Conservation of RNase III Processing Pathways and Specificity in Hemiascomycetes

Guillaume Chanfreau*

Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California Los Angeles, Los Angeles, California 90095-1569

Received 13 May 2003/Accepted 3 July 2003

Rnt1p, the only known Saccharomyces cerevisiae RNase III endonuclease, plays important functions in the processing of precursors of rRNAs (pre-rRNAs) and of a large number of small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs). While most eukaryotic RNases III, including the Schizosaccharomyces pombe enzyme Pac1p, cleave double-stranded RNA without sequence specificity, Rnt1p cleavage relies on the presence of terminal tetraloop structures that carry the consensus sequence AGNN. To search for the conservation of these processing signals, I have systematically analyzed predicted secondary structures of the 3’-external transcribed spacer (ETS) sequences of the pre-rRNAs and of flanking sequences of snRNAs and snoRNAs from sequences available in 13 other Hemiascomycetes species. In most of these species, except in Yarrowia lipolytica, double-stranded RNA regions capped by terminal AGNN tetraloops can be found in the 3’ ETS sequences of rRNA, in the 5’- or 3’-end flanking sequences of sn(o)RNAs, or in the intergenic spacers of polycistronic snoRNA transcription units. This analysis shows that RNase III processing signals and RNase III cleavage specificity are conserved in most Hemiascomycetes species but probably not in the evolutionarily more distant species Y. lipolytica.

RNases III form a family of double-stranded RNA (dsRNA) endonucleases found in both prokaryotes and eukaryotes. These enzymes are involved in the processing of a large number of stable RNAs. Prokaryotic and eukaryotic RNase III-like proteins participate in the maturation of the precursors of rRNAs (pre-rRNAs). In eukaryotes, RNases III are required for processing of the 35S pre-rRNA; in both Saccharomyces cerevisiae and Schizosaccharomyces pombe, RNases III cleave the 3’-external transcribed spacer (ETS) found in the 35S pre-rRNA (2, 3, 13, 23, 25). This processing event is one of the earliest steps in pre-rRNA processing. In addition to having a role in pre-rRNA processing, eukaryotic RNases III play important roles in the processing of several families of stable small RNAs. In S. cerevisiae, spliceosomal small nuclear RNAs (snRNAs) are processed by Rnt1p cleavage in the 3’-extension found in the precursors of U1, U2, U4, and U5 snRNAs (1, 5, 15, 21, 24). This function is conserved in fungi, at least for the U2 snRNA from S. pombe, whose 3’-end processing requires Pac1p, the S. pombe ortholog of RNase III (28). Earlier genetic data suggest that the role of Pac1p is not restricted to the U2 snRNAs but that it is also involved in the processing of other S. pombe snRNAs (20). In addition to the processing of snRNAs, Rnt1p has also been shown to have a major function in the processing of small nucleolar RNAs (snoRNAs). Some of these snoRNAs have been synthesized with a 5’-extension, and Rnt1p cleavage provides an entry site for final trimming to the mature 5’-end by the 5’–3’ exonucleases Xrn1p and Rat1p (6; C. Y. Lee et al., in press). A few other snoRNAs are synthesized as polycistronic precursors, from which Rnt1p cleavage produces monocistronic intermediates that are further trimmed to the mature end by exonuclease digestion (6, 7, 17). It is not clear whether the function of RNase III in snRNA and snoRNA processing is conserved in multicellular eukaryotes. However, plant and metazoan RNases III belonging to the Dicer family are essential for processing a large number of microRNAs (10–12, 16, 18).

