

# Editing Reactions from the Perspective of RNA Structure

Matthias Homann

1	Introduction.....	2
1.1	Principles of RNA Structure and RNA–Protein Interactions.....	3
2	Editing of mRNA Sequences .....	5
2.1	Paramyxovirus .....	5
2.2	<i>Physarum</i> .....	5
2.3	RNA Editing in Plant Organelles.....	7
2.4	Mammalian Editing: apoB.....	9
2.5	RNA Editing by Adenosine Deamination.....	12
2.6	Insertion/Deletion Editing in Kinetoplastids .....	17
3	Future and Conclusions.....	24

*“If something is in me which can be called religious then it is the unbounded admiration for the structure of the world so far as our science can reveal it.”*

Albert Einstein

**Abstract** RNA editing belongs to the large group of processing reactions that are required to convert primary RNA transcripts into mature and functional transcripts. The main determinants of specificity rest in the three-dimensional structures of RNA and protein molecules that act in concert to coordinate and regulate the posttranscriptional steps in gene expression. Many high-resolution structures of RNA–protein complexes, including the ribosome, have become available during the last decade and have offered detailed views of the intracellular RNA world. The focus of this review is to highlight the contributions of RNA structure to the specificity and efficiency of RNA editing. Editing occurs by a variety of mechanisms, but the fidelity of the reactions critically depends on the specific sequences and structures of the RNA molecules involved and on their recognition by trans-acting factors, including proteins and RNA. Hence, the editing machineries, also termed “editosomes”, make

---

Genetics, Darmstadt University of Technology, Schnittspahnstr. 10, 64287 Darmstadt, Germany; mhomann@hrzpub.tu-darmstadt.de

use of RNA–RNA, RNA–protein and protein–protein interactions to achieve specificity and efficiency. High-resolution structures of protein components of various editosomes exist, but reports of RNA structures and RNA–protein complexes are still limited. Progress can be expected in the near future from more efficient purification and crystallization techniques developed in other fields of RNA processing, like RNA interference, splicing and catalysis. Although each structure reveals only a static view of a multistep reaction, they will eventually lead to a better understanding of the dynamic molecular machines involved in RNA editing.

## 1 Introduction

Eukaryotic gene expression involves extensive processing of primary RNA transcripts to generate mature and functional RNAs. A wide variety of RNA binding proteins act in concert with complex three-dimensional structures of RNA molecules to coordinate and regulate posttranscriptional steps in gene expression. High-resolution structures of RNA–protein complexes as well as of single components involved in RNA modification (reviewed in Reichow et al. 2007), RNA splicing (Stark and Lührmann 2006), RNA catalysis (RNase P: Evans et al. 2006; group I introns: Vicens and Cech 2006; small ribozymes: Nelson and Uhlenbeck 2006), RNA transport (Müller et al. 2007) and RNA interference mechanisms (Ma et al. 2004, 2005; Song et al. 2004; Parker et al. 2005; MacRae et al. 2006) have become available during the last decade. Above all, the crystal structures of the ribosome and its subunits have offered unprecedented views of the intracellular RNA world (for reviews, see Moore and Steitz 2003; Noller 2005) and allow predictions as to the dynamics and the molecular interactions that underlie RNA-based gene expression pathways.

The focus of this review is to highlight the contributions of RNA structure to the specificity and efficiency of RNA editing. RNA editing is a widespread phenomenon in which the sequence information of certain transcripts is altered by enzymatic modification or the addition/deletion of nucleotides. Editing occurs by a variety of mechanisms, but a minimal set of components can be defined that are essential for accurate editing. First, the RNA to be edited usually contains sequence and/or structural information needed to specify the editing site (“cis-acting signals”). Second, trans-acting factors, which can be proteins as well as RNA, have to interact with the target RNA at or near the editing site. Third, enzymatic activities have to be directed to the specific site and catalyze the editing reaction(s). All reaction steps are catalyzed within RNP complexes (reviewed in Moore 2005), and the accuracy of editing critically depends on the specific RNA sequences and structures and on their recognition by the protein factors. These RNP complexes have in some cases been termed “editosome”, in

analogy with the two major RNA-based molecular machines, the ribosome and the spliceosome.

### ***1.1 Principles of RNA Structure and RNA–Protein Interactions***

Owing to the wealth of high-resolution structures of the ribosome (e.g. Ban et al. 2000; Wimberly et al. 2000; Hansen et al. 2002; Korostelev et al. 2006), many principles of RNA folding have been uncovered (reviewed in Noller 2005). Local RNA motifs such as U-turns, T-loops, K-turns, A-minor interactions, A platforms and tetraloops have been classified (Leontis and Westhoff 2003; Lescoute et al. 2005; Leontis et al. 2006) and are likely to contribute to RNA editing processes as well. However, only one of the RNAs/RNP complexes involved in RNA editing has been crystallized so far (Schumacher et al. 2006). In addition, the solution structure of the RNA substrate for a specific C-to-U deamination editing event is available (see Sect. 2.4; Maris et al. 2005a).

RNA–protein interactions are governed by a distinct set of protein motifs that are involved in RNA binding (Chen and Varani 2005; see Table 1). The most common domains are the RNA recognition motif (RRM, also called RNA binding domain, RBD, or ribonucleoprotein domain, RNP), the K-homology (KH) domain, and the double-stranded RNA binding domain, dsRBD. In humans, the RRM alone is present in about 500 proteins, or 2% of the human genome, often found in multiple copies or in combination with other domains (reviewed in Maris et al. 2005b). Another very common domain is the zinc-finger motif that accounts for about 3% of the human genome, although only few zinc-fingers of the CCCH and CCHC type are known to be involved in RNA recognition (Lu et al. 2003; Hudson et al. 2004). Table 1 lists several other motifs that are present in specific enzymes acting on RNA (e.g. RNase III) or that function in specific processes such as RNA interference (e.g. PAZ and PIWI domains). Some particularly striking examples of RNA recognition are provided by proteins using helical repeat motifs, such as the Pumilio proteins (Wang et al. 2002), TRAP (Antson et al. 1999) and Sm/Lsm proteins (Kambach et al. 1999). In addition, small domains with highly basic stretches rich in Arg and Lys have been found to be involved in RNA recognition, e.g. in the cases of HIV-1 Tat (Ye et al. 1995) and Rev (Battiste et al. 1996) proteins. These domains, however, are not strictly pre-organized motifs, since they will fold into their active structure only upon binding to RNA (Frankel and Smith 1998). High-resolution structures of the most common motifs have been determined, either in their isolated form or in complex with RNA. These structures generally reveal a highly conserved folding of the respective motif, with little conservation at the amino acid sequence level. Accordingly, these motifs function as platforms capable of high-affinity recognition of practically any RNA sequence. This principle is best illustrated by the large number of proteins using the RRM

**Table 1** Protein motifs involved in RNA recognition

Motif	Appearance	Reference
RRM, RBD, RNP	Sxl	Handa et al. (1999)
	PAB	Deo et al. (1999)
	U1A	Allain et al. (1997)
	U2B/U2A	Price et al. (1998)
	hnRNPA1	Ding et al. (1999)
	Nucleolin	Johansson et al. (2004)
KH	Nova	Lewis et al. (2000)
DsRBD	Staufen	Ramos et al. (2000)
	RNase III	Kharrat et al. (1995)
PAZ	Dicer	MacRae et al. (2006)
	Ago	Ma et al. (2004)
PIWI	Ago	Ma et al. (2005)
Helical repeat	Pumilio	Wang et al. (2002)
	TRAP	Antson et al. (1999)
	Sm/Lsm proteins	Kambach et al. (1999)
Arg-rich	Tat	Puglisi et al. (1995)
	Rev	Battiste et al. (1996)
Zn2+ finger	TFIIIA	Lu et al. (2003)
	TIS11d	Hudson et al. (2004)

domain for RNA recognition (Maris et al. 2005b), which may even be used for protein–protein interactions (Kielkopf et al. 2004).

Most of the RNA binding motifs listed in Table 1 have been identified in proteins known to be involved in diverse RNA editing mechanisms (e.g. Stuart et al. 2005). However, given the large variability of RNA recognition mechanisms, the identification of any of these motifs in putative editing components is of little predictive value to define protein function.

One of the most elegant ways for proteins to recognize RNA sequences of any kind is the use of RNA as co-factors: short RNA molecules with partial complementarity to the respective target RNAs. This principle can be found in the cases of U1, U2 and U6 RNAs that are part of the snRNPs and support splice site definition as well as catalysis (Valadkhan 2005). Other examples are the modification guide snoRNAs that act on pre-rRNAs (Reichow et al. 2007), and siRNAs as well as microRNAs that target corresponding mRNAs for degradation or translational arrest (Valencia-Sanchez et al. 2006). The most complex of all editing mechanisms, the kinetoplastid mitochondrial editing, makes use of such RNA co-factors (see Sect. 2.6 and chapters by Carnes and Stuart, Göringer et al., and Ochsenreiter and Hajduk, all in this volume). Indeed, the term “guide RNA” has been coined for the short 60–75 nt RNA molecules that specify editing sites during kinetoplastid editing (Blum et al. 1990).

## 2 Editing of mRNA Sequences

### 2.1 *Paramyxovirus*

Members of the Paramyxoviridae group are non-segmented negative-strand RNA viruses (NNV) that contain five to ten linked genes on a single RNA genome of 15 to 16 kb. Virus-encoded RNA-dependent RNA polymerases (RdRP) terminate transcription of individual mRNAs by stuttering on short runs of uridylyte residues at the 3' end of each gene, which effectively leads to mRNA polyadenylation (Lamb and Kolakofsky 1996). This stuttering of the RdRP is also responsible for editing of the viral phosphoprotein (P-) mRNA by the insertion of a distinct (and virus-specific) number of G-nucleotides at a single defined site of the mRNA (Vidal et al. 1990). For Sendai virus, the co-transcriptional insertion of a single G-nucleotide occurs along the “slippery” template sequence 3'-UUUUUUC $\overline{\text{C}}^{1052}$ C-5' ( $\text{U}_6\text{C}_3$ ) once the G-nucleotide opposite  $\text{C}^{1052}$  is added to the growing mRNA. The RdRP pauses, allowing the AAAAAAGG message to slip back one nucleotide such that the six template Us pair with five As and the first G-nucleotide.  $\text{C}^{1052}$  is copied again as the RdRP resumes elongation, resulting in the net insertion of one G-nucleotide (Hausmann et al. 1999a).

Two features of the RNA sequence around and upstream (relative to the growing mRNA) of the editing site contribute to the pausing and stuttering of the viral RdRP (Hausmann et al. 1999b). First, the “slippery” polypyrimidine template sequence allows realignment of template and message without interrupting the duplex bound to the active centre of the RdRP. The thermodynamic barrier to this realignment will thus be very low (Kolakofsky and Hausmann 1998). Second, the number of Gs inserted is determined genetically for each paramyxovirus by a six-nucleotide sequence immediately upstream of the  $\text{U}_6\text{C}_3/\text{A}_6\text{G}_3$  duplex. This region supposedly lines the exit channel of the RdRP and, therefore, is likely to be engaged in base-specific RNA-protein interactions with the RdRP (Hausmann et al. 1999a). Although detailed structural information of this interaction is not available, the conservation of the six-nucleotide sequence within the group of the Paramyxoviridae argues for a crucial function of this sequence in determining the target site and the number of G-insertions.

### 2.2 *Physarum*

RNA editing in *Physarum polycephalum* is a mitochondrial process that affects mRNA, rRNA and tRNA transcripts alike. At least two distinct editing mechanisms operate in the mitochondria. The first reaction is a C-to-U conversion at four positions within the transcript of the cytochrome C oxidase subunit 1, *co1* (Gott et al. 1993). This type of editing resembles plant mitochondrial editing, since it usually affects the first or second codon position and restores conserved protein

sequence (Gott 2001). The second mechanism is the insertion of more than 400 C-residues, less than 100 U-residues and several dinucleotide combinations (UU, AA, UA, CU, GU, GC) into nearly all mitochondrial RNA transcripts (Gott et al. 2005). These insertional editing events cause numerous frame shifts that are required to restore correct reading frames in most of the mitochondrial mRNAs. Recently, nucleotide deletion editing was identified in the transcript of *nad2* where three consecutive adenines were missing in the final mRNA (Gott et al. 2005). The insertion (and, potentially, also the deletion?) of nucleotides is clearly a co-transcriptional event that is either a feature of the mitochondrial RNA polymerase itself or is promoted by additional trans-acting factors that associate with the polymerase (Cheng et al. 2001). Mono- and dinucleotide insertions may be directed by a distinct set of specificity factors (Wang et al. 1999; Horton and Landweber 2000; Byrne and Gott 2004).

