ELSEVIER

Biochimie 84 (2002) 775–790

BIOCHIMIE

Review

The expanding snoRNA world

Jean-Pierre Bachellerie [a,*], Jérôme Cavaillé [a], Alexander Hüttenhofer [b]

[a] *Laboratoire de Biologie Moléculaire Eucaryote du CNRS, Université Paul-Sabatier, 118, route de Narbonne, 31062 Toulouse cedex 4, France*
[b] *Universität Münster, Institut für Experimentelle Pathologie/Molekulare Neurobiologie, Von-Esmarch-Str. 56, 48149 Münster, Germany*

Received 25 February 2002; accepted 13 March 2002

Abstract

In eukaryotes, the site-specific formation of the two prevalent types of rRNA modified nucleotides, 2'-*O*-methylated nucleotides and pseudouridines, is directed by two large families of snoRNAs. These are termed box C/D and H/ACA snoRNAs, respectively, and exert their function through the formation of a canonical guide RNA duplex at the modification site. In each family, one snoRNA acts as a guide for one, or at most two modifications, through a single, or a pair of appropriate antisense elements. The two guide families now appear much larger than anticipated and their role not restricted to ribosome synthesis only. This is reflected by the recent detection of guides that can target other cellular RNAs, including snRNAs, tRNAs and possibly even mRNAs, and by the identification of scores of tissue-specific specimens in mammals. Recent characterization of homologs of eukaryotic modification guide snoRNAs in Archaea reveals the ancient origin of these non-coding RNA families and offers new perspectives as to their range of function. © 2002 Société française de biochimie et biologie moléculaire / Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

Keywords: SnoRNA; 2'-*O*-methylation; Pseudouridine; Imprinting; Archaea

1. Introduction

The biogenesis of eukaryotic ribosomes in the nucleolus involves an intricate series of pre-rRNA processing steps. This results in the removal of extended spacer regions from the primary transcript and production of stoichiometric amounts of mature small and large subunit rRNAs. Before its cleavage by endo- and exonucleases, the nascent pre-rRNA undergoes a complex pattern of nucleoside modifications of its mature small subunit (SSU) and large subunit (LSU) sequences. These modifications are of two prevalent types, 2-*O*-ribose methylation or pseudouridylation, each involving about 50–100 sites per eukaryotic ribosome [1–3]. They are exclusively located within the most conserved, functionally important domains of mature RNAs, particularly into the structural elements contributing to the peptidyl-transfer region and its vicinity [4] and their positions are largely (but not perfectly) conserved among distant eukaryotes. While these modifications of elusive role are not absolutely essential, they are likely to fine-tune rRNA folding and interactions with ribosomal proteins, thereby modulating both the biogenesis and activity of the ribosomes. The two types of eukaryotic rRNA modifications are directed by two large families of snoRNAs (small nucleolar RNAs) which specify the sites to be modified, in both cases through the formation of a specific duplex at the rRNA modification site, while the catalytic function is provided by a common protein enzyme, methylase or pseudouridine synthase, associated with the snoRNA. Since a single snoRNA guide can direct one, or at most two rRNA modifications, the number of these RNA species was expected until recently to approach 200 in vertebrates.

However, recent studies show that the complexity of the two snoRNA guide families has been largely underestimated. Their diversity relates not only to their genetic organization and biosynthesis but also to the existence of variant snoRNA structures and multiple cellular RNA targets, reflecting a range of cellular functions beyond ribosome biogenesis. Novel members of the modification guide snoRNA families target spliceosomal snRNAs in vertebrates, tRNAs in Archaea and even probably eukaryotic mRNAs, in addition to rRNAs. Remarkably, an increasing number of “orphan” guides without known RNA targets have been identified. Several of them are expressed in a tissue-specific fashion and submitted to genomic imprinting, adding another level of complexity to the biological roles of snoRNA guides in mammals. Meanwhile, the

* Corresponding author. Tel.: +33-5-61-33-59-34; fax: +33-5-61-33-58-86.
E-mail address: bachel@ibcg.biotoul.fr (J.P. Bachellerie).

© 2002 Société française de biochimie et biologie moléculaire / Éditions scientifiques et médicales Elsevier SAS. All rights reserved.
PII: S0300-9084(02)01402-5

identification of homologs of guide snoRNAs in organisms lacking a nucleus, Archaea, provides further insights into the evolutionary origin and function of these two large families of non-coding RNAs.

Fundamental properties of snoRNA guides for rRNA modification have been reviewed previously [2,3,5–9] and the reader is referred to these articles for further information on their gene organization, biogenesis and guide function, as well as to reviews dealing with pre-rRNA processing and nucleotide modification in eukaryotes [1,10–13]. In this article, following an updated summary of the properties of rRNA modification guides, we will focus on recent breakthroughs revealing the unanticipated structural and functional diversity of the two families of guide snoRNAs in organisms ranging from Archaea to Eukarya.

2. The two families of rRNA modification guides

2.1. Structure and function

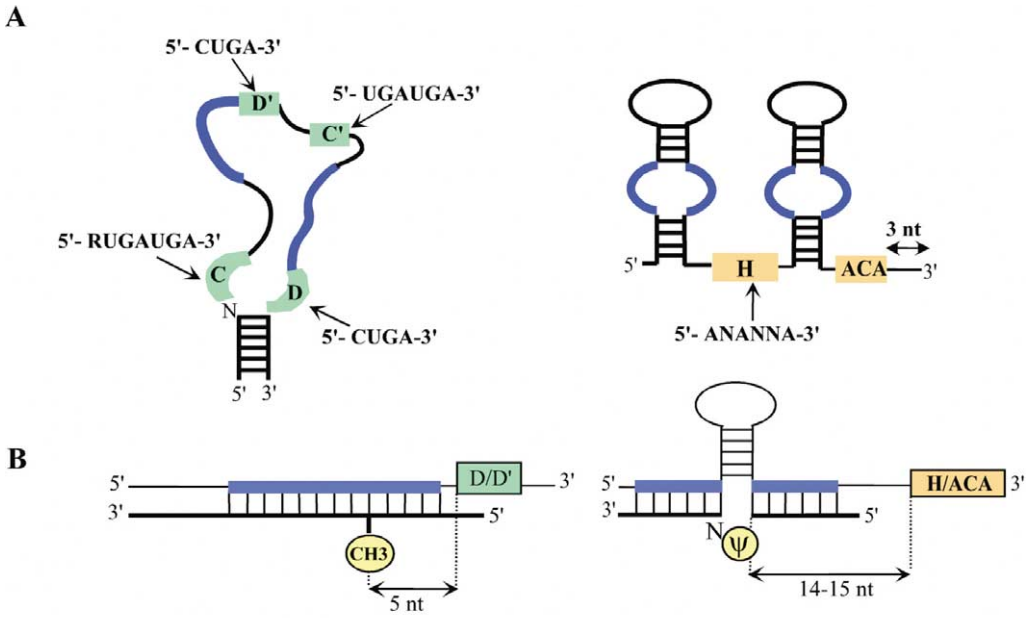
Except for the RNA component for RNase MRP, all snoRNAs to date fall into two major classes, antisense box C/D and box H/ACA snoRNAs, based on the presence of short consensus sequence motifs [14]. Most members of the two snoRNA families guide the 2'-*O*-ribose methylations and pseudouridylations, respectively, of rRNA. It is noteworthy, however, that a handful of them are involved instead in pre-rRNA cleavages [12]. The two snoRNA guide families have been identified in a wide spectrum of eukaryal species, ranging from metazoans to yeasts, plants, and kinetoplastid protozoans.

2.1.1. Methylation guides

Antisense box C/D snoRNAs contain two short sequence motifs, box C (5'PuUGAUGA3') and box D (5'CUGA3'), located only a few nucleotides away from the 5' and 3' ends, respectively (Fig. 1A). The two motifs are generally brought together in a typical 5'-3' terminal stem-box structure, involving the 4–5 nt at both termini, which is critical for snoRNA biogenesis and nucleolar localization [2,15–19]. They also contain another, less conserved copy of the box C motif, box C', in their central portion [20], and an additional box D motif, termed box D', in their 5' half [21]. Finally, they exhibit immediately upstream from box D and/or box D' one or two of the so-called antisense elements, i.e. sequence tracts 10- to 21-nt in length that are complementary to a site of rRNA 2'-*O*-ribose methylation [22–24]. Methylation guide function of box C/D antisense snoRNAs and the essential role of the CUGA box motif in determining the precise nucleotide to be methylated in the RNA duplex, at the fifth position upstream from box D or box D' (Fig. 1B), have been experimentally demonstrated [23,24]. Expression of an artificial box C/D snoRNA carrying an appropriate antisense element is sufficient to target a novel ribose methylation on the predicted pre-rRNA nucleotide and also, to a lesser extent, to RNA-polymerase II transcripts. This shows that the antisense element associated with box D (or D') is the sole determinant of the site of methylation [24]. More recently, based on a set of conserved box motifs and precisely positioned, relatively long rRNA complementarities, computational genomic searches have helped identify a virtually full complement of rRNA methylation guides in two eukaryal species with a compact, completely sequenced genome, yeast *S. cerevisiae* [25] and plant *A. thaliana* [26–28]. Presently, cognate box C/D snoRNAs are known for 51 of the 55 ribose methylated sites in *S. cerevisiae* rRNA [25]. In *A. thaliana*, 66 C/D snoRNAs potentially able to direct a total of 86 rRNA ribose methylations have been identified, but the precise number of 2'-*O*-methylated rRNA nucleotides remains to be assessed in this organism [26–28]. In mammals, while a computer search of databases have helped identify many antisense C/D snoRNAs in the early phase of this research [22,29–31], further progress has essentially resulted from experimental screens [23,32]. A screen for small non-mRNAs in mouse based on an EST-like sequencing approach has recently detected 72 new C/D and 41 new H/ACA snoRNAs, most of them targeting rRNA [32]. As a result, snoRNA guides for 93 of the 105–107 2'-*O*-methylated sites in mammalian rRNAs are now known. Scores of rRNA methylation guides have also been identified in *Drosophila* [33] and *Trypanosoma* [34–36]. Interestingly, a universally conserved LSU rRNA ribose methylation with a predicted snoRNA guide, corresponding to Um4458 in human 28S rRNA, is catalyzed by a site-specific methyltransferase in bacteria and mitochondria [37,38].

2.1.2. Pseudouridylation guides

Box H/ACA snoRNAs share a common secondary structure consisting of two large hairpin domains linked by a hinge and followed by a short tail (Fig. 1A). Conserved motifs termed boxes H (ANANNA, where N stands for any nucleotide) and ACA (a trinucleotide always found three nucleotides away from the 3' end) are located in the hinge and tail, respectively [14,39]. Each H/ACA snoRNA contains an appropriate bipartite guide sequence in the internal loop of one (or both) of the two large hairpin domains [3,40]. The two stems forming the 9–13 bp bipartite guide duplex precisely flank the substrate uridine which remains accessible for isomerization (Fig. 1B). Reminiscent of the target/box-D spacing rule observed for methylation guides, the conserved distance between the target uridine and the downstream H or ACA box of the snoRNA, 14–16 nt, is a critical determinant of the pseudouridylation site [40,41]. Genomic search of pseudouridylation guides, in contrast to methylation guides, is severely limited by their shorter box motifs and bipartite (instead of single, continuous) antisense elements. No fruitful genomic search has been reported so far for yeast *S. cerevisiae* and only 15 yeast H/ACA snoRNAs are known, which collectively target 19 of the 44 *S. cerevisiae* rRNA pseudouridines [40,42]. In mammals, guides for 42 of the 91–93 rRNA pseudouridines [43] have



C

	C/D snoRNPs		H/ACA snoRNPs	
	Eukarya	**Archaea**	**Eukarya**	**Archaea**
Core proteins	Nop1p (Fibrillarin)	Fibrillarin	Cbf5p (Dyskerin)	Cbf5
	Nop56p (p62)	Nop56 /Nop58	Gar1p	Gar1
	Nop58p		Nop10p	Nop10
	snu13p (15.5 kDa)	rpL7Ae	Nhp2p	rpL7Ae
Transiently-associated proteins	Srp40p (Nopp140)		Srp40p (Nopp140p)	
	SMN		SMN	
	p50 (rvb2)	p50/p55		
	p55 (rvb1)			
RNA targets	rRNAs, snRNAs	rRNAs	rRNA, snRNAs	
	mRNAs (?)	tRNAs	SL RNA (trypanosome)	rRNAs
			other RNAs ?	