The identification of a large number of Rnt1p processing signals in rRNAs, snRNAs, and snoRNAs points to conserved structural features of Rnt1p substrates. Most Rnt1p dsRNA substrates are capped by tetraloop sequences carrying the weak consensus sequence AGNN (4, 6). These tetraloops are essential for cleavage activity, and the cleavage site is positioned 14 to 16 bp away from the AGNN tetraloop, suggesting that the enzyme acts as an RNA helical ruler (4). Structural analysis of these tetraloops showed that the enzyme probably recognizes a specific tetraloop conformation and that the strongest sequence requirement resides in the universal G2 position, which must be in the syn conformation (26). In contrast to Rnt1p, most other RNase III-like enzymes cleave dsRNA with only very limited or no sequence specificity. Escherichia coli RNase III specificity relies on antideterminants in the dsRNA region (27), and S. pombe Pac1p as well as human Dicer does not seem to exhibit sequence specificity. These observations raise the question of when the specificity for AGNN tetraloops arose during the evolution of RNase III enzymes.

I took advantage of the sequencing effort of the Généolevures genomic program (22), which selected a set of species representative of the Hemiascomycetes class, and performed the sequencing of a large number of random sequence tags (RSTs) to answer two questions. First, it is not known whether the function of RNase III in processing pre-rRNAs, snRNAs, and snoRNAs is conserved in Hemiascomycetes. Second, the dif-
ference in cleavage specificity between *S. cerevisiae* Rnt1p, which requires AGNN tetraloops, and *S. pombe* Pac1p, which does not seem to require this type of terminal loop, suggests that the Rnt1p specificity for AGNN tetraloops may have been acquired after the divergence of *S. cerevisiae* and *S. pombe*. The analysis of a large number of predicted RNase III substrates in Hemiascomycetes may provide an answer to these questions and help determine when this RNA structure specificity was adopted during evolution. In this study, I show that predicted Rnt1p-type cleavage sites are present in most Hemiascomycetes sequences analyzed and that most of these predicted sites exhibit terminal AGNN tetraloops that are compatible with Rnt1p cleavage specificity. These results demonstrate that RNase III processing pathways, as well as RNase III cleavage specificity, are conserved among most Hemiascomycetes, with the exception of the evolutionarily distant species *Yarrowia lipolytica*.

### MATERIALS AND METHODS

Sequences and secondary-structure analysis. Sequences for rRNA 3′ ETSs and snoRNAs were retrieved from the Genolevures website by using a BLAST analysis of the *S. cerevisiae* RSTs (http://cli.labri.fr/Genolevures/advanced_blast.php3). For identification of the 3′ ETS sequences, the last 250 nucleotides (nt) of the 25S rRNA were used for the BLAST search, and RSTs containing more than 300 nt of sequence downstream from the 25S rRNA were selected. For snoRNAs, the entire mature sn(o)RNA sequence was used. In the case of modification guide C and D box-containing (box C/D) snoRNAs, many hits were obtained; these sequences in some cases included short regions of complementarity with the rRNAs corresponding to the guide sequences. Most of the uncovered sequences whose BLAST E value was above 1.0 were discarded. For other sequences with high E values (0.1 to 0.9), the presence of the box C and box D sequences was manually verified, and sequences that did not show both of these conserved boxes were discarded. RSTs that did not contain more than 120 nt of the flanking region were also discarded. RNA secondary-structure prediction with Mfold (14) was performed by using M. Zuker’s Mfold web server (29) (http://www.bioinfo.rpi.edu/applications/mfold/rna/). For most snoRNA sequences, only one RST was found in the Genolevures database and analyzed. For most 3′ ETS sequences, several hits were obtained for each species. In this case, two to three sequences obtained from independent RSTs were analyzed, and in all cases, the sequences and the secondary structures were found to be identical. Secondary-structure prediction was performed on 120 to 300 nt of regions flanking the 5′ end, 3′ end, or intergenic sequences. One of the 5%–suboptimal structures is shown in the figures.

