Indian Journal of Biochemistry & Biophysics
Vol. 38, June 2001, pp. 129-134

Minireview

Functional diversity of hnRNP proteins

O P Singh*
Department of Zoology, Arunachal University, Rono Hills, Itanagar 791111 (Arunachal Pradesh), India

Received 20 July 2000; revised and accepted 9 November 2000

The hnRNP proteins bind to pre-mRNA concomitant with transcription and form ribonucleoprotein complex essential for post-transcriptional events. More than 20 different groups of hnRNP proteins are known in humans. The number seems to be similar in other organisms too. The hnRNP proteins possess different types of RNA binding domains which directly interact with RNA, most likely in a sequence specific manner. The functions of hnRNP proteins range from mRNA packaging and transport, to mRNA splicing and silencing. Further, indirectly, their role has also been implicated in oogenesis. Hence, hnRNP proteins play diversified roles both in nucleus and cytoplasm, and regulate the gene expression at various levels.

The precursors of messenger RNA (pre-mRNA) represent a small fraction of heterogeneous nuclear RNAs (hnRNA) synthesized by RNA polymerase II in eukaryotic nuclei. Concomitant with their synthesis, the pre-mRNA associate with a group of proteins referred to as heterogenous nuclear ribonucleoprotein (hnRNP) proteins. The hnRNP proteins fold and package the pre-mRNA to form pre-mRNP complexes[1,2]. These complexes appear in the nucleus first as perichromatin fibrils and then perichromatin granules[3].

The formation of mRNP complexes is essential for the accomplishment of post-transcriptional events and transportation of mature mRNA from the nucleus to the outer cytoplasm of the cell[4]. The protein moiety of pre-mRNP complexes constitutes mainly of hnRNP proteins. However, some other proteins such as small nuclear ribonucleoproteins (snRNPs) and splicing factors (SF) have also been found to be associated directly or indirectly, with the processing of pre-mRNA present in the pre-mRNP complex[5-9].

Pre-mRNP complexes are dynamic structures. At different stages of the biogenesis of mRNA, different sets of proteins play specific roles besides providing a general function of mRNA packaging and stability, after which they dissociate to make room for interaction with other proteins (Fig. 1). Hence, mRNA is always found associated with some or other proteins from its inception and biogenesis in the nucleus to intra-nuclear transport, nucleocytoplasmic translocation through nuclear pore complex (NPC) and finally, its translation and degradation in the cytoplasm.

Earlier, the functions of hnRNP proteins were implicated in folding and packaging of the pre-mRNA in order to protect it from degradation from nuclear ribonucleases[1]. However, recent findings suggest highly diversified roles of hnRNP proteins ranging from transcription and processing of pre-mRNA to nucleocytoplasmic transport and translation of mRNA. Consequently, hnRNP proteins seem to play an important role in gene regulation both at pre- and post-transcriptional levels.

General features of hnRNP proteins

The hnRNP proteins are among the most abundant proteins in the nucleus. In humans more than 20 different groups of hnRNP proteins, designated from A1 to U in the molecular range of 34-120 kD, have been identified in *HeLa* cells[1,2,4]. Very similar two dimensional (2D) gel electrophoresis patterns of hnRNP proteins have also been reported in other vertebrates such as rodents, avians and amphibians[10-15]. In invertebrates, comparatively less number of hnRNP proteins are characterized, but the total number seems to be comparable to that of vertebrates[16-18].

Initially, all hnRNP proteins were assumed to be confined to the nucleus[19,20]. However, recently by immunofluorescence microscopy some of them have been localized both in the nucleus and cytoplasm, and

* Tel: 91-0360-277571; Fax: 91-0360-277317
E-mail: ops57@hotmail.com

were found to shuttle between the two compartments of the cell. For instance, human hnRNP proteins C and U remain confined to the nucleus, whereas proteins of A and B groups shuttle between nucleus and cytoplasm[21]. Electron microscopic study revealed that hrp36 of *Chironomus* binds to Balbiani Ring (BR) pre-mRNA during transcription and accompanies mRNA till it reaches the cytoplasm. After being released from the mRNA in the cell's cytoplasm, this protein returns to the nucleus[22]. Such proteins are referred to as 'shuttling' hnRNP proteins. Immunoblotting analysis of different hnRNP proteins has revealed a differential distribution of these proteins among various tissues, and they were also found to occur in different proportions[23].

