

The challenge of target sequence specificity in C→U RNA editing

Nicholas O. Davidson

Departments of Internal Medicine, Molecular Biology and Pharmacology, Washington University, St. Louis, Missouri, USA

Address correspondence to: Nicholas O. Davidson, Division of Gastroenterology, Box 8124, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, Missouri 63110, USA. Phone: (314) 362-2027; Fax: (314) 362-2033; E-mail: nod@im.wustl.edu.

J. Clin. Invest. **109**:291–294 (2002). DOI:10.1172/JCI200214979.

There is increasing awareness that multiple, often overlapping mechanisms exist for amplifying the repertoire of protein products specified through the mammalian genome. An expanding array of processing and targeting mechanisms is now emerging, each representing a potentially important restriction point in the regulation of eukaryotic gene expression, and each expanding the possibilities specified by the literal code of the genome. These co- and posttranscriptional regulatory events include capping, alternative splicing, differential polyadenylation, RNA editing, nuclear export, alternative decay and degradation pathways, as well as alterations in ribosomal loading or translation. Recent advances suggest important links between RNA editing, particularly C→U editing, and other co- and posttranscriptional events regulating gene expression.

Forms of RNA editing in mammalian cells

In RNA editing, the coding sequence of the transcript is altered from its genomically templated version. In mammals, two major types of RNA base-modification have been described, both forms of substitutional editing, in which a single nucleotide is altered posttranscriptionally (reviewed in ref. 1). These include adenosine-to-inosine (A→I) and cytidine-to-uridine (C→U) editing. In all instances where the translation products have been characterized, the product of the edited mRNA acquires functional characteristics distinct from those of the product of the unedited transcript (reviewed in ref. 1).

Mammalian A→I RNA editing occurs on an unspliced mRNA template and requires partial base pairing between exonic and adjacent intronic sequences in a double-stranded RNA conformation. Inosine is translated as guanosine in the edited mRNA, so the effect of editing mimics that of an A→G change. Examples of transcripts that undergo A→I editing include the calcium-gated glutamate receptor, GluR-B, as well as subtypes of the 5HT_{2c} receptor and hepatitis δ virus (1). A family of adenosine

deaminases acting on RNA (ADARs) is responsible for these A→I editing events, all of which operate through hydrolytic deamination of a defined target transcript (2). Members of the ADAR family function as modular enzymes and operate in vitro without obligate cofactor requirements. Each ADAR contains a C-terminal deaminase domain and an N-terminal domain with varying numbers (two or three) of double-stranded RNA-binding motifs (1, 3). These features of A→I RNA editing, namely the *cis*-acting requirements, the range of targets, and the enzymatic machinery, serve as useful reference points in understanding the mechanisms regulating the second form of base-modification editing, C→U RNA editing.

C→U RNA editing involves a site-specific deamination of cytidine in a nuclear transcript that is mediated by an enzyme complex. Physiological targets for mammalian C→U editing are currently confined to nuclear mRNAs encoding apoB and neurofibromatosis type I (NF1) transcripts. C→U RNA editing is mediated by an enzyme complex for which some but not all the component factors are known and which, unlike A→I editing, requires a minimum of two protein components acting on a suitable template to reproduce targeted enzymatic deamination in vitro.

Cis-acting requirements

The original and most fully detailed example of C→U RNA editing is mammalian apoB mRNA, in which a site-specific cytidine deamination introduces a UAA stop codon into the translational reading frame, resulting in synthesis of a truncated protein, apoB48 (reviewed in ref. 4). C→U RNA editing of apoB occurs within enterocytes of the mammalian small intestine, as well as in the liver in some species (reviewed in ref. 4). Under physiological circumstances, C→U editing of apoB mRNA targets a single cytidine out of more than 14,000 nucleotides, a process constrained by stringency in the *cis*-acting elements and by the protein factors responsible for targeted deamination. These *cis*-acting elements, which have been well described in recent reviews (1, 4), include an AU-rich bulk RNA context in which an 11-nucleotide motif occurs four nucleotides downstream of the targeted cytidine. This sequence is flanked by elements both 5' and 3' of a minimal, approximately 50-nucleotide editing cassette. Structural predictions indicate that this region assumes significant secondary structure (5–8) with the targeted cytidine exposed at the apex of a stem-loop bulge. The most plausible interpretation of these collective requirements is to coordinate optimal alignment of the substrate with respect to the active site of the editing enzyme.