Statistical evaluation of the presence of AGNN stem-loops in rRNA and snoRNA flanking sequences. To evaluate the statistical significance of the presence of predicted stem-loops capped by AGNN tetraloops, 111 sequences flanking random genes were selected from the Genolevures database. Two-hundred-nucleotide segments of these flanking sequences were folded using Mfold (14, 29) to detect the presence of predicted stem-loops of at least 4 bp capped by AGNN or XGNN (AGNN, UGNN, CGNN, or GGNN but not YNCG or GNRA) sequences. Four AGNN and 11 XGNN stem-loops were detected in these random sequences. This random sample was used for statistical evaluation with a binomial proportions test. Comparing the frequency of the presence of AGNN stem-loops in the RSTs containing flanking sequences of rRNA and snoRNA to that in the random sample showed a highly significant enrichment of AGNN stem-loops in these sequences (*P* ≤ 1.25 × 10^{-22}). The presence of XGNN stem-loops in the RSTs containing flanking sequences of rRNA and snoRNA was also highly significant (*P* ≤ 9.2 × 10^{-21}).

### RESULTS

Rationale of sequence analysis. To analyze the putative presence of RNase III processing sites in a large number of substrates and Hemiascomycetes species, I searched the Genolevures sequence database for orthologs of rRNA and sn(o)RNA sequences (see Materials and Methods). Because Rnt1p processing sites are usually present less than 300 nt from the mature sn(o)RNA sequences, I restricted the subsequent secondary-structure analysis to RSTs that contained only sufficient (>120 nt) 5′- or 3′-end-flanking sequences. I also searched the Genolevures sequence database for the presence of sequences downstream from the 25S rRNA to obtain putative 3′ ETS sequences. After these sequences were retrieved, predicted RNA secondary structures were obtained by using the Mfold algorithm (14, 29). The list of RSTs analyzed for the presence of predicted secondary structures is shown in Table 1 and 2. Table 1 shows the RSTs used to analyze the predicted secondary structures in the 3′ ETS sequences downstream from the 25S rRNA from all 13 Hemiascomycetes species. Table 2 shows the list of RSTs used to analyze processing signals upstream (5′) or downstream (3′) from snRNAs and snoRNAs or located in the spacer of polycistronic arrays. For these snoRNAs, I searched for ortholog sequences of all known Rnt1p processing substrates. Some known substrates could not be retrieved from the RSTs, or the corresponding RSTs did not contain enough flanking sequence information. Therefore, these snoRNAs are not included in Table 2.

Conservation of RNase III processing sites in the 3′ ETS sequences of the pre-rRNAs. Ribosomal DNA (rDNA) is present in a large number of copies per haploid genome for all the Hemiascomycetes species analyzed (22). Because the Genolevures sequencing program was based on the random sequencing of genomic fragments, rDNA sequences are represented in multiple copies in the RSTs of the Genolevures database for all of these species. This feature allowed the retrieval of several copies of 3′ ETS sequences for all 13 Hemiascomycetes species (Table 1). For all of these species, extensive secondary structures could be predicted immediately.

