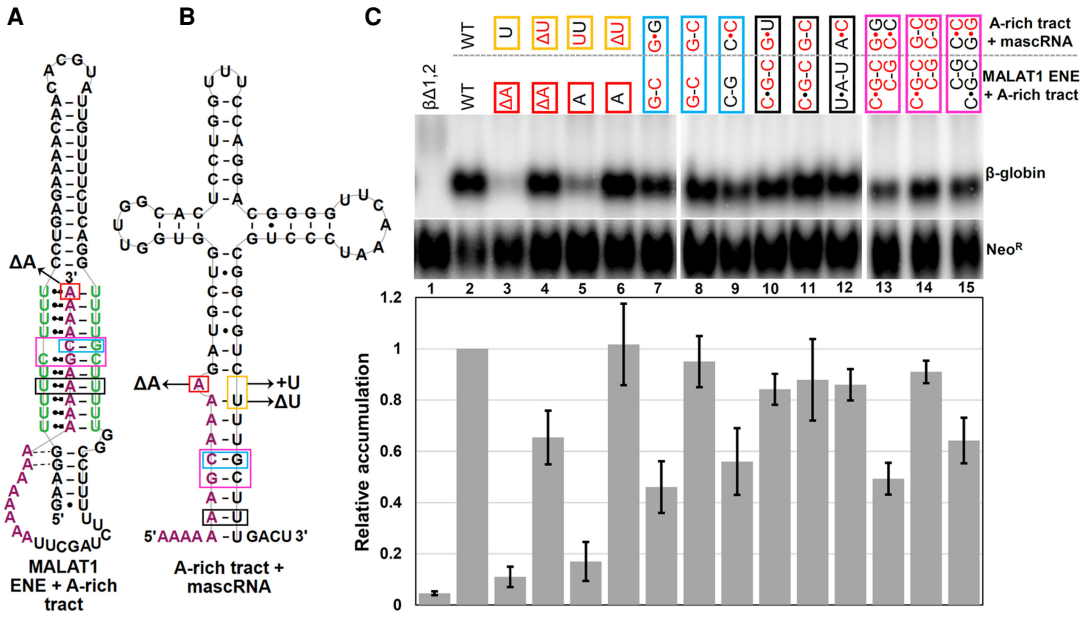

FIGURE 2. Functional significance of the predicted base-pairing interaction between the A-rich tract and the mascRNA 3′ trailer. Schematic diagrams show the 3′-terminal triple helix structure of mature MALAT1 (*A*) and the mascRNA tRNA-like structure with its predicted extended acceptor stem in the MALAT1 precursor (*B*). Colored boxes outline the nucleotides mutated in this study. (*C*) Northern blot analysis of β-globin and Neo^R transcripts (*top*) and quantitation (*bottom*) were carried out as in Figure 1E. Wild-type (black) and mutated (red) nucleotides are shown in colored boxes as in *A* and *B*. The Δ symbol represents a nucleotide deletion. Relative accumulations, which are normalized relative to the wild-type construct, are the average of at least three independent experiments ± SD.