Fig. 1. Structural features of the two families of modification guide snoRNAs. (A) Schematic secondary structures of the C/D and H/ACA classes of eukaryotic snoRNAs, with indication of their conserved box motifs (in green and orange boxes, respectively) and sequence tracts complementary to the cognate RNA target, also termed antisense elements (thick blue lines). The nucleotide targeted for modification is denoted by a yellow circle. For C/D snoRNAs (left), the 5'-3' terminal stem allowing the formation of the box C/D structural motif is represented. (B) Canonical structure of each type of guide RNA duplex. (C) Sets of associated proteins and categories of cellular RNA targets identified so far. Archaeal homologs of the two sets of snoRNP proteins are indicated.

been identified [32,40]. In yeast and mammals, the two hairpin domains are essential for rRNA pseudouridylation, even when the H/ACA snoRNA contains a single guide sequence [41]. In *Trypanosoma*, however, the first characterized H/ACA snoRNA consists of a single-hairpin structure, representing the shortest rRNA pseudouridylation guide described so far in Eukarya [44].

2.2. Biogenesis and assembly of snoRNP particles

2.2.1. SnoRNA gene organization and expression

Both snoRNA families are closely related by their unusual genomic organization and modes of biosynthesis. In vertebrates, they are encoded within introns and are not independently transcribed but processed from the pre-mRNA introns, in most cases by exonucleolytic digestion of the debranched lariat. Outside Metazoa, guide snoRNAs exhibit a more diverse gene organization. In yeast *S. cerevisiae*, only a few of them are intronic and most are synthesized from independent mono-, di-, or polycistronic RNA transcripts processed by endo- and exonucleases [45–47]. Occurrence of clusters of multiple different snoRNA genes, transcribed as polycistronic precursors, sometimes encoding both box C/D and box H/ACA types, from which individual snoRNAs are processed, is also widespread in higher plants and in *Trypanosoma* [26–28,34,44,48,49]. Intriguingly, U86, a novel yeast C/D snoRNA which has a human intronic homolog, is encoded within an open reading frame and its synthesis

appears to be alternative to that of the cotranscribed mRNA [50]. Finally, detection of archaeal homologs of box C/D snoRNAs (see below) also points to the peculiar aspects in their biosynthesis, with the coding sequences of several of them partially overlapping upstream and/or downstream open reading frames.

Most genes hosting intronic snoRNA guides for rRNA modifications code for proteins involved in ribosome biosynthesis or function, which suggests that this particular gene organization might provide a regulatory link between partners in the same biological process. However, the hypothesis cannot be generalized: several intronic snoRNA guides are hosted by different genes in different eukaryotes [51–53] and an increasing number of vertebrate host genes lack any direct relationship with translation, some of them even being apparently devoid of protein-coding potential [21,54–56]. However, all vertebrate genes hosting an intronic snoRNA guide for rRNA modification belong to the family of actively transcribed housekeeping genes termed 5' TOP (terminal oligopyrimidine) genes, which could provide the basis for a coordination of snoRNA biosynthesis at the transcriptional level [55,56]. In yeast too, promoter regions of mono- or polycistronic snoRNA genes and genes hosting intronic snoRNAs share common control elements, pointing to coordinated transcription [47].

Whether the snoRNA genes are independently transcribed or are intron-encoded, correct processing and nucleolar localization of the mature guide snoRNAs are directed by conserved boxes C/D and H/ACA through bound snoRNP proteins [15–18,57–59].

2.2.2. Associated proteins

Both types of guide snoRNAs function as small ribonucleoprotein particles (snoRNPs), each one consisting of a site-specific snoRNA associated with a small set of proteins common to each guide family (Fig. 1C).

Box C/D snoRNPs contain four evolutionarily conserved, essential proteins, fibrillarin (Nop1p), Nop56p, Nop58p and Snu13p. Fibrillarin, which exhibits amino acid sequence motifs characteristic of SAM-dependent methyltransferases [60], is the likely snoRNA-guided modifying enzyme [61], as point mutations in the methylase-like domain disrupt all rRNA methylations [62]. Snu13p protein, the yeast homolog of mammalian 15.5 kD protein, binds specifically to the C/D structural motif, termed a K-turn, in which an asymmetric, 3 nt internal loop is flanked by a regular stem on one side and two non-Watson-Crick, sheared G.A pairs on the other [63–65]. This protein which binds the same structural motif in U4 within the U4/U6.U5 tri-snRNP complex [66] is also an integral component of this spliceosomal complex, raising the interesting possibility that mRNA splicing and snoRNA synthesis may be linked. Two other core proteins of C/D snoRNPs, Nop56p and Nop58p, are highly related to each other in sequence [67,68].

Proteins common to H/ACA snoRNPs include evolutionarily conserved proteins Cbf5p (dyskerin), Gar1p, Nhp2p and Nop10p, all essential for the pseudouridylation reaction. Cbf5p is likely to correspond to the catalytic component of the H/ACA-snoRNA-guided modification, based on the presence of signature motifs for pseudouridine synthase and deleterious effects of point mutations in these motifs [69,70]. Mutations in the human homolog of Cbf5p, dyskerin, are a cause of dyskeratosis congenita, an inherited human disease first proposed to result from ribosome dysfunction [71]. However, another form of the disease linked to an altered telomerase activity has been recently characterized [72], suggesting that the critical effect of dyskerin mutations in the pathology rather reflects the specific binding of the Cbf5p homolog to telomerase RNA, which also contains an H/ACA domain (see below). Based on amino acid sequence homologies, Cbf5p appears closely related to TruB, which catalyzes the pseudouridine formation in the T loops of virtually all tRNAs. The crystal structure of TruB bound to RNA suggests that both enzymes recognize RNA in a similar manner [73]. Nhp2p, Nop10p and Cbf5p constitute the core of H/ACA snoRNPs [74,75].

Recently, a pair of highly conserved nucleoplasmic proteins related to each other which have both DNA helicase activity and are linked with chromatin remodeling and transcription, p50 (Rvb2) and p55, were observed to bind to a model box C/D snoRNA in vitro, suggesting that they function at an early stage of snoRNP biogenesis in the nucleoplasm [67,76]. Remarkably, depletion of p50 impairs assembly or trafficking of C/D snoRNPs and H/ACA snoRNPs as well. In each snoRNP type, two core proteins, Snu13p and Nhp2p, respectively, are strikingly related to each other, particularly in their middle portion, which might provide a common element mediating coordinated biogenesis of the two snoRNA families [77]. In yeast *S. cerevisiae*, Snu13p and Nhp2p exhibit significant similarity to ribosomal protein L30, raising the possibility of a further link between ribosome and snoRNP production. Snu13p and Nhp2p share a common archaeal homolog, ribosomal protein L7Ae, pointing to a common evolutionary origin for the two families of modification guide snoRNAs (see below).

3. A wide range of cellular RNA targets for eukaryotic snoRNA guides

3.1. Modification guides for spliceosomal snRNAs

Mammalian U1, U2, U4, U5 and U6 snRNAs contain a very substantial number of 2'-*O*-methylations and pseudouridylations, amounting collectively to 30 and 24, respectively [78]. Interestingly, these modifications are mainly located in the snRNA segments involved in intermolecular RNA–RNA interactions or conformational switches during spliceosome assembly and function, suggesting that they play an important role in splicing control [78]. In line with this notion, modifications in the 5' terminal region of

vertebrate U2 snRNA are required for snRNP assembly and pre-mRNA splicing [79].

Antisense box C/D snoRNAs able to guide snRNA 2'-*O*-methylations have been first reported in the case of U6 in vertebrates [80,81]. The homolog of one of these, mgU6-47, has been recently identified in *Drosophila* and fission yeast *S. pombe* [82]. The disruption of the *S. pombe* mgU6-47 gene demonstrated this snoRNA is absolutely required for site-specific 2'-*O*-methylation of U6. Although the cognate U6 methylation is not essential, it seems to be involved in a fine-tuning of mRNA splicing [82]. Several snoRNA guides for the formation of additional snRNA ribose methylations not only in U6, but also in U2 and U4 snRNAs have been recently detected through a general screen for small non-mRNAs in mouse [32]. Several H/ACA snoRNAs that are able to guide pseudouridylations in mouse U2 and U6 snRNAs have also been identified in the same experimental screen [32]. Meanwhile, an outstanding C/D-H/ACA chimeric snoRNA, U85, directing 2'-*O*-methylation of position 45 and pseudouridylation of uridine 46 in the invariant loop of U5 snRNA has been identified in human and *Drosophila* [83]. Current observations suggest that other chimeric C/D-H/ACA snoRNAs guides for the 2'-*O*-methylation and pseudouridylation of vertebrates snRNAs might exist [84].

Yeast *S. cerevisiae* snRNAs, the 2'-*O*-methylations of which have not been investigated, contain considerably less pseudouridines than vertebrates snRNAs, with only three pseudouridines in U2 snRNA and no pseudouridine at all in U6 snRNA, in contrast to the 12 and 3 pseudouridines, respectively, present in their vertebrate homologs [78,85]. No snoRNA guide for a yeast snRNA pseudouridylation has been reported so far. The three pseudouridines in yeast U2 are conserved in vertebrates, corresponding to ψ34, ψ41 and ψ43 in human U2. Remarkably, although one of these sites, ψ34, has a cognate H/ACA snoRNA guide in mammals, analysis of yeast mutants showed snRNA pseudouridylations in yeast do not involve Cbf5p, the rRNA-pseudouridine synthase catalyzing the H/ACA-guided reaction [69,85]. This suggests that either mechanisms of ψ34 formation in U2 snRNA are dramatically different between yeast and vertebrates or yeast has a redundant, entirely distinct biosynthetic pathway for this particular modification. Interestingly, the formation of another conserved pseudouridine in yeast U2, corresponding to ψ43 in vertebrate U2, is catalyzed by Pus1p, a pseudouridine synthase which also modifies tRNAs [85].

The Cajal body (or coiled body, CB) is a nuclear organelle frequently located close to the nucleolus which appears involved in the biogenesis of snRNAs and snoRNAs as well [18,86,87]. Newly synthesized snRNAs are proposed to transit through the CBs before and after their cytoplasmic stage of snRNP assembly, before reaching the nuclear structures where splicing is detected, the speckles [88]. Recent in situ hybridizations have shown that U85, the chimeric C/D-H/ACA snoRNA guiding modifications in U5 snRNA, as well several other snoRNAs predicted to guide modifications of snRNAs U1, U2, U4 and U5 colocalize with the CBs in human HeLa cells, raising the possibility that modifications of these snRNAs are taking place in the CBs [151]. However, the results of injection experiments into the Xenopus oocyte suggest that U2 internal modification occurs within the nucleolus [89].

3.2. Orphan guide snoRNAs

The finding that telomerase RNA in vertebrates contains a typical H/ACA domain [90] and that human H/ACA snoRNPs and telomerase share evolutionarily conserved proteins [91,92] expands the structural and functional diversity of the box H/ACA snoRNA motif, suggesting that some of the novel snoRNAs in this group might have unanticipated functions. No presumptive RNA target for pseudouridylation has been identified for the H/ACA domain of vertebrate telomerase RNA. Interestingly, the 3' terminal stem-loops of human telomerase RNA and U17 (see below) have some specific features that distinguish them from other vertebrate H/ACA, since they could by themselves reconstitute a snoRNP in vitro unlike their counterpart in other H/ACA snoRNAs [93]. Moreover, in the context of the detection of a chimeric C/D-H/ACA snoRNA [83], it is noteworthy that a couple of dramatically enlarged C/D snoRNAs have been identified in mouse and yeast, MBI-43 and snR63, respectively, which each guides a single rRNA 2'-*O*-methylation [25,32]. These extended snoRNAs might have a more complex function than the canonical rRNA methylation guides characterized so far.

