Eur. J. Biochem. **267**, 4002–4010 (2000) © FEBS 2000

A *trans*-acting factor, isolated by the three-hybrid system, that influences alternative splicing of the amyloid precursor protein minigene

Andrej Poleev, Annette Hartmann and Stefan Stamm

Max-Planck-Institute for Neurobiology, Munich, Germany

Two clones were isolated in a three-hybrid screen of a rat fetal brain P5 cDNA library with an intronic splicing enhancer of the amyloid precursor protein (APP) gene as RNA bait. These clones represent the rat homologues of the previously described genes CUG-binding protein (CUG-BP) and Siah-binding protein (Siah-BP). Both interact in a sequence-specific manner with the RNA bait used for library screening as well as with the CUG repeat. In contrast, no interactions were observed in the three-hybrid assay with other baits tested. In two-hybrid assays, Siah-BP interacts with U2AF65 as well as with itself. EWS, an RGG-type RNA-binding protein associated with Ewing sarcoma, was identified as an interacting partner for the CUG-BP homologue in a two-hybrid assay for protein–protein interactions performed with various factors involved in RNA metabolism. Splicing assays performed by RT-PCR from cells cotransfected with certain cDNAs and an APP minigene, used as a reporter, indicate exclusion of exon 8 if the CUG-BP homologue is present. We conclude that clone AF169013 and its counterpart in human CUG-BP could be the *trans*-acting factors that interact with the splicing enhancer downstream of exon 8, and in this way influence alternative splicing of the APP minigene.

Keywords: alternative splicing; amyloid precursor protein (APP); CUG-binding protein (CUG-BP); three-hybrid system.

Alternative usage of splice sites requires *cis*-acting elements, such as 5′ and 3′ consensus sequences for the binding of the basic splicing machinery, as well as gene-specific elements for binding of regulatory *trans*-acting factors [1,2]. Consequently, specific exon or intron sequences capable of strongly stimulating or repressing splice-site usage have been designated splicing enhancers or silencers, respectively. Although several proteins have been identified in *Drosophila* that act in this way [3–6], the regulatory molecules that direct alternative splicing in vertebrates remain elusive. Furthemore, reports describing such causative interactions in vertebrates have been from simple biochemical analysis and not genetic functional assays. One of the most extensively clarified cases reported so far describes biochemical interactions within the c-*src* gene. Analysis of mouse c-*src* gene indicates that the neuronal inclusion of the N1 exon requires an intronic splicing-enhancer sequence between +17 and +142 nucleotides downstream of the exon [7]. The central, most conserved, portion of this enhancer sequence (nucleotides 38–70) is called the downstream control sequence [8], which binds to a complex of regulatory proteins that is important in allowing 18-nucleotide exon N1 splicing. A protein containing four K homology RNA-binding domains is shown to be the key component in the co-operative assembly of the multiprotein complex at the splicing enhancer [9]. The latest report demonstrates by genetic and biochemical methods that Nova-1 regulates in a sequence-specific manner alternative splicing of two inhibitory receptor pre-mRNAs, glycine receptor α2 and $GABA_A$ [10].

By the use of splicing assays and minigenes, sequences regulating neuron-specific splicing have been found in vertebrates. Several forms of the amyloid precursor protein (APP) mRNA can be generated by alternative splicing. Of these, the APP695 mRNA, in which exons 7 and 8 are excluded, is expressed specifically in neurons, suggesting that this alternative splicing is regulated in a neuron-specific manner [11,12]. Mutational analysis of an intron region flanking exon 8 shows two distinct *cis*-acting elements that modulate the splicing of this exon of the APP gene in a sequence-dependent manner. Both elements are located upstream of the branchpoint sequence, the first at −42 to −47 and the second at −36 to −38 upstream of exon 8 [13]. Agrin, an extracellular matrix protein expressed in both neurons and muscle cells, is differentially spliced in neuronal tissues, and inclusion of exon Y/28 is regulated by a sequence lying 15–68 nucleotides downstream of the alternatively spliced exon [14]. A potential splicing enhancer also exists in exon EN of the CLB gene [15].