#### TABLE 1. RSTs used to analyze 3′ ETS secondary structures

| RST no.       | Species | Terminal loop |
|---------------|---------|---------------|
| A00AA0300D08D1P1 | *S. bayanus* | AGGA |
| A00AA0060E03TP1  | *S. bayanus* | AGGA |
| A00AA0090C11TP1  | *S. exigua*  | GGGO |
| A00AA0180B07D1   | *S. exigua*  | GGGO |
| A00AA0050A02TP1  | *S. servazzii* | AGAU |
| A00AA0020F06TP1  | *S. servazzii* | AGAU |
| A00AA0170G12CP1  | Z. rouxii | AGGA |
| A00AA0090H12TP1  | Z. rouxii | AGGA |
| A00AA0090A04D1   | S. klayveri | AGGA |
| A00AA0100A09D1   | S. klayveri | AGGA |
| A00AA0130E12TP1  | K. thermotolerans | AGGC |
| A00AA0140E03TP1  | K. thermotolerans | AGGC |
| A00AA0020D03TP1  | K. thermotolerans | AGGC |
| BA0AB0030F06LP1  | K. lactis | AGGU |
| BA0AB0130B09LP1  | K. lactis | AGGU |
| AZ0AA0120F01A1   | K. marxianus | AGGU |
| AZ0AA0050C01D1   | K. marxianus | AGGU |
| BB0A0027A07TP1   | *P. angusta* | GGUA |
| BB0A0170H04D1P1  | *P. angusta* | GGUA |
| XBC0A0020C02D1   | D. hansenii | AGAA |
| BCA0A0140H12T1   | D. hansenii | AGAA |
| AX0AA0370H06TP1  | P. sorbitophila | AGGU |
| AX0AA0080H11CP1  | P. sorbitophila | AGGU |
| BD0A0150E02TP1   | C. tropicalis | UAG |
| BD0A0120C11TP1   | C. tropicalis | UAG |
| BD0A0090A05D1    | C. tropicalis | UAG |
| AW0A0130D09D1    | Y. lipolytica | GAC/GCA |
| AW0A0020C05D1    | Y. lipolytica | GAC/GCA |
| AW0A0180H09T1    | Y. lipolytica | GAC/GCA |
downstream from the mature 3′ end of the 25S rRNA (Fig. 1). For most Hemiascomycetes species, the predicted secondary structures are capped by terminal AGNN tetraloops (Table 1; Fig. 1). The occurrence of the stems capped by AGNN tetraloops is highly significant compared to their occurrence in a random sequence population extracted from the Genolevure database ($P \leq 1.25 \times 10^{-22}$) (see Materials and Methods). The predicted structures suggest that the 3′ ETS sequence is cleaved by an RNase III-like activity in these species and that the Rnt1p orthologs have retained the specificity for AGNN tetraloops. Closely related species have conserved several base pairs in their 3′ ETS secondary structures (compare Saccharomyces kluyveri and Kluyveromyces thermotolerans, Kluyveromyces marxianus and Kluyveromyces lactis, and Pichia sorbitophila and Debaryomyces hansenii), and the sequences and therefore the secondary structures are identical in S. cerevisiae and Saccharomyces bayanus. For Saccharomyces exiguus and Pichia angusta, the terminal loops exhibit a G at the first position.
However, it is known that Rnt1p can tolerate a G at the first position of the tetraloop, as long as the tetraloop does not adopt a GNRA fold (8). The other exceptions were found in Candida tropicalis and in Y. lipolytica, for which the 3′/H11032 ETS sequences did not exhibit this type of terminal tetraloop structure in any of the RST sequences analyzed. Interestingly, two predicted stem-loops structures were found in the 3′/H11032 ETS of Y. lipolytica. These two stems may correspond to genuine processing signals, as a particular sequence has been conserved in both of these stems (Fig. 1). Thus, these two stems may correspond to duplicated processing signals. The stems are not capped by AGNN tetraloops but rather by trinucleotide loops. Overall, analysis of the 3′ ETS sequences suggests that RNase III processing in the 3′ ETS occurs in most Hemiascomycetes species and that the RNase III activities involved have a specificity for dsRNA capped by AGNN tetraloops, with the exception of Y. lipolytica and possibly C. tropicalis (see below). Not surprisingly, Y. lipolytica is the most distant species from S. cerevisiae on the evolutionary scale (22).

Conservation of 5′-end processing signals. A large number of independently transcribed box H/ACA and box C/D snoRNAs are processed by RNase III cleavage in the 5′ extension of the precursor, followed by exonucleolytic digestion (6; Lee et al., in press). I searched for orthologs of these snoRNAs and analyzed the 5′-end-flanking sequences for the presence of predicted secondary structures. A total of 27 orthologous sequences flanked by sufficient 5′-end-flanking sequences were found for a total number of 12 different snoRNAs (two box H/ACA snoRNAs, snR36, and snR46, and 10 box C/D snoRNAs) (Table 2; Fig. 2). In most cases, extensive secondary structures could be predicted at a short distance upstream from the mature snoRNA sequences, and these stems were capped by AGNN tetraloops. Interestingly, the distances between the last nucleotide of the stem and the 5′ end of the