An increasing number of ubiquitously expressed snoRNAs belonging to the box H/ACA or C/D structural families in mammals but devoid of antisense element to rRNA or snRNA have been reported recently [32,94]. A single yeast specimen in this category, a C/D snoRNA, has been detected [50]. Recent EST screen for mouse small non-mRNAs identified 15 C/D and 11 H/ACA orphan snoRNAs, the biological role of which remains elusive [32]. Many RNA species not directly related to ribosome biogenesis, including several mRNAs, have been reported to localize transiently to the nucleolus and the 5'-terminal processing of some yeast pre-tRNAs by RNAse P takes place in the nucleolus [95,96]. This suggests that a host of cellular RNAs might be targeted by these intriguing orphan modification guides. A search for potential target sites in the three stable non-coding RNAs trafficking through the nucleolus, telomerase RNA, RNase P and SRP RNA, was negative for all the orphan snoRNAs reported so far. Alternatively, some orphan snoRNAs could be exclusively involved in pre-rRNA cleavages instead of modification, similar to C/D snoRNAs U3, U8 and U22 [12,97,98] and yeast H/ACA snR10 and snR30 [99,100]. Remarkably, the vertebrate homolog of yeast snR30, U17, also involved in pre-rRNA cleavage, has a 3' domain exhibiting hallmarks of a pseudouridylation guide for a still unidentified target, point-

ing to a dual function of this snoRNA, as previously demonstrated for a few modification guides targeting rRNA, such as H/ACA snoRNAs E2 and E3 and C/D snoRNA U14, which all have an additional role in pre-rRNA cleavage [101,102].

3.3. Looking for mRNA targets

Trans-splicing is an RNA processing pathway where a common, relatively short terminal 5' exon, the spliced leader (SL RNA), is linked to all or to a variable fraction of mRNAs in an organism. The novel process, initially discovered in kinetoplastid protozoans, has been subsequently identified in widely different phyla of metazoans, including nematodes [103]. *Trypanosoma* SL RNA can be cross-linked to another small RNA, SLA1, that folds as an H/ACA snoRNA containing a single-hairpin structure. Recently, SLA1 was found to be potentially able to guide pseudouridine formation at a position of the spliced leader that is always pseudouridylated, position 12 relative to the 5' splice site of SL RNA [44]. Mutations introduced in the SL RNA to disrupt the presumptive canonical guide duplex involving SLA1 abolished pseudouridylation, as expected [104]. The conserved SL RNA pseudouridine might play a role in modulating SL-snRNA interactions during *trans*-splicing or, alternatively, in the export of spliced mRNAs from the nucleus or even during translation. Remarkably, SLA1 RNA is unique so far among modification guides in also binding to the Sm core protein [105]. It localizes both in the nucleolus and the nucleoplasm where it is likely to bind its target SL RNA, probably reflecting the shuttling of SLA1 between distinct nuclear sites.

The possibility that some of the above-mentioned orphan guide snoRNAs, which are ubiquitously expressed in mammals, might target an mRNA must be considered, although the presence of mRNA pseudouridines or 2'-*O*-methylations (external to the 5' cap structure) has not been reported so far. This hypothesis remains difficult to test by the search of appropriate complementarity in complex mammalian genomes, unless unusually long guide duplexes are involved. Screening of myriad candidate targets by comparative search of multiple mammalian genomes, based on conserved guide RNA duplexes, might provide clues on this exciting issue in the near future. Remarkably, among the novel brain-specific snoRNAs (see below), C/D snoRNA MBII-52 displays an unusually long (18 nt), phylogenetically conserved antisense element to an mRNA that is specifically expressed in the brain [106]. This snoRNA could play a key role in the processing of its presumptive RNA target, serotonin receptor 5-HT2c mRNA (Fig. 2). Strikingly, the mRNA nucleotide potentially targeted for 2'-*O*-methylation by snoRNA MBII-52 is also the subject of a physiologically important adenosine-to-inosine editing [107]. In vitro, 2'-*O*-methylation of the adenosine to be edited dramatically inhibits its deamination to inosine [108], suggesting for MBII-52 a role in the regulation of 5-HT2c mRNA editing (see also below). Alternatively or additionally, the same antisense element in snoRNA MBII-52 might also control the alternative splicing of 5-HT2c pre-mRNA at a nearby splice site, through steric occlusion instead of nucleotide modification (Fig. 2).

As discussed below, further study of an intriguing set orphan guide snoRNAs exclusively expressed in the brain might open new insights as to the potential range of functions of these unexpectedly large snoRNA families.

4. Brain-specific snoRNAs and genomic imprinting

In contrast to all known rRNA or snRNA modification guides and unlike most orphan snoRNAs, an increasing number of recently identified snoRNAs, mostly of the C/D

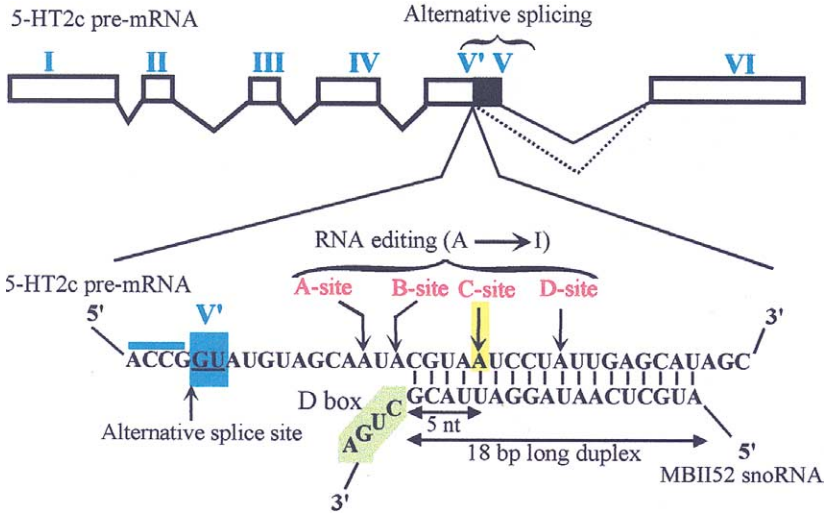


Fig. 2. A potential mRNA target for brain-specific MBII-52 C/D snoRNA. Top: Schematic structure of the serotonin receptor 5-HT2c pre-mRNA. Bottom: Potential base pairing between the antisense element of MBII-52 and a segment of 5-HT2c exon V undergoing adenosine-to-inosine editing at four sites, termed A–D. Based on the RNA duplex structure, RNA editing C-site (denoted in yellow) is predicted to be targeted for 2'-*O*-methylation by snoRNA MBII-52. The location of a nearby site of alternative splicing is also indicated.

family, exhibit a tissue-specific expression pattern, being mainly expressed within the brain [106,109–113]. Intriguingly, the genes of all of them are subjected to genomic imprinting, an epigenetic phenomenon that restricts gene expression to only one chromosome, either the paternal or the maternal allele [114]. An exception from that rule seemed to be the HBI-36/MBI-36 snoRNA which is also exclusively expressed in the brain, but is not subject to imprinting. However, expression of this snoRNA gene occurs still from one chromosome only, since its host gene the serotonin receptor 5-HT2c, locates to the X chromosome. Thus it appears as if all brain-specific snoRNAs detected so far are exclusively transcribed from one chromosome only. Around 50 imprinted genes have been described in mammals so far, most of them involved in the regulation of fetal/placental growth, cell cycle and brain development. While the evolutionary origin and biological significance of genomic imprinting is still under active debate, this process has an evolutionary cost. Since imprinted genes are functionally haploid, genetic alterations such as chromosomal deletions or uniparental disomy can lead to loss of expression of the active allele or abnormal expression of the silent allele, giving rise to various human genetic syndromes (Prader-Willi, Angelmans, Beckwith-Wiedemann), behavioral disorders (autism, Turner and Tourette syndromes) or even cancers (Wilm's tumors). Unexpectedly, imprinted snoRNA genes have been recently detected at two human chromosomal loci: 15q11q13 and 14q32 in which they share the same outstanding genomic organization [106,110,112,152]. The novel snoRNAs are intron-encoded like all guide snoRNAs previously reported in vertebrates. However, the snoRNA-containing intron and flanking exons are tandemly repeated to scores of copies (Fig. 3). The snoRNAs are processed from complex transcripts spanning the entire snoRNA repeat array which also give rise to spliced RNAs lacking open reading frames. The function of the spliced transcripts of the snoRNAs host genes remains elusive.

4.1. Paternally expressed snoRNA genes at human 15q11q13 (mouse 7C)

The first imprinted snoRNA genes, MBII-13, MBII-85 and MBII-52, have been identified through a systematic search for small non-messenger RNAs expressed in mouse brain [106]. MBII-85 has been also characterized by two other independent approaches [110,112]. Their genes map at mouse Chr. 7C and their human homologs, HBII-13, -85 and -52, at the syntenic 15q11q13 region (Fig. 3). HBII-52 and HBII-85 are arranged into two tandem arrays of 27 and 47 snoRNA genes copies, respectively, each one embedded within a ~2 kb long repeat unit [106,113]. HBII-13 snoRNA is structurally related to HBII-52 and encoded by a single gene copy. Within each of the two HBII-85 and HBII-52 clusters, copies of each snoRNA sequence are highly similar to each other (90% and 94% identity, respectively) whereas surrounding sequences in the repeat unit diverge substantially from each other. The snoRNA gene organization at the syntenic mouse locus is similar but its analysis remains incomplete [106,110,112]. While their flanking sequences diverge extensively between human and mouse, the three novel snoRNAs are strongly conserved, pointing to their biological importance. Human 15q11q13 locus is associated with two very different human neurological disorders: Prader-Willi (PWS) or Angelman syndromes (AS) that result from the loss of paternal or maternal gene expression within this region, respectively [115]. The three C/D snoRNAs are not expressed in a PWS patient (with a large paternal deletion of the whole imprinted locus) or in a mouse model mimicking some aspects of the PWS phenotype, indicating that they are only expressed from the paternal allele [106]. They are processed from a very large transcript (>460 kb), an antisense RNA to the maternally expressed UBE3A gene [113] which might regulate UBE3A paternal expression (Fig. 3). While the three imprinted snoRNAs are processed from the same transcript, only HBII-52 is strictly brain-specific suggesting the involve-

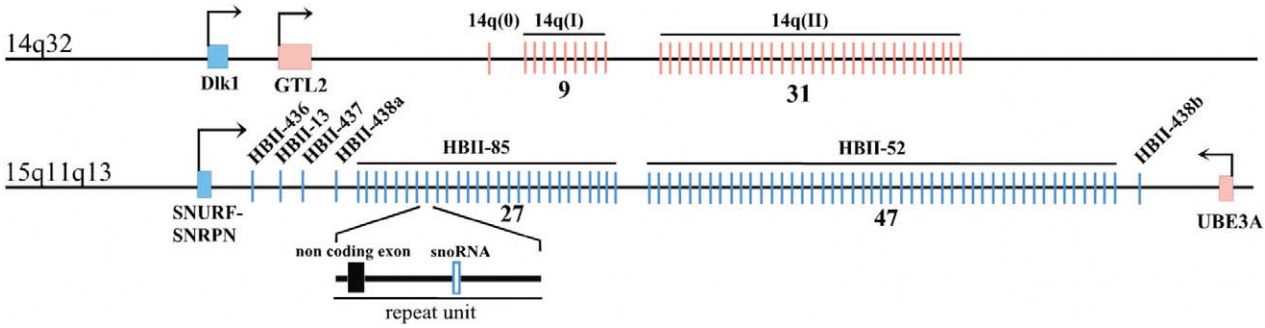


Fig. 3. Schematic representation of clusters of tandemly repeated snoRNA genes at two human imprinted loci. Only a portion of each imprinted locus, 14q32 and 15q11q13, is represented. Imprinted genes are depicted by blue and red boxes, for paternally expressed and maternally expressed genes, respectively, with their transcription start site and direction of transcription indicated by arrows. The paternally or maternally expressed C/D snoRNA genes are depicted by vertical bars (blue or red, respectively) with the name and copy number of each snoRNA in a cluster indicated above and below, respectively. The typical structure of a snoRNA-containing repeat unit is depicted in the case of snoRNA HBII-85 below the repeat array (the non-coding exon is denoted by a black box, and intronic sequences by a thick line).

ment of tissue-specific RNA processing events. Whether all the snoRNA gene copies are expressed is not known but HBII-85 and HBII-52 are among the most abundant C/D snoRNAs detected in mouse brain.