Several methods have been established for finding proteins that bind to given RNA sequences, such as standard molecular methods, including RNAse CV1 footprinting assays, gel-shift assays, conventional biochemical purification (reviewed in [16]), and some new methods, such as systematic evolution of ligands by expontential enrichment (SELEX) [17–21], a new affinity-chromatography strategy called StreptoTag [22], and one that uses a reporter gene system [23]. Recently, the three-hybrid system has been established, in which RNA-binding proteins can be isolated by screening of expression libraries in yeast with a given RNA bait [24,25]. This system has been used succesfully to isolate a number of mRNA-binding proteins, but so far has not identified any factors that influence differential

Correspondence to A. Poleev, GeneCraft, Tresckowstrasse 10, 48163 Münster, Germany. Fax: + 49 251970 9704, E-mail: andrejpoleev@yahoo.com

Abbrevations: APP, amyloid precursor protein; CUG-BP, CUG-binding protein; Siah-BP, Siah-binding protein; CLB, clathrin light chain B gene; SELEX, systematic evolution of ligands by expontential enrichment; RRM, RNA-recognition motif.

Note: the nucleotide sequences reported in this paper have been submitted to GenBank with accession numbers AF165892and AF169013.

(Received 8 February 2000; revised 25 April 2000; accepted 27 April 2000)

splicing [26–30]. In the present study, we have examined various sequences described as splicing enhancers in order to clone the corresponding RNA-binding proteins. Four screens were performed with three different RNA baits expressing such sequences. Two RNA-binding proteins were isolated in a screen with an intronic splicing enhancer from the APP gene. Both clones interact with the RNA bait in a sequence-specific manner. One of them, homologous to the previously described CUG-binding protein (CUG-BP), influences splicing of an APP minigene in transfection experiments. This is also shown for the human CUG-BP gene. We conclude that human CUG-BP as well as its counterparts could be involved in post-transcriptional regulation of the APP gene as the *trans*-acting factors.

EXPERIMENTAL PROCEDURES

Plasmids

Original plasmids pIIIA MS2-1, pIII/IRE MS2, pAD IRP, used to establish and perform the three-hybrid screens, were a gift from M. Wickens and S. Fields. pIIIA MS2-APP contains six copies of the sequence 5′-ATGTTTCTCTTT-3′ derived from the APP gene, cloned into the *Xma*I site in sense orientation, and pIIIA MS2-antiAPP contains three copies of the same sequence in an inverted orientation. pIIIA MS2-CUG and pIIIA MS2-antiCUG contain eight copies of the CUG trinucleotide repeat in sense and antisense orientations, respectively. pIIIA MS2-EN and pIIIA MS2-agrin contain sequences of exon EN of clathrin light chain B (5′-TCGCTGAGTGATACCATTGGC-TATGTGG-3′) and an intronic splicing enhancer downstream of exon Y of the agrin gene (5′-TTACCCAACTCAACTCACC-ATCTTTGTAGCCATTCCTAGAGTAGCCCTTTCCCA-3′) in sense orientations.

Construct RSVA689, containing the APP minigene, was a gift from Y. Sakaki. pEGFP-C2 vector (Clontech) was used to clone cDNAs (human CUG-BP, clones AF169013 and AF165892) in-frame with green fluorescent protein (GFP). These expression vectors were used in transfection experiments. Plasmids used in the two-hybrid system contain the full coding regions of corresponding genes, cloned into GBT9 (bait) or pGAD424 (prey) vectors.

The cDNA library was constructed by Stratagene (catalogue number 838401) and contains about 3.7×10^6 primary plaques. Oligo(dT)-primed cDNAs were isolated from postnatal P5 rat brain mRNA and cloned into *Eco*RI and *Xho*I sites of the HybriZAP vector. The library was then converted by *in vivo* mass excision to a pAD-GAL4 target plasmid library and replicated in a λ-resistant *Escherichia coli* strain, according to the manufacturer's instructions.

Cell transfections and RNA isolation

The calcium phosphate method of transfection was used in all experiments. A 1 μg portion of each plasmid was used to transfect the HEK293 cell line grown in six-well plates. To keep an equal quantity of DNA, the C2EFGP vector was used as carrier. One day later the cells were lysed and total RNA was prepared using an RNeasy mini kit (Qiagen) as described by the manufacturer. The experiment was repeated twice.