Editing site selection is the result of local features of the RNA sequence immediately preceding the editing site (Byrne and Gott 2002) and their interaction with the editing machinery, consisting of – at least – the mitochondrial RNA polymerase. Additional specificity factors such as trans-acting guide RNAs or proteins have not been identified as of yet but may be involved in modulating polymerase activity. Therefore, the detailed mechanisms of editing site recognition and subsequent nucleotide insertion still remain to be clarified. The RNA polymerases itself may play the dominant role, since it is more closely related to the single-chain bacteriophage polymerases (Masters et al. 1987; Ceramiakan et al. 1996). These enzymes exhibit some editing-related activities, such as the addition of non-templated nucleotides within homopolymer tracts (Macdonald et al. 1993) or at the 3' termini of transcripts, e.g. by the T7 RNA polymerase (Milligan et al. 1987).

Editing sites in mRNA transcripts show a number of statistical rules that point to a function of local RNA primary sequence in editing site selection (Gott et al. 2005). On average, insertion events occur every 25 nucleotides on a given mRNA, with two thirds located at the third codon position. Cis-acting RNA signals are restricted to the region less than 15 bases upstream of the editing site (Byrne and Gott 2002). However, statistical analysis of the base composition within this region revealed no significant sequence pattern (Gott et al. 2005). The only conserved bases are found at pos. -1 being uridine in 82% of the editing sites, and pos. -2 being a purine (62% A, 21% G), combining for a purine-U motif at -2/-1 in 69% of all cases. Co-variation analysis of pairs of bases within the region from -15 to +15 surrounding the editing sites revealed no correlations between pairs of bases. Thus, the formation of conserved Watson-Crick base pairs as part of RNA secondary structures in the vicinity of editing sites is unlikely. Since nucleotide insertions are co-transcriptional events, DNA template sequences surrounding the editing site may contribute to editing specificity. Potential cis-acting DNA template determinants were shown to be restricted to a region within 15 bp upstream and 15–20 bp of downstream DNA (Byrne and Gott 2002).

Since the purine/U bias at the -2/-1 position upstream of the editing site is the only cis-acting element common to all editing sites, additional specificity factors must be present at the site of transcription. Either a guiding mechanism involving

guide RNAs (that should involve at least one guide per site with the potential to form at least 7–8 bp) can be postulated, or additional proteins act as specificity factors by interacting with the nascent RNA and/or by modulating polymerase activity and processivity. Identification of such RNA as well as protein co-factors will require the ability to isolate and purify editosome complexes from editing active mitochondrial fractions that could be subjected to mass spectroscopy. Recently, a putative mitochondrial RNA polymerase from *Physarum polycephalum* was identified (Miller et al. 2006) that may serve as a starting protein for the affinity purification of editosome complexes using the Tap-tagging technology (Rigaut et al. 1999).

### 2.3 RNA Editing in Plant Organelles

RNA editing in the mitochondria and chloroplasts of plants occurs by C-to-U deamination that affects a large number of mRNAs encoded by the organellar genomes. In a given plant species, about 20–40 Cs are deaminated in the mRNAs of chloroplasts (Maier et al. 1996; Tsudzuki et al. 2001), contrasting with the 400–1,000 editing sites in mRNAs within the mitochondria. In addition, very few (~0.5%) reverse U-to-C editing events have been identified. About 90% of all editing events affect the coding regions of mRNAs and, in most cases, the editing reaction is essential for the restoration of conserved codons/amino acids.

Given the large number of plant organellar editing events, it seems surprising that the specificity factors are still elusive. Several recent analyses in the mitochondria of pea (Takenaka et al. 2004), wheat (Choury et al. 2004; Choury and Araya 2006) and cauliflower (van der Merwe et al. 2006), and in the chloroplasts of tobacco (Chateigner-Boutin and Hanson 2002; Hayes et al. 2006) and *Arabidopsis* (Hegeman et al. 2005) have narrowed the search for the cis-acting RNA signals to the immediate vicinity of the editing site. The essential elements usually reside within 15–20 nts upstream and 1–5 nucleotides downstream of the editing site. However, these flanking regions do not share consensus elements at the primary or secondary structure level, suggesting that site-specific editing requires several hundreds of specific factors. In a recent study, the editing efficiencies at two sites, C77 and C259, within the wheat mitochondrial *cox2* transcript were compared (Choury et al. 2004). In both cases, essential sequences were identified within a 22-nt region from –16 to +6 relative to the editing site (+1). This 22-nt sequence could be placed at different sites within the large RNA without loss of editing efficiency, suggesting context-independent editing. Replacement of the downstream region (+2 to +6) of C77 with the homologous region from the C259 site led to strongly reduced editing efficiency, whereas replacement of the upstream region (–16 to –1) completely abolished editing at C77. These results confirm that each editing site sequence is recognized by an individual set of specificity factors.



Some of these factors may recognize groups of related sequences (Chateigner-Boutin and Hanson 2002, 2003), and editing sites with identical core determinants (−15 to +2) exist that are likely to be recognized by the same trans-acting factor (Tillich et al. 2005). Recently, two protein candidates for site-specific trans-acting factors were identified by cross-linking technology (Miyamoto et al. 2002, 2004). A third protein factor was identified in *Arabidopsis* chloroplasts by a genetic screening method (Kotera et al. 2005). In this study, mutants were isolated that were defective in a unique RNA editing event that creates the initiation codon of the chloroplast NAD(P)H dehydrogenase subunit D (ndnD). The mutations were assigned to the CRR4 protein, a member of the plant combinatorial and modular protein family (PCMP; Aubourg et al. 2000) that is a subgroup of the pentatricopeptide repeat (PPR) family (Small and Peeters 2000). The PPR family consists of about 450 genes in the *Arabidopsis* genome. Members of the PPR family are involved in RNA processing, stabilization and translation in chloroplasts and mitochondria of yeasts and higher plants (Kotera et al. 2005 and references therein). The PPR repeats have been implicated in RNA binding in the case of the chloroplast PPR protein HCF152 (Nakamura et al. 2003), and the number of PPR repeats was shown to determine the affinity and specificity for RNA. CRR4 itself contains 11 PPR repeats consisting of a degenerate 35-amino acid unit, and Kotera et al. (2005) suggest it functions as a specificity factor for the recognition of the first editing site in the ndhD transcript. The authors propose a model, similar to the apoB editing system (see Sect. 2.4), in which PPR proteins interact with individual editing sites to recruit the editing machinery containing the deaminase activity. Consistent with this model, 189 of the 452 PPR proteins in *Arabidopsis* were predicted to localize to the mitochondria, and 96 to the chloroplasts (Nakamura et al. 2003).

PPR proteins belong to the family of helical repeat proteins, and the modes of RNA recognition may be analogous to those of another helical repeat protein family, the PUF proteins. The five PUF proteins in yeast were recently shown to regulate groups of 100–200 target mRNAs carrying similar 3′ UTR recognition motifs of 9–11 nucleotides (Gerber et al. 2004). The crystal structure of the human PUF protein Pumilio1 in complex with its RNA target revealed that such RNA sequence motifs are bound as single-stranded unstructured RNA. Each base within the sequence motif is contacted by a single Pumilio homology domain (PHD, ~34 amino acids) via 4–5 amino acid side chains. Two of the side chains are required for base-specific hydrogen bonding to the Watson-Crick interface, while two hydrophobic side chains form a “sandwich” with the base stacking in between. Thus, the helical repeats of PPR may functionally resemble those of the PUF proteins and may serve as a platform for interaction with a variety of RNA molecules. The sequence specificity of RNA recognition should rely on distinct amino acid side chains exposed on this platform to contact the Watson-Crick interface of the RNA bases.

Following this model, the PPR proteins would direct an editing deaminase activity to the RNA target site that is bound in a single-stranded and unstructured



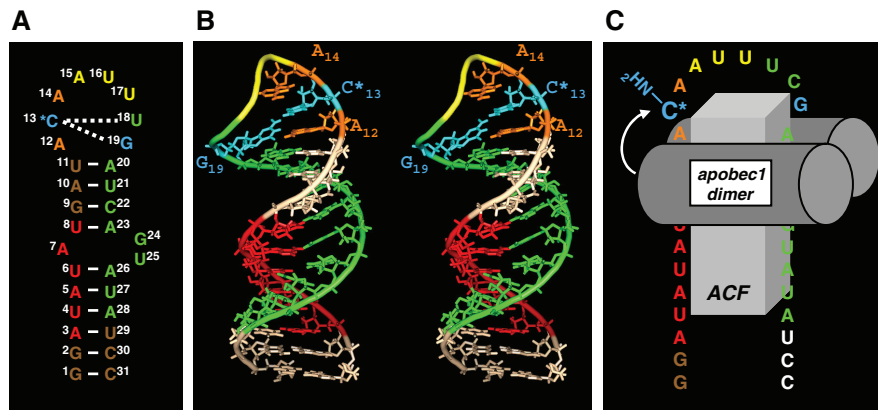
form. The protein might even serve to pre-orient the editing site C towards deamination, in analogy with the function of the ACF protein during apoB-editing described in the following section.

## 2.4 Mammalian Editing: apoB

Editing of apolipoprotein B (apoB) mRNA was the first RNA editing event described in vertebrates (Chen et al. 1987; Powell et al. 1987). The editing reaction is catalyzed by the APOBEC1 enzyme (apoB editing catalytic subunit 1), the first identified mammalian cytidine deaminase (Teng et al. 1993). APOBEC1 is expressed in the small intestine, where it catalyzes the formation of a premature stop codon in the apoB mRNA by deamination of cytidine<sup>666</sup> to uridine. As a result, a truncated form of apoB with 48 kDa (apoB48) is expressed, which has a tissue-specific function in lipid metabolism in the small intestine that differs from that of its full-length counterpart apoB100 expressed in the liver. The editing event thus contributes to genetic plasticity by generating two protein isoforms from a single gene locus. APOBEC1 belongs to a family of 12 cytidine deaminases identified so far in the human genome (reviewed in Turelli and Trono 2005). Most of these are involved in the deamination of dC to dU at the level of DNA, with diverse cellular functions in antibody diversification, retroviral defence and retrotransposon silencing (reviewed in Holmes et al. 2007).

### 2.4.1 APOBEC-1 Target Site Structure and Dynamics

APOBEC-1 is the only member of the deaminase family for which an mRNA target could be identified. The target site of APOBEC1 lies within the 14,000-nt apoB mRNA of which a single cytidine residue at position 6666 is recognized and deaminated by the editosome machinery. The specificity of the reaction relies on sequence motifs surrounding the editing site that are highly conserved within vertebrate species. The minimal editing competent sequence is 26 nucleotides long and folds into a stem-loop secondary structure that has important functions in the mechanism of RNA recognition by the editing factors (Richardson et al. 1998). It contains three cis-acting elements in addition to the C<sup>6666</sup> editing site (Fig. 1). The first element is an 11-nucleotide mooring sequence located downstream of the editing site. The mooring sequence is separated from the C<sup>6666</sup> editing site by a “spacer element” with a variable length of 2–8 nucleotides. The third sequence is an A/U-rich efficiency element upstream of the editing site that modulates editing efficiency. The minimal editing competent complex that recognizes the RNA stem-loop consists of two proteins: APOBEC1 as the catalytic subunit, and APOBEC1 complementation factor, ACF (Mehta et al. 2000). The 64-kDa ACF contains three



**Fig. 1** RNA editing by site-specific deamination of C<sup>666</sup> in the apoB mRNA. **A** Secondary structure of a minimal 31-nt substrate of APOBEC1-dimer/ACF complex (Maris et al. 2005a). The editing site cytidine C\*13 is indicated (*cyan*; corresponds to C<sup>666</sup> of the apoB message). Sequence elements important for binding and site specificity are colour-coded (*green* mooring sequence, *red* efficiency element, *yellow* spacer element, *brown* remaining nts). **B** Stereo-view of the apoB stem-loop structure derived from NMR analysis (Maris et al. 2005a). One of the calculated structures is shown in which the target site C\* is sandwiched between A14 and A12 (both *orange*), and involved in hydrogen bonding to G19 (*cyan*). **C** ACF was shown to melt the apoB stem-loop in order to disrupt the interactions of the target site C\* and expose the NH<sub>2</sub> group for attack by an APOBEC1 dimer that is recruited to the target site by ACF (Maris et al. 2005a)

RNA recognition motifs (RRMs) and a putative double-stranded RNA binding domain (dsRBD). ACF binds to both APOBEC1 and the mooring sequence of the stem-loop RNA with high affinity (K<sub>d</sub>=8 nM, Mehta and Driscoll 2002), and is required to anchor APOBEC1 at the editing site. APOBEC1 itself was shown to interact with the sequence motif UUUN A/U U (Anant and Davidson 2000) 2 nts downstream of C<sup>666</sup> with much lower affinity. The enzyme has catalytic activity on a minimal apoB RNA substrate in the absence of auxiliary factors (Chester et al. 2004). The reaction has a temperature optimum of 45°C that was lowered to physiological temperatures in the presence of ACF. A model was suggested in which ACF promotes a conformational transition in the RNA substrate and stabilizes an editing competent conformation that forms spontaneously at higher temperatures (Chester et al. 2004).