Different hnRNP proteins bind directly to pre-mRNA through their highly conserved RNA binding domains (RBDs) or modules or motifs and interact

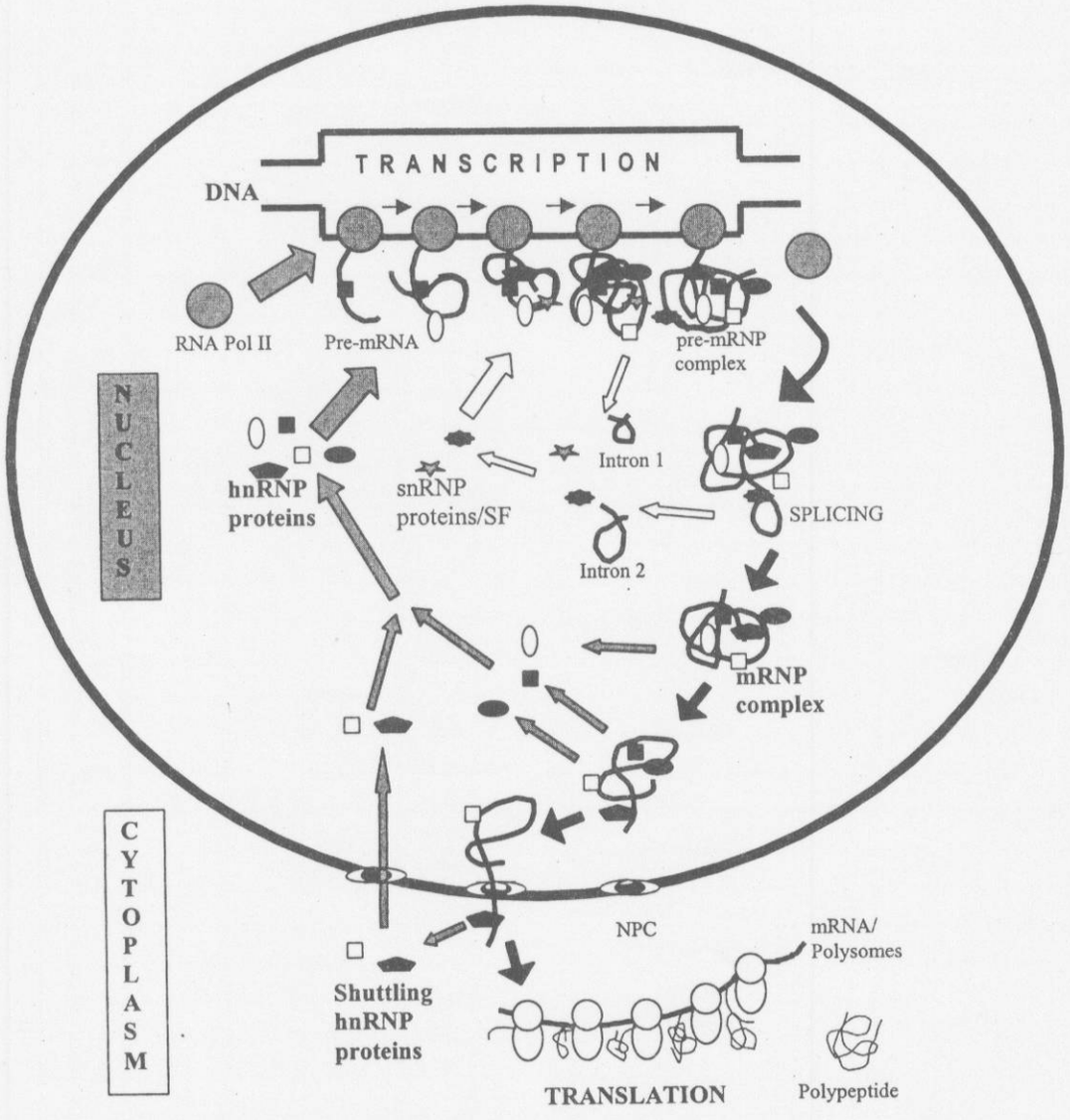

Fig. 1 — Transcription, processing and transport of RNA polymerase II transcript in an eukaryotic cell. [The transcript (pre-mRNA) concomitantly binds with hnRNP and snRNP/SF proteins and forms RNP complex. The processing, such as splicing occurs within the RNP complex. The splicing of pre-mRNA begins cotranscriptionally and continues after the transcription is over. The removal of Intron 1, as shown in figure is the result of cotranscriptional splicing, whereas Intron2 is removed after the release of transcript in the nucleoplasm. Subsequently, snRNP/SF proteins are also released. Simultaneously, the mRNP complex moves towards nuclear periphery and then transported to the cytoplasm through the NPC. Some of the hnRNP proteins are shed in the nucleus before the transport of mRNA, whereas some other (shuttling hnRNP proteins) accompany the mRNA till it reaches the cytoplasm and then return to the nucleus. In this way they keep shuttling between nucleus and cytoplasm]

with other proteins of the RNP complex with their auxiliary domains[24-26]. In general, hnRNP proteins possess one or more RNA binding motifs for RNA-protein interactions, and at least one auxiliary domain for protein-protein interaction. The following four major types of RNA binding motifs have been identified in different hnRNP proteins. They may be found in one or multiple copies in different hnRNP proteins. Some of the major hnRNP proteins along with their structural motifs are listed in Table 1.

Table 1 — Characteristic features of some major hnRNP proteins (Refer to 4,49)

Organisms	Proteins	Molecular mass (kD)	Structural motifs