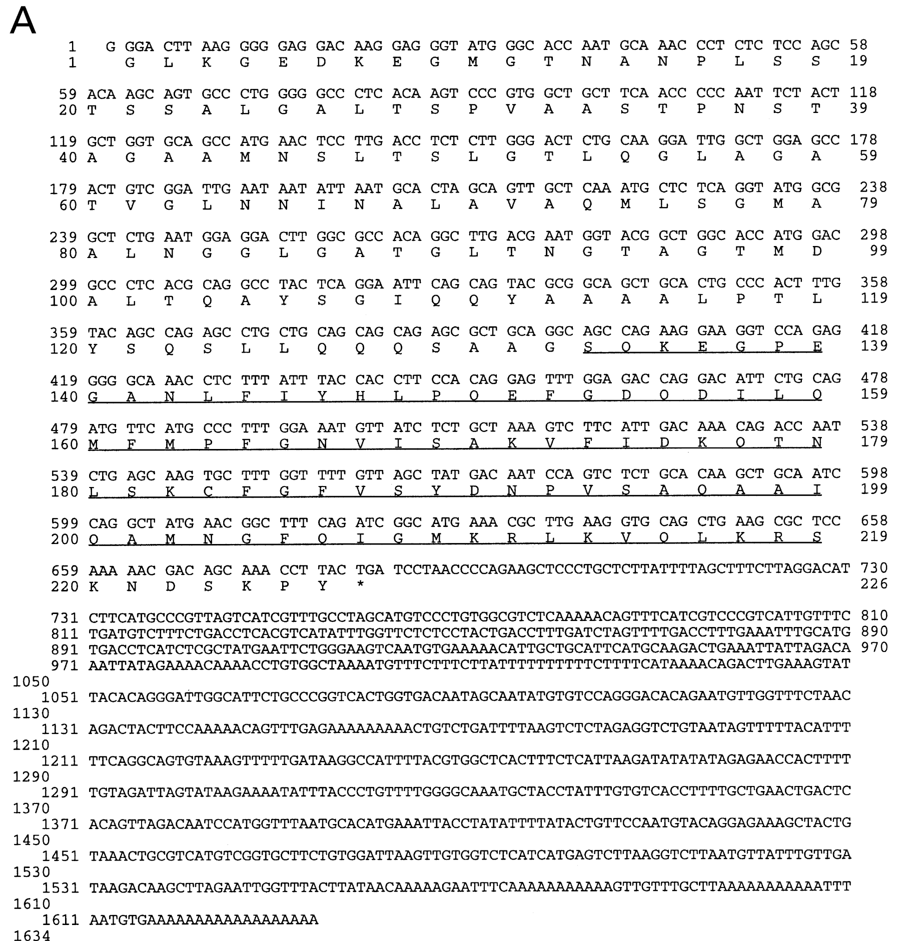

Fig. 1. Partial nucleotide and deduced amino-acid sequences of the clone AF169013, isolated in the three-hybrid screen. (A) The clone homologous to mouse ETR-R3b or human CUG-BP genes (accession number AF169013). The nucleotide and amino-acid sequences are numbered on both sides. The indicated ORF is open at the 5′ end. The stop codon is indicated by an asterisk. The RRM domain is underlined.
(B) Comparison of the predicted amino-acid sequences of the clone AF169013and its closest known homologue, rat ETR-R3b.