---

**FIG. 1.** Predicted secondary structures in the 3′ ETS sequences of the pre-rRNAs in various Hemiascomycetes. Distances in nucleotides to the 3′ ends of the 25S rRNAs are indicated. Distances are only approximate, since the exact 3′ ends have not been experimentally mapped but are estimated from the S. cerevisiae sequence. The sequences of the 3′ ETSs are identical in S. cerevisiae and S. bayanus. Base pairs that have been conserved in closely related species (S. kluveri and K. thermostoletans, K. marxianus and K. lactis, and P. sorbitophilia and D. hansenii) are boxed. The boxed sequence shown for Y. lipolytica corresponds to a sequence conserved in both stems.
FIG. 2. Predicted secondary structures in the 5'-end processing signals of Hemiascomycetes snoRNAs. The distances in nucleotides are from the last nucleotides of the stems to the predicted 5' ends of the mature snoRNAs. Distances are only approximate, since the exact 5' ends have not been experimentally mapped but are estimated from the *S. cerevisiae* sequence. *S.c.*, *S. cerevisiae*; *S.b.*, *S. bayanus*; *S.e.*, *S. exigus*; *S.s.*, *Saccharomyces servazzii*; *K.l.*, *K. lactis*; *K.t.*, *K. thermotolerans*; *K.m.*, *K. marxianus*; *Z.r.*, *Z. rouxii*. 
snoRNA are often comparable for *S. cerevisiae* and the other species. In some cases, the predicted processing signals seem to have diverged from a single stem to two stems which potentially stack onto each other to reconstitute a processing signal, as described previously for the snR40 5'-end processing signal in *S. cerevisiae* (6). This is the case for snR47 in *K. marxianus* and for snR52 in all *Kluyveromyces* species. For some very close species, for example, *S. cerevisiae* and *S. bayanus*, the sequence of the stems has diverged, but compensatory mutations have maintained the overall architecture of the stem-loop (e.g., in snR47). In some other cases, the upper part of the stem near the tetraloop has conserved some sequence similarities between *S. cerevisiae* and other species (e.g., in snR47, snR62, snR71, and snR79).

This analysis did not reveal any ortholog of snoRNA sequences from the *Pichia* genus or from *D. hansenii* and *Y. lipolytica*. Since these organisms are the most evolutionarily distant from *S. cerevisiae*, the sequences of these snoRNAs may have diverged enough so that orthologous sequences can hardly be found with standard BLAST procedures. Several matches of very short sequences with high *P* values were obtained for some snoRNAs in some of these species, but they were not analyzed further. However, some of the matches may represent true orthologous sequences. Interestingly, a true ortholog of the snR47 snoRNA was identified in *C. tropicalis*, and a predicted AGAA stem-loop was found in the 5'-end-flanking sequence of this snoRNA. This finding suggests that even though AGNN stem-loops are absent from the 3' ETS of the rDNA, an RNase III activity that recognizes AGNN tetraloops is present in *C. tropicalis*. It is unknown whether the same activity processes the rRNA on the 3' ETS or whether a duplicated RNase III activity specialized in processing the rRNA 3' ETS exists without specificity for dsRNA capped by AGNN tetraloops.

**Conservation of polycistronic box C/D snoRNA processing signals.** In *S. cerevisiae*, several box C/D snoRNAs are expressed as polycistronic transcription units and processed by Rnt1p cleavage followed by exonuclease processing (6, 7, 17). In these cases, the Rnt1p cleavage sites are often made of short stems capped by AGNN tetraloops, which stack onto longer stems where cleavage occurs (17). Partial orthologous sequences were found for the operon expressing seven snoRNAs (snR78 to -72 [snR78–72]) in *S. bayanus*, *K. lactis*, and *P. sorbitophilia* (Table 2; Fig. 3). The *S. bayanus* RST contained all snoRNAs but snR78, the *K. lactis* sequences contained snR78–75, and *P. sorbitophilia* contained snR75–73. Since the distances between the mature snoRNAs in these RSTs are usually 120 nt or less, these species have probably conserved the polycistronic mode of expression of these snoRNAs. For *S. bayanus* and *K. lactis*, RNase III processing signals similar to those observed in *S. cerevisiae* could be predicted (Fig. 3). For simplicity, the longer stems onto which the AGNN short stems potentially stack were not represented in Fig. 3. No secondary structures capped by AGNN tetraloops could be identified for the *P. sorbitophilia* sequences.