4.2. Maternally expressed snoRNA genes at human 14q32 (mouse distal 12)

Another brain-specific snoRNA, RBII-36, has been isolated by the screening of a rat C/D snoRNA library [109]. It is encoded within a previously described non-protein coding gene, Bsr, that spans an array of ~100 snoRNA-containing tandemly repeated units of 0.9 kb [116], strongly reminiscent of the snoRNA gene organization observed at the PWS locus. RBII-36 snoRNA is exclusively expressed in neurons, in which it exhibits a nucleolar localization. RBII-36 is generated by the superimposition of two mutually exclusive processes, the classical exonucleolytic processing of the debranched lariat and a splicing-independent pathway involving endonucleolytic cleavages within pre-mRNA. While RBII-36 and its Bsr host gene could only be detected in rat [109,116], the human and mouse loci syntenic to the rat Bsr-containing 6q32 locus, human 14q32 and mouse distal 12, respectively, also contain tandemly repeated arrays of novel, imprinted tissue-specific C/D snoRNAs [152]. In contrast to PWS-encoded snoRNAs, these intronic snoRNAs are only expressed from the maternal allele in mouse (Fig. 3). Interestingly, their antisense elements exhibit a substantial level of inter-copy divergence in each repeat array, supporting the notion that the tandemly repeated gene organization could provide the basis for a functional diversification of some C/D snoRNAs.

4.3. Origin and function of imprinted C/D snoRNAs

The various tandemly repeated C/D snoRNA families share a common 5'-3' terminal stem sequence (consensus: 5'GGACC....GGTCC 3') distinct from that of any of the previously reported, non-repeated C/D snoRNAs. This might indicate that they have all evolved from a common snoRNA ancestor gene in the early stages of mammalian evolution, possibly after retrotransposition of the ancestral snoRNA gene followed by a series of tandem duplications at the insertion locus. In agreement with the notion that retrotransposition events may have played a major role in the mobility of functional snoRNA genes [117], several typical C/D or H/ACA snoRNA retrogenes are found in the human and mouse genomes (J.P. Bachellerie, unpublished data). They exhibit the hallmark 3' poly(A)-tailed snoRNA coding sequence flanked by a pair of 10–15-nt direct repeats, reflecting the insertion of a likely cDNA intermediate at staggered nicks in the genome (Fig. 4). Interestingly, one of them is closely related to one of the repeated, imprinted C/D snoRNAs mentioned above, MBII-48. Many imprinted loci share common features including the presence of repeated sequences and non-coding RNA genes [118]. Non-coding RNAs are involved in multiple aspects of gene expression including dosage compensation phenomenon [119], and the snoRNA host genes, or even the snoRNAs themselves could play a direct role in the imprinting mechanism. Since protein components of C/D snoRNPs, p55 and p50, are linked with chromatin remodeling and with transcription complexes [76] and Nop56p/Nop58p interact with matrix-attached regions (MARs) in plants [120], these proteins could be targeted to a particular imprinted chromosomal locus by a high local concentration of C/D snoRNAs, independent of their antisense elements. Alternatively, since the two imprinted loci are associated with neurological defects and HBII-85 snoRNA is considered as a PWS candidate gene [121], the repeated snoRNAs might affect brain function and/or development by specific pairing to yet unknown RNA targets. However, except for HBII-52, which contains an 18-nt long antisense element to the brain-specific mRNA encoding the serotonin receptor 2C (see above), such targets remain to be identified. The only brain-specific H/ACA snoRNA reported so far [106], MBI-36 or HBI-36 (unrelated to the above-mentioned rat RBII-36) is also an orphan modification guide, but is not, however, imprinted and expressed from repeated genes. It is encoded in the second intron of the brain-specific 5-HT2c serotonin receptor gene, which is curiously the presumptive target of imprinted, brain-specific HBII-52 C/D snoRNA. Interestingly, HBI-36 and HBII-52 have an opposite localization in the brain: HBI-36 is exclusively expressed within choroid plexus (highly vascular structures arising from the wall of ventricles), the sole brain area in which expression of HBII-52 is not detected.

5. Archaeal modification guides, tRNA targets and archaeal splicing

In contrast to eukaryotes, the rRNA of typical bacterium *Escherichia coli* contains only four 2'-*O*-methylations and 10 pseudouridines and each of these modifications appears to be catalyzed by a site-specific protein enzyme, ribose methylase or pseudouridine synthase, without any RNA cofactor [38,122,123]. Among prokaryotic organisms, Archaea appear more closely related to Eukarya than to Bacteria by multiple aspects of the macromolecular machineries involved in DNA replication, transcription and translation [124]. A first assay of the extent of rRNA 2'-*O*-methylations in an archaeon, *Sulfolobus solfataricus*, has revealed a high number of methylations, 67, very similar to eukaryal rRNAs [125], pointing to the potential existence of an RNA-guided site-selection system in Archaea too.

5.1. Archaeal C/D guides for rRNA 2'-O-methylation

Consistent with this notion, homologs of eukaryotic C/D snoRNP core proteins, fibrillarin, Nop56/Nop/58 and Snu13p are present in the genomes of the two major lines of archaeal descent, Crenarchaeaotes and Euryarchaeotes

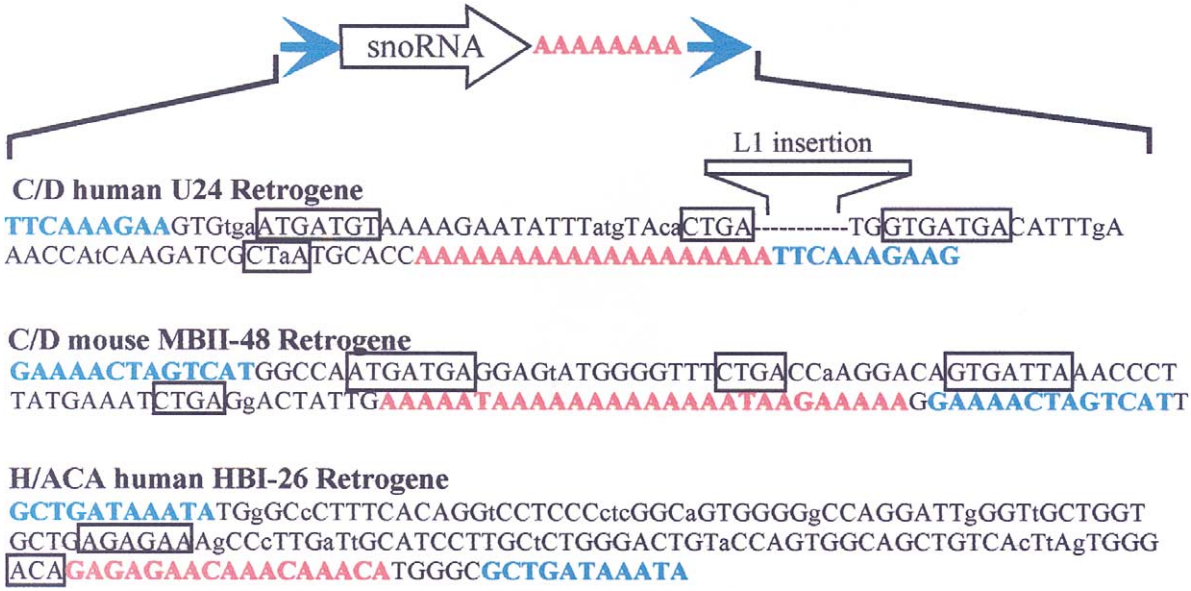


Fig. 4. Retrogenes of modification guide snoRNAs in mammalian genomes. The A-rich tract (in red) downstream from the snoRNA coding region and the pair of flanking directs repeats (in blue) are denoted. Within the snoRNA sequence, the hallmark box motifs are boxed and the nucleotide differences as compared to the functional gene in the same species denoted by lower case.

[63,126,127]. Archaeal homologs of box C/D snoRNAs have been recently discovered by two independent approaches involving genomic searches [128,129]. Although only a minority of the predicted 2'-*O*-methylations have been verified so far, current evidence suggests that most if not all archaeal C/D sRNAs (small RNAs) function as guides for rRNA ribose methylation obeying the same spacing rule as eukaryotic C/D snoRNAs. Through the computer screen of the three complete, closely related *Pyrococcus* genomes for hallmark box motifs combined with interspecies comparative analysis, 46 C/D sRNA species conserved among the three hyperthermophilic euryarchaea have been detected and experimentally identified [128]. In an independent study, 18 sRNAs harboring C/D hallmarks have been characterized in crenarchaeote *S. acidocaldarius*, through analysis of a cDNA library generated from small RNAs immunoprecipitated from total cell extracts with antibodies raised against cloned archaeal homologs of fibrillarin and Nop56p [129]. In the next step, these authors retrained a previously developed eukaryotic search program [25] based on the structural traits of experimentally identified *S. acidocaldarius* sRNAs. They detected 13 high-scoring candidates, 10 of which were experimentally verified, in closely related archaeon *S. solfataricus*, and predicted more than 50 candidates in each of the three *Pyrococcus*, most of which were confirmed independently [128]. Omer et al. [129] also predicted 23 untested candidates in *A. pernix* but only 8 and 4 (also untested) candidates in euryarchaea *M. jannaschii* and *A. fulgidus*, respectively. The very small number of C/D sRNAs predicted by genomic search in some Archaea might merely reflect the limitations of the search program and the presence of substantially different C/D sRNA structures in these species. However, a recent experimental screen for small non-messenger RNAs in hyperthermophilic euryarchaeote *A. fulgidus* could detect only one additional presumptive methylation guide for rRNA [153], suggesting that the extent of rRNA 2'-*O*-methylation may widely vary among Archaea, even among hyperthermophiles. The notion that high levels of rRNA methylation reflect the need for an increased thermodynamical stability of the ribosome [130] may therefore not apply to all archaeal groups.

Archaeal C/D sRNAs appear as paradigms of minimalist antisense C/D snoRNA, probably reflecting the size constraints on compact archaeal genomes. Slightly shorter than their eukaryotic counterparts, they exhibit an outstanding structural homogeneity, with almost identical sizes, extended consensus motifs for box C and C' and the quasi-systematic presence of two (instead of one) rRNA antisense elements [128–130]. Archaeal box C/D sRNAs are able to assemble into functional RNA-protein complexes in the eukaryotic nucleus and direct site-specific 2'-*O*-methylation of eukaryotic rRNA [131]. Remarkably, in each archaeal C/D sRNA, the pair of rRNA antisense elements always matches two target sites very close to each other in rRNA structure, suggesting the simultaneous formation of two guide duplexes which could reflect a chaperone function in the control of pre-rRNA folding. Likewise, many plant box C/D snoRNAs have dual antisense elements matching pairs of close rRNA sites [26]. Known genes of archaeal C/D sRNAs are not clustered but widely distributed in the genome. They are encoded in both DNA strands, usually mapping within the small portion of the archaeal genomes corresponding to the inter-ORF segments. A few C/D sRNAs, however, slightly overlap the upstream and/or downstream protein-coding region [128–130]. Although a part of these could reflect wrongly annotated ORFs, the possibility that some sRNAs are generated from indepen-

dent sRNA promoters embedded within an ORF or through differential processing of chimeric sRNA/mRNA transcripts remains to be tested. At least two closely linked *Pyrococcus* C/D sRNAs, sR12 and sR39, seem to be produced by the processing of a common sRNA precursor [128].

5.2. Methylation guides for tRNAs

A few predicted, and experimentally verified C/D sRNAs lacking the usual 9–12 nt long antisense element to 16S or 23S rRNA have been detected in several Archaea [129,130,132]. Most of them exhibit instead an antisense element to a tRNA in the same species. The presence of a 2'-*O*-methylation at the predicted tRNA position remains untested in most cases. However, the tRNA guide function of some of the novel C/D sRNAs seems most likely.