B

```
CUG/BP      1                           MNGTLDHPDQPDLDAIKMFVGQVPRTWSEKDLREL  35
                                        ||| ||| |||| ||||||||| || |||| | ||
ETR-R3B     1                           MNGALDHSDQPDPDAIKMFVGQIPRSWSEKELKEL  35

CUG/BP     36 FEQYGAVYEINVLRDRSQNPPQSKGCCFVTFYTRKAALEAQNALHNMKVLPGMHHPIQMK  95
              || ||||| ||||||||||||||||||||||||||||||||||||| | |||||||||||
ETR-R3B    36 FEPYGAVYQINVLRDRSQNPPQSKGCCFVTFYTRKAALEAQNALHNIKTLPGMHHPIQMK  95

CUG-BP     96 PADSEKNNAVEDRKLFIGMISKKCTENDIRVMFSSFGQIEECRILRGPDGLSRGCAFVTF 155
              |||||| |||||||||||| |||| ||||||||| |||||||||||||||||||||||||
ETR-R3B    96 PADSEKSNAVEDRKLFIGMVSKKCNENDIRVMFSPFGQIEECRILRGPDGLSRGCAFVTF 155

CUG-BP    156 TTRAMAQTAIKAMHQAQTMEGCSSPMVVKFADTQKDKEQKRMAQQLQQQMQQISAASVWGN 216
               |||||| ||||||| ||||||||| ||||||||||||| |  |||  |||||   |  |||
ETR-R3B   156 STRAMAQNAIKAMHQSQTMEGCSSPIVVKFADTQKDKEQRRLQQQLAQQMQQLNTAT-WGN 215

CUG-BP    217 LAGLNTLGPQYLALLQQTASSGNLNTLSSLHPMGGLNAMQLQNLAALAAAASAAQNTPSG 276
              | ||  | |||||||||  || ||   |    | | || |||||| ||||| |||
ETR-R3B   216 LTGLGGLTPQYLALLQQATSSSNLGAFSGIQQMAGMNALQLQNLATLAAAAAAAQTSATS 275

AF169013    1                                                  ...GLKGEDKEGMG  11

CUG-BP    277 TNALTTSSSPLSV--LTSSG--SSPSSSSSNSVNPIASLGALQTLAGATAGLN-VGSLA- 330
              |||   ||      |||   | | |     |   ||| || ||||| |||    ||
AF169013   12 TNANPLSSTSSALGALTSPVAASTPNSTAGAAMNSLTSLGTLQGLAGATVGLNNINALAV  71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ETR-R3B   276 TNANPLSSTSSALGALTSPVAASTPNSTAGAAMNSLTSLGTLQGLAGATVGLNNINALAV 335

CUG-BP    331 -----GMAALNGGLGSSGLSNGTGSTMEALTQAYSGIQQYAAAALPTLYNQNLLTQQSIGAA 387
                  |||||||||||||||||||| ||||||||||||||||||| | || |||  ||
AF169013   72 AQMLSGMAALNGGLGATGLTNGTAGTMDALTQAYSGIQQYAAAALPTLYSQSLLQQQS--AA 131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
ETR-R3B   336 AQMLSGMAALNGGLGATGLTNGTAGTMDALTQAYSGIQQYAAAALPTLYSQSLLQQQS--AA 395

CUG-BP    388 GSQKEGPEGANLFIYHLPQEFGDQDLLQMFMPFGNVVSAKVFIDKQTNLSKCFGFVSYDN 447
              ||||||||||||||||||||||||| |||||||||| |||||||||||||||||||||||
AF169013  132 GSQKEGPEGANLFIYHLPQEFGDQDILQMFMPFGNVISAKVFIDKQTNLSKCFGFVSYDN 191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ETR-R3B   396 GSQKEGPEGANLFIYHLPQEFGDQDILQMFMPFGNVISAKVFIDKQTNLSKCFGFVSYDN 455

CUG-BP    448 PVSAQAAIQSMNGFQIGMKRLKVQLKRSKNDSKPY 482
              ||||||||| |||||||||||||||||||||||||
AF169013  192 PVSAQAAIQAMNGFQIGMKRLKVQLKRSKNDSKPY 226
              |||||||||||||||||||||||||||||||||||
ETR-R3B   456 PVSAQAAIQAMNGFQIGMKRLKVQLKRSKNDSKPY 490
```

Fig. 1. continued.

RT-PCR analysis

Total RNA (1–3 μg) was mixed with 25 pM reverse primer derived from exon 9 of the APP gene, and incubated for 5 min at 70 °C. cDNA synthesis was performed for 1 h at 42 °C with SuperScript reverse transcriptase according to the manufacturer's instructions (Life Technologies), and terminated at 58 °C for 15 min. A 5-μL portion of the cDNA pool was taken for PCR with oligonucleotides derived from exon 6 and exon 9. PCR was performed in a Perkin–Elmer 9600 thermocycler for 30 cycles (58 °C for 0.5 min, 72 °C for 0.5 min, 93 °C for 0.3 min). The primers used for the amplification were as follows: forward, 5′-TGAAGACAAAGTAGTAGAAGTAGCAG-3′; reverse, 5′-CTGGGACATTCTCTCTCGGTGCTTG-3′.