To further investigate the contribution of the base-pairing interactions between the MALAT1 A-rich tract and the mascRNA 3′ trailer to the generation of a correct 3′ end and thereby a functional triplex, we introduced mutations into other regions of the triplex A-rich tract and into predicted complementary nucleotides in the 3′ trailer (Fig. 2A,B), including (i) C of the C-G Watson–Crick base pair separating the two stacks of base triples in the triple helical structure (blue box), (ii) A of a U•A-U base triple in the lower stack of base triples (black box), and (iii) G of the C•G-C triple (magenta box). As shown in Figure 2C (lanes 7–8), reduction in the accumulation of the reporter transcript caused by the C-G to G-C mutation was rescued by a compensatory mutation in the 3′ trailer that restores its predicted base-pairing interaction with the A-rich tract. The same compensatory mutation in the 3′ trailer (G to C mutation) reduced accumulation of the reporter transcript containing the wild-type MALAT1 triplex to ~55% (Fig. 2C, lane 9). Interestingly, none of the mutations in the U•A-U triple located in the middle of the lower stack of triples in the MALAT1 triplex structure (Fig. 2A, boxed in black) or in the complementary U nucleotide in the mascRNA 3′ trailer (Fig. 2B boxed in black) yielded a substantial change in reporter accumulation (Fig. 2C, lanes 10–12). These results are in agreement with a previous mutational study using βΔ1,2 reporter assays (Brown et al. 2012), which showed that mutations in a U•A-U base triple located in the upper triplex region (triplex II in Supplemental Fig. S3A) invariably have more deleterious effects compared to the same type of mutations located in the lower triplex region (triplex I in Supplemental Fig. S3A). Exceptions are those mutants that maintain the wild-type base-pairing interactions between the A-rich tract and mascRNA 3′ trailer (Supplemental Fig. S3). Finally, we investigated the effect of changing the order of the C•G-C triplet and C-G doublet on reporter accumulation. Inversion resulted in ~50% drop (Fig. 2C, lane 13), which is consistent with previously reported data (Brown et al. 2014). As predicted, compensatory mutations in the 3′ trailer of mascRNA recover the accumulation of the reporter containing the inverted C•G-C triplet and C-G doublet (Fig. 2C, lane 14), further supporting interaction between the A-rich tract and the 3′ trailer of mascRNA. Together, these observations suggest that formation of a tRNA-like structure with an extended stem formed by the A-rich tract and the 3′ trailer of mascRNA, as well as a bulged A located immediately upstream of the RNase P cleavage site, are important for the accumulation of reporter transcripts containing the MALAT1 triplex.

To gain further insight into the contribution of the interaction between the A-rich tract and the 3′ trailer of mascRNA to the 3′ processing by RNase P of MALAT1 in the reporter transcript, we generated a series of intron-containing β-globin (β-WT) reporters by inserting either the wild-type or a mutant MALAT1 ENE + A-rich tract + mascRNA into the 3′-UTR (Fig. 3A,B) and assaying β-globin transcript accumulation and 3′-end processing in human