Human	A1	34	2xRBD, RGG
	A2/B1	36 & 38	2xRBD, RGG
	B2	39	-
	C1/C2	41& 43	RBD – AspGlu
	D	44 - 48	-
	E	36 - 43	RBDs
	G	43	RBDs, RGG
	F/H	53 & 56	RBDs
	I	59	4xRBD
	K/J	62 & 68	KH motif
	L	68	4xRBD
	M	68	4xRBD
	N	70	-
	P	72	-
	Q	76 - 77	-
	R	82	-
	S	105	-
	T	113	-
	U	120	RGG Box
Drosophila	hrp 36	36	2xRBD, RGG
	hrp 40	40	2xRBD, RGG
	hrp 48	48	2xRBD
	hrb 87F	87	2xRBD, RGG
	hrb 98DE	98	2xRBD, RGG

RNP motif

The most common RNA binding motif is the RNP consensus sequence, the RNA binding domain (cs-RBD), also known as RNA recognition motif (RRM) or RNP motif. The cs-RBD/RRM/RNP motif is a 90 amino acid sequence containing two short highly conserved sequences of 8 and 6 amino acids referred to as RNP1 and RNP2, respectively. RNP1 and RNP2 are separated by about 30 amino acid sequence. These short peptides form a part of the ß sheet in the secondary structure of an hnRNP protein. It is believed that RNP1 and RNP2 sequences interact directly with RNA. The structure of RNP motif has been found to be highly conserved in different groups of hnRNP proteins. For example, the RNP motif of hnRNP C protein of humans possesses more than 40% similarity to that of hnRNP Nab3p protein of yeast[27-29].

RGG box

Another type of RNA binding domain comprises of a 26 amino acid long sequence containing several Arg-Gly-Gly (RGG) tripeptide repeats with several interspersed aromatic residues. RGG box is found either alone or in combination with other RNA binding motifs in different hnRNP proteins. For instance, in hnRNP U protein the RGG box is the only domain responsible for RNA binding activity whereas, in hnRNP A1 protein the RGG box is found in combination with RRM, and most likely both the modules (i.e. RGG box and RRM) mediate protein-RNA interactions in this protein[4,30].