Orthologous sequences for the tricistronic snR57-snR55-snR61 snoRNA transcription unit were identified in *K. marxianus*, *Zygosaccharomyces rouxii*, and *P. sorbitophilia*. Although no AGNN stem-loop can be predicted from the *S. cerevisiae* sequence (unpublished data), polycistronic processing sites could be predicted in all of these three species (Fig. 3), either between snR57 and snR55 or between snR55 and snR61, strongly suggesting that these three snoRNAs are also expressed from a polycistronic precursor. Orthologs of the
snR41-snR70-snR51 tricistronic expression unit were identified in *K. lactis* and *K. marxianus* (Table 2; Fig. 3), with short GGNN stem-loops between each species. While this sequence diverges from the AGNN consensus, it is known that Rnt1p can tolerate a G at the first position (8). Overall, these results suggest that the polycistronic mode of expression of some of the box C/D snoRNAs has been conserved in Hemiascomycetes and that the corresponding orthologs of RNase III play a role in the processing of these species. For some of the snoRNAs, no processing signal could be identified (e.g., between snR55 and snR61 for *K. lactis* and *P. sorbitophilia*). Despite the absence of processing signals between snR55 and snR61, it is possible that a single processing signal between snR55 and snR55 is sufficient to separate the three snoRNAs; that is, if the AGNN stem-loop located between snR55 and snR55 may be sufficient to separate the three snoRNAs from each other, as was observed for snR75 (17). For the other snoRNA species, the absence of the predicted AGNN tetraloop does not necessarily mean that these structures do not exist; previous computer searches for secondary structures in the polycistronic arrays revealed that the processing signals are usually harder to predict than are rRNA or 5′-end processing signals, even if they have been identified in vivo (4). Therefore, the absence of predicted sites does not necessarily mean that they do not exist in vivo.

**Conservation of 3′-end processing signals.** The 3′-end processing of snRNAs and snoRNAs by RNase III is clearly a dispensable processing pathway. This conclusion is suggested by the fact that the 3′ ends of most snoRNAs that are processed by Rnt1p can be generated by exonucleases or by the polyadenylation machinery in the absence of Rnt1p (1, 5, 15, 21, 24). For example, the U5 snRNA is present in *S. cerevisiae* with two forms; the longer one is processed through a Rnt1p-dependent pathway, while the shorter form is processed through a Rnt1p-independent pathway (5). *S. exiguus* also exhibits these two forms, while *S. kluveyeri, K. lactis,* and *Y. lipolytica* exhibit only one form (19). These observations suggest that for these three species, the dependence on RNase III for processing has been lost. Eleven RSTs were obtained with sufficient 3′-end-flanking sequences for spliceosomal snRNAs or the U3 snoRNA. The conservation of stem-loop structures capped by AGNN tetraloops was not as strong for these 3′-end processing signals as those described previously for 5′-end processing signals (Table 2; Fig. 4). Only one AGUU stem-loop was found for *Z. rouxii* U5 snRNA, while no AGNN stem-loop could be predicted in RSTs from *K. marxianus* and *K. lactis*. This negative result is consistent with the observation that only one form of U5 is present in *K. lactis*, probably processed through an RNase III-independent pathway. Similarly, only one species out of three showed an AGNN stem-loop for the

**FIG. 4.** Predicted secondary structures in the 3′ extensions of snRNAs and of the U3 snoRNA. The estimated distances in nucleotides to the mature 3′ ends of the sno(o)RNAs are indicated.