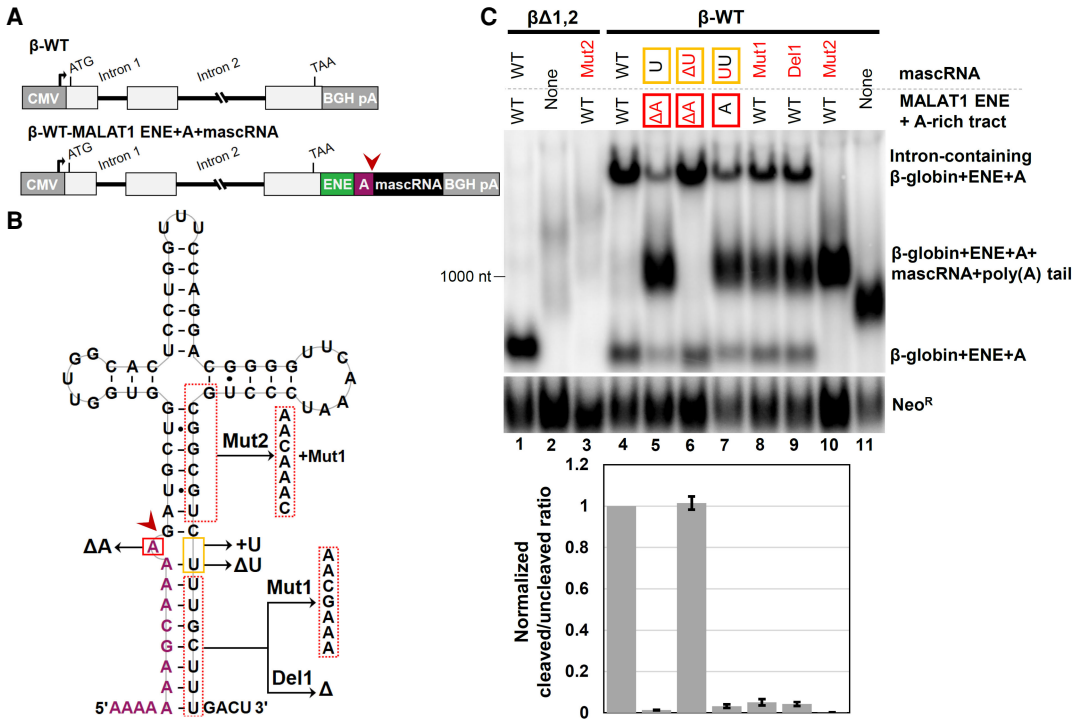

FIGURE 3. Predicted base-pairing interactions between the A-rich tract and the mascRNA 3′ trailer contribute to the 3′-end maturation of reporter transcripts. (*A*) Schematics of the intron-containing β-globin (β-WT) reporter constructs with or without the MALAT1 ENE (green), the A-rich tract (purple), and the mascRNA (black). Introns are represented by black lines; exons are represented by gray boxes; the RNase P cleavage site in the wild-type MALAT1 is marked by an arrowhead. (*B*) Predicted tRNA-like secondary structure for mascRNA with an extended acceptor stem formed by the A-rich tract (purple) and the mascRNA 3′ trailer. Boxed nucleotides were mutated. Mut2 was created from Mut1 by mutating the additional nucleotides shown in the acceptor stem. (*C*) Northern blot analysis of β-globin and Neo^R RNAs (*top*) and quantitation (*bottom*) examined the effect of the extended stem on reporter transcript maturation. Disruption of the extended stem or elimination of the bulged nucleotide upstream of the RNase P cleavage site abolished efficient maturation of the reporter transcript. Maturation efficiency is the ratio of the cleaved to uncleaved transcripts normalized to the wild-type ratio. Cleaved transcripts are the sharp β-globin bands (*top* and the *bottom*), while uncleaved transcripts are the more diffuse polyadenylated β-globin bands appearing in the *middle*. Representative data are the average of at least three independent experiments ± SD.

HEK293T cells via northern blot analysis. While the intronless β-globin transcript is retained in the nucleus and degraded unless stabilized by triplex forming elements, the β-WT reporter transcript is exported to the cytoplasm and accumulates to levels detectable by northern blot (Collis et al. 1990; Conrad and Steitz 2005). Reporter transcripts that are processed by RNase P are expected to appear as sharp bands. In contrast, unprocessed reporter transcripts should become polyadenylated and appear as diffuse bands migrating slower than processed bands. Therefore, we used the β-WT reporter to investigate mutants that might be deficient in 3′-end processing by RNase P. In Mut2 (Fig. 3B) the acceptor stem of mascRNA is completely disrupted and expected to abolish the formation of a tRNA-like structure. Therefore, Mut2 provides a negative control for 3′-end processing by RNase P (Wilusz et al. 2008). Mut2 was unable to stabilize the βΔ1,2 reporter transcript (Fig. 3C, lane 3), consistent with previous observations (Brown et al. 2012; Zong et al. 2016). As shown in Figure 3C, lane 4, the β-WT reporter transcript containing the wild-type MALAT1 ENE + A-rich tract + mascRNA accumulated as two different isomers: (i) an isomer that comigrates with the RNase P-processed intronless reporter transcript (Brown et al. 2012) produced by the βΔ1,2 reporter system, and (ii) an isomer (upper band) that migrates more slowly. The appearance of the band corresponding to each isomer is sharp, suggesting that they are not polyadenylated. The accumulation of the two isomers was drastically decreased in the ΔA mutant (Fig. 3C, lane 5), while a third isomer (middle band) accumulated as an intense smeary band. Restoration of the bulged A in the ΔA–ΔU mutant recovered wild-type processing activity (Fig. 3C, lane 6). Consistently, addition of an extra U to the mascRNA 3′ trailer, which is predicted to prevent formation of the bulged A, yielded a processing pattern resembling the ΔA mutant (Fig. 3C, lane 7). To identify the different RNA isomers produced by the β-WT reporter variants in Figure 3, we probed the northern blots with DNA probes targeting the second β-globin intron or the region between the RNase P cleavage site and the poly(A) signal. DNA oligonucleotides

targeting the second intron detected only the upper band (Supplemental Fig. S4A,B), whereas DNA probes specific to the region downstream from the RNase P cleavage site hybridized only to the middle band (Supplemental Fig. S4A,C). These data suggest that RNase P processing does not occur efficiently in the absence of the bulged A (Supplemental Fig. S2, structure 3 and 5) resulting in the formation of a polyadenylated transcript (Fig. 3C, lanes 5 and 7).