For *Pyrococcus* sR47, sR48 and sR49, the notion is supported by phylogenetic evidence [132]. Based on the box D + 5 nt spacing rule, each of them can target the wobble position in the anticodon of three different tRNA species, tRNA-Leu(CAA), tRNA-Leu(UAA) and tRNA-Met(elong), through 11-nt long antisense elements conserved in the three *Pyrococcus* species. Moreover, guide duplexes involving sR47 and sR49 are supported by the compensatory changes in distantly related Archaea, *A. fulgidus* and *M. jannaschii*, respectively, in which the cognate sRNA homolog has been identified [132]. As for sR48, it harbors a pair of tRNA antisense elements targeting two different sites, position 58 and the first position of the anticodon, in the same tRNA-Leu(UAA), similar to double-guide archaeal sRNAs typical of rRNA methylation. In contrast to sR47 and sR48, sR49 targets the precursor, instead of the mature tRNA [132]. Its guide duplex spans the 5' exon/intron junction in pre-tRNA-Met, suggesting that methylation of this tRNA takes place before splicing, like some eukaryotic tRNA modifications [133]. For a fourth tRNA methylation guide, sR50, which also targets a tRNA precursor, the predicted function is supported by both comparative and experimental evidence, as detailed below. A search of the yeast *S. cerevisiae* genome for candidate C/D snoRNAs able to target yeast tRNA 2'-*O*-methylations was negative. No likely tRNA target could be identified either for any of the mammalian C/D snoRNAs devoid of rRNA or snRNA complementarities mentioned above (C. Gaspin, J.P. Bachellerie, unpublished data). Archaeal C/D sRNAs targeting rRNAs or tRNAs bind directly and specifically ribosomal protein L7Ae, the archaeal homolog of eukaryotic C/D snoRNP protein Snu13p, apparently by recognizing same elementary features in the box C/D structural motif [77] (V. Segault, B. Charpentier, B.C. d'Orval, M.L. Bortolin, J.P. Bachellerie, C. Branlant, unpublished data).

5.3. Novel C/D sRNAs linked to archaeal splicing

Archaeal sR50, the novel C/D guide for a tRNA 2'-*O*-methylation, is outstanding because it corresponds to the intron of its presumptive target, tRNA-Trp [132]. Based on the box D + 5 nt spacing rule, it can direct in *cis* two distinct 2'-*O*-methylations in the unspliced pre-tRNA-Trp, on position 34, the first position of the anticodon, in the 5' exon and on position 39 in the 3' exon, through intramolecular guide duplexes both spanning an exon–intron junction. The presence of 2'-*O*-methylations at these tRNA-Trp positions had been previously reported in halophile *H. volcanii* [134]. The unusually large tRNA-Trp intron is so far the sole archaeal intron displaying box C/D sRNA hallmarks. A search of yeast tRNA introns for box C/D motifs was negative. In archaeal pre-tRNA-Trp, box motifs are conserved and the two intramolecular duplexes maintained through compensatory changes in a wide range of euryarchaea, pointing to the biological importance of the two 11- and 8-bp long interactions. The guide function of the box C/D intron has been verified in *H. volcanii*, through mutagenesis of the box motifs, using an in vitro splicing/RNA modification assay in which the two 2'-*O*-methylations were faithfully reproduced [132]. Molecular mechanisms of the intramolecularly guided methylations and their potential roles in a control of tRNA splicing remain to be assessed. In vitro, pre-tRNA-Trp splicing does not seem to depend on nearby 2'-*O*-methylations at positions 34 and 39 [132]. However, an independent analysis of various pre-tRNA-Trp mutants in *H. volcanii* suggests that these modifications might affect splicing efficiency in vivo [135]. Interestingly, the hallmark BHB (bulge-helix-bulge) structural motif required for the splicing of archaeal introns cannot form in pre-tRNA-Trp until the pair of intramolecular guide duplexes involving the intron is dissociated. Subsequent to ribose methylation of positions 34 and 39, major rearrangements of the pre-rRNA structure must therefore occur before splicing can take place. This outstanding biological system should help illuminate a still elusive aspect of the function of modification guides in Archaea and Eukarya as well, i.e. their additional, intrinsic role of RNA chaperones for the folding of cognate RNA targets [2,9]. Addressing this experimentally challenging point could provide insights on the biological significance of the nucleotide modification itself. In this context, it is noteworthy that an increasing number of RNA modifying enzymes, including 2'-*O*-methylases and pseudouridine synthases, appear essential for growth although the modifications they catalyze are dispensable, possibly reflecting a role of the modification process as a quality control mechanism [136–139].

Another intriguing archaeal C/D sRNA associated with splicing has been identified in a study revealing a further link between intron splicing and pre-rRNA processing in Archaea [140]. In the archaeal pre-rRNA primary transcript, the BHB motifs found in the long processing stems flanking pre-16S and pre-23S processing intermediates are cleaved by the splicing endonuclease. It was recently found that this cleavage is followed by ligation of the pre-rRNA spacers surrounding each pre-rRNA intermediate in the primary transcript, similar to religation of spliced archaeal exons

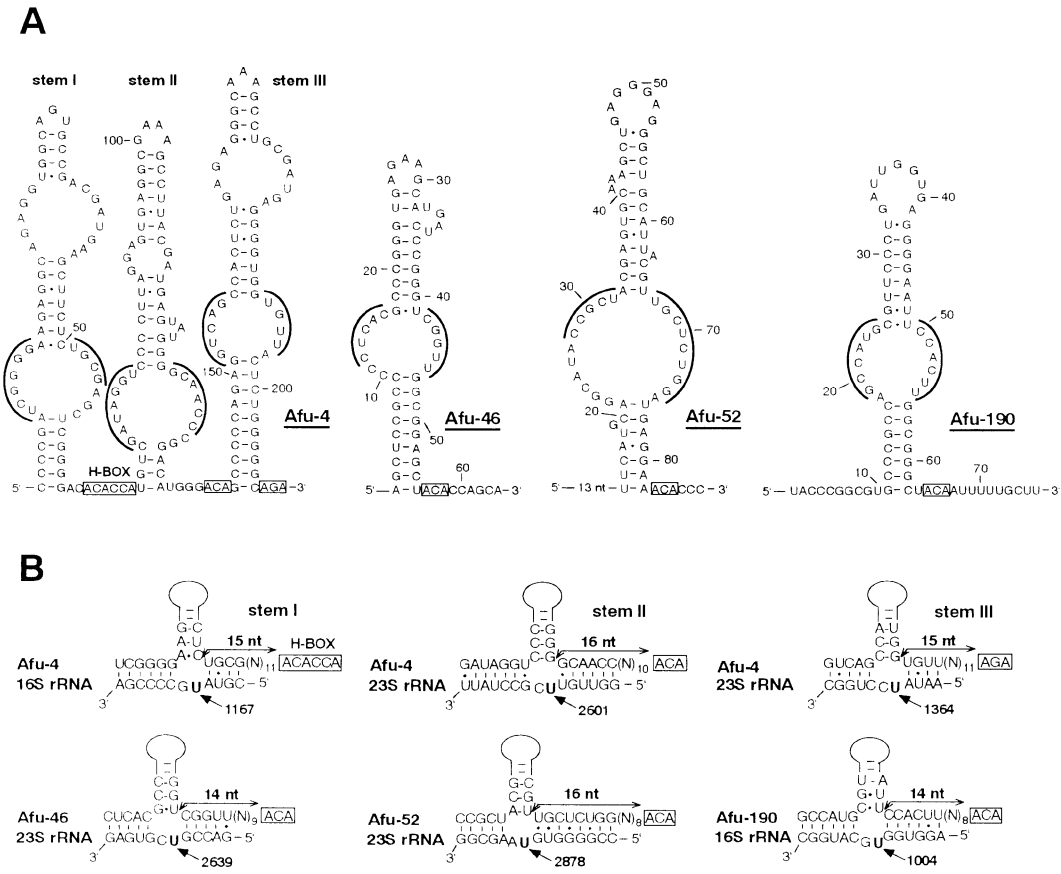


Fig. 5. Four presumptive pseudouridylation guide sRNAs from archaeon *A. fulgidus* [153]. (A) Proposed secondary structures of Afu-4, Afu-46, Afu-52 and Afu-190. The H or ACA/AGA motifs are boxed and in each pseudouridylation pocket the two tracts of complementarity to the rRNA target are overlined. (B) Potential base-pairing interaction with 16S or 23S rRNA involving each pseudouridylation pocket. Each predicted site of pseudouridylation is denoted by an arrow and its location within the cognate rRNA indicated by numbering. Its distance from the ACA/AGA (or H) box (from 14 and 16 nt) is also indicated. The guide sRNA sequence in a 5' to 3' orientation is shown in the upper strand, with the apical part of the long hairpin domain schematized by a solid line.

[140]. Surprisingly, the spliced RNAs resulting from religation at the BHB motif at pre-16S RNA or pre-23S RNA boundaries in euryarchaeon *A. fulgidus* or crenarchaeon *S. solfataricus*, respectively, harbor the box C/D motifs. The spliced RNAs, designated as 16S-D or 16S/23S-D, respectively, depart from the above-mentioned C/D sRNAs guiding rRNA or tRNA methylation by their larger size and more complex structure. However, their C and D motifs form a K-turn structural motif typical of C/D snoRNAs and bind the L7Ae protein with the same affinity as archaeal guides for rRNA or tRNA methylations, but no known RNA target site consistent with their methylation guide function was identified [140]. The novel religated archaeal sRNAs might belong to a subset of C/D small RNAs controlling pre-rRNA folding and processing, such as essential eukaryotic snoRNAs U3, U8 or U22 [12,97,98], but which could act in *cis* instead of *trans*. However, one of them, *A. fulgidus* 16S-D RNA, exists as a stable, separate entity after excision from the rRNA operon transcript, pointing to a more complex role than that of a mere *cis*-acting RNA chaperone.

5.4. Pseudouridylation guides

The few archaeal rRNAs analyzed so far contain only a very small number of pseudouridines, similar to eubacterial but in marked contrast to eukaryotic rRNAs [123,141]. However, genes encoding homologs of three H/ACA snoRNP core proteins, Gar1p, Nop10p and pseudouridine synthase Cbf5p, are present in archaeal genomes, raising the possibility that the pseudouridine formation in archaeal rRNAs also involves homologs of eukaryotic box H/ACA snoRNAs [142]. In a recent experimental screen for small non-messenger RNAs in *A. fulgidus*, four candidate H/ACA small RNAs have been identified for the first time [153], three of them strongly reminiscent of the *Trypanosoma* single-hairpin pseudouridylation guides [44]. Together, the four H/ACA sRNAs are predicted to direct a total of six rRNA pseudouridylations, through the formation of bipartite guide duplexes typical of eukaryotic H/ACA snoRNAs and obeying the same target-H/ACA-motif spacing rule (Fig. 5). In line with the notion that these sRNAs represent

bona fide pseudouridylation guides, the predicted pseudouridines have been experimentally verified in 16S and 23S *A. fulgidus* rRNAs.

