CHAPTER 1

RNA INTERACTIONS

Manja Marz[a,b], and Peter F Stadler[a,c,d,e],

[a]Bioinformatics Group, Department of Computer Science; and Interdisciplinary Center for Bioinformatics, University of Leipzig, Härtelstrasse 16-18, D-01407, Leipzig, Germany
[b]RNA-Bioinformatics Group, Institut für Pharmazeutische Chemie, Philipps-Universität Marburg, Marbacher Weg 6, 35037 Marburg
[c]Institute for Theoretical Chemistry, University of Vienna, Währingerstraße 17, A-1090 Wien, Austria
[d]Fraunhofer Institute for Cell Therapy und Immunology, Perlickstrasse 1, D-04103 Leipzig, Germany
[e]Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, USA

1.1 INTRODUCTION

Formation of base pairs between complementary nucleic acids are the key for both the structure formation of individual RNAs as well as for interactions between RNA and/or DNA molecules. The patterns of base pairing constitute the *secondary structures*. They characterize functional classes of ncRNAs and are often well-conserved over large evolutionary time-scales, reviewed by Bompfünewerer *et al.* (2005, 2007).

It should not be ignored, that specific base pairing also contibutes to the tertiary structure of the RNA and is recognized by proteins interacting specifically with a particular RNA species.

The importance of specific base pairing is not limited to the structure of a single RNA, however. On the one hand, it has a crucial impact on the tertiary (3D) structure of the RNA and thus way in which it is recognized by its protein partners; on the other hand, secondary structures are also formed by interacting RNAs, thus

determining both strength and exact position of hybridization between two or more partners.

In this brief survey we focus on how ncRNAs interact with their partners in a large number of different molecular and functional contexts. Base-pairing patterns play a central role because of their large impact on the structure of binding sites for proteins, or – more directly – because the interaction consists of intermolecular base pairs. We organize our contribution by the composition of the players: RNA with DNA, RNA with RNA, and RNA with proteins. Our presentation does not strive to be exhaustive and attempts to give a broad overview rather than an in-depth discussion of specific examples. Wherever possible, we thus cite recent reviews rather than the original literature.

1.2 NCRNA-DNA INTERACTION

RNA:DNA hybrids play a crucial role in transcription termination in bacteria: the formation of the terminator hairpin in the nascent RNA transcripts shortens the RNA:DNA duplex associated with the polymerase complex and facilitates dissociation (Komissarova et al., 2002). Another important case are the RNA primers of the Okazaki fragment on the lagging strand during DNA replication (MacNeill, 2001).

Despite their use in biotechnology (Nadal et al., 2005; Suzuki, 2008), very little is known about the potential of triple-helices as form of direct RNA:DNA interaction.

1.2.1 Chromatin Regulation

Comprehensive surveys such as the ENCODE and FANTOM projects (The ENCODE Project Consortium, 2007; The FANTOM Consortium, 2005) demonstrated that the genomes of higher Eukaryotes are pervasively transcribed. In mammals, many of these large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression (Khalil et al., 2009; Sheik et al., 2010; Mattick et al., 2009). Apparently, there are diverse roles for lincRNAs in processes from embryonic stem cell pluripotency to cell proliferation, based on the observation that these transcripts are differentially regulated by key transcription factors such as p53, NFkappaB, or Nanog. Similar mechanisms are at work also in the yeast *Saccharomyces cerevisiae* (Harrison et al., 2009) and in plants (Swiezewski et al., 2009; Matzke et al., 2009). A significant fraction of long non-coding RNAs are subject to at least moderate stabilizing selection on the exon (Marques and Ponting, 2009), corroborating the functionality of these transcripts. Even transcripts with little or no sequence conservation may be functional as shown by the deep conservation of the gene structure itself (Hiller et al., 2009).

At present, multiple models for the mode of action of these long ncRNA transcripts have been proposed, reviewed e.g. in (Hekimoglu and Ringrose, 2009), and there is at least circumstantial evidence that different transcripts may be governed by different mechanisms, Fig. 1.1.

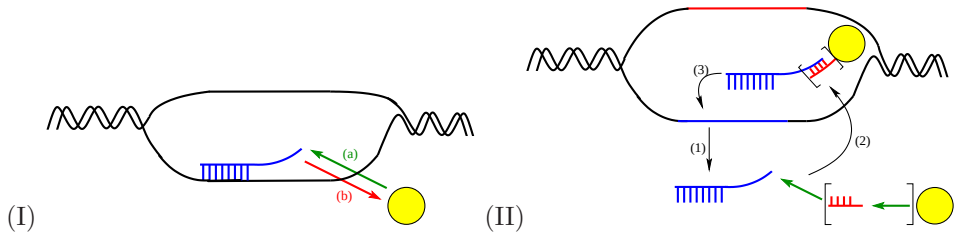

Figure 1.1 Putative mechanisms of chromatin-associated RNAs (modified from (Hekimoglu and Ringrose, 2009)): **(I)** Direct RNA mediated chromatin regulation during transcription. Epigenetic regulator proteins (yellow) are (a, green) recruited directly from transcribed ncRNA (blue) or released from acting site during transcription (b, red). **(II)** Indirect association to DNA regulatory elements after transcription. After transcription of long ncRNAs (blue) from DNA (1), regulative epigenetic regulators (yellow) may interact directly (green) or via antisense RNA mediator (red, brackets) with ncRNAs (2) in order to bind to DNA (3). With this mechanism the recruitment of epigenetic regulatory proteins, switches of gene expression states, and maintenance of epigenetic memory are performed.

1.2.2 Y RNA

Y RNAs are believed to play two intracellular roles. Y RNAs were originally discovered as an RNA component of Ro RNPs, binding to Ro60 for RNA quality control (Stein et al., 2005). However, recently it was shown to be essential for chromosomal DNA replication (Christov et al., 2006). For chromosomal DNA replication a cluster of Y RNAs is associated with the protein RPA and the sliding clamp PCNA and other proteins (Krude, 2010; Szüts et al., 2005, 2003). Y RNAs are hereby the essential fraction for replication (Y1, Y3, Y4 and Y5 located in a narrow cluster). These ncRNAs are over expressed in human cancer cells. Y RNAs are known nowadays in mammals and nematodes.

1.3 RNA-RNA INTERACTIONS

1.3.1 General Properties

RNA-RNA interactions provide one of the fundamental mechanisms of cellular regulation. Single-stranded nucleic acids readily form complex interaction structures (co-folds) stabilized by complementary base pairing, thereby achieving a high sequence specificity. This recognition principle is utilized for wide variety of biological functions, including the decoding of the genetic code. Each codon is recognized by a complementary anti-coding presented by the corresponding tRNA. In this case, the interaction covers only three nucleotides, which is stabilized by the surrounding ribosomal machinery (Aitken et al., 2010). The direct binding of small RNAs to each other also plays a crucial role for the catalytic activity of the spliceosome, where snRNAs and the pre-mRNA co-fold (Valadkhan, 2005).

A present, we lack high-throughput methods to assess RNA-RNA interactions both *in vitro* and *in vivo*. The direct proof RNA-RNA binding *in vivo* is at least very difficult so that most if not all experimental reports on RNA-RNA cofolds come from *in vitro* experiments. To this end chemical probing (Brunel and Romby, 2000)

is used to determine the secondary structure of both the isolated binding partners and their interaction structures. In practice, chemical probing, which determines by means of context-specific degradation whether a nucleotide is paired or unpaired only provides constraints which in some cases imply an unique RNA secondary structure model, while in other cases ambiguities remain.