Mismatched nucleotides and bulges at position −1 preceding the mature tRNA 5′ end have been reported to be important for efficient processing of yeast pre-tRNAs that contain >6-bp extensions of the acceptor stem formed by complementarity between the 5′-leader and 3′-trailer sequences (Lee et al. 1997). Additionally, examination of human pre-tRNA sequences revealed the presence of −1/+1 bulges or internal loops in the majority of pre-tRNA sequences possessing extended acceptor stems (Gogakos et al. 2017). Therefore, our finding that a bulged A in the MALAT1-mascRNA sequence is required for efficient RNase P processing is consistent with earlier findings. Apparently, continuous and strong coaxial extension of the acceptor stem of mascRNA in the ΔA and + U mutants (Supplemental Fig. S2, structure 3 and 5) negatively affects mascRNA recognition and subsequent processing by RNase P, consistent with previous tRNA studies (Ziehler et al. 2000; Marvin et al. 2011). Moreover, disruption or lack of mRNA polyadenylation has been reported to result in decreased terminal intron removal (Niwa and Berget 1991; Cooke et al. 1999; Rigo and Martinson 2008). Therefore, formation of the intron-containing β-globin transcript by the reporter that contains the wild-type MALAT1-mascRNA is consistent with efficient RNase P processing, which precedes polyadenylation and results in ablation of the poly(A) signal. Finally, our results demonstrate that disruption of base-pairing interactions between the A-rich tract and the 3′ trailer in the Mut1 and Del1 mutants reduces RNase P processing, resulting in the formation of the polyadenylated isomer (Fig. 3C, lanes 8–9).

Our data demonstrate that disruption of the extended acceptor stem and/or bulged nucleotide upstream of the RNase P cleavage site hamper robust maturation of the nascent MALAT1 3′ end and should result in lower level of mascRNA as well. Thus, we tested Mut1, Del1, Del2, Ins, ΔA, and ΔA-ΔU mutants for the accumulation of mascRNA using northern blot analysis (Supplemental Figs. S5, S6). Accumulation of the βΔ1,2 reporter transcript changes proportionally with changes in the level of mascRNA (Supplemental Fig. S6B). Therefore, northern blots probed for mascRNA provide direct evidence and confirm that the observed reduction in the level of reporter transcript is caused by inhibition of RNase P cleavage and inefficient 3′-end maturation of MALAT1.

Formation of 3′-blunt-ended triplex structures is one way to protect the extreme 3′ end of an RNA transcript from rapid deadenylation-dependent decay in vivo. 3′-blunt-ended triple helices can be formed through a steric mechanism (Torabi et al. 2021a,b) or by fixing the register of the 3′ end so that it does not extend beyond the triplex (Brown et al. 2014). The latter mechanism is used in MALAT1, relying on the intervening GC dinucleotides in the A-rich tract. Participation of the GC dinucleotide in the C•G-C base triple and the C-G base pair results in formation of a blunt-ended structure by fixing the register of the A-rich tract (Fig. 1C; Brown et al. 2014). Our data suggest yet another function for the GC dinucleotides in the A-rich tract, which impacts robust maturation of the nascent MALAT1 3′ end. Formation of an extended stem with a bulged nucleotide between the MALAT1 A-rich tract and the mascRNA 3′ trailer depends on the presence of the GC dinucleotides in the A-rich tract and 3′ trailer, respectively (Fig. 1D). It has been shown previously that substitution of the C•G-C triple with a U•A-U triple and the C-G doublet with an A-U doublet in the MALAT1 triplex (all U/A mutant) results in approximately fivefold reduction in the accumulation of the reporter transcript (Brown et al. 2012). As shown in Supplemental Figure S7, compensatory mutations in the mascRNA 3′ trailer of the all U/A mutant does not confer rescue (relative activity: 23 ± 3% versus 24 ± 2%). Although the compensatory mutations are expected to restore base-pairing interactions between the A-rich tract and the mascRNA 3′ trailer, they cannot implement formation of the bulged nucleotide upstream of the RNase P cleavage site. Therefore, our data suggest that the GC dinucleotides are required not only for forming a blunt-ended triplex in the mature MALAT1 but also for the precise alignment of the A-rich tract with the mascRNA 3′ trailer to force formation of the bulged A near the RNase P cleavage site.