Yeast two-hybrid and three-hybrid systems

The three-hybrid screens were performed as described previously [24]. The yeast strain L40-coat was transformed with bait and cDNA library plasmids, and plated on to synthetic dextrose (SD)-agar plates without leucine, histidine and uracil, and with 5 mM 3-aminotriazole. After one week, colonies were picked and restreaked on SD-agar without leucine and with fluorooratic acid (5-FOA) to eliminate the bait plasmid. These clones were restreaked on SD-agar without leucine, and assayed for bait-independent growth. Selected clones were sequenced and, if they contained any sensible ORFs, used for further assays. To test RNA-binding specificity, several different RNA plasmids were reintroduced into selected clones by mating. For this the second yeast strain, R40-coat, with the opposite mating type of L40-coat, was used. The two-hybrid assay was performed according to the Clontech mating protocol, using yeast strains HF7c and Y187.

RESULTS

Cloning factors that interact with the APP splicing enhancer

For three-hybrid screens, a cDNA library from postnatal P5 rat brain was used. Accordingly, sequences derived from potential splicing enhancers of genes differentially spliced in the brain were selected. Three such sequences, derived from the CLB, agrin and APP genes were cloned downstream from the MS2 sequence in pIIIAMS-2 vector to construct hybrid RNAs. These bait constructs were cotransformed with cDNA prey plasmids in yeast and screened for interactions. Two screens performed with the EN exon of CLB as bait and one screen with sequences from an intronic splicing enhancer of the agrin gene were not successful. The fourth screen was performed with a potential splicing enhancer of the APP gene and yielded about 65 clones. Of the 10 clones selected for sequencing, only two had ORFs and both were by homology putative RNA-binding proteins (Figs 1 and 2).

One of them, designated AF169013, is almost identical in the coding region with the previously described rat Elav-type ETR-R3b and mouse NAPOR-3 genes, the homologue of