RNA splicing precedes C→U RNA editing and helps account for the specificity of the latter processing event. ApoB RNA editing occurs within the nucleus, and its preferred substrate is spliced, polyadenylated RNA. The

relationship between RNA splicing and C→U editing has been explored directly in recent studies by Smith and colleagues, who demonstrated that chimeric apoB RNA substrates containing an intron either 5' or 3' to the minimal editing cassette are edited in cultured cells with lower efficiency than a spliced template of otherwise identical sequence (9, 10). These authors also explored effects of conditional export of an unspliced apoB RNA chimera in which an editing cassette had been inserted within an intron, upstream of a Rev response element (RRE). Cells cotransfected with this chimeric plasmid and with a Rev expression construct that allowed for nuclear export of the unspliced transcript as a result of Rev-RRE binding (10). Under these conditions, the unspliced apoB RNA chimera was efficiently edited, suggesting but not formally proving that editing takes place in the nucleus (10). These findings indicate that targeting of pre-mRNA to the splicing pathway may be inhibitory to C→U RNA editing, and they suggest that pre-mRNA is not the preferred substrate. Hence, C→U editing apparently follows RNA splicing. In this respect, it differs from A→I, which appears to be tightly coupled to splicing (11).

It remains to be determined whether the binding of splicing factors in proximity to the splice donor and acceptor sites of pre-mRNA inhibits the binding or other regulatory functions of the proteins involved in C→U editing. This question is of importance since apoB RNA editing occurs in the middle of a large exon (>7 kb), the targeted cytidine being flanked by >2 kb of exonic sequence, presumably a site remote from the region to which the splicing machinery attaches.

RNA splicing plays a distinct but no less central role in the C→U editing of NF1. Earlier descriptions of NF1 RNA editing indicated that a subset of tumors manifested C→U RNA editing, changing an arginine CGA to a UGA termination codon (12). Two distinguishing features of tumors that edit the NF1 RNA have emerged. First, they produce an alternatively spliced NF1 RNA in which an exon downstream of the edited base is included. Second, they express the mRNA for apobec-1, the catalytic deaminase of the apoB RNA editing enzyme (13). The presence of the alternatively spliced exon (23A) in NF1 transcripts that undergo C→U editing is predicted through molecular modeling (D. Mukhophadyay et al., unpublished observations) to position the targeted cytidine in a favorable configuration relative to apobec-1. Thus, important functional interactions between local and distant *cis*-acting elements, in conjunction with the regulated expression of *trans*-acting factors, together modulate C→U editing activity and specificity. The functional and clinical significance of C→U editing of NF1 RNA in these tumors is not resolved, but the findings suggest a new level of genetic heterogeneity in relation to the loss of tumor suppressor function of neurofibromin.

Trans-acting factors

C→U RNA editing is mediated by an enzyme complex that includes the RNA-specific cytidine deaminase apobec-1 (4) and apobec-1 complementation factor (ACF), a novel protein that likely serves as the RNA recognition component of the core enzyme complex (14, 15).

RNA-specific cytidine deaminases. Apobec-1, the catalytic subunit of the apoB RNA-editing holoenzyme, is a developmentally regulated RNA-specific cytidine deaminase expressed almost exclusively in the luminal gastrointestinal tract of humans (reviewed in ref. 4). The recent finding that apobec-1 mRNA occurs in a subset of tumors in patients with neurofibromatosis provides the first example of its expression beyond the gastrointestinal tract in humans (13). The expression and localization pattern of apobec-1 in the human small intestine coincides with that of its presumed target, apoB mRNA, but apobec-1 expression in peripheral tissues of rodents, including sites that express virtually no apoB mRNA, remains enigmatic. The targeted deletion of the murine *apobec-1* confirmed that this protein is required for C→U RNA editing of apoB and for production of plasma apoB48 but yielded no obvious clues to an auxiliary or covert function for apobec-1 (reviewed in refs. 1, 4).