Overall, our data argue that extension of the acceptor stem of mascRNA is required for efficient 3′-end processing of MALAT1. Similar interactions are predicted to form between the A-rich tract of MENβ, which is a lncRNA containing a 3′-end triplex-forming motif (Brown et al. 2012; Wilusz et al. 2012), and the 3′ trailer of its downstream tRNA-like structure (menRNA) (Supplemental Fig. S8; Sunwoo et al. 2009). Although a 5-bp extension of an acceptor stem was shown to enhance the apparent cleavage rate by yeast RNase P (Hsieh et al. 2009), a recent cryo-EM structure of yeast RNase P complexed with a tRNA precursor suggests that the 3′ trailer is single-stranded in a catalytically active form of the complex (Lan et al. 2018). We speculate that RNase P-dependent processing and maturation of triple helical structures such as MALAT1 require interactions beyond formation of the acceptor stem of the tRNA-like structures. Formation of the additional stem formed between the A-rich tract and the mascRNA 3′ trailer may enhance folding of a tRNA-like structure while preventing the premature formation of the triple helix prior to the action of RNase P. This conclusion is in agreement with earlier

molecular dynamic simulations predicting formation of a partially folded ENE structure that is independent of the A-rich tract before triple helix formation (Yonkunas and Baird 2019). The exact mechanism by which extension of the acceptor stems in tRNA-like structures contributes to human RNase P cleavage remains unclear. Yet, base-pairing interactions between the A-rich tract and the mascRNA 3′ trailer does explain why RNase P cleavage occurs temporally before RNase Z cleavage: formation of a double-stranded structure involving the 3′ trailer is known to inhibit the activity of RNase Z (Nashimoto et al. 1999).

In conclusion, our data suggest that efficient processing of the MALAT1 3′ end by human RNase P requires extension of the acceptor stem of mascRNA and the presence of a bulged A 5′ to the cleavage site. Recently, an antisense oligonucleotide (ASO)-based strategy has proved promising for targeting MALAT1 and inhibition of breast and lung cancer progression in animal models (Arun et al. 2016; Gong et al. 2019). Based on our results, we propose that inhibiting MALAT1 maturation by targeting the 3′ region of mascRNA (including the 3′-trailer sequence) using ASOs may provide efficient knockdown of this lncRNA. Consistently, it was previously shown that targeting the 3′ region of mascRNA with ASOs efficiently inhibits mascRNA biogenesis (Wilusz et al. 2008). Therefore, our findings may benefit therapeutic interventions targeting MALAT1.

MATERIALS AND METHODS

Plasmids construction and mutagenesis

The intronless (βΔ1,2) and intron-containing (β-WT) β-globin plasmids use pcDNA3 as the vector. These plasmids contain a human β-globin gene flanked by an upstream cytomegalovirus promoter and a downstream bovine growth hormone polyadenylation signal (Conrad and Steitz 2005). The sequence of the MALAT1 ENE + A-rich tract + mascRNA was inserted into an ApaI site in the 3′-untranlated region of the β-globin gene using a standard restriction enzyme digestion and ligation protocol. Nucleotide deletions, insertions or substitutions were introduced into the ENE-containing constructs using standard PCR-based mutagenesis methods.