In contrast to the experimental difficulties, RNA secondary structures and RNA-RNA interactions can be predicted efficiently by computational methods. The folding problem is posed as a combinatorial matching problem with certain constraints. The nucleotides of the two sequences are represented as vertices of graph, whose edges encode the logically allowed Watson-Crick and GU base pairs. We then search for a matching in this graph (i.e., a subset of edges such that every vertex/nucleotides take part in only one edge) that satisfies further structural constraints and maximizes an energy function. For RNA secondary structure prediction, for instance, one requires that base-pairs do not cross each other. The energy function accounts for base-pair stacking and loop entropies, i.e., it favors parallel base pairs and discounts long unpaired regions (Mathews et al., 1999). Although the RNA-RNA interaction problem (RIP) and the closely related RNA folding problem with arbitrary pseudoknots are NP-complete in their most general forms (Alkan et al., 2006), efficient polynomial-time dynamic programming algorithms can be derived by restricting the space of allowed configurations.

The simplest approaches concatenate the two interacting sequences and subsequently employ a slightly modified standard secondary structure folding algorithm (Hofacker et al., 1994), possibly allowing some pseudoknots (Rivas and Eddy, 1999). The resulting model, however, still does not generate all relevant interaction structures. Alternatively, internal base-pairs in the interaction partners are neglected (Rehmsmeier et al., 2004). `RNAup` (Mückstein et al., 2006) and `intaRNA` (Busch et al., 2008) restrict interactions to a single contact interval; this model class has proved particularly useful for bacterial sRNA/mRNA interactions. To-date only a handful of interaction structures are known that are even more complex, some of which we will encounter later in this section.

1.3.2 MicroRNAs and RNA Interference

MicroRNAs were the first *small regulatory RNAs* found in animals, but turned out not to be the only ones. During the past few years, a variety of additional RNA classes associated with the RNA interference (RNAi) pathways (Kim and Rossi, 2008) were discovered, many of which share functional properties and processing machinery, see Carthew and Sontheimer (2009); Verdel et al. (2009) for recent reviews. Their size of about 20-30 nt and their final destination define these RNAs as a reasonably homogeneous group. They guide large protein complexes to their targets, thus comprising the "RNA sensor" allowing sequence specific binding of the proteins. Both miRNAs and siRNAs form subclasses of this large class of small ncRNAs involved predominantly in gene silencing. MiRNAs stand out from the other small RNAs in several ways: they are processed from an extremely stable hairpin-structured precursor and they are typically highly conserved over long evolutionary timescales (Hertel et al., 2006), indicating crucial regulatory functions.

Although microRNAs clearly interact by basepairing with their targets, our understanding of the rules that govern the recognition of microRNA target sites. Surprisingly, even exact complementarity is insufficient in some cases, while rather poor

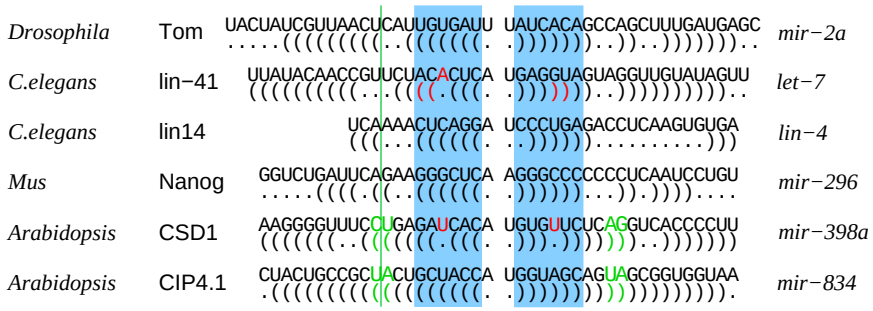

Figure 1.2 Structures of a few miRNA-target interaction, (Brodersen and Voinnet, 2009).

interactions are still functional in some cases. For animal microRNAs, the "seed rule", for instance, postulates that exact complementarity between nucleotides 2-7 with the mRNAs is necessary. There are several well-documented exceptions, however (Brodersen and Voinnet, 2009).

MicroRNAs may act in two distinct ways to downregulate protein expression either by translational repression or by *slicing*, i.e., endonuclease cleavage of the target mRNA. The latter pathway is the principle mode of action of siRNAs and possibly other sub-classes of short RNAs. MicroRNAtarget complementarity, Fig. 1.2, does not necessarily predict the regulatory output of the interaction. In particular complementarity does not imply slicing, while central mismatches prevent slicing, reviewed in (Brodersen and Voinnet, 2009). This observation is consistent with structural models that suggest that the mRNA faces the Ago RNase H active site about 10 nt from the beginning of the miRNA-mRNA duplex (Song et al., 2004). Imperfectly paired target sites in animal transcripts often occur multiple times in the same mRNA. Cooperative action of multiple sites has also been observed.

1.3.3 Small Bacterial RNAs and Translational Control

Small non-coding RNAs (sRNAs) form an important class of regulators of gene expression. In bacteria, their role in the cell was not fully appreciated until the discovery of hundreds of sRNAs in *E. coli* and many other species (Vogel and Sharma, 2005; Waters and Storz, 2009). While the chaperone protein Hfq plays an important role in facilitating sRNA/mRNA interactions, sRNAs are also abundant in species such as *Helicobacter pylori* that lack Hfq (Sharma et al., 2010), Fig. 1.3.

The exact location and structure of the sRNA/target duplex has a decisive impact on the function of the sRNA. While most sRNAs act to downregulate protein expression exceptions are well known. Typically, the sRNAs regulate translation by influencing the accessibility of the Shine-Dalgarno (SD) sequence. If the SD is contained in a stable secondary structure, translation is inhibited. Most sRNA achieve this by binding directly to the SD, thus covering the SD by sRNA/mRNA interaction (Sharma et al., 2007). Several sRNAs, however, (among them DsrA, GlmZ, RNAIII, RprA, RyhB, and Qrr) activate translation by binding the 5' mRNA region in such a way that the mRNA refolds to liberate a sequestered ribosome binding

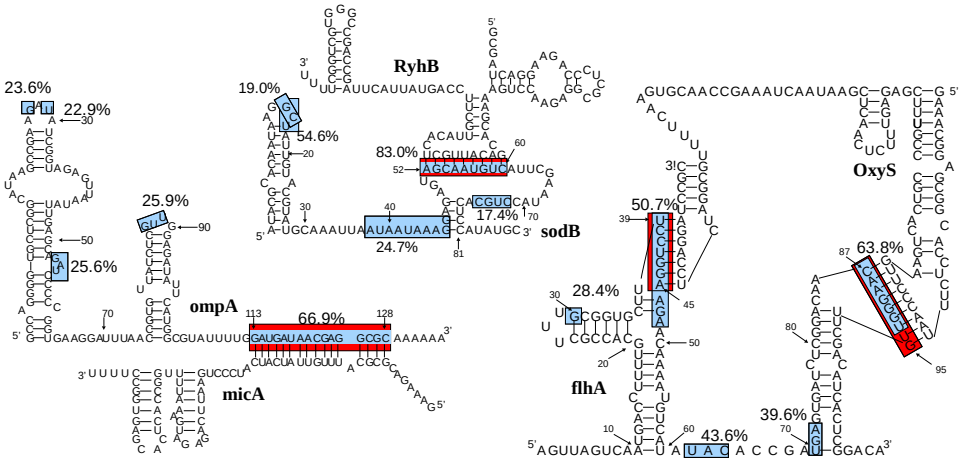

Figure 1.3 Three examples of interaction between bacterial sRNA and mRNAs. The primary interaction region(s), corresponding to the published structures are highlighted in red, hybridization probabilities computed by `rip2` (Huang et al., 2010) are annotated by shaded boxes, identifying additional hybridization regions that may further stabilize the interaction.

site (Fröhlich and Vogel, 2009). The pairing of GadY sRNA to the 3'-end alters processing and increases mRNA stability.