**FIG. 5.** Phylogenetic distribution of AGNN stem-loop processing signals in Hemiascomycetes. The cladogram of Hemiascomycetes was constructed from rDNA sequences of the 25S rRNAs by using phylogenetic distances obtained from reference 22 and with *S. pombe* as an outgroup. Asterisks indicate the presence of at least one predicted AGNN stem-loop processing signal. Open squares indicate that no RST was found for this category of processing signal in the particular species. Open circles indicate that while at least one RST was available, no AGNN stem-loop could be detected.
U4 snRNA (Table 2; Fig. 4). The conservation of the stem-loop structure seems stronger for the U3 snoRNA, for which both K. lactis and Z. rouxii exhibited AGNN stem-loops (Fig. 4). Overall, these phylogenetic observations strengthen previous biochemical data suggesting that 3'-end processing of snRNAs by RNase III is a redundant pathway that may have been lost during the recent evolution of yeast species.

**DISCUSSION**

In this study, I show that RNA processing signals obeying S. cerevisiae RNase III specificity rules are present in most Hemiascomycetes species analyzed in the Généoleures program. Predicted RNase III cleavage sites are present in all four categories of processing signals described so far: 3' ETS, 5' end, 3' end, and polycistronic (Fig. 5). The most complete phylogenetic picture was obtained for the rRNA 3' ETS sequences, since they were found in abundance in the Généoleures RSTs and since they are easily identified with a standard BLAST analysis due to the strong conservation of the 25S rRNA sequence. Strikingly, Rnt1p-type stem-loops were present in all species but C. tropicalis and Y. lipolytica (Table 1; Fig. 1). In the case of snoRNAs, orthologous sequences were harder to obtain for evolutionarily distant species such as S. cerevisiae when a cladogram is established by using rRNA sequences (Fig. 5) and seems to have diverged early from other Hemiascomycetes species. Based on these observations, two scenarios are possible. It is possible that the specificity of RNase III for dsRNA capped by AGNN tetraloops was present in the common ancestor of all these Hemiascomycetes but that the Y. lipolytica RNase III somehow evolved and lost the specificity for the tetraloops even though this specificity was conserved in all other species. Alternatively, it is possible that the specificity for AGNN tetraloops was acquired after the divergence between Y. lipolytica and all other Hemiascomycetes species analyzed. In both cases, the events require a rapid coevolution of the enzyme specificity and the substrate sequences. While the second scenario seems more parsimonious since S. pombe Pac1p does not rely on AGNN tetraloop recognition, further analysis of RNase III processing signals in more distant Ascomycetes species will possibly reveal which is the more likely evolutionary scenario.

**ACKNOWLEDGMENTS**

I thank the Généoleures program for their sequencing effort and C. Collinier for critical reading of the manuscript. This work was supported by NIGMS grant R01 GM-61518 from the NIH.