Cell culture, transfection, and northern blots

Human embryonic kidney cells (HEK293T) were grown at 37°C and 5% CO_2 in Dulbecco's modified Eagle medium supplemented with 10% heat-inactivated fetal bovine serum, 1% penicillin/streptomycin, and 2 mM L-glutamine. HEK293T cells were transfected with 0.5 µg β-globin reporter and 1.5 µg pBluescript plasmid using Mirus TransIT-293 according to the manufacturer's protocol. Cells were harvested ∼40 h after transfection and RNA was isolated using TRIzol reagent (Life Technologies). Up to 10 µg of total RNA for each sample was separated on a 1.4%–1.8% agarose/6.5% formaldehyde gel. To analyze mascRNA levels, RNA samples were resolved on a gradient urea PAGE (4.5% top part and 12% bottom part). Northern blot analyses and the sequence of the oligonucleotides used to probe for Neo^R and β-globin transcripts were described previously (Torabi et al. 2021b). To detect β-globin transcripts containing intron 1, blots were probed using three DNA probes spanning this intron: 5′-GTCTCCT TAAACCTGTCTTGTAACCT-3′, 5′-GAGTCTTCTCTGTCTCCAC ATGC-3′ and 5′-GGAAAATAGACCAATAGGCAGAGAG-3′. To probe for β-globin transcripts containing intron 2, blots were probed using three DNA probes spanning intron 2: 5′-GGAGATT ATGAATATGCAAATAAGCACACA-3′, 5′-GAGGTATGAACATG ATTAGCAAAAGGG-3′ and 5′-GTTATTCTTTAGAATGGTGCAA AGAGGC-3′. DNA probes specific to the region 12 to 162 nucleotides downstream from the RNase P cleavage site, which detect only β-globin transcripts containing the unprocessed MALAT1 3′ end, include: 5′-GAGGGGCAAACAACAGATGGCTGG-3′, 5′-GC TGATCAGCGAGCTCTAGCATTTA-3′ and 5′-GTCCTGGAAACC AGGAGTGCCAA-3′. All DNA probes were 5′ ^{32}P-labeled by T4 polynucleotide kinase in the presence of [γ^{32}P]ATP before use. Bands were detected using a Storm 860 (GE Healthcare), and quantitated using ImageQuant software.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank Dr. Kazimierz Tycowski for critical review of the manuscript, Angela Miccinello for editorial work, and all Steitz laboratory members for helpful discussions. The work was supported by the Damon Runyon Cancer Research Foundation DRG225716 (to S.-F.T.) and National Institutes of Health P01CA16038 (to J.A.S.). J.A.S. is an investigator of the Howard Hughes Medical Institute.

Received April 21, 2021; accepted July 7, 2021.

REFERENCES

Ali MM, Akhade VS, Kosalai ST, Subhash S, Statello L, Meryet-Figuiere M, Abrahamsson J, Mondal T, Kanduri C. 2018. PAN-cancer analysis of S-phase enriched lncRNAs identifies oncogenic drivers and biomarkers. *Nat Commun* **9:** 883. doi:10.1038/s41467-018-03265-1

Arun G, Diermeier S, Akerman M, Chang KC, Wilkinson JE, Hearn S, Kim Y, MacLeod AR, Krainer AR, Norton L, et al. 2016. Differentiation of mammary tumors and reduction in metastasis upon MALAT1 lncRNA loss. *Genes Dev* **30:** 34–51. doi:10.1101/gad.270959.115

Arun G, Diermeier SD, Spector DL. 2018. Therapeutic targeting of long non-coding RNAs in cancer. *Trends Mol Med* **24:** 257–277. doi:10.1016/j.molmed.2018.01.001

Arun G, Aggarwal D, Spector DL. 2020. MALAT1 long non-coding RNA: functional implications. *Noncoding RNA* **6:** 22.