The molecular basis of ACF function was addressed by NMR spectroscopy of the 31-nucleotide apoB-mRNA stem-loop containing the essential cis-acting elements (Maris et al. 2005a). The calculated structures confirmed a previously proposed model (Chester et al. 2004) and suggested the requirement for several conformational rearrangements of the RNA structure during the editing process. The authors proposed a model in which the rearrangements are induced by the sequential binding of ACF and APOBEC1.

The calculated structures for the unedited (C<sup>666</sup>) as well as the edited version (U<sup>666</sup>) of the apoB mRNA target reveal an extended stem-loop conformation in

which the editing site nucleotide (C/U13) is located within a flexible octaloop (Richardson et al. 1998). The base is sandwiched between two adenines (A12 and A14) and involved in H-bond interactions with one of the two 3'-terminal loop-bases (G19 or U18; dotted lines in Fig. 1A). The remaining nucleotides of the octaloop (A15–G19) were disordered in the NMR structures and interpreted to be flexible (Maris et al. 2005a). Thus, the N4 position of C6666 is buried within the octaloop and a conformational change is required in order to expose the amino group for deamination.

The structural dynamics of the mooring sequence and its recognition by the ACF protein were proposed to contribute to this loop rearrangement (Maris et al. 2005a). In all known species, the mooring sequence is embedded in a structurally dynamic environment consisting of an irregular stem-loop structure with bulges and internal loops. In the NMR structure, the first two nucleotides of the mooring sequence (U18 and G19) are part of the octaloop and appear to be flexible. The remaining residues (A20 to A28), together with U4 to U11, form an irregular A-form helix interrupted by an internal loop composed of G24, U25 and A7. This internal loop has a flexible conformation that also destabilizes the surrounding AU base pairs and even the hydrogen bond interactions of the upper stem base pairs. Thus, most of the mooring sequence is very dynamic (U18, G19) to moderately dynamic (A20 to A26). This flexibility plays a critical role in RNA recognition, since a mooring sequence completely annealed to its antisense strand (Mehta and Driscoll 2002) is not recognized by ACF. The unpaired nucleotides U18, G19, G24 and U25 could serve as a primary recognition site for ACF. Thus, it seems that editing of apoB mRNA requires a flexible mooring sequence, a flexible spacer and no strong base pairing of the editing site C with surrounding base pairs.

In order to analyse the role of ACF in promoting the necessary conformational changes, RNA binding studies were performed with ACF34, the N-terminal part of ACF containing the three essential RRM s (Maris et al. 2005a). NMR studies indicated that ACF34 bound to U18 and U25 at 25 °C without denaturing the stem-loop. At 37–42 °C, however, the stem-loop was melted, indicating that ACF34 bound the RNA as a single strand. By melting the RNA stem-loop, ACF34 also disrupts the structure around C13 and renders its amino group accessible for deamination by APOBEC-1. In addition, interaction site mapping by UV cross-linking studies revealed that binding of ACF to the RNA restricted the access of APOBEC-1 to a single site (C13) for editing (Maris et al. 2005a). These experimental findings support the conclusions drawn from homology modelling of APOBEC-1 using the coordinates of the yeast cytidine deaminase CCD1 crystal structure as a template (Xie et al. 2004). The resulting APOBEC-1 dimer model revealed that the RNA substrate would only be accommodated in its single-stranded form, supporting an essential function of ACF in disrupting the RNA hairpin secondary structure.

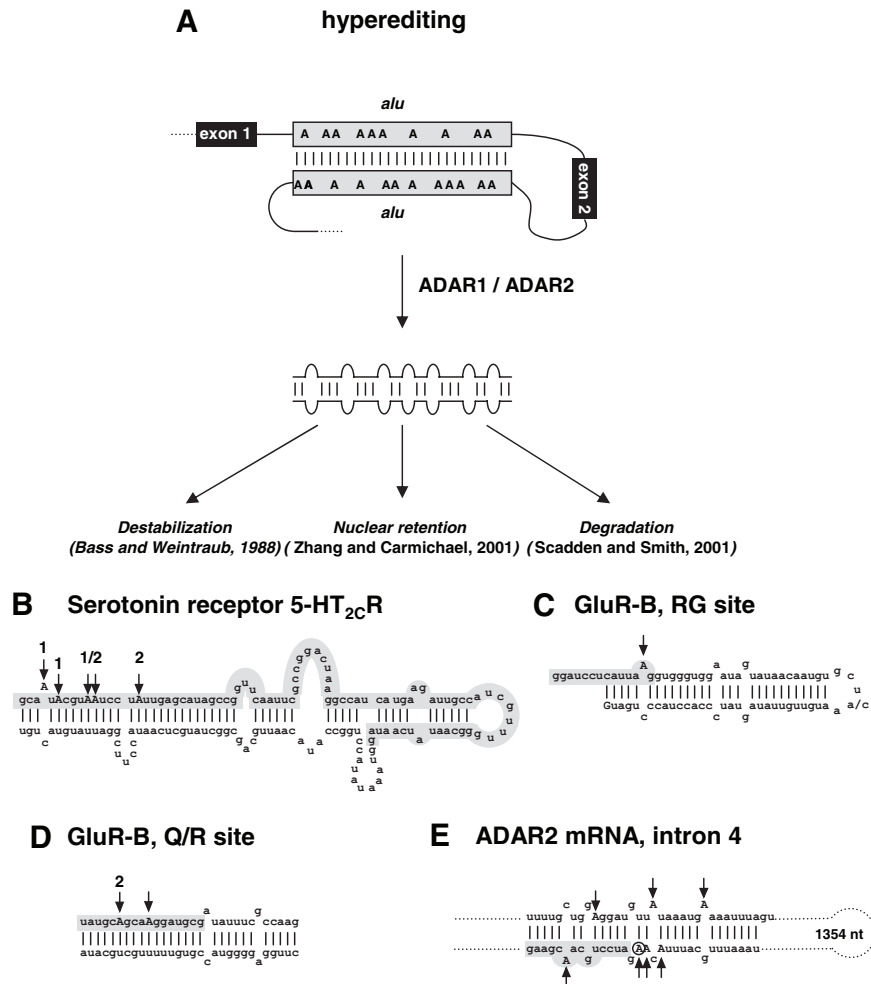
The sequence of events outlined above can be viewed as a hallmark example for the structural and conformational changes that occur during a given editing event: formation of a well-defined and, at the same time, flexible RNA structure; site-specific binding of a protein factor; creation of a nucleation site for melting of the stem-loop structure; recruitment of the enzymatic editing activity and restricting its

access to a single defined site; catalysis, dissociation and re-formation of the thermodynamically more stable, folded RNA structure.

## 2.5 RNA Editing by Adenosine Deamination

RNA editing by the family of adenosine deaminases that act on RNA (ADAR enzymes) involves the conversion of adenosine to inosine by hydrolytic N6 deamination. This modification can change RNA structure, creates new splice sites and alters codon identities, since inosine pairs most stably with cytidine. Editing of adenines within coding regions may therefore cause amino acid substitutions and create multiple protein isoforms from single genes. Adenosine deamination appears to be the most widespread type of editing in higher eukaryotes, and ADAR genes are present in most metazoan genomes (for a review, see Bass 2002). ADAR activities were first identified in *Xenopus laevis* as an activity that destabilizes double-stranded RNA (dsRNA) by promiscuous deamination of adenosines within RNA duplexes (Bass and Weintraub 1988). Since then, several cases of site-specific deamination within coding regions of mRNAs have been reported in mammals, insects, molluscs, worms and viruses (reviewed in Valente and Nishikura 2005).

ADAR enzymes act on RNA substrates that are either completely or largely double-stranded. Recently, bioinformatics and experimental approaches have revealed extensive deamination activities in humans mainly within non-coding regions of mRNAs (reviewed in Levanon et al. 2005). Most of these A-to-I substitutions are clustered within Alu-repetitive elements that often form extended RNA duplexes when expressed in close proximity and in opposite orientations (Fig. 2A). Up to 50% of the As of both strands of an RNA duplex may be converted into inosines, referred to as “hyper-editing” (DeCervo and Carmichael 2005). A-to-I editing is especially prevalent in the central nervous system, where it may contribute to neuronal plasticity and function. Direct sequencing of human brain cDNA revealed an editing frequency of 1:1,000 nucleotides in non-coding regions (Blow et al. 2004). Since the human genome contains more than a million Alu repeats that account for ~10% of the genome, it has been suggested that most human primary transcripts are subject to A-to-I editing (Athanasiadis et al. 2004; Levanon et al. 2004). The immediate consequences of hyper-editing are thermodynamic destabilization of RNA duplexes (Bass and Weintraub 1988), retention of the transcripts in the nucleus (Kumar and Carmichael 1997; Zhang and Carmichael 2001), and I-specific cleavage and degradation of RNA (Scadden and Smith 2001; see Fig. 2A). Thus, the predominant functions of ADAR activities lie in the control of dsRNA levels that arise from repetitive sequences and viral replication. Furthermore, ADARs counteract dsRNA-dependent RNA interference pathways and are likely to mediate regulatory functions of antisense transcripts in the human genome that were recently shown to be much more prevalent than previously thought (Katayama et al. 2005).



**Fig. 2** RNA editing by deamination of adenines. **A** Hyper-editing of perfectly double-stranded RNA substrates frequently occurs within non-coding regions of RNA transcripts when repetitive sequences such as Alu elements are transcribed in opposite directions. ADAR1 and ADAR2 both promiscuously deaminate adenines to inosines with different cellular consequences. **B–E** Site-specific deamination within the coding regions (grey) of distinct mRNA targets in mammals. **B** ADAR1 and ADAR2 target sites within the serotonin 5-HT<sub>2C</sub>R message. Several neighbouring As of an imperfect RNA duplex are edited, indicating that ADARs recognize adenines in different sequence context. **C** A single adenine is edited at the GluR-B message. The two downstream mismatches (a/c and g/g) and the target site A/C mismatch contribute to specificity and efficiency of editing at the RG site. **D** Editing at the Q/R site is catalyzed exclusively by ADAR2. **E** ADAR2 mRNA contains a highly conserved sequence within intron 4 that folds into a conserved secondary structure. Several As within this imperfect duplex are targeted by ADAR2 enzyme, and editing at one of these positions (*circled*) creates a non-canonical splice acceptor site, leading to altered splicing and to the synthesis of a truncated and inactive version of ADAR. Thus, a negative regulatory feedback loop is created

### 2.5.1 A-to-I Editing Within Coding Regions

Site selective A-to-I editing in mRNA coding regions has been found in mammals, *Xenopus*, *Drosophila*, squid and several RNA viruses (reviewed in Valente and Nishikura 2005). Editing primarily affects ion channel and receptor mRNAs of the nervous systems, and often leads to functionally important changes in codon identities and amino acid sequences. Compared to the massive hyper-editing in non-coding regions described above, the few site-selective deamination events may be interpreted as “by-products” of ADAR activities that became established during evolution due to advantageous effects within a specific cell type or tissue of the organism (Maas et al. 2003).