The binding of a small RNA to an mRNA can have long-range effects through a refolding of the mRNA. The binding affinity of *HuR* protein to human mRNAs, for example, can be tuned by artificial "opener" and "closer" RNAs that interact far away from the *HuR* binding motif (Meisner et al., 2004; Hackermüller et al., 2005). So far, it remains open to what extent such induced RNA refolding plays a role in eukaryotes as well. Changes in RNA secondary structure can also be induced by temperature changes or metabolite binding. A wide variety of procaryotic riboswitches, reviewed in (Henkin, 2008), utilizes this mechanism to control either translation (by hiding or exposing the SD) or transcription (by forming an pre-mature terminator signal upstream of the coding sequence). Complex sensors consisting of elaborate RNA structures also play a role in recently-evolved eukaryotic mRNAs (Kaempfer, 2003; Ray et al., 2009).

A recent study of the primary transcriptome of *Helicobacter pylori* (Sharma et al., 2010) demonstrated an unexpectedly high number of anti-sense transcripts and alternative transcription start sites, indicating that there may be an additional layer of regulatory complexity based on these novel transcripts.

1.3.4 Small Nucleolar RNAs and Chemical Modifications

Chemical modifications of RNAs are an evolutionarily old phenomenon. In particular, tRNAs are heavily edited. In total, over 100 structurally distinguishable modified nucleosides have been encountered so far in different types of RNAs from many diverse organisms of all domains of life (Limbach et al., 1994; Dunin-Horkawicz et al., 2006). These modifications increase the chemical diversity of RNA and are beneficial or even critical for proper folding and function of the RNA molecule.

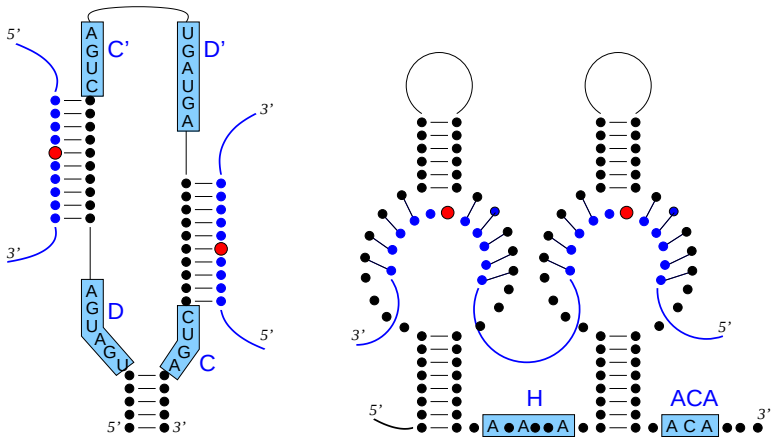

Figure 1.4 The two classes of snoRNA have different structures and different interactions with their targets. Box C/D snoRNAs (left) guide 2'O-methylation to the fifth nucleotide from the D and D' boxes. Some C/D snoRNAs lack a discernible D' box and modify only a single target. In contrast, the target site of box H/ACA snoRNAs is determined by the structure. The target uracil (red ball) is positioned by two specific interactions of the flanking target RNA sequence with the complementary sequence of the recognition loop of the snoRNA (Ni et al., 1997).

Most types of modifications are carried out by specialized, site-specific enzymes. A large number of pseudouridine and 2'-O-methylated nucleotides however, are synthesized by generic RNP complexes that recognize their target sites in in rRNAs, snRNAs, and tRNAs through their RNA components, the H/ACA and C/D snoRNAs, respectively (Bachellerie et al., 2002). These two classes of enzymatically active RNPs are ubiquitous in eukaryotic and archaeal organisms but are not found in bacteria. The RNA components are evolutionarily and structurally unrelated, Fig. 1.4.

A small class of hybrid snoRNAs contain an H/ACA and a box C/D motif. An example is U85, guiding both the pseudouridylation of base U46 and the 2'O-ribose methylation of base C45 of the U5 snRNA (Henras et al., 2004). Several snoRNAs, in particular those of hybrid structure but also several otherwise canonical ones, contain a CAB box (`AGAG`, typically located in a hairpin loop). The signal directs the snoRNAs to the Cajal body, where they guide modifications of the RNA polymerase II-transcribed snRNAs (U1, U2, U4 and U5). In Archaea, the snoRNAs also target tRNA precursors (Singh et al., 2008).

An increasing number of orphan snoRNAs, i.e., snoRNAs lacking targets in rRNAs or snRNAs, has been described in different eukaryotes. In particular, a subgroup of snoRNAs expressed in the mammalian brain does not appear to be involved in modifications of rRNAs or snRNAs. Instead, some of them target specific mRNAs and appear to interfere with A-to-I editing (Vitali et al., 2005; Rogelj, 2006).

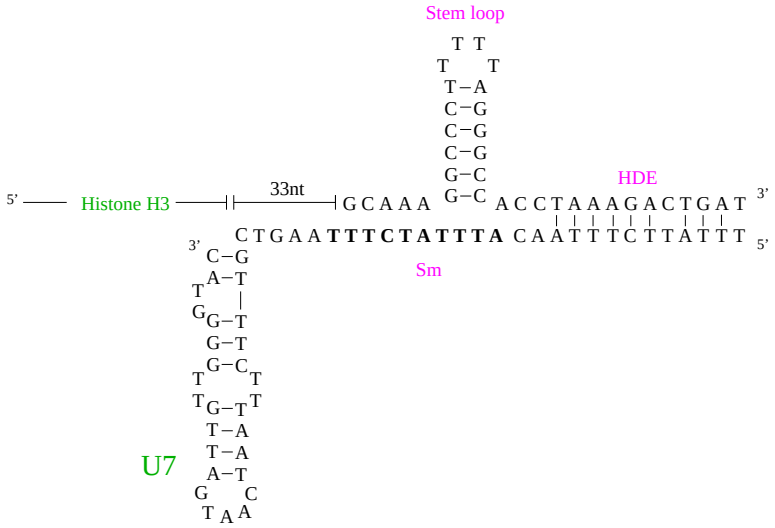

Figure 1.5 U7-directed processing of histone 3' ends. U7 snRNA interacts upstream of the Sm protein binding site with histone downstream element (HDE), about 30 nt downstream of histone-pre-mRNA.

1.3.5 RNAs in RNA End-Processing

The overwhelming majority of protein-coding pol-II transcripts ends in a poly(A) tail that is generated by endonucleolytic cleavage followed by polyadenylation (Proudfoot, 2004). This is not an absolute rule, however.

The mRNAs of the replication-dependent histones lack poly(A) and instead feature a highly conserved stemloop structure in their 3' UTRs. It binds the stemloop binding protein that ensures RNA stability and enhances translational efficiency. The 3' end is determined by base-pairing of the histone-down-stream element (HDE) with the U7 snRNA, Fig. 1.5, which directs endonucleolytic cleavage (Marzluff, 2005).

In contrast, the 3' ends of the pol-II transcribed snRNAs are produced by the *Integrator*, a specialized protein complex (Baillat et al., 2005), while the 3' of the telomerase RNA is processed by the spliceosome in *Schizosaccharomyces pombe* (Box et al., 2008), and RNase P RNA processes the 3' ends of the two long ncRNAs MALAT-1 and MENϵ, utilizing a tRNA-like element to attract the tRNA processing machinery. Besides tRNAs, RNase P RNA in addition cleaves the yeast HRA1 ncRNAs, bacterial riboswitches, and possibly also some box C/D snoRNAs, see (Wilusz and Spector, 2010) for a review.