KH motif

It is a sequence of about 45 amino acids, containing a conserved core sequence namely, Val-Ile-Gly-X-X-Gly-X-X-Ile (where X may be any amino acid) flanked by a few interspersed conserved residues. The K homology (KH) motif was first identified in hnRNP K protein[31]. Like other RNA binding motifs, the KH motif is found in one or multiple copies. In hnRNP K protein, three KH domains are found.

The Zinc finger motif

This is another RNA binding motif reported to be present in NAB2 hnRNP protein of yeast[32].

Functional diversity of hnRNP proteins

Role in alternative splicing

Alternative splicing is a major mechanism for controlling the expression of cellular and viral genes. The splicing factors, such as SF2/ASF and other members of the SR (Serine Arginine) protein family play an important role in the general splicing *in vitro,* and also help in regulating alternative splicing by promoting the use of proximal 5' splice site[33,34]. In contrast to this, hnRNP A1 promotes the use of distal 5' splice sites[35]. Consequently, the antagonizing activities of SR proteins and hnRNP A1 are the key determinants of alternative 5' splice site selection *in vitro.*

To know the effect of changes in relative amount of SF2/ASF and hnRNP A1 *in vivo*, the SF2/ASF or hnRNP A1 complementary DNA (cDNA) were transiently over expressed in *HeLa* cells, and the effect on alternative splicing of several cotransfected reporter genes was measured[36]. It was found that increased expression of SF2/ASF activated proximal 5' splice sites, promoted inclusion of a neuron-specific exon and prevented abnormal exon skipping in the case of rat clathrin light chain B. Whereas, over expression of hnRNP A1 activated distal 5' splice sites. Therefore, it was concluded that variation in the intracellular levels of antagonistic splicing factors influence different modes of alternative splicing *in vivo* and may be a natural mechanism for tissue-specific or developmental regulation of gene expression.

The generally expressed 53 kD hnRNP F protein has also been found to be involved in a neural-specific pre-mRNA splicing event by positive control of splice site choice[37]. In mammalian nervous system, alternative splicing controls the production of a number of neuron-specific proteins, important for neuronal development and function. For example, in mouse the *src* gene produces different products in neurons and in non-neuronal cells by alternative splicing. The *src* gene contains an 18-nucleotide exon, N1, that is inserted between constitutive exons 3 and 4 in neurons (producing n-*src*) but skipped in other cells (producing c-*src*) [38,39]. Mutational analysis identified an intronic sequence between 38 and 70 nucleotides downstream of N1 5' splice site as an essential element to yield normal levels of N1 splicing *in vivo*. Min *et al.* [37] identified a neural specific protein complex that binds very specifically to this downstream regulatory sequence, and the hnRNP F protein is a component of that complex. It was also demonstrated that hnRNP F protein together with other neural-specific factors are critical to N1 exon splicing *in vitro*. Further, antibody inhibition experiments indicate that the hnRNP F protein is a functional part of this complex.

Role in mRNA silencing and regulation of translation

Various hnRNP proteins play important functions not only in the nucleus but also in the cytoplasm. The cytoplasmic function of hnRNP K and hnRNP E1 have been demonstrated by Ostareck *et al.*[40]. They showed that hnRNP K and hnRNP E1 function as specific cytoplasmic regulatory proteins of LOX mRNA translation.

Erythroid 15-lipoxygenase (LOX) is a key enzyme involved in erythroid cell differentiation. It can attack intact phospholipids and seems to participate in the breakdown of internal membranes (such as mitochondrial membranes) during the late stages of reticulocyte maturation[41]. The LOX mRNA is expressed but translationally silenced in early erythroid cells in the bone marrow and the peripheral blood. Study by Ostareck *et al*[40] revealed that LOX mRNA is specifically associated with hnRNP K and hnRNP E1 in translationally inactive mRNPs of peripheral blood reticulocytes and is involved in translational silencing of LOX mRNA during erythroid differentiation. A differentiation control element in its 3' untranslated region confers translational silencing until late stage erythropoiesis by associating with hnRNP K and hnRNP E1. Further, transfection of hnRNP K and hnRNP E1 into *HeLa* cells specifically silenced the translation of reporter mRNAs bearing a differentiation control element in their 3' untranslated region. In a reconstituted cell-free translation system, addition of recombinant hnRNP K and hnRNP E1 recapitulates this regulation via a specific inhibition of 80S ribosomes assembly on LOX mRNA. It was proposed that both proteins control cap-dependent and internal ribosome entry site-mediated translation by binding to differentiation control elements.