ADAR enzymes are dsRNA binding proteins, the vertebrate family having three members, i.e. ADAR1, ADAR2 and ADAR3. ADAR1 and ADAR2 are ubiquitously expressed in many tissues, whereas ADAR3 expression is restricted to the brain (Chen et al. 2000). All ADAR enzymes have an N-terminal dsRNA binding domain with either two (ADAR2) or three (ADAR1) dsRNA binding motifs (dsRBM), whereas the C-terminal sequences encode the catalytic deaminase domain. The dsRBM is a 65-aa motif with a characteristic  $\alpha\beta\beta\beta\alpha$  fold (Kharrat et al. 1995). Apart from the ADAR enzymes, the dsRBM has been found in a wide range of dsRNA binding proteins including reverse transcriptase, PKR and dicer nucleases (reviewed in Beal 2005; see Table 1). The crystal structures of the dsRBM in complex with a short RNA double strand reveal numerous contacts spanning 16 base pairs of an A-type RNA helix, including two minor grooves and the intervening major groove (Ryter and Schultz 1998; Ramos et al. 2000). Most of the contacts are sequence unspecific, involving hydrogen bonding and ionic contacts to the 2'OH groups and phospho-diester. This explains the lack of a strict sequence requirement for dsRNA binding by the ADAR enzymes: human ADAR1 and ADAR2 both promiscuously deaminate dsRNA. Perfectly double-stranded RNAs with a minimum of 16 base pairs can function as substrates, although optimal deamination requires at least 25–30bp and preferably >100bp (Bass and Weintraub 1988). At the same time, ADAR1 and ADAR2 have distinct but overlapping specificities for single adenosines in only a very few transcripts. Since no auxiliary factors are needed – at least in vitro – to achieve specificity of ADARs for distinct sites, the question arises as to which factors contribute to directing the ADAR enzymes to these unique sites within the few pre-mRNA targets known to date.

### 2.5.2 Structural Determinants of Site Selectivity of ADAR Activity

The best characterised examples of site-selective deamination in mammals occur in transcripts of the glutamate receptor subunits gluR-B, -C and -D and gluR-5 and -6 (Fig. 2C und D), and in the transcript of the serotonin receptor subunit 2C (5-HT<sub>2C</sub>R; Fig. 2B). Editing of several codons within these transcripts leads to altered

amino acids and generates receptors with altered functions (reviewed in Seeburg and Hartner 2003). In all cases, the edited adenosines are embedded within imperfect RNA duplex structures that are formed between exon sequences and sequences with partial complementarity from downstream introns (see Fig. 2).

Structural characteristics of the exon/intron duplex as well as certain sequence preferences contribute to the selection of distinct adenosines by the ADAR enzymes. The RNA duplexes are interrupted by bulges and internal loops that were shown to separate and uncouple long RNA duplexes into shorter targets that are independently bound and modified by the ADAR enzymes (Lehmann and Bass 1999; Dawson et al. 2004). In a reciprocal approach, removal of two mismatches downstream of the R/G site in the GluR-B transcript (see Fig. 2C) did not alter the binding affinity of ADAR2 for its RNA substrate (Öhman et al. 2000). However, the editing efficiency at the R/G site was decreased due to the fact that mismatch removal allowed promiscuous ADAR binding anywhere along the resulting RNA duplex. In an extension of this study, Källman et al. (2003) could show that – although binding occurred anywhere on the RNA duplex – editing was still specific for the R/G site. This indicated that additional structural and/or sequence characteristics at or near the edited adenine contribute to editing site selection.

There is no discrete sequence motif common to all ADAR target sites and, in the case of the 5-HT<sub>2C</sub> mRNA editing by ADAR1 and ADAR2, five adenines in different sequence context (covering 13 nt) are edited (see Fig. 2B). However, distinct 5'-nearest neighbour preferences have been identified for ADAR1 (U=A>C>G) and ADAR2 (U=A>C=G), and the 3' preferences for ADAR2 were U=G>C=A (Lehmann and Bass 2000). In addition, the identity of the base opposite to the editing site is crucial, with a preference for A–C mismatches or A–U base pairs, whereas editing is strongly disfavoured with an opposite purine (Källman et al. 2003). Since the editing site and the surrounding bases are contacted by the catalytic domain of the ADARs during deamination, a structural complementarity of editing site structure and the catalytic domain is required for A-site selection (Yi-Brunozzi et al. 2001). Indeed, ADAR2 binding was reported to induce conformational changes at the editing site, consistent with a model in which the adenine that is initially buried in the RNA duplex as an AU base pair or an AC mismatch is flipped out and accommodated in the catalytic centre (Stephens et al. 2000). Flipping of the adenine was dependent on the presence of the N-terminal catalytic domain, and did not occur upon dsRBM binding to the duplex (Yi-Brunozzi et al. 2001). On the contrary, dsRBM binding increased the conformational flexibility of duplex nucleotides opposite the editing site, showing that both functional domains act in concert to lower the activation energy for base flipping. The catalytic domain of ADAR2 contributes significantly to A selectivity, as demonstrated by a protein construct containing the ADAR2 catalytic domain and the ADAR1 dsRBMs: this chimeric protein retained the substrate selectivity observed for ADAR2 (Wong et al. 2001). In a recent study, sequence conservation was reported in the case of ADAR2 editing of its own mRNA that creates a non-canonical AI 3'-splice acceptor



site (see Fig. 2E; Dawson et al. 2004). In addition to the 5'- and 3'-nearest neighbour preferences, seven positions within 18 nt upstream and 15 nt downstream of the editing site were conserved, and the conservation of three of these sites correlated with *in vitro* editing efficiency.

Taken together, site-specific adenosine deamination by the ADAR enzymes occurs by a series of dynamic interactions and structural changes. Initial binding of the dsRBMs to regions of duplexed RNA may be sequence-unspecific but these will be needed to direct the catalytic domain to the specific editing site. Sequence-specific interactions of the catalytic domain as well as unspecific binding of the dsRBMs contribute to destabilization of the local duplex and lower the activation energy for flipping the adenine out of its helical context. The adenine is accommodated in a binding pocket to allow nucleophilic attack by enzyme-bound water. The recently published crystal structure of the N-terminal half of ADAR2 supports this model, in which the adenine that is initially buried in the RNA duplex is flipped out and accommodated in the catalytic centre to allow nucleophilic attack by an enzyme-bound water molecule (Macbeth et al. 2005). However, the structure does not reveal how the catalytic domain may be engaged in sequence-specific interactions with a duplexed RNA target, and more detailed structural data will become available only from high-resolution structures of ADAR2 crystals in complex with duplex RNA.

To that end, the solution structure of an N-terminal fragment of ADAR2 encompassing the two dsRBDs was determined in complex with one of its natural substrates, the 71-nt RG site stem-loop (compare with Fig. 2C; Stefl et al. 2006). Both dsRBDs adopt the conserved  $\alpha\beta\beta\beta\alpha$  fold common to all dsRBD family members. While dsRBD1 interacts with the conserved pentaloop (Stefl and Allain 2005), dsRBD2 recognizes the two AC mismatches at the editing site and adjacent to it. The two bulged C-residues experienced the largest chemical shift changes upon dsRBD2 binding, indicating direct interactions with the protein. The model calculated from the NMR analyses revealed that recognition of the RNA by both RBDs is structure-specific, rather than base-specific. dsRBD1 interacts with the stem-loop structure by contacting the minor groove of the GCUCA pentaloop, and dsRBD1 recognizes the RNA helix containing two A-C mismatches separated by ten base pairs (Stefl et al. 2006).

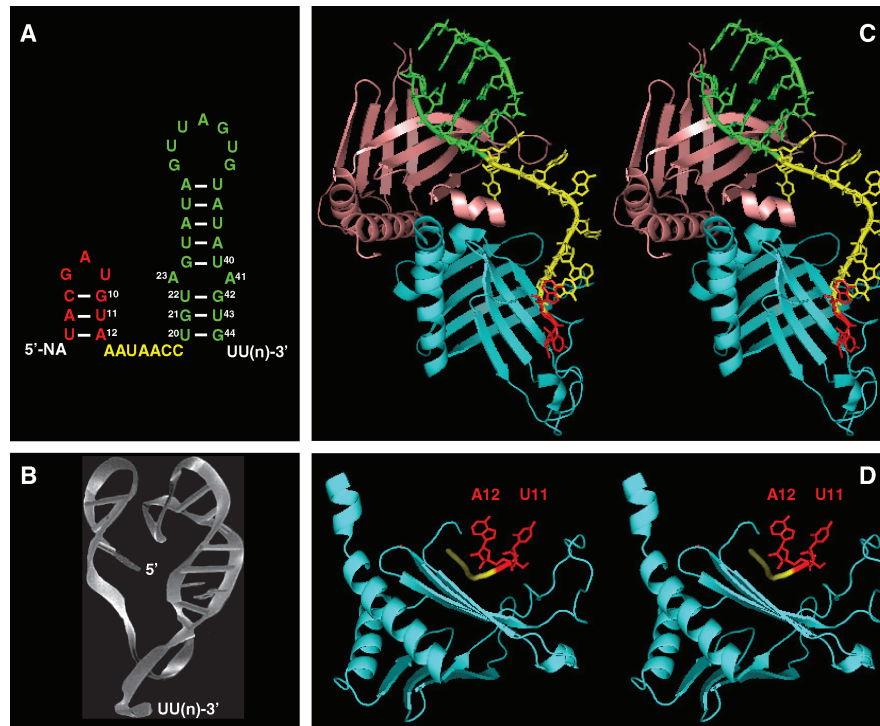
One has to keep in mind that *in vivo*, ADAR enzymes compete with several nuclear RNA binding and modifying proteins that are working on any primary RNA transcript. These include hnRNP proteins that interact with any RNA transcript once it emerges from the polymerase complex. In addition, components of exon definition complexes, snRNPs, SR proteins, and dsRNA binding proteins such as drosha and dicer interact with specific sites on the RNAs. All these proteins may restrict access of the ADAR enzymes to only a few sites, and the observed target site specificity of ADAR enzymes may result rather from effective competition by these proteins for binding of ADARs to other target sites that – *per se* – would have a similar affinity for the ADAR enzymes. It remains to be seen whether specificity determinants *in vivo* rest in target site RNA structure alone and to what extent competitive binding of other proteins to potential target sites contributes to

restricting access of ADARs to very few, select sites of the coding parts of the transcriptome.

## 2.6 Insertion/Deletion Editing in Kinetoplastids

Kinetoplastids are a group of protozoan pathogens that include *Trypanosoma* and *Leishmania* parasites responsible for severe human diseases such as African sleeping sickness and Chagas disease (Barrett et al. 2003). RNA editing in kinetoplastids is a mitochondrial process that creates functional mRNAs by the guide RNA-directed insertion and deletion of uridylate residues into most mitochondrial transcripts. It is characterised by an enzymatic reaction cycle involving the coordinated activities and interactions of gRNAs, pre-edited mRNAs and numerous protein components organized in a high molecular mass ribonucleoprotein complex, often termed the “editosome”. Several purification protocols have confirmed the existence of at least 20 editing-associated proteins (see the chapter by Carnes and Stuart, this volume). Editing is initiated by the formation of a short duplex structure of 8–10 base pairs between the pre-edited mRNA and a cognate gRNA. This short duplex precisely defines the editing site: it functions as a substrate for an endoribonucleolytic activity that cleaves the pre-mRNA immediately 5′ of the helical segment. During deletion-type editing, uridylate residues are exonucleolytically removed from the 3′ end of the cleavage product, whereas insertion-type editing involves the addition of U-nucleotides. The exact number of uridylate residues to be added or deleted at each site is controlled by the gRNA sequence adjacent to the anchor region. This region, called the “information sequence”, functions as a template for the removal or the addition of Us until complementarity (allowing also G:U wobble base pairs) between gRNA and pre-mRNA is achieved at each editing site. The editing cycle is completed by the ligation of the processed 5′ fragment to the 3′ fragment of the pre-mRNA.