**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. 17:3738–3746.
2. Abou Elela, S., H. Igel, and M. Ares, Jr. 1996. RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. Cell 85:115–124.
3. Allmang, C., and D. Tollervey. 1998. The role of the 3' external transcribed spacer in yeast pre-rRNA processing. J. Mol. Biol. 278:57–78.
4. 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. USA 97:3142–3147.
5. Chanfreau, G., S. A. Elela, M. Ares, Jr., and C. Guthrie. 1997. Alternative 3'-end processing of U5 snRNA by RNase III. Genes Dev. 11:2741–2751.
6. Chanfreau, G., P. Legrain, and A. Jacquier. 1998. Yeast RNase III as a key processing enzyme in small nuclear RNAs metabolism. J. Mol. Biol. 284: 975–988.
7. Chanfreau, G., G. Rotondo, P. Legrain, and A. Jacquier. 1998. Processing of a dicistronic small nuclear RNA precursor by the RNA endonuclease Rnt1. EMBO J. 17:3726–3737.
8. Danin-Kreiselman, M., C. Y. Lee, and G. Chanfreau. 2003. RNase III-mediated degradation of unspliced pre-mRNAs and lariat introns. Mol. Cell 11:1279–1289.
9. 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.
10. Grishok, A., A. E. Pasquinelli, D. Conde, N. Li, S. Parrish, I. Ha, D. L. Baillie, A. Fire, G. Ruvkun, and C. C. Melo. 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell 106:23–34.
11. Hutvagner, G., J. McLachlan, A. E. Pasquinelli, E. Balint, T. Tuschi, and P. D. Zamore. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293:834–838.
12. Ketting, R. F., S. E. Fischer, E. Bernstein, T. Sijen, G. J. Hannon, and R. H. Plasterk. 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. 15:2654–2659.
13. 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 5:909–917.
14. Mathews, D. H., J. Sabina, M. Zuker, and D. H. Turner. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. 288:911–940.
15. Morlando, M., P. Greco, B. Dichtl, A. Fatica, W. Keller, and I. Bozzoni. 2002. Functional analysis of yeast snoRNA and snRNA 3’-end formation mediated by uncoupling of cleavage and polyadenylation. Mol. Cell. Biol. 22:1379–1389.
16. Park, W., J. Li, R. Song, J. Messing, and X. Chen. 2002. CARPELFAC- TORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr. Biol. 12:1484.
17. 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. Buchelerie. 1999. Seven novel methylation guide small nucleolar RNAs are processed from a common polycistrionic transcript by Rat1p and RNase III in yeast. Mol. Cell. Biol. 19:1144–1158.
18. Reinhart, B. J., E. G. Weinstein, M. W. Rhoudes, B. Bartel, and D. P. Bartel. 2002. MicroRNAs in plants. Genes Dev. 16:1616–1626.
19. Roiha, H., E. O. Shuster, D. A. Brow, and C. Guthrie. 1989. Small nuclear RNAs from budding yeasts: phylogenetic comparisons reveal extensive size variation. Gene 82:137–144.
20. Rotondo, G., M. Gillespie, and D. Frendewey. 1995. Rescue of the fission yeast snRNA synthesis mutant snm1 by overexpression of the double-strand-specific PacI ribonuclease. Mol. Gen. Genet. 249:758–765.
21. Seidel, R. L., B. Zheng, A. Asuru, and B. C. Rymond. 1999. U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway. Nucleic Acids Res. 27:587–595.
22. Souciet, J.-L., M. Aigle, F. Artiguenave, G. Blandin, M. Bolotin-Fukuhara, E. Bon, P. Brothier, S. Casaregola, J. de Montigny, B. Dujon, P. Durrens, C. Gaillardin, A. Lépingle, B. Llorente, A. Malpertuy, C. Neuveglise, O. Ozier-Kalogeropoulos, S. Potier, W. Saurin, F. Tekai, C. Toffano-Noche, M. Wesołowski-Louvé, P. Wincker, and J. Weissman. 2000. Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies. FEBS Lett. 487:3–12.
23. Spasov, K., L. I. Perdomo, E. Evukine, and R. N. Nazar. 2002. RAC protein directs the complete removal of the 3’ external transcribed spacer by the PacI nuclease. Mol. Gen. Genet. 269:441–452.
24. 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. 20:441–452.
25. Wu, H., H. Xu, L. J. Miraglia, and S. T. Crooke. 2000. Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. J. Biol. Chem. 275:36957–36965.
26. 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. 20:7240–7249.
27. Zhang, K., and A. W. Nicholson. 1997. Regulation of ribonuclease III processing by double-helical sequence antideterminants. Proc. Natl. Acad. Sci. USA 94:13437–13441.
28. Zhou, D., D. Frendewey, and S. M. Lobo Ruppert. 1999. Pac1p, an RNase III homolog, is required for formation of the 3’end of U2 snRNA in Schizosaccharomyces pombe. RNA 5:1083–1098.
29. Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31:3406–3415.