6. General conclusions

Methylation of 2'-hydroxyl groups may protect RNA from hydrolytic degradation, enhance hydrophobic surfaces and stabilize helical stems. Pseudouridines, through their flexible C–C glycosyl bonds and increased capacity, relative to uridines, to form H-bonds, may significantly contribute to RNA tertiary structure. Nucleotide modifications directed by snoRNA guides appear in most cases dispensable for cell viability or growth. However, they are likely to have an important biological role by fine-tuning a wide range of RNA–RNA and RNA–protein interactions. This might occur not only in the assembly and function of the ribosome and splicing apparatus but in other fundamental cellular processes as well, as suggested by the detection in a wide range of organisms of an increasing number of snoRNA guides devoid of rRNA or snRNA antisense elements. Both families of snoRNA guides are of ancient origin, largely predating the evolution of a morphologically distinct nucleolar entity in the nucleus. Furthermore, the presence in archaeal guide RNPs of a core protein which is also a ribosomal protein, rpL7Ae, suggests that both guide families, which have a related RNP organization, might have a common evolutionary origin in primordial ribosomes [77]. Underlining their biological importance, modification guide sRNAs are present in a wide spectrum of archaeal organisms whose genomes are severely constrained in size, and detection of novel specimens closely linked to the splicing of archaeal introns open new perspectives as to their range of function. These ancient RNA families show an outstanding potential for structural and functional diversification in the evolution, particularly in complex multicellular organisms, as reflected by the identification of scores of ubiquitous orphan C/D and H/ACA snoRNAs, in addition to the intriguing tissue-specific, imprinted snoRNAs present at multiple tandemly repeated copies in the mammalian genomes. The two snoRNA guide families thus appear as a paradigm of non-coding RNAs highly adapted for the modern RNA world, in which a wide range of posttranscriptional controls are achieved by steric occlusion of sites through sequence- specific recognition of target precursor or mature RNAs [143–146]. Modification guides have intrinsic properties of RNA chaperones, and the discovery of numerous orphan specimens opens the possibility that these two properties may be dissociated in some cases. Distinguishing between the modification guide and chaperone function could illuminate the biological significance of the nucleotide modification itself.

Guide snoRNAs are short, metabolically stable RNAs whose largely modular structure provides the basis for tools in fundamental or applied research. Methylation or pseudouridylation can be targeted to novel rRNA sites by expressing an artificial snoRNA guide carrying an appropriate antisense element [24,41]. Site-directed modification of RNA polymerase II and RNA polymerase III transcripts transiting through the nucleolus is also effective [9,81]. To test systematically the importance in cell growth of individual naturally unmodified rRNA nucleotides, particularly in 25S rRNA peptidyltransferase center, a method for constructing libraries of yeast C/D snoRNA genes that can introduce novel methylations into any rRNA segment of interest has been developed [147]. The C/D box structural motif has been successfully used to target to the nucleolus a chimeric hammerhead ribozyme, highly efficient in the selective cleavage of a nucleolar RNA [148], or demonstrate nucleolar transit of HIV-1 RNA and the Rev protein [149,150]. Guide variants targeted to different intracellular compartments through the utilization of appropriate localization signals may now be envisioned. A few years ago, identification of the two large families of snoRNAs and discovery of their unanticipated guide function have contributed to reveal the amazing diversity of gene controls mediated by non-coding RNAs. Recent results in the field indicate that we may expect more surprises and exciting developments in this respect in the next future.

Acknowledgements

The work performed in J.P.B. laboratory was supported by laboratory funds from the Centre National de la Recherche Scientifique and Université Paul Sabatier, Toulouse, by grants from Association pour la Recherche sur le Cancer, the Toulouse Genopole and the Ministère de l'Education Nationale, de la Recherche et de la Technologie (Programme de Recherche Fondamentale en Microbiologie et Maladies Infectieuses et Parasitaires, 2001–2002) to J.P.B., and a grant from the Programme Interdisciplinaire du C.N.R.S. "Dynamique et réactivité des Assemblages Biologiques" to J.C. Work in the A.H. laboratory was supported by the German Human Genome Project through the BMBF (#01KW9966) and an IZKF grant (Teilprojekt F3, Münster).

References

[1] B.E. Maden, The numerous modified nucleotides in eukaryotic ribosomal RNA, Prog. Nucl. Acid Res. Mol. Biol. 39 (1990) 241–303.

[2] J.P. Bachellerie, J. Cavaillé, Small nucleolar RNAs guide the ribose methylations of eukaryotic rRNAs, in: H. Grosjean, R. Benne (Eds.), Modification and Editing of RNA: The Alteration of RNA Structure and Function, ASM Press, Washington, DC, 1998, pp. 255–272.

[3] J. Ofengand, M.J. Fournier, The pseudouridine residues of rRNA: number, location, biosynthesis and function, in: H. Grosjean, R. Benne (Eds.), Modification and Editing of RNA: The Alteration of RNA Structure and Function, ASM Press, Washington, DC, 1998, pp. 229–253.

[4] R. Brimacombe, P. Mitchell, M. Osswald, K. Stade, D. Bochkariov, Clustering of modified nucleotides at the functional center of bacterial ribosomal RNA, Faseb. J. 7 (1993) 161–167.

[5] J.P. Bachellerie, J. Cavaille, Guiding ribose methylation of rRNA, Trends Biochem. Sci. 22 (1997) 257–261.

[6] C.M. Smith, J.A. Steitz, Sno storm in the nucleolus: new roles for myriad small RNPs, Cell 89 (1997) 669–672.

[7] L.B. Weinstein, J.A. Steitz, Guided tours: from precursor snoRNA to functional snoRNP, Curr. Opin. Cell. Biol. 11 (1999) 378–384.

[8] T. Kiss, M.L. Bortolin, W. Filipowicz, Characterization of the intron-encoded U19 RNA, a new mammalian small nucleolar RNA that is not associated with fibrillarin, Mol. Cell. Biol. 16 (1996) 1391–1400.

[9] J.P. Bachellerie, J. Cavaillé, L.H. Qu, Nucleotide modifications of eukaryotic rRNAs: the world of small nucleolar RNAs revisited, in: R. Garrett, S. Douthwaite, A. Liljas, A. Matheson, P.B. Moore, H.F. Noller (Eds.), The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions, ASM Press, Washington, DC, 2000, pp. 191–203.

[10] B.G. Lane, J. Ofengand, M.W. Gray, Pseudouridine and O2'-methylated nucleosides. Significance of their selective occurrence in rRNA domains that function in ribosome-catalyzed synthesis of the peptide bonds in proteins, Biochimie 77 (1995) 7–15.

[11] J. Ofengand, A. Bakin, J. Wrzesinski, K. Nurse, B.G. Lane, The pseudouridine residues of ribosomal RNA, Biochem. Cell. Biol. 73 (1995) 915–924.

[12] D. Tollervey, *Trans*-acting factors in ribosome synthesis, Exp. Cell. Res. 229 (1996) 226–232.

[13] B. Sollner-Webb, K. Tyc, J.A. Steitz, Ribosomal RNA processing in eukaryotes, in: R.A. Zimmermann, A.E. Dahlberg (Eds.), Ribosomal RNA. Structure, Evolution, Processing, and Function in Protein Biosynthesis, Telford, Caldwell, New Jersey, 1995, pp. 469–490.

[14] A.G. Balakin, L. Smith, M.J. Fournier, The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions, Cell 86 (1996) 823–834.

[15] J. Cavaille, J.P. Bachellerie, Processing of fibrillarin-associated snoRNAs from pre-mRNA introns: an exonucleolytic process exclusively directed by the common stem-box terminal structure, Biochimie 78 (1996) 443–456.

[16] E. Caffarelli, A. Fatica, S. Prislei, E. De Gregorio, P. Fragapane, I. Bozzoni, Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA, Embo. J. 15 (1996) 1121–1131.

[17] T.S. Lange, A. Borovjagin, E.S. Maxwell, S.A. Gerbi, Conserved boxes C and D are essential nucleolar localization elements of U14 and U8 snoRNAs, Embo. J. 17 (1998) 3176–3187.

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

[19] T. Villa, F. Ceradini, I. Bozzoni, Identification of a novel element required for processing of intron-encoded box C/D small nucleolar RNAs in *Saccharomyces cerevisiae*, Mol. Cell. Biol. 20 (2000) 1311–1320.

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

[21] K.T. Tycowski, M.D. Shu, J.A. Steitz, A mammalian gene with introns instead of exons generating stable RNA products, Nature 379 (1996) 464–466.

[22] M. Nicoloso, L.H. Qu, B. Michot, J.P. Bachellerie, Intron-encoded, antisense small nucleolar RNAs: the characterization of nine novel species points to their direct role as guides for the 2'-*O*-ribose methylation of rRNAs, J. Mol. Biol. 260 (1996) 178–195.

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

[24] J. Cavaille, M. Nicoloso, J.P. Bachellerie, Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides, Nature 383 (1996) 732–735.

[25] T.M. Lowe, S.R. Eddy, A computational screen for methylation guide snoRNAs in yeast, Science 283 (1999) 1168–1171.

[26] F Barneche, C. Gaspin, R. Guyot, M. Echeverria, Identification of 66 box C/D snoRNAs in *Arabidopsis thaliana*: extensive gene duplications generated multiple isoforms predicting new ribosomal RNA 2'-*O*-methylation sites, J. Mol. Biol. 311 (2001) 57–73.

[27] L.H. Qu, Q. Meng, H. Zhou, Y.-Q. Chen, Identification of 10 novel snoRNA gene clusters from *Arabidopsis thaliana*, Nucl. Acids Res. 29 (2001) 1623–1630.

[28] J.W. Brown, G.P. Clark, D.J. Leader, C.G. Simpson, T. Lowe, Multiple snoRNA gene clusters from *Arabidopsis*, RNA 7 (2001) 1817–1832.

[29] M. Nicoloso, M. Caizergues-Ferrer, B. Michot, M.C. Azum, J.P. Bachellerie, U20, a novel small nucleolar RNA, is encoded in an intron of the nucleolin gene in mammals, Mol. Cell. Biol. 14 (1994) 5766–5776.

[30] L.H. Qu, M. Nicoloso, B. Michot, M.C. Azum, M. Caizergues-Ferrer, M.H. Renalier, J.P. Bachellerie, U21, a novel small nucleolar RNA with a 13 nt complementarity to 28S rRNA, is encoded in an intron of ribosomal protein L5 gene in chicken and mammals, Nucl. Acids Res. 22 (1994) 4073–4081.

[31] L.H. Qu, Y. Henry, M. Nicoloso, B. Michot, M.C. Azum, M.H. Renalier, M. Caizergues-Ferrer, J.P. Bachellerie, U24, a novel intron-encoded small nucleolar RNA with two 12 nt long, phylogenetically conserved complementarities to 28S rRNA, Nucl. Acids Res. 23 (1995) 2669–2676.

[32] A. Huttenhofer, M. Kiefmann, S. Meier-Ewert, J. O'Brien, H. Lehrach, J.P. Bachellerie, J. Brosius, RNomics: an experimental approach that identifies 201 candidates for novel, small, non-messenger RNAs in mouse, Embo. J. 20 (2001) 2943–2953.

[33] K.T. Tycowski, J.A. Steitz, Non-coding snoRNA host genes in *Drosophila*: expression strategies for modification guide snoRNAs, Eur. J. Cell. Biol. 80 (2001) 119–125.

[34] Y. Xu, L. Liu, C. Lopez-Estrano, S. Michaeli, Expression studies on clustered trypanosomatid box C/D small nucleolar RNAs, J. Biol. Chem. 276 (2001) 14289–14298.

[35] D.A. Dunbar, A.A. Chen, S. Wormsley, S.J. Baserga, The genes for small nucleolar RNAs in *Trypanosoma brucei* are organized in clusters and are transcribed as a polycistronic RNA, Nucl. Acids Res. 28 (2000) 2855–2861.

[36] D.A. Dunbar, S. Wormsley, T.M. Lowe, S.J. Baserga, Fibrillarin-associated box C/D small nucleolar RNAs in *Trypanosoma brucei*. Sequence conservation and implications for 2'-*O*-ribose methylation of rRNA, J. Biol. Chem. 275 (2000) 14767–14776.

[37] H. Bugl, E.B. Fauman, B.L. Staker, F. Zheng, S.R. Kushner, M.A. Saper, J.C. Bardwell, U. Jakob, RNA methylation under heat shock control, Mol. Cell. 6 (2000) 349–360.

[38] T. Caldas, E. Binet, P. Bouloc, G. Richarme, Translational defects of *Escherichia coli* mutants deficient in the Um(2552) 23S ribosomal RNA methyltransferase RrmJ/FTSJ, Biochem. Biophys. Res. Commun. 271 (2000) 714–718.

[39] P. Ganot, M. Caizergues-Ferrer, T. Kiss, The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation, Genes Dev. 11 (1997) 941–956.

[40] P. Ganot, M.L. Bortolin, T. Kiss, Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs, Cell 89 (1997) 799–809.