RNA–RNA interactions play crucial roles during all steps of the editing cycle. The first step of the editing cycle is the annealing of gRNAs with pre-mRNAs via short stretches of complementarity. The efficiency of the base-pairing interactions will be influenced by the RNA structures and by proteins associated with the RNAs. The secondary structure of gRNAs has been analysed by structure probing and UV melting spectroscopy of different gRNAs (Schmid et al. 1995), and a 3D model has been developed on the basis of these results (see Fig. 3B; Hermann et al. 1997). The gRNAs fold into two imperfect hairpin loops of low thermodynamic stability, with the anchor sequence identified as part of the 5′ hairpin and both terminal ends, including the poly-U-tail, being single-stranded (Fig. 3B). This higher-order tertiary structure may permit recognition by components of the editing machinery, while its low thermodynamic stability allows efficient annealing and unfolding during the editing cycle. The base-pairing interaction between gRNA and pre-mRNA was shown to be catalyzed by the RNA binding proteins gBP21 (Köller et al. 1997; Müller et al. 2001) and possibly gBP25 (Blom et al. 2001)



**Fig. 3** Structures of gRNA alone and in complex with MRP1/2. **A** Secondary structure of gND7-506. Functional elements are colour-coded (*red* anchor sequence/stem-loop I, *green* stem-loop II, *yellow* spacer sequence). *Numbers* correspond to the nucleotides visible in the X-ray structure of the gRNA/MRP1/2 complex displayed in **C** (Schumacher et al. 2006). **B** gRNA 3D model developed by Hermann et al. (1997). **C** X-ray structure of gND7-506 in complex with MRP1 (*cyan*)/MRP2 (*pink*) heterodimer. **D** U11 and A12 (*red*) of the anchor sequence are exposed towards the solvent and are pre-oriented for the interaction with the target pre-mRNA

in *T. brucei* (also named MRP1 and MRP2; Simpson et al. 2004), and by their orthologs gbp27/29 in *C. fasciculata* (Blom et al. 2001) and Ltp26/28 in *L. tarentolae* (Aphasizhev et al. 2003a). GBP21 was shown to bind to the 3'-hairpin element of gRNAs with nanomolar affinity (Hermann et al. 1997; Köller et al. 1997). In addition, gBP21 exhibits a preference for single-stranded RNA to which it binds in a sequence-unspecific way by making use of up to six ionic contacts (Müller and Göringer 2002). On the basis of these in vitro analyses, a “matchmaker”-model for gBP21-catalyzed gRNA-pre-mRNA annealing was proposed. (1) The protein binds to the 3' hairpin of gRNAs. (2) The 5' end of the gRNA is unfolded in order to present the anchor region for annealing to the pre-mRNA. Up to six ionic bonds may participate in this step, thereby reducing electrostatic repulsion. (3) The anchor sequence hybridizes to the complementary target sequence on the pre-mRNA, thereby forming a partially double-stranded RNA. A 30-fold rate enhancement of gBP21-catalyzed annealing was determined

with an apparent association rate constant of  $3.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$  (Müller et al. 2001). GBP21 was found stably associated with editing complexes (Lambert et al. 1999) where it may be engaged in protein–protein interactions – rather than binding to the editing product, since it exhibits a much lower affinity for double-stranded RNAs (Müller and Göringer 2002).

GBP21 is not essential for RNA editing, as shown by genetic knockout mutants (Lambert et al. 1999). However, it contributes to the efficiency of the editing reaction by a 30-fold acceleration of the very first step of editing that might otherwise be rate limiting at an uncatalyzed rate of  $1.2 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ . Comparable in vitro rate enhancements were determined for a 100-kDa Ltp26/28 tetrameric complex in *L. tarentolae* (Aphasizhev et al. 2003b).

The matchmaker model of gBP21 action that was proposed on the basis of RNA modelling (Herrmann et al. 1997) and biochemical data (Müller and Göringer 2002) was essentially confirmed by the recently determined crystal structures of gBP21/25 (MRP1/MRP2) alone and in complex with gRNA (Schumacher et al. 2006). These structures revealed a novel “whirly” fold for the protein heterodimer complex (see Fig. 3C). gRNA binding occurs to a highly basic  $\beta$ -sheet surface present on both molecules of the protein dimer, via sequence-non-specific electrostatic contacts and hydrogen bonds. Exactly as predicted by Müller and Göringer (2002), the MRP dimer binds to the duplexed base of stem-loop II (the first four base pairs of the 3′ hairpin) while it unfolds stem-loop I and orients the first two bases of the anchor sequence towards the solvent (Fig. 3C and D). Altogether, nine positively charged amino acids are involved in gRNA binding, contacting seven backbone phosphate moieties. Specifically, the phosphate moieties of the first two anchor bases plus the adjacent nucleotide are neutralized by three arginines and a hydrogen-bonding network involving three additional amino acids. Thus, the protein dimer decreases the electrostatic repulsion between the two RNA substrates, favouring gRNA–pre-mRNA duplex formation. The two binding platforms originally proposed in the matchmaker model by Müller and Göringer (2002) can now be attributed to the two individual MRP monomers of the dimer, one (MRP2) being responsible for binding of the base of stem-loop II, while the other (MRP1) unfolds stem-loop I by binding the anchor bases and exposing them to the solvent in a conformation suitable for duplex formation (Fig. 3C and D; Schumacher et al. 2006).

Apart from gBP21 and its orthologs, two additional proteins have been described that may affect the gRNA–pre-mRNA interaction. Both the Y-box protein RBP16 (Pelletier and Read 2003) and the RGG-box protein TbRGG1 (Vanhamme et al. 1998) are able to interact with poly-U stretches present at the 3′ terminus of all gRNAs. This activity may either affect the initial annealing reaction or it may play a role during one of the later steps of the editing cycle.

The second step of editing involves the recognition of the gRNA–mRNA complex by an endonucleolytic activity that specifically cleaves the P-diester 5′ of the duplex. Recognition of the anchor duplex and the adjacent unpaired nucleotide of the pre-mRNA has to occur structure specifically, rather than sequence specifically, since hundreds of different target sites have to be processed by

the endonucleolytic activities. However, cleavage efficiency is not independent of the identity of the adjacent mismatch pair (Lawson et al. 2001), implying more complex recognition and possibly contributions of editosome components other than the endonuclease(s) to cleavage site selection. The structure of the gRNA–mRNA complex resulting from the first step of editing was analysed *in vitro* to address the substrate structural requirements for the subsequent endonucleolytic cleavage step (Leung and Koslowsky 1999, 2001a). For three gRNA–pre-mRNA pairs, the anchor sequences were correctly positioned and base-paired. The U-tail was shown to interact with a purine-rich region upstream of and close to the editing site, whereas the information sequence was predicted to form a stem-loop between the two duplex regions. Thus, the annealing step produces gRNA–pre-mRNA duplex structures with common secondary structure features that are recognized by proteins from the editing machinery (Leung and Koslowsky 2001b). These initial studies were confirmed by solution structure probing of protein-free gRNA–pre-mRNA complexes (Yu and Koslowsky 2006). In order to gain insight into the contributions of RNA structure to each step of editing, it will be important to extend these studies by probing RNA structures in the presence of single proteins as well as editosome preparations. Thus, analysis of the structural requirements for cleavage site recognition requires the identification of the endonuclease(s) involved. Current editosome purification protocols revealed five proteins with RNase III motifs and two proteins with exonuclease motifs (Stuart et al. 2005). Three of these proteins are likely candidates for editosome endonucleases, since they contain conserved RNase III domains plus a U1-like zinc-finger and a dsRNA binding domain (KREPB1, 2 and 3). KREPB3 (kinetoplast RNA editing protein B3) was shown to be essential *in vivo*, and its repression inhibited cleavage of insertion editing sites *in vitro* (Carnes et al. 2005). Cleavage of deletion editing sites was blocked by repression of KREPB1B (Trotter et al. 2005), indicating that both proteins act as endonucleases or at least are involved in endonucleolytic cleavage. However, the substrate specificity of the two candidates has not been determined, since *in vitro* endonuclease activity could not be shown to date.

One of the editosome components, named TbMP42, was shown to have endoribonuclease activity *in vitro* (Brecht et al. 2005). The protein contains two zinc-fingers and a potential OB-fold but lacks a characteristic nuclease motif. Apart from the endonuclease activity, it exhibits exoribonuclease activity with a preference for U-nucleotides. *In vitro* cleavage studies with recombinant TbMP42 and short model substrates revealed that the protein recognizes looped-out nucleotides with a preference for cleavage 5' of the second mismatch nucleotide. This shift of cleavage specificity may be explained by the inaccessibility of the first mismatched nucleotide due to helical stacking on the preceding duplex in the model substrate. Thus, faithful cleavage and editing *in vivo* may rely on additional proteins that modify local RNA structures and help to loop out the first base for recognition by TbMP42 or by the endonuclease candidates described above.

The following steps of the editing cycle include the removal of 3'-terminal Us by exoribonuclease activities (exoUase) or the addition of Us by 3'-terminal

uridylyl transferase (TUTase). Finally, RNA ligases have to join the processed ends of the pre-mRNA cleavage fragments. All enzymatic activities have been identified as parts of editing active complexes (reviewed by Stuart et al. 2005). Candidate proteins for the exoUase (Brecht et al. 2005; Kang et al. 2005), for the TUTase (Aphasizhev et al. 2003b; Ernst et al. 2003) and for the ligase(s) (Huang et al. 2001; McManus et al. 2001; Schnauffer et al. 2001) have been identified in *T. brucei* and *L. tarentolae*. Most of these proteins occur as pairs or sets with sequence similarities (Panigrahi et al. 2003), and the functions of most of the editosome proteins have been addressed by epi-genetic and/or biochemical studies using purified editosome fractions. On the basis of these studies, the existence of separate editosome complexes with different specificities for insertion or deletion editing has been proposed (Schnauffer et al. 2003; Panigrahi et al. 2006). Alternatively, the composition of the editosome may be highly dynamic, allowing the interchange of related proteins with the proposed specificities at deletion or insertion editing sites.

The candidate proteins can be classified according to their catalytic and structural motifs, such as the dsRBDs and RNase III-motifs in putative endonucleases (Worthey et al. 2003; Simpson et al. 2004; Stuart et al. 2005). However, due to the lack of in vitro activity assays based on purified proteins in most cases, little is known about the specific RNA target structures and recognition modes that are required for each individual step of the editing cycle. The presence of characteristic motifs per se is not sufficient to allow conclusions regarding protein activity, as shown for promiscuous activities (RNA binding and protein interaction) for RRM, dsRBDs and Zn fingers (reviewed in Chen and Varani 2005). Hence, most of the RNA structural requirements have been assessed within the context of functional editosome complexes, and a number of recent studies have shed light on several general principles.

To start with, each editing site (ES) contains all the information needed to commit editing complexes to full-round insertion or deletion editing (Cifuentes-Rojas et al. 2005). An insertion ES can be converted into a deletion site by simply mutating the pre-mRNA mismatches at the ES and changing the gRNA information sequence accordingly. Hence, editing active fractions contain all enzymatic activities needed for insertion as well as deletion editing, and the ES sequence itself is the only determinant for recruitment of the necessary enzymes. While the pre-mRNA and gRNA sequences outside the ES do not affect the outcome of editing, they may influence editing efficiencies (Cifuentes-Rojas et al. 2005). It was shown that the length and thermodynamic stability of the anchor duplex directly correlate with editing efficiencies in the case of COII pre-mRNA editing directed by an engineered trans-acting gRNA (Golden and Hajduk 2006). In addition, a single gRNA point mutation in a region outside of the two duplex-forming elements flanking the editing site dramatically changes the gRNA tertiary structure and causes a strong decrease in editing rates. The mutation may inhibit a step prior to cleavage, i.e. gRNA binding/unfolding by MRP1/2 and annealing to the pre-mRNA substrate (Golden and Hajduk 2006). Conclusions will remain ambiguous, since the editing assays usually include several steps and enzymatic



activities, and each of these could be rate-limiting. Therefore, more refined assays with single components are needed to dissect the RNA structural requirements for each single step of editing.