Role as transactivator

The role of hnRNP K has been implicated in trans-activation of a variety of RNA polymerase II promoters. The hnRNP K protein possesses selective single stranded DNA (ssDNA)[42] and RNA[43] binding activities. The deletion analysis showed that hnRNP K possesses several non-overlapping, DNA binding domains, each capable of specific binding with pyrimidine-rich strand of CT element found in the human *c-myc* gene[42]. However, the RNA binding activity of this protein depends upon the KH domains as well as the arginine/glycine-rich regions of the protein[43].

The hnRNP K protein was shown to stimulate chloramphenicol acetyltransferase (CAT) gene expression[41, 44]. To investigate the trans-activation function of the hnRNP K protein, Lee *et al.*[43] carried out transient transfection experiments using several chimeric reporter gene constructs. It was observed

that trans-activation function depends on the sequences of hnRNP K that are also necessary for RNA binding. However, the RNA binding motifs are not sufficient for trans-activation. Further, it has been clearly demonstrated that hnRNP K protein increased the steady level of the reporter mRNA without altering its decay rate. In fact, hnRNP K protein trans-activates the reporter genes by increasing the level of transcription.

Role in oogenesis

The hrp40 proteins of *Drosophila melanogaster* are counterparts of hnRNP A/B proteins of vertebrates, and have been reported to play essential role in oogenesis[45]. The hrp40 is encoded by the squid gene which is required for dorsoventral axis formation during oogenesis. Females homozygous for the squid1 allele (first mutant allele of the squid gene identified in screening), which specifically disrupts oogenesis, lay dorsalized eggs that do not survive to adulthood. However, complete deletion of the squid gene results in lethality indicating that hrp40 plays an essential function in non-ovarian tissues also[46].

Immunolocalization of hrp40 in wild-type and squid1 mutant-type developing egg chambers has revealed that hrp40 is present in nurse cells, oocyte and follicle cells of wild-type egg chambers. In contrast, hrp40 is present only in the follicle cells of squid1 egg chambers as it was absent from its nurse cells and oocyte. These findings reveal the dynamic patterns of expression and localization of hrp40 during development and provide evidence for an essential role of hnRNP proteins in oogenesis[45].

Role in pre-mRNA packaging

In general, different hnRNP proteins bind to, fold and package the hnRNA or pre-mRNA to form hnRNP complexes or particles. The formation of hnRNP complexes protects the pre-mRNA from degradation by the action of ribonucleases. At the same time, formation of hnRNA-hnRNP protein complexes is essential for proper processing and transport of the associated pre-mRNA.

Initially, it was believed that hnRNP proteins bind to pre-mRNA indiscriminately, similar to histones that bind to DNA in a regular fashion[4]. Later on, it was found that hnRNP proteins possess differential affinity for specific RNA sequences. Moreover, the hnRNP complexes are not static structures as they have been found to be dynamic in the sense that their composition changes at different stages of transcript processing. As a result, a concept of transcript-specific hnRNP complexes has emerged. In a preliminary study, the distribution of two hnRNP proteins, hrp36 and hrp45 was analysed on polytene chromosome puffs, which represent sites of nascent hnRNP complexes. This study revealed that all puffs do not contain both the proteins. The localization of these two proteins differs both in terms of their quantity and quality on different puffs. Thus, it was concluded that the hnRNP proteins do not bind to all pre-mRNA molecules randomly, rather each transcript is probably associated with a specific subset of hnRNP proteins[18].