Homologs of apobec-1 have been identified in both human and murine databases, and two such homologs, activation induced deaminase (AID) and apobec-2/ARCD-1, have recently been characterized in detail. AID was identified in a genetic screen of differentially expressed products detected during lymphocyte activation and Ig class switch recombination. Targeted disruption of AID in mice leads to elevated levels of serum IgM concentration and defective class switch recombination, as well as alterations in somatic hypermutation (16). AID exhibits cytidine deaminase activity on a monomeric substrate, but no authentic RNA target has yet emerged for its activity in vivo, and its cofactor requirements are unknown. Its chromosomal location, close to that of *APOBEC1* on 12p13, suggests that it may have arisen as a gene duplication. Another homolog of apobec-1 has been identified that also manifests cytidine deaminase activity on a monomeric substrate. This homolog, the product of a locus on chromosome 6p21, referred to as apobec-2 (17) or apobec-1-related cytidine deaminase 1 (ARCD-1) (18), has no C→U editing activity on apoB RNA (hence its designation as a *related* deaminase) but rather inhibits apobec-1-mediated RNA editing through its ability to interact with apobec-1 and with ACF (18). The inhibition of apobec-1-mediated C→U RNA editing by ARCD-1 is reminiscent of the inhibition of ADAR1- and ADAR2a-mediated A→I editing of 5HT_{2c} receptor RNAs by a related homolog, ADAR3 (3). These findings considered together raise the possibility that some members of the gene family of RNA-specific deaminases play a tissue-specific regulatory role in modulating target specificity and enzyme activity, perhaps through interactions with other members.

The RNA-binding subunit. Using independent affinity enrichment strategies, two groups simultaneously identified a second component of the apoB RNA-editing holoenzyme (14, 15) that, together with apobec-1, is both necessary and sufficient to mediate C→U RNA editing in vitro. This novel 65-kDa protein, ACF (14) or apobec-1-stimulating protein (15), represents the regulatory subunit of the core enzyme; recombinant apobec-1 and ACF are sufficient to recapitulate efficient editing of a synthetic apoB RNA template in vitro (14, 15, 19). ACF mRNA is developmentally regulated

and ubiquitously expressed in normal human tissues (albeit at low levels) and is detectable in tumors from NF1 patients (13–15, 20). Apobec-1 and ACF together represent the minimal core of the holoenzyme, although their physiological stoichiometry within the native holoenzyme is not fully understood. Apobec-1, for instance, functions as a homodimer of about 27 kDa monomers and, along with ACF, constitutes the core of a holoenzyme complex that sediments at 27S and likely includes numerous other proteins (21, 22).

ACF contains three nonidentical RNA recognition motifs as well as a carboxyl-terminus arginine-glycine-rich region and a double-stranded RNA-binding domain (14, 15). The presence of these various RNA-binding motifs led to predictions, experimentally confirmed (14, 15, 19), that ACF would bind apoB RNA with high specificity. Recent studies have defined more precisely the functional domains of ACF and have revealed that the amino-terminal 380 residues of ACF contain the apoB RNA-binding domains as well as the domain responsible for physical interaction with apobec-1 (23). These domains overlap in a region spanning the second and third RNA recognition motif (23). The carboxyl-terminus arginine-glycine-rich domain, and also the double-stranded RNA-binding domain, each appear dispensable for C→U editing complementation (23), and more recent studies suggest that virtually complete in vitro complementation activity is retained in the amino-terminal 380 residues of ACF (V. Blanc et al., unpublished observations).