The rRNA operon is transcribed by pol-I. The maturation of the ribosomal RNAs involves a complex cascade of processing steps. In Eukarya and possibly Archaea, this involves also the use of the atypical box C/D U3 snoRNA (Marz and Stadler, 2009) that acts as an RNA-chaperone mediating the correct structure conformations of the pre-rRNA for endonuclease cleavage (Atzorn et al., 2004).

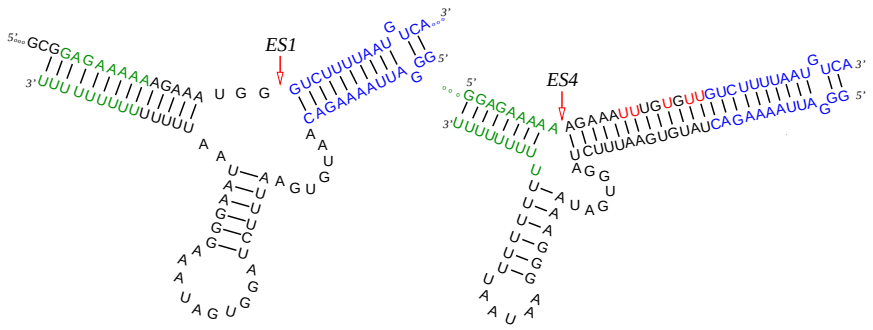

Figure 1.6 U insertion editing of the CYb mRNA of *Trypanosoma* (Yu and Koslowsky, 2006). A single guide RNA (bottom strand) encodes multiple editing sites. The gRNA and its mRNA target (upper strand) form a 3-stem structure that is refolded as the processing moves from one editing site to the next.

1.3.6 Guide RNAs and RNA Editing

Many of the mRNAs of the kinetoplasts (mitochondria) of Trypanosomes and Leishmanias are edited by inserting and/or deletion of uridines, a process that depends on a specialized class on ncRNAs, the guide RNAs (gRNAs). In contrast to the ncRNA-guided editing system in Kinetoplastida, the C→U editing in plant organelles is directed by *cis*-acting elements and a plethora of PPR enzymes, reviewed in (Shikanai, 2006).

The gRNAs of kinetoplastids are typically 50-70nt long and contain three functional elements. A 5-21 nt region on their 5' side acts as anchor specifically recognizing the target mRNA. The "guide region" in the middle of the molecule serves a template for editing. It is complementary (allowing GU pairs, however) to the mature, edited mRNA. The 3'-tail consists of post-transcriptionally added poly-U tail. The editing process takes place in the editosome, a complex structure comprising more than 20 proteins arranged around the mRNA/gRNA pair (Simpson et al., 2004). The editosome typically performs several successive round of of enzymatic reactions templated by a gRNA, Fig. 1.6. Most of the gRNAs are not encoded in the major maxicircle DNA, which also contains the mitochondrial rRNA and protein-coding genes. Instead, large numbers of gRNA genes are be located on the minicircles (Ochsenreiter et al., 2008; Madej et al., 2008).

The U insertions and deletions can be extensive. The 1246 nts ND7 mRNA of *Trypanosoma brucei*, for instance, is processed by inserting 551 and deleting 86 uridines. It is no surprise, therefore, that the U in/dels typically change the open reading frames of the mRNAs. Intriguingly, the pre-mRNAs of the *Trypanosoma brucei* mitochondrion are edited in sometimes alternative ways to yield distinctive protein sequences. Alternative gRNAs are utilized (Ochsenreiter et al., 2008) to expands the diversity of mitochondrial proteins by alternative editing.

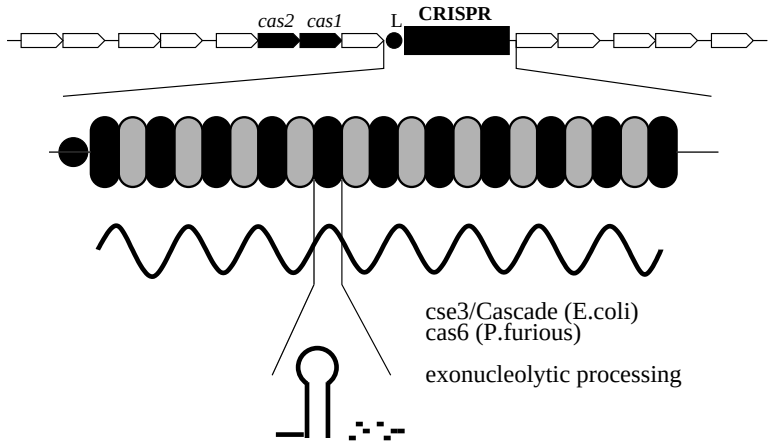

Figure 1.7 Processing of CRISPR content into crRNAs (redrawn from (Horvath and Barrangou, 2010; Karginov and Hannon, 2010)). cas genes (black constrained, white optionally) upstream and downstream of CRISPR cluster. The Leader sequence (L) directly upstream of CRISPR. Transcribed CRISPRs are cleaved within the repeat by cse3 or cas6. Additionally, pre-crRNA is exonucleolytic processed into mature crRNA consisting of a repeat tag (folding into a single hairpin) and the spacer sequence.

1.3.7 CRISPRs

Immunity against viruses and plasmids in 40% of Eubacteria and 90% of Archaea is connected to clustered regularly interspaced short palindromic repeat sequences (CRISPRs). Transcripts from these loci target foreign nucleic acids in by specific-sequence detection. In addition to other well-described defense strategies, such as prevention of adsorption, blocking of injection, and abortive infection (Horvath and Barrangou, 2010), the CRISPR clusters store invading genetic information in preparation for an "immune response" during a second infection. Immunization is established in three phases (van der Oost et al., 2009; Marraffini and Sontheimer, 2010): (1) incorporation of new spacers into CRISPR arrays, (2) expression and processing of CRISPR RNAs (crRNA), and (3) CRISPR interference.

Various *cas* proteins are located upstream and optionally downstream of CRISPR cluster, Fig. 1.7. Six core genes *cas1* to *cas6*, located <1kb around CRISPR have been identified to interact beside repeat-associated mysterious proteins (RAMP) with mature processed CRISPR (crRNA). *cas1* acts as endonuclease within immunization process and *cas2 and cas6* are a sequence specific endoribonuclease. The CRISPR leader located upstream of repeat tags is defined by a low-complexity A-T-rich region. This area acts as promoter for CRISPR transcript. The spacer addition provides novel phage resistance, whereas spacer deletion was shown to result in a loss of phage resistance (Barrangou et al., 2007). With the CRISPR/Cas system horizontal gene transfer between distant organisms was shown by resistance of bacteria that where never invaded by a certain phage before (Godde and Bickerton, 2006).

Recently, a possible relationship between CRISPR interference and eukaryotic RNA silencing has been discussed (Marraffini and Sontheimer, 2010), since there

are many obvious similarities: (1) RNA guiding effector apparatus to the target, (2) the gene function is affected in a programmable and sequence directed manner, (3) adaptive and heritable components used to establish recoverable genomic records of past invasions. On the other hand, the two systems clearly are not homologous: both the protein machinery is completely different and crRNAs are not amplified post-transcriptionally from single stranded precursors. For more details, we refer to the reviews (Marraffini and Sontheimer, 2010; Horvath and Barrangou, 2010; Karginov and Hannon, 2010).