Brown JA, Valenstein ML, Yario TA, Tycowski KT, Steitz JA. 2012. Formation of triple-helical structures by the 3′-end sequences of MALAT1 and MENβ noncoding RNAs. *Proc Natl Acad Sci* **109:** 19202–19207. doi:10.1073/pnas.1217338109

Brown JA, Bulkley D, Wang J, Valenstein ML, Yario TA, Steitz TA, Steitz JA. 2014. Structural insights into the stabilization of

MALAT1 noncoding RNA by a bipartite triple helix. *Nat Struct Mol Biol* **21:** 633–640. doi:10.1038/nsmb.2844

Chen AA, Garcia AE. 2013. High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations. *Proc Natl Acad Sci* **110:** 16820–16825. doi:10.1073/pnas.1309392110

Collis P, Antoniou M, Grosveld F. 1990. Definition of the minimal requirements within the human β-globin gene and the dominant control region for high level expression. *EMBO J* **9:** 233–240. doi:10.1002/j.1460-2075.1990.tb08100.x

Conrad NK, Steitz JA. 2005. A Kaposi's sarcoma virus RNA element that increases the nuclear abundance of intronless transcripts. *EMBO J* **24:** 1831–1841. doi:10.1038/sj.emboj.7600662

Cooke C, Hans H, Alwine JC. 1999. Utilization of splicing elements and polyadenylation signal elements in the coupling of polyadenylation and last-intron removal. *Mol Cell Biol* **19:** 4971–4979. doi:10.1128/MCB.19.7.4971

Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J, et al. 2019. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* **47:** D766–D773. doi:10.1093/nar/gky955

Gogakos T, Brown M, Garzia A, Meyer C, Hafner M, Tuschl T. 2017. Characterizing expression and processing of precursor and mature human tRNAs by hydro-tRNAseq and PAR-CLIP. *Cell Rep* **20:** 1463–1475. doi:10.1016/j.celrep.2017.07.029

Gong N, Teng X, Li J, Liang XJ. 2019. Antisense oligonucleotide-conjugated nanostructure-targeting lncRNA MALAT1 inhibits cancer metastasis. *ACS Appl Mater Interfaces* **11:** 37–42. doi:10.1021/acsami.8b18288

Gutschner T, Hammerle M, Diederichs S. 2013. MALAT1—a paradigm for long noncoding RNA function in cancer. *J Mol Med (Berl)* **91:** 791–801. doi:10.1007/s00109-013-1028-y

Hsieh J, Walker SC, Fierke CA, Engelke DR. 2009. Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release. *RNA* **15:** 224–234. doi:10.1261/rna.1309409

Kuhn CD, Wilusz JE, Zheng Y, Beal PA, Joshua-Tor L. 2015. On-enzyme refolding permits small RNA and tRNA surveillance by the CCA-adding enzyme. *Cell* **160:** 644–658. doi:10.1016/j.cell.2015.01.005

Lan P, Tan M, Zhang Y, Niu S, Chen J, Shi S, Qiu S, Wang X, Peng X, Cai G, et al. 2018. Structural insight into precursor tRNA processing by yeast ribonuclease P. *Science* **362:** eaat6678. doi:10.1126/science.aat6678

Lee Y, Kindelberger DW, Lee JY, McClennen S, Chamberlain J, Engelke DR. 1997. Nuclear pre-tRNA terminal structure and RNase P recognition. *RNA* **3:** 175–185.

Levinger L, Bourne R, Kolla S, Cylin E, Russell K, Wang X, Mohan A. 1998. Matrices of paired substitutions show the effects of tRNA D/T loop sequence on *Drosophila* RNase P and 3′-tRNase processing. *J Biol Chem* **273:** 1015–1025. doi:10.1074/jbc.273.2.1015

Lin R, Maeda S, Liu C, Karin M, Edgington TS. 2007. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene* **26:** 851–858. doi:10.1038/sj.onc.1209846

Lin J, Hu Y, Zhao JJ. 2018. Repression of multiple myeloma cell growth in vivo by single-wall carbon nanotube (SWCNT)-delivered MALAT1 antisense oligos. *J Vis Exp* **142:** e58598. doi:10.3791/58598

Marvin MC, Walker SC, Fierke CA, Engelke DR. 2011. Binding and cleavage of unstructured RNA by nuclear RNase P. *RNA* **17:** 1429–1440. doi:10.1261/rna.2633611