Formation of the C→U editing complex

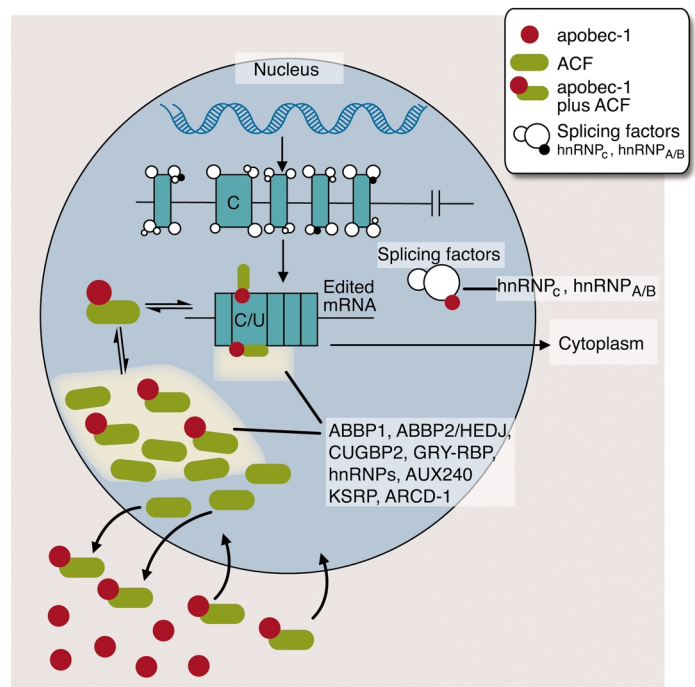
The data from in vitro, reconstituted systems indicate that ACF and apobec-1 interact physically with each other and with a synthetic apoB RNA template to coordinate C→U editing (14, 15, 19, 24) within the nucleus. Both ACF and apobec-1 are low-abundance proteins and localization of their endogenous products has proven challenging, but ACF has indeed been localized to the nucleus in transfected cells and, upon cotransfection with apobec-1, results in nuclear colocalization and redistribution of apobec-1 from a predominantly cytoplasmic location (19, 23, 24). These findings suggest that apobec-1 translocates into the nucleus, in association with ACF, raising the possibility of a shuttling function for one or both proteins.

Other proteins also interact with apobec-1 and/or apoB RNA. For instance, we and others have identified a homolog of ACF, GRY-RBP, within an S100 fraction of intestinal extracts that is enriched for C→U RNA editing activity (19, 25). GRY-RBP contains homologous RNA-binding domains to ACF, binds short apoB RNA templates, and also binds and colocalizes with both apobec-1 and ACF (19). Unlike ACF, GRY-RBP exhibits no C→U RNA editing complementation activity (19). Several other proteins, identified through two-hybrid or other genetic screens, have been demonstrated to interact with apobec-1 and to bind apoB RNA; many of these exhibit modulatory activity. Indeed, the original descriptions of apoB RNA-editing activity within 27S particles isolated from rat liver strongly implied the presence of a large multi-component complex (21). One such protein, CUGBP2, cofractionates in a complex enriched in C→U RNA-editing complementation activity and binds apoB RNA, colocalizing in the nucleus with apobec-1 and ACF (22). Like GRY-RBP, CUGBP2 inhibits C→U RNA editing of apoB, most plausibly by forming a regulatory complex with each of the essential components of the core enzyme complex, apobec-1, ACF, and apoB RNA (22).

It bears emphasis that apobec-1 itself binds apoB RNA, albeit with low affinity, and conservative mutations within apobec-1 that eliminate its RNA binding without compromising its cytidine deaminase activity abrogate C→U RNA editing of apoB (26, 27). Thus, the RNA-binding activity of apobec-1 may in itself constrain target specificity. Accordingly, a considerable focus of research activity is directed toward identifying *trans*-acting components whose presence within the apoB RNA-editing holoenzyme may regulate target RNA binding and restrict cytidine deamination of other transcripts. Remarkably, forced overexpression of apobec-1 in the livers of transgenic mice and rabbits is associated with promiscuous RNA editing of multiple cytidines,

Figure 1

Posttranscriptional C→U RNA editing. Following transcription, nascent pre-RNA associates with hnRNPs and splicing factors. Spliced mRNA is shown with the minimal core complex of the apoB RNA holoenzyme (apobec-1 plus ACF) in proximity to the targeted cytidine. Note that apobec-1 functions as a dimer and that the stoichiometry of binding (ACF/apobec-1) is currently unknown. The activity of the core enzyme complex is regulated through protein-protein interaction of both apobec-1 and ACF, each with a number of proteins that may together represent the larger holoenzyme complex. ACF is proposed to function in the translocation of apobec-1, as well, and is illustrated here shuttling between the nucleus and cytoplasm.



beyond the canonical site in apoB RNA, and leads to hepatic dysplasia and hepatocellular carcinoma (28, 29).