1.4 RIBOZYMES

A ribonucleic acid enzyme (ribozyme) is an RNA molecule that, like the much more familiar protein-based enzymes, catalyzes a chemical reaction. Natural ribozymes include for example the peptidyl transferase activity of the 23S rRNA, the hammerhead and hairpin ribozyme, RNase P RNA and RNase MRP RNA. Similar to the ribosome, the spliceosome is in essence also a ribozyme (Valadkhan, 2007). The efficiency of the catalysis, however, depends to a certain extent on secondary structure features of the mRNA to be catalyzed (Warf and Berglund, 2010).

Ribonuclease P (RNase P) and mitochondrial RNA processing (RNase MRP) are ribozymes acting in processing tRNA and rRNA, respectively. The RNA subunits are evolutionarily related. RNase MRP is eukaryotic specific, where as RNase P is present in all kingdoms of life but not known in plants and heterokonts (Piccinelli et al., 2005). The high similarity of P and MRP RNA secondary structures (Collins et al., 2000) and similarity of the protein contents and interactions of RNase P and MRP (Aspinall et al., 2007; Walker and Engelke, 2006) suggest that P and MRP RNAs are paralogs. Both enzymes consist of an RNA molecule and 9-12 protein subunits, most of which are part of both enzymes. An *in silico* relationship was shown for Pop8 and Rpp14/Pop5 as well as Rpp25 and Pop6 (Rosenblad et al., 2006). The RNA subunit and its interacting proteins build an coevoluting network.

Beyond the relatively small collection of natural ribozymes, it is quite easy to "breed" catalytically active RNAs (and even single-stranded DNAs) in *in vitro* selection (SELEX) experiments, see e.g. the recent review (Talini et al., 2009).

1.5 NCRNA-PROTEIN INTERACTION: RIBONUCLEOPROTEINS

1.5.1 7SK

7SK RNA is known for deuterostomes (Krüger and Benecke, 1987; Murphy et al., 1987; Gruber et al., 2008b,a; Marz et al., 2009). It negatively controls Transcription by regulating the activity of the Positive Transcription Elongation Factor b (P-TEFb) for Polymerase II (Peterlin and Price, 2006; Egloff et al., 2006). In this process HEXIM1/2 proteins dissolve from P-TEFb and bind to the polymerase III transcript 7SK RNA (Yik et al., 2003; Michels et al., 2003, 2004; Blazek et al., 2005). Furthermore the La-related protein 7 (LARP7) was proven to regulate the stability of 7SK RNA (He et al., 2008; Krueger et al., 2008; Markert et al., 2008).

The network of 7SK RNA and its specific protein partners is a metazoan invention (Gürsoy et al., 2000; Gruber et al., 2008b,a; Marz et al., 2009). The well characterized LARP7 homologs (Bousquet-Antonelli and Deragon, 2009) are found

in most metazoan clades HEXIM was found in all clades of metazoans excluding Platyhelminthes. Two copies are known for eutheria (Byers et al., 2005; Marz et al., 2009).

In contrast, another protein known to interact with 7SK RNA, MePCE/BCDIN3 has a much broader phylogenetic distribution, indicating that it has other important functions beyond its interaction with 7SK RNA.

1.5.2 SRP RNA

The signal recognition particle (RNP) targets ribosomes to the endoplasmatic reticulum (ER) in order to translate proteins into the ER. The SRP complex consists of a highly conserved ∼300 nt SRP RNA (or 7S RNA) and six proteins (9, 14, 19, 54, 68 and 72 SRPs) in eukaryotes. In archaeas four protein particles are absent and SRP19 and SRP54 exists compared to one Ffh protein (homolog to SRP54) and a RNA molecule (also known as 4.5S RNA or 6S RNA in *Bacillus*) (Rosenblad et al., 2003). Although the network of SRP RNA and corresponding proteins is highly conserved, an evolving network between the three main kingdoms of life is clearly visible (Andersen et al., 2006).

Eukaryotic cells have been shown to decrease protein synthesis and increasing the expression of protein quality control mechanisms, such as chaperones and proteases. The bacterial SRP receptor, FtsY, inhibits the translation of both SRP-dependent and SRP-independent proteins (Bürk et al., 2009).

In higher plants two different SRP-dependent mechanisms are known: once proteins are post-translational transferred to chloroplasts and on the other hand proteins are co-translational encoded by the plastid genome (Rosenblad and Samuelsson, 2004). Although for the first system in general no RNA component seems to be necessary. However in single plastids from red algae and chlorophyta a reminiscent eubacterial SRP was identified. The co-translational protein requires the SRP RNA, which accelerates the interaction between the SRP and SRP receptor 200-fold (Jaru-Ampornpan et al., 2009). This SRP RNA is missing in the chloroplast SRP (cpSRP) pathway. Instead, the cpSRP and cpSRP receptor (cpFtsY) by themselves can interact 200-fold faster than their bacterial homologous (Jaru-Ampornpan et al., 2009).

1.5.3 Telomerase RNA

Linear chromosomal ends are replicated by a telomerase enzyme. This consists of the catalytic protein component telomerase reverse transcriptase (TERT) and the core functional unit telomerase RNA (Box et al., 2008). Telomerase dates back to the origin of eukaryotes. Notable exceptions are diptera including *Anopheles* and *Drosophila*, which use retrotransposons or unequal recombination instead of a telomerase enzyme. Although TERT proteins are experimentally validated for a wide range of eukaryotes, the knowledge of RNA:protein interacting networks implies that telomerase RNA exists for more organisms than the experimentally validated small phylogenetic groups within vertebrates, yeasts, ciliates and plasmodia (Podlevsky et al., 2008; Xie et al., 2008).

1.5.4 Spliceosomal snRNAs, SL RNA, SmY RNA

Post-transcriptional modifications by connecting different genomic parts are known in different variations and for diverse phylogenetic groups.

Small nuclear RNAs act with up to 200 proteins as a large RNP (spliceosome) in eukaryotes to splice exons of protein-coding genes together (Nilsen, 2003). We find two homolog systems for splicing in eukaryotes: The major spliceosome splices with U1, U2, U4, U5 and U6 snRNA more than 90% of the protein-coding genes, whereas the minor spliceosome with U11, U12, U4atac, U5 and U6atac plays an ancillary role in the nucleus. Although the splicing reactions seem to date back until the last unknown common ancester (LUCA), the interaction to proteins seems to be an eukaryotic innovation (Dávila López et al., 2008). Contrary to *cis*-splicing, in eight eukaryotic phyla a short leader sequence derived from small SL RNA is added to the 5' end of a mRNA by *trans*-splicing (Hastings, 2005; Nilsen, 2001). The Sm protein binds to SL RNAs in order to act in this spliceosome-catalyzed process. In nematoda SmY RNA are hypothesized to recycle the spliceosome proteins after SL RNAs are consumed in the *trans*-splicing reaction (MacMorris et al., 2007). Direct interactions are not known.

1.5.5 vault RNA

The vault ribonucleoprotein complex consists of 8-16 non-coding vault RNAs, a major vault protein (MVP) and two minor vault proteins (VPARP and TEP1) (Smith, 2001). Vault RNAs are polymerase III transcripts with length of about 100 nt which are known throught the deuterostomes (Stadler et al., 2009) and in a few protostomes (Mosig et al., 2009). It is known that they are not necessary for the structural assembly of vault particles (van Zon et al., 2003).