The role of RNA structure and base pairing in reactions occurring after the endonucleolytic step was addressed by a pre-cleaved editing assay (Igo et al. 2002a, 2002b). The addition of Us to the 5'-cleavage fragment by TUTase activity was not affected by the identity of the guiding nucleotide. However, the added Us are stabilized by guiding As, probably by blocking the exoUase through base pairing. Thus, base pairing determines the number of Us added and is a prerequisite for efficient ligation (see below). The position of an upstream duplex between the pre-mRNA purine stretch and the gRNA U-tail affects the efficiency of full-round editing. A 15-bp duplex immediately adjacent to an insertion site with only two A-mismatches in the gRNA was not edited, since the initial endonucleolytic cleavage step was blocked (Igo et al. 2002b). Moving the duplex at least four nucleotides further away from the ES enabled efficient cleavage, irrespective of the actual distance. This indicates that the endonuclease activity recognizes an internal loop with a minimal size and number of mismatched bases in the range of 3–6 nucleotides. However, the sum of U addition and ligation was more efficient when the duplex was moved closer to the ES, indicating that one or both of these activities prefer substrates that resemble perfectly helical duplexes. Two ligases have been identified in the *T. brucei* editosome, named TbREL1 and TbREL2 (Huang et al. 2001; McManus et al. 2001; Schnauffer et al. 2001). In vitro studies with purified and recombinant proteins have shown that both enzymes ligate RNA termini and require a bridging RNA molecule, which in vivo is performed by a gRNA (Palazzo et al. 2003). In addition, both enzymes preferentially ligate splinted RNA with no gaps, a structure that resembles a perfect RNA duplex with a nick at the ES. This preference contributes to the accuracy of editing by “counting” the correct number of Us at the ES. Unfortunately, the recently solved crystal structure of the N-terminal catalytic domain of one of the *T. brucei* ligases, TbREL1 (Deng et al. 2004), does not reveal an obvious RNA binding mode, since patches of positive electrostatic potential around the ATP binding site are missing. The only obvious determinants of RNA binding are three completely conserved aromatic F-residues that are exposed to the solvent and may be engaged in stacking interactions with RNA bases. Possibly, other editosome proteins interact with the ligases, e.g. via protein–protein interaction motifs in the C-terminal domain and, in this way, contribute substrate specificity of the *T. brucei* ligases.

Two 3'-terminal uridylyl transferases (TUTases) have been identified in *T. brucei*, one of which is part of the editosome complexes (Ernst et al. 2003). This enzyme, named TbMP57, primarily adds a single U to the 3' end of ssRNA, with a preference for a 3'-terminal A or G. It also adds the specified number of Us to a pre-cleaved double-stranded RNA editing substrate, reflecting the characteristics of natural insertion editing events (Ernst et al. 2003). Thus, the substrate RNA structure is a major determinant of TbMP57 activity. The crystal structure of TbMP57 revealed a large, concave basic electrostatic surface that



potentially interacts with the substrate RNA (Deng et al. 2005). In the absence of a co-crystallized model RNA substrate, however, predictions as to the binding modes of single-stranded or double-stranded RNA can not be made.

For deletion editing, a specified number of non-matched Us has to be deleted from the 3' end of the 5'-cleavage fragment. Such exonuclease activities are part of the editosome and have been characterised (Aphasizhev and Simpson 2001; Igo et al. 2002b). Two candidate proteins with exonuclease motifs, TbMP100 (also: REX1) and TbMP99 (also: REX2) are part of the editosome (reviewed in Stuart et al. 2005). One of these, REX1, has been expressed as a recombinant protein and has been shown to exhibit the expected trimming activity on 3'-overhanging Us (Kang et al. 2005). A third protein, named TbMP42, which lacks any obvious exonuclease motifs was shown to have endo- and exoribonuclease activity when expressed as a recombinant protein. The protein displayed exonuclease activity *in vitro* with a preference for Us and also showed the specificity for 3'-overhanging single-stranded Us that is required during deletion editing (Brecht et al. 2005). The 3D structure of TbMP100/REX1 has been addressed by homology modelling (Mian et al. 2006). Apart from a potential hydrophobic pocket that would accommodate an extrahelical U-residue, no RNA binding sites were identified and thus, predictions as to the binding modes of editing substrates can not be made. Recently, reconstitution of deletion editing was achieved in a pre-cleaved editing assay with two recombinant enzymes, rREX1 (TbMP100) and rREL1 (Kang et al. 2005). Both enzymes were expressed from insect cells and, *in vitro*, were able to faithfully process and ligate two editing fragments in the presence of a guide RNA. Although both protein preparations were not absolutely pure, this experimental system faithfully reproduces the characteristics of editing and should allow a detailed analysis of structural requirements for exonuclease activity (rREX1) as well as ligase activity (REL1).

Taken together, a processive editosome will contain components for its structural integrity, for substrate binding, for each of the catalytic steps (cleavage, U addition or deletion, ligation) and for the translocation of the complex along a pre-edited mRNA. The sheer number of components and activities points to a complex and dynamic nature of the editing reaction that resembles the dynamics of other RNA machines in the cell, the ribosome and the spliceosome. It is of greatest interest to elucidate the structure of the editosome and its RNA components. However, the isolation of structurally defined editosomes may be impossible due to their heterogeneity and dynamic nature (e.g. Panigrahi et al. 2006). Many editosome activities were found in pairs, and the existence of distinct editosomes with distinct functions in deletion or insertion editing has been suggested (reviewed in Stuart et al. 2005). Alternatively, the composition of the editosome may be dynamic, allowing the interchange of related proteins with different activities according to the specific needs of deletion or insertion editing sites.

Thus, progress in the elucidation of RNA and RNP structures will rely on more efficient purification methods and on the ability to "freeze" the editosome at

distinct steps during the editing cycle. This could be achieved by protein-specific inhibitors or by antibiotics and related compounds that interact with defined RNA elements or proteins within the editing RNP complex. In this respect, combinatorial technologies may prove helpful for the isolation of high-affinity ligands and inhibitors in the form of nucleic acids (aptamers), proteins (microproteins, peptides) or small molecules.

### 3 Future and Conclusions

The three-dimensional structures of some of the most important players in the intracellular RNA world have become available during the last decade. High-resolution crystal structures of the small ribozymes as well as the large ribozymes, including the ribosome, have provided deep insights into the principles of RNA folding and catalysis. In addition, many RNA protein complexes involved in regulatory mechanisms and posttranscriptional processing reactions have been crystallized (as summarized in Table 1). Compared to the catalytic RNAs, the appreciation of RNA structural roles during editing processes is still in its infancy, mainly due to the lack of high-resolution structures of editing substrates and the complexes of RNA and editing factors. Elucidation of such structures may well suffer from the same inherent difficulties that also hamper spliceosome structure analysis, namely high flexibility and dynamic composition of the catalytic complexes. Already the most simple of all editing reactions, the A-to-I deamination catalyzed by the ADAR enzymes, consists of a series of events including binding, base flipping, nucleophilic attack and hydrolysis, and each step is influenced by RNA structural changes during the reaction. Thus, even with the elucidation of the three-dimensional architecture of RNA structures and RNA complexes involved in editing reactions, this will capture only frozen pictures of the molecular machines that require dynamic flexibility to accomplish their task. Even in the case of the ribosome where rapid progress from structural, biochemical, biophysical and genetic studies has been achieved, we still lack a full understanding of the mechanisms of tRNA selection and accommodation, translocation and catalysis of peptide bond formation (Korostelev et al. 2006). A detailed, static structure of editing complexes is thus only a starting point for studies that will ultimately need to explain the molecular dynamics of the different editing mechanisms.

Nevertheless, experimentally determined three-dimensional structures as well as 3D models derived from structure calculations will provide insights into the function and mechanisms of editing reactions and components. More importantly, they will inspire efforts to design new experiments that will eventually lead to an understanding of the diverse mechanisms of RNA editing in eukaryotic cells.

**Acknowledgements** Work in the author's laboratory is supported by the Dr. Illing Foundation.

## References

- Allain FH, Howe PW, Neuhaus D, Varani G (1997) Structural basis of the RNA-binding specificity of human U1A protein. *EMBO J* 16:5764–5772
- Anant S, Davidson NO (2000) An AU-rich sequence element (UUUN[A/U]U) downstream of the edited C in apolipoprotein B mRNA is a high-affinity binding site for Apobec-1: binding of Apobec-1 to this motif in the 3' untranslated region of c-myc increases mRNA stability. *Mol Cell Biol* 20:1982–1992
- Antson AA, Dodson EJ, Dodson G, Greaves RB, Chen X-P, Gollnick P (1999) Structure of the trp RNA-binding attenuation protein, TRAP, bound to RNA. *Nature* 401:235–242
- Aphasizhev R, Simpson L (2001) Isolation and characterization of a U-specific 3' 5'-exonuclease from mitochondria of *Leishmania tarentolae*. *J Biol Chem* 276:21280–21284
- Aphasizhev R, Aphasizhev I, Nelson RE, Simpson L (2003a) A 100-kD complex of two RNA-binding proteins from mitochondria of *Leishmania tarentolae* catalyzes RNA annealing and interacts with several RNA editing components. *RNA* 9:62–76
- Aphasizhev R, Aphasizhev I, Simpson L (2003b) A tale of two TUTases. *Proc Natl Acad Sci USA* 100:10617–10622
- Athanasiadis A, Rich A, Maas S (2004) Widespread A-to-I RNA editing of Alu containing mRNAs in the human transcriptome. *PLOS Biol* 2:e391
- Aubourg S, Boudet N, Kreis M, Lecharny A (2000) In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. *Plant Mol Biol* 42:603–613
- Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289:905–920
- Barrett MP, Burchmore RJ, Stich A, Lazzari JO, Frasch AC, Cazzulo JJ, Krishna S (2003) The trypanosomiasis. *Lancet* 362:1469–1480
- Bass BL (2002) RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* 71:817–846
- Bass BL, Weintraub H (1988) An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell* 55:1089–1098
- Battiste JL, Mao H, Rao NS, Tan R, Muhandiram DR, Kay LE, Frankel AD, Williamson JR (1996)  $\alpha$ -Helix-RNA major groove recognition in an HIV-1 Rev peptide RRE RNA complex. *Science* 273:1547–1551
- Beal PA (2005) Duplex RNA-binding enzymes: headlines from neurobiology, virology, and development. *ChemBioChem* 6:257–266
- Blom D, Van den Burg J, Breek CK, Speijer D, Muijsers AO, Benne R (2001) Cloning and characterization of two guide RNA-binding proteins from mitochondria of *Crithidia fasciculata*: gBP27, a novel protein, and gBP29, the orthologue of *Trypanosoma brucei* gBP21. *Nucleic Acids Res* 29:2950–2962
- Blow M, Futreal PA, Wooster R, Stratton MR (2004) A survey of RNA editing in human brain. *Genome Res* 14:2379–2387
- Blum B, Bakalara N, Simpson L (1990) A model for RNA editing in kinetoplastid mitochondria: guide RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell* 60:189–198
- Brecht M, Niemann M, Schlüter E, Müller UF, Stuart K, Göringer HU (2005) TbMP42, a protein component of the RNA editing complex in African trypanosomes has endo-exoribonuclease activity. *Mol Cell* 17:621–630
- Byrne EM, Gott JM (2002) Cotranscriptional editing of *Physarum* mitochondrial RNA requires local features of the native template. *RNA* 8:1174–1185
- Byrne EM, Gott JM (2004) Unexpectedly complex editing patterns at dinucleotide insertion sites in *Physarum* mitochondria. *Mol Cell Biol* 24:7821–7828
- Carnes J, Trotter JR, Ernst NL, Steinberg AG, Stuart K (2005) An essential RNase III insertion editing endonuclease in *Trypanosoma brucei*. *Proc Natl Acad Sci USA* 102:16614–16619