[41] M.L. Bortolin, P. Ganot, T. Kiss, Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs, Embo. J. 18 (1999) 457–469.

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

[43] J. Ofengand, A. Bakin, Mapping to nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts, J. Mol. Biol. 266 (1997) 246–268.

[44] X.H. Liang, L. Liu, S. Michaeli, Identification of the first trypanosome H/ACA RNA that guides pseudouridine formation on rRNA, J. Biol. Chem. 276 (2001) 40313–40318.

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

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

[47] L.H. Qu, A. Henras, Y.J. Lu, H. Zhou, W.X. Zhou, Y.Q. Zhu, J. Zhao, Y. Henry, M. Caizergues-Ferrer, J.P. Bachellerie, 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 (1999) 1144–1158.

[48] D.J. Leader, G.P. Clark, J. Watters, A.F. Beven, P.J. Shaw, J.W. Brown, Clusters of multiple different small nucleolar RNA genes in plants are expressed as and processed from polycistronic pre-snoRNAs, Embo. J. 16 (1997) 5742–5751.

[49] D.J. Leader, G.P. Clark, J. Boag, J.A. Watters, C.G. Simpson, N.J. Watkins, E.S. Maxwell, J.W. Brown, Processing of vertebrate box C/D small nucleolar RNAs in plant cells, Eur. J. Biochem. 253 (1998) 154–160.

[50] D. Filippini, F. Renzi, I. Bozzoni, E. Caffarelli, U86, a novel snoRNA with an unprecedented gene organization in yeast, Biochem. Biophys. Res. Commun. 288 (2001) 16–21.

[51] M.H. Renalier, M. Nicoloso, L.H. Qu, J.P. Bachellerie, SnoRNA U21 is also intron-encoded in *Drosophila melanogaster* but in a different host-gene as compared to warm-blooded vertebrates, FEBS Lett. 379 (1996) 212–216.

[52] L. Xia, J. Liu, C. Sage, E.B. Trexler, M.T. Andrews, E.S. Maxwell, Intronic U14 snoRNAs of *Xenopus laevis* are located in two different parent genes and can be processed from their introns during early oogenesis, Nucl. Acids Res. 23 (1995) 4844–4849.

[53] M. Cervelli, F. Cecconi, M. Giorgi, F. Annesi, M. Oliverio, P. Mariottini, Comparative structure analysis of vertebrate U17 small nucleolar RNA (snoRNA), J. Mol. Evol. 54 (2002) 166–179.

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

[55] P. Pelczar, W. Filipowicz, The host gene for intronic U17 small nucleolar RNAs in mammals has no protein-coding potential and is a member of the 5'-terminal oligopyrimidine gene family, Mol. Cell. Biol. 18 (1998) 4509–4518.

[56] C.M. Smith, J.A. Steitz, Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes, Mol. Cell. Biol. 18 (1998) 6897–6909.

[57] T.S. Lange, M. Ezrokhi, F. Amaldi, S.A. Gerbi, Box H and box ACA are nucleolar localization elements of U17 small nucleolar RNA, Mol. Biol. Cell. 10 (1999) 3877–3890.

[58] X. Darzacq, T. Kiss, Processing of intron-encoded box C/D small nucleolar RNAs lacking a 5',3'-terminal stem structure, Mol. Cell. Biol. 20 (2000) 4522–4531.

[59] A. Henras, C. Dez, J. Noaillac-Depeyre, Y. Henry, M. Caizergues-Ferrer, Accumulation of H/ACA snoRNPs depends on the integrity of the conserved central domain of the RNA-binding protein Nhp2p, Nucl. Acids Res. 29 (2001) 2733–2746.

[60] A. Niewmierzycka, S. Clarke, *S*-Adenosylmethionine-dependent methylation in *Saccharomyces cerevisiae*. Identification of a novel protein arginine methyltransferase, J. Biol. Chem. 274 (1999) 814–824.

[61] H. Wang, D. Boisvert, K.K. Kim, R. Kim, S.H. Kim, Crystal structure of a fibrillarin homologue from *Methanococcus jannaschii*, a hyperthermophile, at 1.6 A resolution, Embo. J. 19 (2000) 317–323.

[62] D. Tollervey, H. Lehtonen, R. Jansen, H. Kern, E.C. Hurt, Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly, Cell 72 (1993) 443–457.

[63] N.J. Watkins, V. Segault, B. Charpentier, S. Nottrott, P. Fabrizio, A. Bachi, M. Wilm, M. Rosbash, C. Branlant, R. Luhrmann, A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP, Cell 103 (2000) 457–466.

[64] I. Vidovic, S. Nottrott, K. Hartmuth, R. Luhrmann, R. Ficner, Crystal structure of the spliceosomal 15.5 kD protein bound to a U4 snRNA fragment, Mol. Cell. 6 (2000) 1331–1342.

[65] D.J. Klein, T.M. Schmeing, P.B. Moore, T.A. Steitz, The kink-turn: a new RNA secondary structure motif, Embo. J. 20 (2001) 4214–4221.

[66] S. Nottrott, K. Hartmuth, P. Fabrizio, H. Urlaub, I. Vidovic, R. Ficner, R. Luhrmann, Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA, Embo. J. 18 (1999) 6119–6133.

[67] D.R. Newman, J.F. Kuhn, G.M. Shanab, E.S. Maxwell, Box C/D snoRNA-associated proteins: two pairs of evolutionarily ancient proteins and possible links to replication and transcription, RNA 6 (2000) 861–879.

[68] D. Filippini, I. Bozzoni, E. Caffarelli, p62, a novel *Xenopus laevis* component of box C/D snoRNPs, RNA 6 (2000) 391–401.

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

[70] Y. Zebarjadian, T. King, M.J. Fournier, L. Clarke, J. Carbon, Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA, Mol. Cell. Biol. 19 (1999) 7461–7472.

[71] N.S. Heiss, D. Bachner, R. Salowsky, A. Kolb, P. Kioschis, A. Poustka, Gene structure and expression of the mouse dyskeratosis congenita gene, dkc1, Genomics 67 (2000) 153–163.

[72] T. Vulliamy, A. Marrone, F. Goldman, A. Dearlove, M. Bessler, P.J. Mason, I. Dokal, The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita, Nature 413 (2001) 432–435.

[73] C. Hoang, A.R. Ferre-D'Amare, Cocrystal structure of a tRNA Psi55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme, Cell 107 (2001) 929–939.

[74] A. Henras, Y. Henry, C. Bousquet-Antonelli, J. Noaillac-Depeyre, J.P. Gelugne, M. Caizergues-Ferrer, Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs, Embo. J. 17 (1998) 7078–7090.

[75] N.J. Watkins, A. Gottschalk, G. Neubauer, B. Kastner, P. Fabrizio, M. Mann, R. Luhrmann, 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 (1998) 1549–1568.

[76] T.H. King, W.A. Decatur, E. Bertrand, E.S. Maxwell, M.J. Fournier, A well-connected and conserved nucleoplasmic helicase is required for production of box C/D and H/ACA snoRNAs and localization of snoRNP proteins, Mol. Cell. Biol. 21 (2001) 7731–7746.

[77] J.F. Kuhn, E.J. Tran, E.S. Maxwell, Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein, Nucl. Acids Res. 30 (2002) 931–941.

[78] S. Massenet, A. Mougin, C. Branlant, Posttranscriptional modifications in the U snRNAs, in: H. Grosjean, R. Benne (Eds.), Modification and Editing of RNA: The Alteration of RNA Structure and Function, ASM Press, Washington, DC, 1998, pp. 201–228.

[79] Y.T. Yu, M.D. Shu, J.A. Steitz, Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing, Embo. J. 17 (1998) 5783–5795.

[80] K.T. Tycowski, Z.H. You, P.J. Graham, J.A. Steitz, Modification of U6 spliceosomal RNA is guided by other small RNAs, Mol. Cell. 2 (1998) 629–638.

[81] P. Ganot, B.E. Jady, M.L. Bortolin, X. Darzacq, T. Kiss, Nucleolar factors direct the 2'-*O*-ribose methylation and pseudouridylation of U6 spliceosomal RNA, Mol. Cell. Biol. 19 (1999) 6906–6917.

[82] H. Zhou, Y.Q. Chen, Y.P. Du, L.H. Qu, The *Schizosaccharomyces pombe* mgU6-47 gene is required for 2'-*O*-methylation of U6 snRNA at A41, Nucl. Acids Res. 30 (2002) 894–902.

[83] B.E. Jady, T. Kiss, A small nucleolar guide RNA functions both in 2'-*O*-ribose methylation and pseudouridylation of the U5 spliceosomal RNA, Embo. J. 20 (2001) 541–551.

[84] T. Kiss, Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs, Embo. J. 20 (2001) 3617–3622.

[85] S. Massenet, Y. Motorin, D.L. Lafontaine, E.C. Hurt, H. Grosjean, C. Branlant, Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA, Mol. Cell. Biol. 19 (1999) 2142–2154.

[86] A.G. Matera, Nuclear bodies: multifaceted subdomains of the interchromatin space, Trends Cell. Biol. 9 (1999) 302–309.

[87] A. Narayanan, W. Speckmann, R. Terns, M.P. Terns, Role of the box C/D motif in localization of small nucleolar RNAs to coiled bodies and nucleoli, Mol. Biol. Cell 10 (1999) 2131–2147.

[88] J.E. Sleeman, A.I. Lamond, Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway, Curr. Biol. 9 (1999) 1065–1074.

[89] Y.T. Yu, M.D. Shu, A. Narayanan, R.M. Terns, M.P. Terns, J.A. Steitz, Internal modification of U2 small nuclear (sn)RNA occurs in nucleoli of *Xenopus* oocytes, J. Cell. Biol. 152 (2001) 1279–1288.

[90] J.R. Mitchell, J. Cheng, K. Collins, A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end, Mol. Cell. Biol. 19 (1999) 567–576.

[91] V. Pogacic, F. Dragon, W. Filipowicz, H. Human, H/CA small nucleolar RNPs and telomerase share evolutionarily conserved proteins NHP2 and NOP10, Mol. Cell. Biol. 20 (2000) 9028–9040.

[92] C. Dez, A. Henras, B. Faucon, D. Lafontaine, M. Caizergues-Ferrer, Y. Henry, Stable expression in yeast of the mature form of human telomerase RNA depends on its association with the box H/ACA small nucleolar RNP proteins Cbf5p, Nhp2p and Nop10p, Nucl. Acids Res. 29 (2001) 598–603.

[93] F. Dragon, V. Pogacic, W. Filipowicz, In vitro assembly of human H/ACA small nucleolar RNPs reveals unique features of U17 and telomerase RNAs, Mol. Cell. Biol. 20 (2000) 3037–3048.

[94] B.E. Jady, T. Kiss, Characterisation of the U83 and U84 small nucleolar RNAs: two novel 2'-*O*-ribose methylation guide RNAs that lack complementarities to ribosomal RNAs, Nucl. Acids Res. 28 (2000) 1348–1354.

[95] T. Pederson, The plurifunctional nucleolus, Nucl. Acids Res. 26 (1998) 3871–3876.

[96] E. Bertrand, F. Houser-Scott, A. Kendall, R.H. Singer, D.R. Engelke, Nucleolar localization of early tRNA processing, Genes Dev. 12 (1998) 2463–2468.

[97] J.M. Hughes, Functional base-pairing interaction between highly conserved elements of U3 small nucleolar RNA and the small ribosomal subunit RNA, J. Mol. Biol. 259 (1996) 645–654.

[98] K. Sharma, D. Tollervey, Base pairing between U3 small nucleolar RNA and the 5' end of 18S rRNA is required for pre-rRNA processing, Mol. Cell. Biol. 19 (1999) 6012–6019.

[99] J.P. Morrissey, D. Tollervey, Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis, Mol. Cell. Biol. 13 (1993) 2469–2477.

[100] D. Tollervey, A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA, Embo. J. 6 (1987) 4169–4175.

[101] C.A. Enright, E.S. Maxwell, G.L. Eliceiri, B. Sollner-Webb, 5'ETS rRNA processing facilitated by four small RNAs: U14, E3, U17, and U3, RNA 2 (1996) 1094–1099.