1.5.6 Long ncRNA as Coat-Hanger?

A handful of long non-coding RNAs appears to play a crucial role in the organization of complex RNPs, acting like a "coat-hanger" to bring together different proteins. An example is the huge telomerase RNA of yeasts (Zappulla and Cech, 2006). Another example is the heat-shock ncRNAs hsrω of *Drosophila*. Different isoforms of hsrω are expressed nearly ubiquitously in a developmentally regulated patterns. Upon heat shock, the nuclear-retained longest isoform sequesters a variety of in particular nuclear RNA processing factors (Jolly and Lakhotia, 2006). Similarly, NEAT1 (also known as MENϵ) may act as organizing factor of the paraspeckles (Bond and Fox, 2009; Sasaki et al., 2009), ribonucleoprotein bodies found in the interchromatin space of mammalian cell nuclei.

At present, it is unknown to what extent other long ncRNAs function as stabilizing backbones of large RNPs, bringing together specific combinations of proteins, and what role such complexed might play.

Acknowledgements. This work was supported in part by the State of Saxony through a "Landesstipendium".

Bibliography

Aitken, C. E., Petrov, A., and Puglisi, J. D., 2010. Single ribosome dynamics and the mechanism of translation. *Annu Rev Biophys.* **39**. Doi: 10.1146/annurev.biophys.093008.131427.

Alkan, C., Karakoc, E., Nadeau, J. H., Sahinalp, S. C., and Zhang, K. Z., 2006. RNA-RNA interaction prediction and antisense RNA target search. *J. Comput. Biol.* **13**: 267–282.

Andersen, E. S., Rosenblad, M. A., Larsen, N., Westergaard, J. C., Burks, J., Wower, I. K., Wower, J., Gorodkin, J., Samuelsson, T., and Zwieb, C., 2006. The tmRDB and SRPDB resources. *Nucleic Acids Res* **34**: 163–168.

Aspinall, T. V., Gordon, J. M. B., Bennet, H. J., Karahalios, P., Bukowski, J.-P., Walker, S. C., Engelke, D. R., and Avis, J. M., 2007. Interactions between subunits of *Saccharomyces cerevisiae* RNase MRP support a conserved eukaryotic RNase P/MRP architecture. *Nucleic Acids Res.* **35**: 6439–6450.

Atzorn, V., Fragapane, P., and Kiss, T., 2004. U17/snR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. *Mol Cell Biol.* **24**: 1769–1778.

Bachellerie, J. P., Cavaillé, J., and Hüttenhofer, A., 2002. The expanding snoRNA world. *Biochimie* **84**: 775–790.

Baillat, D., Hakimi, M. A., Näär, A. M., Shilatifard, A., Cooch, N., and Shiekhattar, R., 2005. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**: 265–276.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709–1712.

Blazek, D., Barboric, M., Kohoutek, J., Oven, I., and Peterlin, B. M., 2005. Oligomerization of HEXIM1 via 7SK snRNA and coiled-coil region directs the inhibition of P-TEFb. *Nucleic Acids Res.* **33**: 7000–7010.

Bompfünewerer, A. F., et al., 2005. Evolutionary patterns of Non-Coding RNAs. *Th. Biosci.* **123**: 301–369.

Bond, C. S. and Fox, A. H., 2009. Paraspeckles: nuclear bodies built on long noncoding RNA. *J. Cell Biol.* **186**: 637–644.

Bousquet-Antonelli, C. and Deragon, J. M., 2009. A comprehensive analysis of the La-motif protein superfamily. *RNA* **15**: 750–64.

Box, J. A., Bunch, J. T., Tang, W., and Baumann, P., 2008. Spliceosomal cleavage generates the 3' end of telomerase RNA. *Nature* **456**: 910–914.

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

Brunel, C. and Romby, P., 2000. Probing RNA structure and RNA-ligand complexes with chemical probes. *Methods Enzymol.* **318**: 3–21.

Bürk, J., Weiche, B., Wenk, M., Boy, D., Nestel, S., Heimrich, B., and Koch, H. G., 2009. Depletion of the signal recognition particle receptor inactivates ribosomes in *Escherichia coli*. *J Bacteriol* **191**: 7017–7026.

Busch, A., Richter, A., and Backofen, R., 2008. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**: 2849–2856.

Byers, S. A., Price, J. P., Cooper, J. J., Li, Q., and Price, D. H., 2005. HEXIM2, a HEXIM1-related protein, regulates positive transcription elongation factor b through association with 7SK. *J Biol Chem.* **280**: 16360–16377.

Carthew, R. W. and Sontheimer, E. J., 2009. Origins and mechanisms of miRNAs and siRNAs. *Cell* **136**: 642–655.

Christov, C. P., Gardiner, T. J., Szüts, D., and Krude, T., 2006. Functional requirement of noncoding Y RNAs for human chromosomal DNA replication. *Mol Cell Biol* **26**: 6993–7004. doi:10.1128/MCB.01060-06.

Collins, L. J., Moulton, V., and Penny, D., 2000. Use of RNA secondary structure for studying the evolution of RNase P and RNase MRP. *J Mol Evol.* **51**: 194–204.

Dávila López, M., Rosenblad, M. A., and Samuelsson, T., 2008. Computational screen for spliceosomal RNA genes aids in defining the phylogenetic distribution

of major and minor spliceosomal components. *Nucleic Acids Res* **36**: 3001–10. doi:10.1093/nar/gkn142.

Dunin-Horkawicz, S., Czerwoniec, A., Gajda, M. J., Feder, M., Grosjean, H., and Bujnicki, J. M., 2006. MODOMICS: a database of RNA modification pathways. *Nucleic Acids Res.* **34**: D145–D149.

Egloff, S., Van Herreweghe, E., and Kiss, T., 2006. Regulation of polymerase II transcription by 7SK snRNA: two distinct RNA elements direct P-TEFb and HEXIM1 binding. *Mol. Cell. Biol.* **26**: 630–642.

Fröhlich, K. S. and Vogel, J., 2009. Activation of gene expression by small RNA. *Curr Opin Microbiol.* **12**: 674–682.

Godde, J. S. and Bickerton, A., 2006. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J Mol Evol* **62**: 718–729.

Gruber, A. R., Kilgus, C., Mosig, A., Hofacker, I. L., Hennig, W., and Stadler, P. F., 2008a. Arthropod 7SK RNA. *Mol Biol Evol* **25**: 1923–1930.

Gruber, A. R., Koper-Emde, D., Marz, M., Tafer, H., Bernhart, S., Obernosterer, G., Mosig, A., Hofacker, I. L., Stadler, P. F., and Benecke, B. J., 2008b. Invertebrate 7SK snRNAs. *J Mol Evol* **66**: 107–115.

Gürsoy, H.-C., Koper, D., and Benecke, B.-J., 2000. The vertebrate 7S K RNA separates hagfish (*Myxine glutinosa*) and lamprey (*Lampetra fluviatilis*). *J. Mol. Evol.* **50**: 456–464.

Hackermüller, J., Meisner, N.-C., Auer, M., Jaritz, M., and Stadler, P. F., 2005. The effect of RNA secondary structures on RNA-ligand binding and the modifier RNA mechanism: A quantitative model. *Gene* **345**: 3–12.

Harrison, B. R., Yazgan, O., and Krebs, J. E., 2009. Life without RNAi: noncoding RNAs and their functions in *Saccharomyces cerevisiae*. *Biochem Cell Biol* **87**: 767–779.

Hastings, K. E., 2005. SL trans-splicing: easy come or easy go? *Trends Genet.* **21**: 240–247.