- Cermiakan N, Ikeda TM, Cedergren R, Gray MW (1996) Sequences homologous to yeast mitochondrial and bacteriophage T7 and T3 RNA polymerase are widespread throughout the eukaryotic lineage. *Nucleic Acids Res* 24:648–654
- Chateigner-Boutin AL, Hanson MR (2002) Cross-competition in transgenic chloroplasts expressing single editing sites reveals shared cis elements. *Mol Cell Biol* 22:8448–8456
- Chateigner-Boutin AL, Hanson MR (2003) Developmental co-variation of RNA editing extent of plastid editing sites exhibiting similar cis-elements. *Nucleic Acids Res* 31:2586–2594
- Chen Y, Varani G (2005) Protein families and RNA recognition. *FEBS J* 272:2088–2097
- Chen SH, Habib G, Yang CY, Gu ZW, Lee BR, Weng SA, Silberman SR, Cai SJ, Deslypere JP, Rosseneu M (1987) Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. *Science* 238:363–366
- Chen CX, Cho DS, Wang Q, Lai F, Carter KC, Nishikura K (2000) A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA* 6(5):755–767
- Cheng YW, Visomirski-Robic LM, Gott JM (2001) Nontemplated addition of nucleotides to the 3'-end of nascent RNA during RNA editing in *Physarum*. *EMBO J* 20:1405–1414
- Chester A, Weinreb V, Carter CW Jr, Navaratnam N (2004) Optimization of apolipoprotein B mRNA editing by APOBEC1 apoenzyme and the role of its auxiliary factor, ACF. *RNA* 10:1399–1411
- Choury D, Araya A (2006) RNA editing site recognition in heterologous plant mitochondria. *Curr Genet* 50:405–416
- Choury D, Farre JC, Jordana X, Araya A (2004) Different patterns in the recognition of editing sites in plant mitochondria. *Nucleic Acids Res* 32:6397–6406
- Cifuentes-Rojas C, Halbig K, Sacharidou A, De Nova-Ocampo M, Cruz-Reyes J (2005) Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes. *Nucleic Acids Res* 33:6610–6620
- Dawson TR, Sansam CL, Emeson RB (2004) Structure and sequence determinants required for the RNA editing of ADAR2 substrates. *J Biol Chem* 279:4941–4951
- DeCervo J, Carmichael GG (2005) Retention and repression: fates of hyperedited RNAs in the nucleus. *Curr Opin Cell Biol* 17:302–308
- Deng J, Schnauffer A, Salavati R, Stuart KD, Hol WG (2004) High resolution crystal structure of a key editosome enzyme from *Trypanosoma brucei*: RNA editing ligase 1. *J Mol Biol* 343:601–613
- Deng J, Ernst NL, Turley S, Stuart KD, Hol WG (2005) Structural basis for UTP specificity of RNA editing TUTases from *Trypanosoma brucei*. *EMBO J* 24:4007–4017
- Deo RC, Bonanno JB, Sonenberg N, Burley SK (1999) Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* 98:835–845
- Ding J, Hayashi MK, Zhang Y, Manche L, Krainer AR, Xu RM (1999) Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev* 13:1102–1115
- Ernst NL, Panicucci B, Igo RP Jr, Panigrahi AK, Salavati R, Stuart K (2003) TbMP57 is a 39 terminal uridylyl transferase (TUTase) of the *Trypanosoma brucei* editosome. *Mol Cell* 11:1525–1536
- Evans D, Marquez SM, Pace NR (2006) RNase P: interface of the RNA and protein worlds. *Trends Biochem Sci* 31:333–341
- Frankel AD, Smith CA (1998) Induced folding in RNA-protein recognition: more than a simple molecular handshake. *Cell* 92:149–151
- Gerber AP, Herschlag D, Brown PO (2004) Extensive association of functionally and cytologically related mRNAs with Puf family RNA-binding proteins in yeast. *PLOS Biol* 2:e79
- Golden DE, Hajduk SL (2006) The importance of RNA structure in RNA editing and a potential proofreading mechanism for correct guide RNA:pre-mRNA binary complex formation. *J Mol Biol* 359:585–596
- Gott JM (2001) RNA editing in *Physarum polycephalum*. In: Bass B (ed) RNA editing: frontiers in molecular biology. Oxford University Press, Oxford, UK, pp 20–37

- Gott JM, Visomirski LM, Hunter JL (1993) Substitutional and insertional RNA editing of the cytochrome c oxidase subunit 1 mRNA of *Physarum polycephalum*. *J Biol Chem* 268:25483–25486
- Gott JM, Parimi N, Bundschuh R (2005) Discovery of new genes and deletion editing in *Physarum* mitochondria enabled by a novel algorithm for finding edited mRNAs. *Nucleic Acids Res* 33:5063–5072
- Handa N, Nureki O, Kurimoto K, Kim I, Sakamoto H, Shimura Y, Muto Y, Yokoyama S (1999) Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* 398:579–585
- Hansen JL, Ippolito JA, Ban N, Nissen P, Moore PB, Steitz TA (2002) The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Mol Cell* 10:117–128
- Hausmann S, Garcin D, Delenda C, Kolakofsky D (1999a) The versatility of paramyxovirus RNA polymerase stuttering. *J Virol* 73:5568–5576
- Hausmann S, Garcin D, Morel AS, Kolakofsky D (1999b) Two nucleotides immediately upstream of the essential A6G3 slippery sequence modulate the pattern of G insertions during Sendai virus mRNA editing. *J Virol* 73:343–351
- Hayes ML, Reed ML, Hegeman CE, Hanson MR (2006) Sequence elements critical for efficient RNA editing of a tobacco chloroplast transcript *in vivo* and *in vitro*. *Nucleic Acids Res* 34:3742–3754
- Hegeman CE, Hayes ML, Hanson MR (2005) Substrate and cofactor requirements for RNA editing of chloroplast transcripts in *Arabidopsis* *in vitro*. *Plant J* 42:124–132
- Hermann T, Schmid B, Heumann H, Göringer HU (1997) A three-dimensional working model for a guide RNA from *Trypanosoma brucei*. *Nucleic Acids Res* 25:2311–2318
- Holmes RK, Malim MH, Bishop KN (2007) APOBEC-mediated viral restriction: not simply editing? *Trends Biochem Sci* 32:118–218
- Horton TL, Landweber LF (2000) Evolution of four types of RNA editing in myxomycetes. *RNA* 6:1339–1346
- Huang CE, Cruz-Reyes J, Zhelonkina AG, O'Hearn S, Wirtz E, Sollner Webb B (2001) Roles for ligases in the RNA editing complex of *Trypanosoma brucei*: Band IV is needed for U-deletion and RNA repair. *EMBO J* 20:4694–4703
- Hudson BP, Martinez-Yamout MA, Dyson HJ, Wright PE (2004) Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol* 11:257–264
- Igo RP Jr, Lawson SD, Stuart K (2002a) RNA sequence and base pairing effects on insertion editing in *Trypanosoma brucei*. *Mol Cell Biol* 22:1567–1576
- Igo RP Jr, Weston DS, Ernst NL, Panigrahi AK, Salavati R, Stuart K (2002b) Role of uridylyate specific exoribonuclease activity in *Trypanosoma brucei* RNA editing. *Eukaryot Cell* 1:112–118
- Johansson C, Finger LD, Trantirek L, Mueller TD, Kim S, Laird-Offringa IA, Feigon J (2004) Solution structure of the complex formed by the two N-terminal RNA-binding domains of nucleolin and a pre-rRNA target. *J Mol Biol* 337:799–816
- Källman AM, Sahlin M, Ohman M (2003) ADAR2 A→I editing: site selectivity and editing efficiency are separate events. *Nucleic Acids Res* 31:4874–4881
- Kambach C, Walke S, Young R, Avis JM, de la Fortelle E, Raker VA, Lührmann R, Li J, Nagai K (1999) Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* 96:375–387
- Kang X, Rogers K, Gao G, Falick AM, Zhou S, Simpson L (2005) Reconstitution of uridine-deletion precleaved RNA editing with two recombinant enzymes. *Proc Natl Acad Sci USA* 102:1017–1022
- Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J, Suzuki H, Carninci P, Hayashizaki Y, Wells C, Frith M, Ravasi T, Pang KC, Hallinan J, Mattick J, Hume DA, Lipovich L, Batalov S, Engstrom PG, Mizuno Y, Faghihi MA, Sandelin A, Chalk AM, Mottagui-Tabar S, Liang Z, Lenhard B, Wahlestedt C, RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Project Core Group), FANTOM Consortium (2005) Antisense transcription in the mammalian transcriptome. *Science* 309:1564–1566

- Kharrat A, Macias MJ, Gibson TJ, Nilges M, Pastore A (1995) Structure of the dsRNA-binding domain of *E. coli* RNase III. *EMBO J* 14:3572–3584
- Kielkopf CL, Lucke S, Green MR (2004) U2AF homology motifs: protein recognition in the RRM world. *Genes Dev* 18:1513–1526
- Kolakofsky D, Hausmann S (1998) Cotranscriptional paramyxovirus mRNA editing: a contradiction in terms? In: Grosjean H, Benne R (eds) *Modification and editing of RNA*. ASM Press, Washington, DC, pp 413–420
- Köller J, Müller UF, Schmid B, Missel A, Kruft V, Stuart K, Göringer HU (1997) *Trypanosoma brucei* gBP21. An arginine-rich mitochondrial protein that binds to guide RNA with high affinity. *J Biol Chem* 272:3749–3757
- Korostelev A, Trakhanov S, Laurberg M, Noller HF (2006) Crystal structure of a 70S ribosome tRNA complex reveals functional interactions and rearrangements. *Cell* 126:1065–1077
- Kotera E, Tasaka M, Shikanai T (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* 433:326–330
- Kumar M, Carmichael GG (1997) Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc Natl Acad Sci USA* 94:3542–3547
- Lamb RA, Kolakofsky D (1996) Paramyxoviridae: the viruses and their replication. In: Fields BN, Knipe DM, Howley PM (eds) *Fields virology*, 3rd edn. Raven Press, New York, NY, pp 1177–1204
- Lambert L, Müller UF, Souza AE, Göringer HU (1999) The involvement of gRNA-binding protein gBP21 in RNA editing – an *in vitro* and *in vivo* analysis. *Nucleic Acids Res* 27:1429–1436
- Lawson SD, Igo RP Jr, Salavati R, Stuart KD (2001) The specificity of nucleotide removal during RNA editing in *Trypanosoma brucei*. *RNA* 7:1793–802
- Lehmann KA, Bass BL (1999) The importance of internal loops within RNA substrates of ADAR1. *J Mol Biol* 291:1–13
- Lehmann KA, Bass BL (2000) Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. *Biochemistry* 39:12875–12884
- Leontis NB, Westhof E (2003) Analysis of RNA motifs. *Curr Opin Struct Biol* 13:300–308
- Leontis NB, Lescoute A, Westhof E (2006) The building blocks and motifs of RNA architecture. *Curr Opin Struct Biol* 2006 16:279–287
- Lescoute A, Leontis NB, Massire C, Westhof E (2005) Recurrent structural RNA motifs, isostericity matrices and sequence alignments. *Nucleic Acids Res* 33:2395–2409
- Leung SS, Koslowsky DJ (1999) Mapping contacts between gRNA and mRNA in trypanosome RNA editing. *Nucleic Acids Res* 27:778–787
- Leung SS, Koslowsky DJ (2001a) Interactions of mRNAs and gRNAs involved in trypanosome mitochondrial RNA editing: structure probing of an mRNA bound to its cognate gRNA. *RNA* 7:1803–1816
- Leung SS, Koslowsky DJ (2001b) RNA editing in *Trypanosoma brucei*: characterization of gRNA U-tail interactions with partially edited mRNA substrates. *Nucleic Acids Res* 29:703–709
- Levanon EY, Eisenberg E, Yelin R, Nemzer S, Hallegger M, Shemesh R, Fligelman ZY, Shoshan A, Pollock SR, Sztybel D, Olshansky M, Rechavi G, Jantsch MF (2004) Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol* 22:1001–1005
- Levanon K, Eisenberg E, Rechavi G, Levanon EY (2005) Letter from the editor. Adenosine-to inosine RNA editing in Alu repeats in the human genome. *EMBO Rep* 6:831–835
- Lewis HA, Musunuru K, Jensen KB, Edo C, Chen H, Darnell RB, Burley SK (2000) Sequence-specific RNA binding by a nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell* 100:323–332
- Lu D, Searles MA, Klug A (2003) Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. *Nature* 426:96–100
- Ma JB, Ye K, Patel DJ (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* 429:318–322



- Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ (2005) Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* 434:666–670
- Maas S, Rich A, Nishikura K (2003) A-to-I RNA editing: recent news and residual mysteries. *J Biol Chem* 278:1391–1394
- Macbeth MR, Schubert HL, Vandemark AP, Lingam AT, Hill CP, Bass BL (2005) Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science* 309:1534–1539
- Macdonald LE, Zhou Y, McAllister WT (1993) Termination and slippage by bacteriophage T7 RNA polymerase. *J Mol Biol* 232:1030–1047
- MacRae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA (2006) Structural basis for double-stranded RNA processing by Dicer. *Science* 311:195–198
- Maier RM, Zeltz P, Kossel H, Bonnard G, Gualberto JM, Grienemberger JM (1996) RNA editing in plant mitochondria and chloroplasts. *Plant Mol Biol* 32:343–365
- Maris C, Masse J, Chester A, Navaratnam N, Allain FH (2005a) NMR structure of the apoB mRNA stem-loop and its interaction with the C to U editing APOBEC1 complementary factor. *RNA* 11:173–186
- Maris C, Dominguez C, Allain FH (2005b) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* 272:2118–2131
- Masters BS, Stohl LL, Clayton AD (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell* 51:98–99
- McManus MT, Shimamura M, Grams J, Hajduk SL (2001) Identification of candidate mitochondrial RNA editing ligases from *Trypanosoma brucei*. *RNA* 7:167–175
- Mehta A, Driscoll DM (2002) Identification of domains in apobec-1 complementation factor required for RNA binding and apolipoprotein-B mRNA editing. *RNA* 8:69–82
- Mehta A, Kinter MT, Sherman NE, Driscoll DM (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol Cell Biol* 20:1846–1854
- Mian IS, Worthey EA, Salavati R (2006) Taking U out, with two nucleases? *BMC Bioinformatics* 16:305
- Miller ML, Antes TJ, Qian F, Miller DL (2006) Identification of a putative mitochondrial RNA polymerase from *Physarum polycephalum*: characterization, expression, purification, and transcription *in vitro*. *Curr Genet* 49:259–271
- Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res* 15:8783–8798
- Miyamoto T, Obokata J, Sugiura M (2002) Recognition of RNA editing sites is directed by unique proteins in chloroplasts: biochemical identification of cis-acting elements and trans-acting factors involved in RNA editing in tobacco and pea chloroplasts. *Mol Cell Biol* 22:6726–6734
- Miyamoto T, Obokata J, Sugiura M (2004) A site-specific factor interacts directly with its cognate RNA editing site in chloroplast transcripts. *Proc Natl Acad Sci USA* 101:48–52
- Moore MJ (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science* 309:1514–1518
- Moore PB, Steitz TA (2003) The structural basis of large ribosomal subunit function. *Annu Rev Biochem* 72:813–850
- Müller UF, Göringer HU (2002) Mechanism of the gBP21-mediated RNA/RNA annealing reaction: matchmaking and charge reduction. *Nucleic Acids Res* 30:447–455
- Müller UF, Lambert L, Göringer HU (2001) Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21. *EMBO J* 20:1394–1404
- Müller M, Heuck A, Niessing D (2007) Directional mRNA transport in eukaryotes: lessons from yeast. *Cell Mol Life Sci* 64:171–180
- Nakamura T, Meierhoff K, Westhoff P, Schuster G (2003) RNA-binding properties of HCF152, an *Arabidopsis* PPR protein involved in the processing of chloroplast RNA. *Eur J Biochem* 270:4070–4081



- Nelson JA, Uhlenbeck OC (2006) When to believe what you see. *Mol Cell* 23:447–450
- Noller HF (2005) RNA structure: reading the ribosome. *Science* 309:1508–1514
- Öhman M, Kallman AM, Bass BL (2000) *In vitro* analysis of the binding of ADAR2 to the pre-mRNA encoding the GluR-B R/G site. *RNA* 6:687–697
- Palazzo SS, Panigrahi AK, Igo RP, Salavati R, Stuart K (2003) Kinetoplastid RNA editing ligases: complex association, characterization, and substrate requirements. *Mol Biochem Parasitol* 127:161–167
- Panigrahi AK, Schnauffer A, Ernst NL, Wang B, Carmean N, Salavati R, Stuart K (2003) Identification of novel components of *Trypanosoma brucei* editosomes. *RNA* 9:484–492
- Panigrahi AK, Ernst NL, Domingo GJ, Fleck M, Salavati R, Stuart KD (2006) Compositionally and functionally distinct editosomes in *Trypanosoma brucei*. *RNA* 12:1038–1049
- Parker JS, Roe SM, Barford D (2005) Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* 434:663–666
- Pelletier M, Read LK (2003) RBP16 is a multifunctional gene regulatory protein involved in editing and stabilization of specific mitochondrial mRNAs in *Trypanosoma brucei*. *RNA* 9:457–468
- Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J (1987) A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell* 50:831–840
- Price SR, Evans PR, Nagai K (1998) Crystal structure of the spliceosomal U2B''-U2A' protein complex bound to a fragment of U2 small nuclear RNA. *Nature* 394:645–650
- Puglisi JD, Chen L, Blanchard S, Frankel AD (1995) Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. *Science* 270:1200–1203
- Ramos A, Grunert S, Adams J, Micklem DR, Proctor MR, Freund S, Bycroft M, St Johnston D, Varani G (2000) RNA recognition by a Staufen double stranded RNA binding domain. *EMBO J* 19:997–1009
- Reichow SL, Hamma T, Ferre-D'Amare AR, Varani G (2007) The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res* 35:1452–1464
- Richardson N, Navaratnam N, Scott J (1998) Secondary structure for the apolipoprotein B mRNA editing site. Au-binding proteins interact with a stem loop. *J Biol Chem* 273:31707–31717
- Rigaut G, Shevchenko A, Rutz B, Wilm M, Mann M, Seraphin B (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 17:1030–1032
- Ryter JM, Schultz SC (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J* 17:7505–7513
- Scadden AD, Smith CW (2001) Specific cleavage of hyper-edited dsRNAs. *EMBO J* 20:4243–4252
- Schmid B, Riley GR, Stuart K, Göringer HU (1995) The secondary structure of guide RNA molecules from *Trypanosoma brucei*. *Nucleic Acids Res* 23:3093–3102
- Schnauffer A, Panigrahi AK, Panicucci B, Igo R Jr, Salavati R, Stuart K (2001) An RNA ligase essential for RNA editing and survival of the bloodstream form of *Trypanosoma brucei*. *Science* 291:2159–2162
- Schnauffer A, Ernst N, O'Rear J, Salavati R, Stuart K (2003) Separate insertion and deletion sub-complexes of the *Trypanosoma brucei* RNA editing complex. *Mol Cell* 12:307–319
- Schumacher MA, Karamooz E, Zikova A, Trantirek L, Lukes J (2006) Crystal structures of *T. brucei* MRP1/MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism. *Cell* 126:701–711
- Seeburg PH, Hartner J (2003) Regulation of ion channel/neurotransmitter receptor function by RNA editing. *Curr Opin Neurobiol* 13:279–283
- Simpson L, Aphasizhev R, Gao G, Kang X (2004) Mitochondrial proteins and complexes in *Leishmania* and *Trypanosoma* involved in U insertion/deletion RNA editing. *RNA* 10:159–170
- Small ID, Peeters N (2000) The PPR motif – a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci* 25:46–47

- Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* 305:1434–1437
- Stark H, Lührmann R (2006) Cryo-electron microscopy of spliceosomal components. *Annu Rev Biophys Biomol Struct* 35:435–457
- Stefl R, Allain FH (2005) A novel RNA pentaloop fold involved in targeting ADAR2. *RNA* 11:592–597
- Stefl R, Xu M, Skrisovska L, Emeson RB, Allain FH (2006) Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. *Structure* 14(2):345–355
- Stephens OM, Yi-Brunozzi HY, Beal PA (2000) Analysis of the RNA editing reaction of ADAR2 with structural and fluorescent analogues of the GluR-B R/G editing site. *Biochemistry* 39:12243–12251
- Stuart KD, Schnauffer A, Ernst NL, Panigrahi AK (2005) Complex management: RNA editing in trypanosomes. *Trends Biochem Sci* 30:97–105
- Takenaka M, Neuwirt J, Brennicke A (2004) Complex cis-elements determine an RNA editing site in pea mitochondria. *Nucleic Acids Res* 32:4137–4144
- Teng B, Burant CF, Davidson NO (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* 260:1816–1819
- Tillich M, Funk HT, Schmitz-Linneweber C, Poltnigg P, Sabater B, Martin M, Maier RM (2005) Editing of plastid RNA in *Arabidopsis thaliana* ecotypes. *Plant J* 43:708–715
- Trotter JR, Ernst NL, Carnes J, Panicucci B, Stuart K (2005) A deletion site editing endonuclease in *Trypanosoma brucei*. *Mol Cell* 20:403–412
- Tsudzuki T, Wakasugi T, Sugiura M (2001) Comparative analysis of RNA editing sites in higher plant chloroplasts. *J Mol Evol* 53:327–332
- Turelli P, Trono D (2005) Editing at the crossroad of innate and adaptive immunity. *Science* 307:1061–1065
- Valadkhan S (2005) snRNAs as the catalysts of pre-mRNA splicing. *Curr Opin Chem Biol* 9:603–608
- Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev* 20:515–524
- Valente L, Nishikura K (2005) ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. *Prog Nucleic Acid Res Mol Biol* 79:299–338
- van der Merwe JA, Takenaka M, Neuwirt J, Verbitskiy D, Brennicke A (2006) RNA editing sites in plant mitochondria can share cis-elements. *FEBS Lett* 580:268–272
- Vanhamme L, Perez-Morga D, Marchal C, Speijer D, Lambert L, Geuskens M, Alexandre S, Ismaili N, Göringer U, Benne R, Pays E (1998) *Trypanosoma brucei* TBRGG1, a mitochondrial oligo(U)-binding protein that co-localizes with an *in vitro* RNA editing activity. *J Biol Chem* 273:21825–21833
- Vicens Q, Cech TR (2006) Atomic level architecture of group I introns revealed. *Trends Biochem Sci* 31:41–51
- Vidal S, Curran J, Kolakofsky D (1990) A stuttering model for paramyxovirus P mRNA editing. *EMBO J* 9:2017–2022
- Wang SS, Mahendran R, Miller DL (1999) Editing of cytochrome b mRNA in *Physarum* mitochondria. *J Biol Chem* 274:2725–2731
- Wang X, McLachlan J, Zamore PD, Hall TM (2002) Modular recognition of RNA by a human pumilio-homology domain. *Cell* 110:501–512
- Wimberly BT, Brodersen DE, Clemons WM Jr, Morgan-Warren RJ, Carter AP, Vonrhein C, Hartsch T, Ramakrishnan V (2000) Structure of the 30S ribosomal subunit. *Nature* 407:327–339
- Wong SK, Sato S, Lazinski DW (2001) Substrate recognition by ADAR1 and ADAR2. *RNA* 7:846–858
- Worthey EA, Schnauffer A, Mian IS, Stuart K, Salavati R (2003) Comparative analysis of editosome proteins in trypanosomatids. *Nucleic Acids Res* 31:6392–6408
- Xie K, Sowden MP, Dance GS, Torelli AT, Smith HC, Wedekind JE (2004) The structure of a yeast RNA-editing deaminase provides insight into the fold and function of activation induced deaminase and APOBEC-1. *Proc Natl Acad Sci USA* 101:8114–8119

- Ye X, Kumar RA, Patel DJ (1995) Molecular recognition in the bovine immunodeficiency virus Tat peptide-TAR RNA complex. *Chem Biol* 2:827–840
- Yi-Brunozzi HY, Stephens OM, Beal PA (2001) Conformational changes that occur during an RNA-editing adenosine deamination reaction. *J Biol Chem* 276:37827–37833
- Yu LE, Koslowsky DJ (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
- Zhang Z, Carmichael GG (2001) The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* 106:465–475

RNA Editing

Göringer, H.U. (Ed.)

2008, XI, 231 p. 41 illus., 13 illus. in color., Hardcover

ISBN: 978-3-540-73786-5