[102] W.Q. Liang, J.A. Clark, M.J. Fournier, The rRNA-processing function of the yeast U14 small nucleolar RNA can be rescued by a conserved RNA helicase-like protein, Mol. Cell. Biol. 17 (1997) 4124–4132.

[103] H. Nielsen, H. Orum, J. Engberg, A novel class of nucleolar RNAs from *Tetrahymena*, FEBS Lett. 307 (1992) 337–342.

[104] X.H. Liang, Y.X. Xu, S. Michaeli, The spliced leader-associated RNA is a trypanosome-specific sn(o)RNA that has the potential to guide pseudouridine formation on the SL RNA, RNA 8 (2002) 237–246.

[105] Z. Palfi, G.L. Xu, A. Bindereif, Spliced leader-associated RNA of trypanosomes. Sequence conservation and association with protein components common to *trans*-spliceosomal ribonucleoproteins, J. Biol. Chem. 269 (1994) 30620–30625.

[106] J. Cavaille, K. Buiting, M. Kiefmann, M. Lalande, C.I. Brannan, B. Horsthemke, J.P. Bachellerie, J. Brosius, A. Huttenhofer, Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization, Proc. Natl. Acad. Sci. USA 97 (2000) 14311–14316.

[107] C.M. Burns, H. Chu, S.M. Rueter, L.K. Hutchinson, H. Canton, E. Sanders-Bush, R.B. Emeson, Regulation of serotonin-2C receptor G-protein coupling by RNA editing, Nature 387 (1997) 303–308.

[108] H.Y. Yi-Brunozzi, L.M. Easterwood, G.M. Kamilar, P.A. Beal, Synthetic substrate analogs for the RNA-editing adenosine deaminase ADAR-2, Nucl. Acids Res. 27 (1999) 2912–2917.

[109] J. Cavaille, P. Vitali, E. Basyuk, A. Huttenhofer, J.P. Bachellerie, A novel brain-specific box C/D small nucleolar RNA processed from tandemly repeated introns of a noncoding RNA gene in rats, J. Biol. Chem. 276 (2001) 26374–26383.

[110] T. los Santos de, J. Schweizer, C.A. Rees, U. Francke, Small evolutionarily conserved RNA, resembling C/D box small nucleolar RNA, is transcribed from PWCR1, a novel imprinted gene in the Prader-Willi deletion region, which is highly expressed in brain, Am. J. Hum. Genet. 67 (2000) 1067–1082.

[111] W. Filipowicz, Imprinted expression of small nucleolar RNAs in brain: time for RNomics, Proc. Natl. Acad. Sci. USA 97 (2000) 14035–14037.

[112] M. Meguro, K. Mitsuya, N. Nomura, M. Kohda, A. Kashiwagi, R. Nishigaki, H. Yoshioka, M. Nakao, M. Oishi, M. Oshimura, Large-scale evaluation of imprinting status in the Prader-Willi syndrome region: an imprinted direct repeat cluster resembling small nucleolar RNA genes, Hum. Mol. Genet. 10 (2001) 383–394.

[113] M. Runte, A. Huttenhofer, S. Gross, M. Kiefmann, B. Horsthemke, K. Buiting, The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A, Hum. Mol. Genet. 10 (2001) 2687–2700.

[114] W. Reik, J. Walter, Genomic imprinting: parental influence on the genome, Natl. Rev. Genet. 2 (2001) 21–32.

[115] R.D. Nicholls, J.L. Knepper, Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes, Annu. Rev. Genom. Hum. Genet. 2 (2001) 153–175.

[116] Y. Komine, N.K. Tanaka, R. Yano, S. Takai, S. Yuasa, T. Shiroishi, K. Tsuchiya, T. Yamamori, A novel type of non-coding RNA expressed in the rat brain, Brain Res. Mol. Brain Res. 66 (1999) 1–13.
[117] J.P. Bachellerie, M. Nicoloso, L.H. Qu, B. Michot, M. Caizergues-Ferrer, J. Cavaille, M.H. Renalier, Novel intron-encoded small nucleolar RNAs with long sequence complementarities to mature rRNAs involved in ribosome biogenesis, Biochem. Cell. Biol. 73 (1995) 835–843.
[118] B. Neumann, P. Kubicka, D.P. Barlow, Characteristics of imprinted genes, Natl. Genet. 9 (1995) 12–13.
[119] J.S. Mattick, M.J. Gagen, The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms, Mol. Biol. Evol. 18 (2001) 1611–1630.
[120] D. Hatton, J.C. Gray, Two MAR DNA-binding proteins of the pea nuclear matrix identify a new class of DNA-binding proteins, Plant J. 18 (1999) 417–429.
[121] J. Wirth, E. Back, A. Huttenhofer, H.G. Nothwang, C. Lich, S. Gross, C. Menzel, A. Schinzel, P. Kioschis, N. Tommerup, et al., A translocation breakpoint cluster disrupts the newly defined 3' end of the SNURF-SNRPN transcription unit on chromosome 15, Hum. Mol. Genet. 10 (2001) 201–210.
[122] C. Gustafsson, R. Reid, P.J. Greene, D.V. Santi, Identification of new RNA modifying enzymes by iterative genome search using known modifying enzymes as probes, Nucl. Acids Res. 24 (1996) 3756–3762.
[123] J. Ofengand, K. Rudd, Bacterial, archaea, and organellar RNA pseudouridines and methylated nucleosides and their enzymes, in: R. Garrett, S. Douthwaite, A. Liljas, A. Matheson, P.B. Moore, H.F. Noller (Eds.), Ribosome: Structure, Function, Antibiotics, and Cellular Interactions, ASM Press, Washington, DC, 2000, pp. 175–190.
[124] G.J. Olsen, C.R. Woese, Archaeal genomics: an overview, Cell 89 (1997) 991–994.
[125] K.R. Noon, E. Bruenger, J.A. McCloskey, Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile *Sulfolobus solfataricus*, J. Bacteriol. 180 (1998) 2883–2888.
[126] K.A. Amiri, Fibrillarin-like proteins occur in the domain Archaea, J. Bacteriol. 176 (1994) 2124–2127.
[127] D.L. Lafontaine, D. Tollervey, Birth of the snoRNPs: the evolution of the modification-guide snoRNAs, Trends Biochem. Sci. 23 (1998) 383–388.
[128] C. Gaspin, J. Cavaille, G. Erauso, J.P. Bachellerie, Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the *Pyrococcus* genomes, J. Mol. Biol. 297 (2000) 895–906.
[129] A.D. Omer, T.M. Lowe, A.G. Russell, H. Ebhardt, S.R. Eddy, P.P. Dennis, Homologs of small nucleolar RNAs in Archaea, Science 288 (2000) 517–522.
[130] P.P. Dennis, A. Omer, T. Lowe, A guided tour: small RNA function in Archaea, Mol. Microbiol. 40 (2001) 509–519.
[131] W.A. Speckmann, Z.H. Li, T.M. Lowe, S.R. Eddy, R.M. Terns, M.P. Terns, Archaeal guide RNAs function in rRNA modification in the eukaryotic nucleus, Curr. Biol. 12 (2002) 199–203.
[132] B. Clouet d'Orval, M.L. Bortolin, C. Gaspin, J.P. Bachellerie, Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNATrp intron guides the formation of two ribose-methylated nucleosides in the mature tRNATrp, Nucl. Acids Res. 29 (2001) 4518–4529.
[133] H. Grosjean, Z. Szweykowska-Kulinska, Y. Motorin, F. Fasiolo, G. Simos, Intron-dependent enzymatic formation of modified nucleosides in eukaryotic tRNAs: a review, Biochimie 79 (1997) 293–302.
[134] R. Gupta, *Halobacterium volcanii* tRNAs. Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs, J. Biol. Chem. 259 (1984) 9461–9471.
[135] D.T. Nieuwlandt, M.B. Carr, C.J. Daniels, In vivo processing of an intron-containing archael tRNA, Mol. Microbiol. 8 (1993) 93–99.
[136] N.S. Gutgsell, M.D. Del Campo, S. Raychaudhuri, J. Ofengand, A second function for pseudouridine synthases: a point mutant of RluD unable to form pseudouridines1915 and1917in *Escherichia coli* 23S ribosomal RNA restores normal growth to an RluD-minus strain, RNA 7 (2001) (1911) 990–998.
[137] D.L. Lafontaine, T. Preiss, D. Tollervey, Yeast 18S rRNA dimethylase Dim1p: a quality control mechanism in ribosome synthesis? Mol. Cell.Biol. 18 (1998) 2360–2370.
[138] B.C. Persson, C. Gustafsson, D.E. Berg, G.R. Bjork, The gene for a tRNA modifying enzyme, m5U54-methyltransferase, is essential for viability in *Escherichia coli*, Proc. Natl. Acad. Sci. USA 89 (1992) 3995–3998.
[139] T.L. Mason, C. Pan, M.E. Sanchirico, K. Sirum-Connolly, Molecular genetics of the peptidyl transferase center and the unusual Var1 protein in yeast mitochondrial ribosomes, Experientia 52 (1996) 1148–1157.
[140] T.H. Tang, T.S. Rozhdestvensky, B. Clouet d'Orval, M.L. Bortolin, H. Huber, B. Charpentier, C. Branlant, J.P. Bachellerie, J. Brosius, A. Huttenhofer, RNomics in Archaea reveals a further link between splicing of archaeal introns and rRNA processing, Nucl. Acids Res. 30 (2002) 921–930.
[141] J.A. Kowalak, E. Bruenger, P.F. Crain, J.A. McCloskey, Identities and phylogenetic comparisons of posttranscriptional modifications in 16 S ribosomal RNA from *Haloferax volcanii*, J. Biol. Chem. 275 (2000) 24484–24489.
[142] Y. Watanabe, M.W. Gray, Evolutionary appearance of genes encoding proteins associated with box H/ACA snoRNAs: cbf5p in *Euglena gracilis*, an early diverging eukaryote, and candidate Gar1p and Nop10p homologs in archaebacteria, Nucl. Acids Res. 28 (2000) 2342–2352.
[143] R.C. Lee, V. Ambros, An extensive class of small RNAs in *Caenorhabditis elegans*, Science 294 (2001) 862–864.
[144] N.C. Lau, L.P. Lim, E.G. Weinstein, D.P. Bartel, An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*, Science 294 (2001) 858–862.
[145] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, Identification of novel genes coding for small expressed RNAs, Science 294 (2001) 853–858.
[146] S.R. Eddy, Non-coding RNA genes and the modern RNA world, Natl. Rev. Genet. 2 (2001) 919–929.
[147] B. Liu, J. Ni, M.J. Fournier, Probing RNA in vivo with methylation guide small nucleolar RNAs, Methods 23 (2001) 276–286.
[148] D.A. Samarsky, G. Ferbeyre, E. Bertrand, R.H. Singer, R. Cedergren, M.J. Fournier, A small nucleolar RNA: ribozyme hybrid cleaves a nucleolar RNA target in vivo with near-perfect efficiency, Proc. Natl. Acad. Sci. USA 96 (1999) 6609–6614.
[149] A. Michienzi, L. Cagnon, I. Bahner, J.J. Rossi, Ribozyme-mediated inhibition of HIV 1 suggests nucleolar trafficking of HIV-1 RNA, Proc. Natl. Acad. Sci. USA 97 (2000) 8955–8960.
[150] S.B. Buonomo, A. Michienzi, F.G. De Angelis, I. Bozzoni, The Rev protein is able to transport to the cytoplasm small nucleolar RNAs containing a Rev binding element, RNA 5 (1999) 993–1002.
[151] X. Darzacq, B.E. Jady, C. Verheggen, T. Kiss, Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs, EMBO J. 21 (2002) 2746–2756.
[152] J. Cavaille, H. Seitz, M. Paulsen, A.C. Ferguson-Smith, J.P. Bachellerie, Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region, Hum. Mol. Genet. 11 (2002) 1527–1538.
[153] T.H. Tang, J.P. Bachellerie, T. Rozhdestvensky, M.L. Bortolin, T. Elge, J. Brosius, A. Huttenhofer, Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*, Proc. Natl. Acad. Sci. USA 99 (2002) 7536–7541.