He, S., Su, H., Liu, C., Skogerbø, G., He, H., He, D., Zhu, X., Liu, T., Zhao, Y., and Chen, R., 2008. MicroRNA-encoding long non-coding RNAs. *BMC Genomics* **21**: 236.

Hekimoglu, B. and Ringrose, L., 2009. Non-coding RNAs in polycomb/trithorax regulation. *RNA Biol* **6**: 129–137.

Henkin, T. M., 2008. Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev.* **22**: 3383–3390.

Henras, A. K., Capeyrou, R., Henry, Y., and Caizergues-Ferrer, M., 2004. Cbf5p, the putative pseudouridine synthase of H/ACA-type snoRNPs, can form a complex with Gar1p and Nop10p in absence of Nhp2p and box H/ACA snoRNAs. *RNA* **10**: 1704–1712.

Hertel, J., Lindemeyer, M., Missal, K., Fried, C., Tanzer, A., Flamm, C., Hofacker, I. L., Stadler, P. F., and The Students of Bioinformatics Computer Labs 2004 and 2005, 2006. The expansion of the metazoan microRNA repertoire. *BMC Genomics* **7**: 15 [epub].

Hiller, M., et al., 2009. Conserved Introns Reveal Novel Transcripts in *Drosophila melanogaster*. *Genome Res.* **19**: 1289–1300.

Hofacker, I. L., Fontana, W., Stadler, P. F., Bonhoeffer, L. S., Tacker, M., and Schuster, P., 1994. Fast folding and comparison of RNA secondary structures. *Monatsh. Chem.* **125**: 167–188.

Horvath, P. and Barrangou, R., 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**: 167–170.

Huang, F. W. D., Qin, J., Reidys, C. M., and Stadler, P. F., 2010. Target prediction and a statistical sampling algorithm for RNA-RNA interaction. *Bioinformatics* **26**: 175–181.

Jaru-Ampornpan, P., Nguyen, T. X., and Shan, S. O., 2009. A distinct mechanism to achieve efficient signal recognition particle (SRP)-SRP receptor interaction by the chloroplast srp pathway. *Mol Biol Cell* **20**: 3965–3973.

Jolly, C. and Lakhotia, S. C., 2006. Human sat III and *Drosophila* hsrω transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells. *Nucleic Acids Res.* **34**: 5508–5514.

Kaempfer, R., 2003. RNA sensors: novel regulators of gene expression. *EMBO Rep.* **4**: 1043–1047.

Karginov, F. V. and Hannon, G. J., 2010. The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol Cell* **37**: 7–19.

Khalil, A. M., et al., 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* **106**: 11675–11680.

Kim, D. and Rossi, J., 2008. RNAi mechanisms and applications. *Biotechniques* **44**: 613–616.

Komissarova, N., Becker, J., Solter, S., Maria, K., and Kashlev, M., 2002. Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Molecular Cell* **10**: 11511162.

Krude, T., 2010. *Genome Stability and Human Diseases*, chapter Non-coding RNAs: New players in the Field of Eukaryotic DNA Replication. Springer.

Krueger, B. J., et al., 2008. LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res* **36**: 2219–2229.

Krüger, W. and Benecke, B. J., 1987. Structural and functional analysis of a human 7SK RNA gene. *J. Mol. Biol.* **195**: 31–41.

Limbach, P. A., Crain, P. F., and McCloskey, J. A., 1994. Summary: the modified nucleosides of RNA. *Nucleic Acids Res.* **22**: 2183–2196.

MacMorris, M., Kumar, M., Lasda, E., Larsen, A., Kraemer, B., and Blumenthal, T., 2007. A novel family of *C. elegans* snRNPs contains proteins associated with *Trans*-splicing. *RNA* **13**: 511–520.

MacNeill, S. A., 2001. DNA replication: Partners in the Okazaki two-step. *Current Biology* **11**: R842–R844.

Madej, M. J., Niemann, M., Hüttenhofer, A., and Göringer, H. U., 2008. Identification of novel guide RNAs from the mitochondria of *Trypanosoma brucei*. *RNA Biology* **5**: 84–91.

Markert, A., Grimm, M., Martinez, J., Wiesner, J., Meyerhans, A., Meyuhas, O., Sickmann, A., and Fischer, U., 2008. The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep.* **9**: 569–575.

Marques, A. C. and Ponting, C. P., 2009. Catalogues of mammalian long noncoding RNAs: modest conservation and incompleteness. *Genome Biol.* **10**: R124.

Marraffini, L. A. and Sontheimer, E. J., 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* **11**: 181–190.

Marz, M., Donath, A., Verstraete, N., Nguyen, V. T., Stadler, P. F., and Bensaude, O., 2009. Evolution of 7SK RNA and its protein partners in metazoa. *Mol Biol Evol* **26**: 2821–2830.

Marz, M. and Stadler, P. F., 2009. Comparative analysis of eukaryotic U3 snoRNA. *RNA Biol* **6**: 503–507.

Marzluff, W. F., 2005. Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts. *Curr. Opin. Cell. Biol.* **17**: 274–280.

Mathews, D., Sabina, J., Zuker, M., and Turner, D., 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**: 911–940.

Mattick, J. S., Amaral, P. P., Dinger, M. E., Mercer, T. R., and Mehler, M. F., 2009. RNA regulation of epigenetic processes. *Bioessays* **31**: 51–59.

Matzke, M., Kanno, T., Daxinger, L., Huettel, B., and Matzke, A. J., 2009. RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol.* **21**: 367–376.

Meisner, N. C., Hackermüller, J., Uhl, V., Aszódi, A., Jaritz, M., and Auer, M., 2004. mRNA openers and closers: modulating AU-rich element-controlled mRNA stability by a molecular switch in mRNA secondary structure. *Chembiochem* **5**: 1432–1447.

Michels, A. A., Nguyen, V. T., Fraldi, A., Labas, V., Edwards, M., Bonnet, F., Lania, L., and Bensaude, O., 2003. MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. *Mol Cell Biol* **23**: 4859–4869.

Michels, A. A., et al., 2004. Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J.* **23**: 2608–2619.

Mosig, A., Zhu, L., and Stadler, P. F., 2009. Customized strategies for discovering distant ncRNA homologs. *Brief. Funct. Genomics Proteomics* **8**: 451–460.

Mückstein, U., Tafer, H., Hackermüller, J., Bernhard, S. B., Stadler, P. F., and Hofacker, I. L., 2006. Thermodynamics of RNA-RNA binding. *Bioinformatics* **22**: 1177–1182.

Murphy, S., Di Liegro, C., and Melli, M., 1987. The *in vitro* transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. *Cell* **51**: 81–87.

Nadal, A., Eritja, R., Esteve, T., and Pla, M., 2005. "parallel" and "antiparallel tail-clamps" increase the efficiency of triplex formation with structured DNA and RNA targets. *Chembiochem.* **6**: 1034–1042.

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

Nilsen, T. W., 2001. Evolutionary origin of SL-addition *trans*-splicing: still an enigma. *Trends Genet* **17**: 678–680.

Nilsen, T. W., 2003. The spliceosome: the most complex macromolecular machine in the cell? *Bioessays* **25**: 1147–1149.

Ochsenreiter, T., Cipriano, M., and Hajduk, S. L., 2008. Alternative mRNA editing in trypanosomes is extensive and may contribute to mitochondrial protein diversity. *PLoS ONE* **3**: e1566.

Peterlin, B. M. and Price, D. H., 2006. Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell.* **23**: 297–305.

Piccinelli, P., Rosenblad, M. A., and Samuelsson, T., 2005. Identification and analysis of ribonuclease P and MRP RNA in a broad range of eukaryotes. *Nucleic Acids Res.* **33**: 4485–4495.