A final consideration is the extent to which protein components of the RNA-editing machinery may participate in other aspects of RNA processing. One component with such overlapping functions is KSRP, an approximately 75-kDa KH-type splicing regulatory protein. KSRP binds to a cluster of intronic regulatory elements downstream of exon N1 of *c-src* and participates in the assembly of a heterogeneous nuclear ribonuclear binding protein (hnRNP) complex that regulates alternative splicing in neural tissues (30). In a rat liver nuclear extract highly enriched in C→U editing activity, KSRP copurifies with ACF (15). KSRP also demonstrates high-affinity apoB RNA binding (15). Its role in C→U RNA editing awaits further study, but even independent of its role in RNA splicing and C→U RNA editing, KSRP was recently identified as a component of the mammalian exosome, a complex of exonucleases believed to regulate the turnover of short-lived, AU-containing mRNA species (31). The cumulative evidence thus points to a wide range of potential targets for KSRP and suggests that components of one macromolecular complex might participate in other higher-order complexes that act on newly synthesized mRNAs. In this regard, it is interesting to note that a number of hnRNPs have also been identified as apobec-1-interacting proteins, although their role in modulating nuclear mRNA metabolism in general and C→U RNA editing in particular is unresolved (Figure 1). The emerging information points to a growing complexity in the composition, regulation, and function of the C→U RNA-editing machinery. The ability of these various constituents to associate interchangeably with different multicomponent complexes within the nucleus and to discriminate target RNAs through subtle alterations in their secondary and tertiary structures implies that other functions will emerge through selective genetic manipulation in selected model systems.

Acknowledgments

The author gratefully acknowledges valuable discussions with all past and present members of the laboratory, in particular Shrikant Anant, Valerie Blanc, Jeff Henderson, and Debnath Mukhopadhyay. Support for the work cited from the author's laboratory was provided by grants HL-38180 and DK-56260 from the NIH.