Role in mRNA transport

The transport of mRNA from its site of synthesis in the nucleus to the site of translation in the cytoplasm can be divided, at least in two steps: (¡) Intranuclear transport from the gene to the nuclear periphery (ii) Nucleocytoplasmic translocation through the NPC. As demonstrated in the case of Balbiani Ring (BR) transcripts, concomitant with transcription the pre-mRNA associates with different hnRNP proteins and forms RNP fibre, which is subsequently folded into compact RNP particles[47]. The completed pre-mRNA/mRNA is delivered in the nucleoplasm as RNP complex which travels to reach the nuclear periphery, most likely in a random fashion[48]. Although the specific role of hnRNP proteins during this stage of movement is not well understood, there is now a direct evidence that RNA moves as an RNA-protein complex[22].

The shuttling hnRNP proteins (proteins which shuttle rapidly between the nucleus and cytoplasm) play a direct role in nucleocytoplasmic translocation of mature mRNA through the NPC. The hnRNP protein A1 of humans and hrp36 of *Chironomus* accompany mRNA up to the cytoplasm and have a role in mRNA transport[22]. Recently, a transport signal, M9 has been identified in hnRNP A1 protein. The M9 is a 38-amino acid sequence that acts as Nuclear Export Signal (NES). It is due to this sequence that the hnRNP A1 protein is able to act as a carrier of the accompanying mRNA from nucleus to the cytoplasm through the NPC. Several point mutations in M9 sequence have been found which disrupt the function of hnRNP A1 protein as a carrier of mRNA[21].

Besides, many other functions of hnRNP proteins have been proposed, such as presenting pre-mRNA in proper way for processing, retention of non-spliced

mRNA in the nucleus and anchoring the mRNA to the nuclear matrix. It seems that many more functions are yet to be revealed. Thus, further studies are required to understand the comprehensive role of hnRNP proteins in regulation of gene expression and its implications.