Podlevsky, J. D., Bley, C. J., Omana, R. V., Qi, X., and Chen, J. J., 2008. The telomerase database. *Nucleic Acids Res.* **36**: D339–D343.

Proudfoot, N., 2004. New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr Opin Cell Biol* **16**: 272278.

Ray, P. S., Jia, J., Yao, P., Majumder, M., Hatzoglou, M., and Fox, P. L., 2009. A stress-responsive RNA switch regulates VEGFA expression. *Nature* **457**: 915–919.

Rehmsmeier, M., Steffen, P., Höchsmann, M., and Giegerich, R., 2004. Fast and effective prediction of microRNA/target duplexes. *Gene* **10**: 1507–1517.

Rivas, E. and Eddy, S. R., 1999. A dynamic programming algorithms for RNA structure prediction including pseudoknots. *J. Mol. Biol.* **285**: 2053–2068.

Rogelj, B., 2006. Brain-specific small nucleolar RNAs. *J Mol Neurosci* **28**: 103–109.

Rosenblad, M. A., Gorodkin, J., Knudsen, B., Zwieb, C., and Samuelsson, T., 2003. SRPDB: Signal Recognition Particle Database. *Nucleic Acids Res* **31**: 363–364.

Rosenblad, M. A., López, M. D., Piccinelli, P., and Samuelsson, T., 2006. Inventory and analysis of the protein subunits of the ribonucleases P and MRP provides further evidence of homology between the yeast and human enzymes. *Nucleic Acids Res* **34**: 5145–5156.

Rosenblad, M. A. and Samuelsson, T., 2004. Identification of chloroplast signal recognition particle RNA genes. *Plant Cell Physiol* **45**: 1633–1639.

Sasaki, Y. T. F., Ideue, T., Sano, M., Mituyama, T., and Hirose, T., 2009. MENϵ/β noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci USA* **106**: 2525–2530.

Sharma, C. M., Darfeuille, F., Plantinga, T. H., and Vogel, J., 2007. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev* **21**: 2804–2817.

Sharma, C. M., et al., 2010. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**: 250–255.

Sheik, M. J., Gaughwin, P. M., Lim, B., Robson, P., and Lipovich, L., 2010. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA* **16**: 324–337.

Shikanai, T., 2006. RNA editing in plant organelles: machinery, physiological function and evolution. *Cell. Mol. Life Sci.* **63**: 698–708.

Simpson, L., Aphasizhev, R., Gao, G., and Kang, X., 2004. Mitochondrial proteins and complexes in *Leishmania* and *Trypanosoma* involved in U-insertion/deletion RNA editing. *RNA* **10**: 159–170.

Singh, S. K., Gurha, P., and Gupta, R., 2008. Dynamic guide-target interactions contribute to sequential 2'-O-methylation by a unique archaeal dual guide box C/D sRNP. *RNA* **14**: 1411–1423.

Smith, S., 2001. The world according to PARP. *Trends Biochem Sci* **26**: 174–179.

Song, J. J., Smith, S. K., Hannon, G. J., and Joshua-Tor, L., 2004. Crystal structure of argonaute and its implications for RISC slicer activity. *Science* **305**: 1434–1437.

Stadler, P. F., et al., 2009. Evolution of vault rnas. *Mol. Biol. Evol.* **26**: 1975–1991.

Stein, A. J., Fuchs, G., Fu, C., Wolin, S. L., and Reinisch, K. M., 2005. Structural insights into RNA quality control: the Ro autoantigen binds misfolded RNAs via its central cavity. *Cell* **121**: 529–539.

Suzuki, T., 2008. Targeted gene modification by oligonucleotides and small DNA fragments in eukaryotes. *Front Biosci.* **13**: 737–744.

Swiezewski, S., Liu, F., Magusin, A., and Dean, C., 2009. Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature* **462**: 799–802.

Szüts, D., Christov, C., Kitching, L., and Krude, T., 2005. Distinct populations of human PCNA are required for initiation of chromosomal DNA replication and concurrent DNA repair. *Exp Cell Res* **311**: 240–250.

Szüts, D., Kitching, L., Christov, C., Budd, A., Peak-Chew, S., and Krude, T., 2003. RPA is an initiation factor for human chromosomal DNA replication. *Nucleic Acids Res* **31**: 1725–1734.

Talini, G., Gallori, E., and Maurel, M. C., 2009. Natural and unnatural ribozymes: back to the primordial RNA world. *Res Microbiol.* **160**: 457–465.

The Athanasius F. Bompfünewerer RNA Consortium:, et al., 2007. RNAs everywhere: Genome-wide annotation of structured RNAs. *J. Exp. Zool. B: Mol. Dev. Evol.* **308B**: 1–25.

The ENCODE Project Consortium, 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**: 799–816.

The FANTOM Consortium, 2005. The transcriptional landscape of the Mammalian genome. *Science* **309**: 1159–1563.

Valadkhan, S., 2005. snRNAs as the catalysts of pre-mRNA splicing. *Curr. Op. Chem. Biol.* **9**: 603–608.

Valadkhan, S., 2007. The spliceosome: a ribozyme at heart? *Biol Chem.* **388**: 693–697.

van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M., and Brouns, S. J., 2009. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem Sci* **34**: 401–407.

van Zon, A., Mossink, M. H., Scheper, R. J., Sonneveld, P., and Wiemer, E. A. C., 2003. The vault complex. *Cell. Mol. Life Sci.* **60**: 1828–1837.

Verdel, A., Vavasseur, A., Le Gorrec, M., and Touat-Todeschini, L., 2009. Common themes in siRNA-mediated epigenetic silencing pathways. *Int J Dev Biol.* **53**: 245–257.

Vitali, P., Basyuk, E., Le Meur, E., Bertrand, E., Muscatelli, F., Cavaillé, J., and Huttenhofer, A., 2005. ADAR2-mediated editing of RNA substrates in the nucleolus is inhibited by C/D small nucleolar RNAs. *J Cell Biol* **169**: 745–753.

Vogel, J. and Sharma, C. M., 2005. How to find small non-coding RNAs in bacteria. *Biol Chem* **386**: 1219–1238.

Walker, S. C. and Engelke, D. R., 2006. Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol.* **41**: 77–102.

Warf, M. B. and Berglund, J. A., 2010. Role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem Sci.* **35**: 169–178.

Waters, L. S. and Storz, G., 2009. Regulatory RNAs in bacteria. *Cell* **136**: 615–628.

Wilusz, J. E. and Spector, D. L., 2010. An unexpected ending: noncanonical 3' end processing mechanisms. *RNA* **16**: 259–266.

Xie, M., Mosig, A., Qi, X., Li, Y., Stadler, P. F., and Chen, J. J.-L., 2008. Size variation and structural conservation of vertebrate telomerase RNA. *J. Biol. Chem.* **283**: 2049–2059.

Yik, J. H., Chen, R., Nishimura, R., Jennings, J. L., Link, A. J., and Zhou, Q., 2003. Inhibition of P-TEFb (CDK9/Cyclin T) kinase and RNA polymerase II transcription by the coordinated actions of HEXIM1 and 7SK snRNA. *Mol Cell* **12**: 971–982.

Yu, L. E. and Koslowsky, D. J., 2006. Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: Structure probing of a gRNA bound to its cognate mRNA. *RNA* **12**: 1050–1060.

Zappulla, D. C. and Cech, T. R., 2006. RNA as a flexible scaffold for proteins: yeast telomerase and beyond. *Cold Spring Harb Symp Quant Biol.* **71**: 217–224.