1. Maas, S., and Rich, A. 2000. Changing genetic information through RNA editing. *Bioessays*. **22**:790–802.
2. Smith, H.C., and Sowden, M.P. 1996. Base-modification mRNA editing through deamination: the good, the bad and the unregulated. *Trends Genet.* **12**:418–424.
3. Chen, C.X., et al. 2000. A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA*. **6**:755–767.
4. Anant, S., and Davidson, N.O. 2001. Molecular mechanisms of apolipoprotein B mRNA editing. *Curr. Opin. Lipidol.* **12**:159–165.
5. Hersberger, M., Patarroyo-White, S., Arnold, K.S., and Innerarity, T.L. 1999. Phylogenetic analysis of the apolipoprotein B mRNA-editing region. Evidence for a secondary structure between the mooring sequence and the 3' efficiency element. *J. Biol. Chem.* **274**:34590–34597.
6. Hersberger, M., and Innerarity, T.L. 1998. Two efficiency elements flanking the editing site of cytidine 6666 in the apolipoprotein B mRNA support mooring-dependent editing. *J. Biol. Chem.* **273**:9435–9442.
7. Anant, S., and Davidson, N.O. 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.
8. Richardson, N., Navaratnam, N., and 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.
9. Sowden, M., Hamm, J.K., Spinelli, S., and Smith, H.C. 1996. Determinants involved in regulating the proportion of edited apolipoprotein B RNAs. *RNA*. **2**:274–288.
10. Sowden, M.P., and Smith, H.C. 2001. Commitment of apolipoprotein B RNA to the splicing pathway regulates cytidine-to-uridine editing-site utilization. *Biochem. J.* **359**:697–705.
11. Raitskin, O., Cho, D.S., Sperling, J., Nishikura, K., and Sperling, R. 2001. RNA editing activity is associated with splicing factors in hnRNP particles: the nuclear pre-mRNA processing machinery. *Proc. Natl. Acad. Sci. USA*. **98**:6571–6576.
12. Skuse, G.R., Cappione, A.J., Sowden, M., Metheny, L.J., and Smith, H.C. 1996. The neurofibromatosis type I messenger RNA undergoes base-modification RNA editing. *Nucleic Acids Res.* **24**:478–485.
13. Mukhopadhyay, D., et al. 2002. C→U editing of neurofibromatosis 1 mRNA occurs in tumors that express both the type II transcript and apobec-1, the catalytic subunit of the apolipoprotein B mRNA-editing enzyme. *Am. J. Hum. Genet.* **70**:38–50.
14. Mehta, A., Kinter, M.T., Sherman, N.E., and Driscoll, D.M. 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.
15. Lellek, H., et al. 2000. Purification and molecular cloning of a novel essential component of the apolipoprotein B mRNA editing enzyme-complex. *J. Biol. Chem.* **275**:19848–19856.
16. Muramatsu, M., et al. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*. **102**:553–563.
17. Liao, W., et al. 1999. APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochem. Biophys. Res. Commun.* **260**:398–404.
18. Anant, S., et al. 2001. ARCD-1, an apobec-1-related cytidine deaminase, exerts a dominant negative effect on C to U RNA editing. *Am. J. Physiol. Cell Physiol.* **281**:C1904–C1916.
19. Blanc, V., et al. 2001. Identification of GRY-RBP as an apolipoprotein B RNA-binding protein that interacts with both apobec-1 and apobec-1 complementation factor to modulate C to U editing. *J. Biol. Chem.* **276**:10272–10283.
20. Henderson, J.O., Blanc, V., and Davidson, N.O. 2001. Isolation, characterization and developmental regulation of the human apobec-1 complementation factor (ACF) gene. *Biochim. Biophys. Acta.* **1522**:22–30.
21. Harris, S.G., et al. 1993. Extract-specific heterogeneity in high-order complexes containing apolipoprotein B mRNA editing activity and RNA-binding proteins. *J. Biol. Chem.* **268**:7382–7392.
22. Anant, S., et al. 2001. Novel role for RNA-binding protein CUGBP2 in mammalian RNA editing. CUGBP2 modulates C to U editing of apolipoprotein B mRNA by interacting with apobec-1 and ACF, the apobec-1 complementation factor. *J. Biol. Chem.* **276**:47338–47351.
23. Blanc, V., Henderson, J.O., Kennedy, S., and Davidson, N.O. 2001. Mutagenesis of apobec-1 complementation factor reveals distinct domains that modulate RNA binding, protein-protein interaction with apobec-1, and complementation of C to U RNA-editing activity. *J. Biol. Chem.* **276**:46386–46393.
24. Yang, Y., Sowden, M.P., and Smith, H.C. 2000. Induction of cytidine to uridine editing on cytoplasmic apolipoprotein B mRNA by overexpressing APOBEC-1. *J. Biol. Chem.* **275**:22663–22669.
25. Lau, P.P., Chang, B.H., and Chan, L. 2001. Two-hybrid cloning identifies an RNA-binding protein, GRY-RBP, as a component of apobec-1 editosome. *Biochem. Biophys. Res. Commun.* **282**:977–983.
26. Navaratnam, N., et al. 1998. Escherichia coli cytidine deaminase provides a molecular model for ApoB RNA editing and a mechanism for RNA substrate recognition. *J. Mol. Biol.* **275**:695–714.
27. MacGinnitie, A.J., Anant, S., and Davidson, N.O. 1995. Mutagenesis of apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, reveals distinct domains that mediate cytosine nucleoside deaminase, RNA binding, and RNA editing activity. *J. Biol. Chem.* **270**:14768–14775.
28. Yamanaka, S., Poksay, K.S., Arnold, K.S., and Innerarity, T.L. 1997. A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. *Genes Dev.* **11**:321–333.
29. Yamanaka, S., et al. 1995. Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc. Natl. Acad. Sci. USA*. **92**:8483–8487.
30. Markovtsov, V., et al. 2000. Cooperative assembly of an hnRNP complex induced by a tissue-specific homolog of polypyrimidine tract binding protein. *Mol. Cell. Biol.* **20**:7463–7479.
31. Chen, C.Y., et al. 2001. AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell*. **107**:451–464.