References

1 Dreyfuss G (1986) *Annu Rev Cell Biol* 2, 459-498
2 Dreyfuss G, Swanson M S & Pinol-Roma S (1988) *Trends Biochem Sci* 13, 86-91
3 Fakan S (1994) *Trends Cell Biol* 4, 86-90
4 Dreyfuss G, Matunis M J, Pinol-Roma S & Burd C G (1993) *Annu Rev Biochem* 62, 289-321
5 Roth M B, Zahler A M & Stolk J A (1991) *J Cell Biol* 115, 587-596
6 Amero S A, Raychaudhury G, Cass C L, van Venrooij W J, Habets W J, Krainer A R & Beyer A L (1992) *Proc Natl Acad Sci USA* 89, 8409-8413
7 Matunis E L, Matunis, M J & Dreyfuss G (1993) *J Cell Biol* 121, 219-228
8 Champlin D T & Lis J T (1994) *Mol Biol Cell* 5, 71-79
9 Alzhanova-Ericsson A T, Sun X, Visa N, Kiseleva E, Wurtz T & Daneholt B (1996) *Genes & Dev* 10, 2881-2893
10 Beyer A L, Christensen M E, Walker B W & LeStourgeon W M (1977) *Cell* 11, 127-138
11 Karn J, Vidali G, Boffa L C & Allfrey V G (1977) *J Biol Chem,* 252, 7307-7322
12 Brunel C & Lelay M N (1979) *Eur J Biochem* 99, 273-283
13 Choi Y D & Dreyfuss G (1984) *J Cell Biol* 99, 1997-2004
14 Wilk H E, Werr H, Friedrich D, Kiltz H H & Schafer K P (1985) *Eur J Biochem* 46, 71-81
15 Leser G P, Wscara-Wilke J & Martin T E (1984) *J Biol Chem* 259, 1827-1833
16 Haynes S R, Johnson D, Raychaudhuri G & Beyer A L (1991) *Nucleic Acids Res* 19, 25-31
17 Matunis M J, Matunis, E L & Dreyfuss G (1992) *J Cell Biol* 116, 245-255
18 Wurtz T, Kiseleva E, Nacheva G, Alzhanova-Ericsson A, Rosen A & Daneholt B (1996) *Mol Cell Biol* 16, 1425-1435
19 Jones K E, Okamura C S & Martin T E (1980) *J Cell Biol* 86, 235-243
20 Pinol-Roma S, Swanson M S, Gall J G & Dreyfuss G (1989) *J Cell Biol* 109, 25752587
21 Pinol-Roma S & Dreyfuss G (1992) *Nature, (London)* 355, 730-732
22 Visa N, Alzhanova-Ericsson A, Sun X, Kiseleva E, Bjorkroth B, Wurtz T & Daneholt B (1996) *Cell* 84, 253-264
23 Faura M, Renau-Piqueras J, Bachs O & Bosser R (1995) *Biochem Biophys Res Commun* 217, 554-560
24 Cobianchi F, Karpel R L, Williams K R, Notario V & Wilson S H (1988) *J Biol Chem* 263, 1063-1071
25 Wilusz J, Feig D I & Shenk T (1988) *Mol Cell Biol* 8, 4477-4483
26 Kim J H, Hahn B, Kim Y K, Choi M & Jang S K (2000) *J Mol Biol* 298, 395-405
27 Merrill B M, Stone K L, Cobianchi F, Wilson S M & Williams K R (1988) *J Biol Chem* 263, 3307-3313
28 Wittekind M, Gorlach M, Friedrichs M & Dreyfuss G (1992) *Biochemistry* 31, 6254-6265
29 Nagai K, Oubridge C, Jessen T H, Li J & Evans P R (1990) *Nature, (London)* 346, 515-520
30 Kiledjian M & Dreyfuss G (1992) *EMBO J* 11, 2655-2664
31 Siomi H, Matunis M J, Michael W M & Dreyfuss G (1993) *Nucleic Acids Res* 21, 1193-1198
32 Matunis M J, Michael W M & Dreyfuss G (1992) *Mole Cell Biol* 12, 257-269
33 Krainer A R, Conway G C & Kozak D (1990) *Cell* 62, 35-42
34 Zahler A M, Lane W S, Stolk J A & Roth M B (1992) *Genes & Dev* 6, 837-847
35 Mayeda A & Krainer A R (1992) *Cell* 68, 365-375
36 Caceres J F, Stamm S, Helfman & Krainer A R (1994) *Science* 265, 1706-1709
37 Min H, Chan R C & Black D L (1995) *Genes & Dev* 9, 2659-2671
38 Levy J B, Dorai T, Wang L H & Brugge J S (1987) *Mol Cell Biol* 7, 4142-4145
39 Martinez R, Mathey P B, Bernards A & Baltimore D (1987) *Science,* 237, 411-415
40 Ostareck D H, Ostareck-Lederer A, Wilm M, Thiele B J, Mann M & Hentze M W (1997) *Cell* 89, 597-606
41 Rapoport S M & Schewe T (1986) *Biochim Biophys Acta* 864, 471-477
42 Tomonaga T & Levens D (1995) *J Biol Chem* 270, 4875-4881
43 Lee M H, Mori S & Raychaudhury P (1996) *J Biol Chem* 271, 3420-3427
44 Takimoto M, Tomonaga T, Matunis M, Avigan M, Krutzsch H, Dreyfuss G & Levens D (1993) *J Biol Chem* 268, 18249-18258
45 Matunis E L, Kelley R & Dreyfuss G (1994) *Proc Natl Acad Sci USA* 91, 2781-2784
46 Kelley R L (1993) *Genes & Dev* 7, 948-960
47 Daneholt B (1997) *Cell* 88, 585-588
48 Singh O P, Bjorkroth B, Masich S, Wieslander L & Daneholt B (1999) *Exp Cell Res* 251, 135-146
49 Swanson, M S (1995) in *Pre-mRNA processing* (Lamond, A I, ed), pp. 17-33, R G Landes Co- Springer-Verlag, Berlin.