Nashimoto M. 2000. Anomalous RNA substrates for mammalian tRNA 3′ processing endoribonuclease. *FEBS Lett* **472:** 179–186. doi:10.1016/S0014-5793(00)01462-9

Nashimoto M, Wesemann DR, Geary S, Tamura M, Kaspar RL. 1999. Long 5′ leaders inhibit removal of a 3′ trailer from a precursor tRNA by mammalian tRNA 3′ processing endoribonuclease. *Nucleic Acids Res* **27:** 2770–2776. doi:10.1093/nar/27.13.2770

Niwa M, Berget SM. 1991. Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev* **5:** 2086–2095. doi:10.1101/gad.5.11.2086

Pan J, Bian Y, Cao Z, Lei L, Pan J, Huang J, Cai X, Lan X, Zheng H. 2020. Long noncoding RNA MALAT1 as a candidate serological biomarker for the diagnosis of non-small cell lung cancer: a meta-analysis. *Thorac Cancer* **11:** 329–335. doi:10.1111/1759-7714.13265

Quinn JJ, Chang HY. 2016. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* **17:** 47–62. doi:10.1038/nrg.2015.10

Rigo F, Martinson HG. 2008. Functional coupling of last-intron splicing and 3′-end processing to transcription in vitro: the poly(A) signal couples to splicing before committing to cleavage. *Mol Cell Biol* **28:** 849–862. doi:10.1128/MCB.01410-07

Statello L, Guo CJ, Chen LL, Huarte M. 2021. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol* **22:** 96–118. doi:10.1038/s41580-020-00315-9

Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. 2009. MEN ε/β nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* **19:** 347–359. doi:10.1101/gr.087775.108

Torabi SF, Chen YL, Zhang K, Wang J, DeGregorio SJ, Vaidya AT, Su Z, Pabit SA, Chiu W, Pollack L, et al. 2021a. Structural analyses of an RNA stability element interacting with poly(A). *Proc Natl Acad Sci* **118:** e2026656118. doi:10.1073/pnas.2026656118

Torabi SF, Vaidya AT, Tycowski KT, DeGregorio SJ, Wang J, Shu MD, Steitz TA, Steitz JA. 2021b. RNA stabilization by a poly(A) tail 3′-end binding pocket and other modes of poly(A)-RNA interaction. *Science* **371:** eabe6523. doi:10.1126/science.abe6523

Wilusz JE. 2016. Long noncoding RNAs: re-writing dogmas of RNA processing and stability. *Biochim Biophys Acta* **1859:** 128–138. doi:10.1016/j.bbagrm.2015.06.003

Wilusz JE, Freier SM, Spector DL. 2008. 3′ end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* **135:** 919–932. doi:10.1016/j.cell.2008.10.012

Wilusz JE, JnBaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. 2012. A triple helix stabilizes the 3′ ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* **26:** 2392–2407. doi:10.1101/gad.204438.112

Yonkunas MJ, Baird NJ. 2019. A highly ordered, nonprotective MALAT1 ENE structure is adopted prior to triplex formation. *RNA* **25:** 975–984. doi:10.1261/rna.069906.118

Zhang B, Mao YS, Diermeier SD, Novikova IV, Nawrocki EP, Jones TA, Lazar Z, Tung CS, Luo W, Eddy SR, et al. 2017. Identification and characterization of a class of MALAT1-like genomic loci. *Cell Rep* **19:** 1723–1738. doi:10.1016/j.celrep.2017.05.006

Ziehler WA, Day JJ, Fierke CA, Engelke DR. 2000. Effects of 5′ leader and 3′ trailer structures on pre-tRNA processing by nuclear RNase P. *Biochemistry* **39:** 9909–9916. doi:10.1021/bi000603n

Zong X, Nakagawa S, Freier SM, Fei J, Ha T, Prasanth SG, Prasanth KV. 2016. Natural antisense RNA promotes 3′ end processing and maturation of MALAT1 lncRNA. *Nucleic Acids Res* **44:** 2898–2908. doi:10.1093/nar/gkw047