A

```
   1                   GGGAGCTGATGCTTGTGGCGCGCGCGGCGGCCGCGACGGACGCAAG ATG GCG ACG GCA   58
   1                                                                  M   A   T   A      4

  59 ACT ATA GCT CTC CAG GTC AAT GGC CAA CAA GGA GGG GGG TCG GAG CCA GCA GCA GCG GCT  118
   5 T   I   A   L   Q   V   N   G   Q   Q   G   G   G   S   E   P   A   A   A   A    24

 119 GCA GCG GCG GCG GCG GCA GTG GTG GCA GCA GGA GAC AAA TGG AAA CCT CCA CAG GGC ACA  178
  25 A   A   A   A   A   A   V   V   A   A   G   D   K   W   K   P   P   Q   G   T    44

 179 GAA TCC ATC AAG ATG GAA AAT GGG CAA AGC ACA GGC ACC AAG CTG GGG CTG CCT CCC CTG  238
  45 E   S   I   K   M   E   N   G   Q   S   T   G   T   K   L   G   L   P   P   L    64

 239 ACG CCC GAG CAG CAG GAG GCC CTC CAG AAG GCC AAG AAA TAT GCA ATG GAG CAG AGC ATC  298
  65 T   P   E   Q   Q   E   A   L   Q   K   A   K   K   Y   A   M   E   Q   S   I    84

 299 AAG AGT GTG CTG GTG AAG CAG ACC ATC GCC CAC CAG CAG CAG CAG CTC ACC AAC CTG CAG  358
  85 K   S   V   L   V   K   Q   T   I   A   H   Q   Q   Q   Q   L   T   N   L   Q   104

 359 ATG GCA GCT CAG CGG CAG CGG GCA CTG GCT ATC ATG TGC CGG GTG TAT GTG GGT TCC ATC  418
 105 M   A   A   Q   R   Q   R   A   L   A   I   M   C   R   V   Y   V   G   S   I   124

 419 TAC TAT GAG CTG GGA GAA GAC ACT ATT CGC CAG GCC TTT GCT CCC TTT GGC CCC ATC AAG  478
 125 Y   Y   E   L   G   E   D   T   I   R   Q   A   F   A   P   F   G   P   I   K   144

 479 AGC ATT GAT ATG TCC TGG GAC TCC GTT ACC ATG AAG CAT AAG GGC TTT GCC TTC GTG GAG  538
 145 S   I   D   M   S   W   D   S   V   T   M   K   H   K   G   F   A   F   V   E   164

 539 TAT GAG GTC CCA GAA GCT GCA CAG CTG GCT TTG GAG CAG ATG AAC TCT GTG ATG CTT GGG  598
 165 Y   E   V   P   E   A   A   Q   L   A   L   E   Q   M   N   S   V   M   L   G   184

 599 GGC AGG AAC ATC AAG GTG GGA AGA CCT AGC AAC ATC GGA CAA GCC CAA CCC ATC ATA GAC  658
 185 G   R   N   I   K   V   G   R   P   S   N   I   G   Q   A   Q   P   I   I   D   204

 659 CAG CTG GCT GAG GAG GCT AGG GCT TTC AAC CGA ATA TAC GTG GCC TCC GTA CAT CAG GAC  718
 205 Q   L   A   E   E   A   R   A   F   N   R   I   Y   V   A   S   V   H   Q   D   224

 719 CTG TCT GAT GAT GAT ATC AAG AGT GTG TTT GAA GCC TTC GGC AAG ATC AAG TCT TGT ACG  778
 225 L   S   D   D   D   I   K   S   V   F   E   A   F   G   K   I   K   S   C   T   244

 779 CTG GCC CGG GAC CCC ACA ACT GGC AAG CAC AAG GGC TAT GGT TTT ATC GAA TAT GAG AAG  828
 245 L   A   R   D   P   T   T   G   K   H   K   G   Y   G   F   I   E   Y   E   K   264

 839 GCC CAG TCG TCC CAG GAT GCT GTG TCC TCC ATG AAC CTC TTT GAT CTG GGT GGC CAG TAC  898
 265 A   Q   S   S   Q   D   A   V   S   S   M   N   L   F   D   L   G   G   Q   Y   284

 899 TTG AGG GTG GGC AAG GCC GTC ACA CCC CCA ATG CCC CTG CTA ACA CCT GCC ACG CCT GGA  958
 285 L   R   V   G   K   A   V   T   P   P   M   P   L   L   T   P   A   T   P   G   304

 959 GGT CTC CCG CCT GCT GCT GCT GTG GCC GCA GCT GCA GCC ACA GCC AAG ATT ACA GCT CAG 1018
 305 G   L   P   P   A   A   A   V   A   A   A   A   A   T   A   K   I   T   A   Q   324

1019 GAA GCA GTG GCT GGA GCT GCA GTG CTG GGT ACT TTA GCC ACA CCA GGA CTA GTG TCC CCA 1078
 325 E   A   V   A   G   A   A   V   L   G   T   L   A   T   P   G   L   V   S   P   344

1079 GCA CTG ACA CTG GCC CAG CCC TTA GGG GCC CTA CCC CAG GCT GTC ATG GCT GCC CAA GCC 1138
 345 A   L   T   L   A   Q   P   L   G   A   L   P   Q   A   V   M   A   A   Q   A   364

1139 CCT GGA GTC ATC ACA GGT GTG ACA CCA GCC CGC CCT CCT ATT CCG GTC ACC ATC CCC TCT 1198
 365 P   G   V   I   T   G   V   T   P   A   R   P   P   I   P   V   T   I   P   S   384

1199 GTG GGA GTG GTG AAC CCC ATC CTA GCC AGC CCG CCA ACG CTG GGT CTG TTG GAA CCC AAG 1258
 385 V   G   V   V   N   P   I   L   A   S   P   P   T   L   G   L   L   E   P   K   404

1259 AAG GAG AAG GAA GAG GAG GAG CTG TTT CCA GAG TCA GAG CGG CCA GAG ATG TTG AGT GAG 1318
 405 K   E   K   E   E   E   E   L   F   P   E   S   E   R   P   E   M   L   S   E   424

1319 CAG GAG CAC ATG AGC ATC TCT GGC AGC AGT GCT CGA CAC ATG GTC ATG CAA AAG CTA CTC 1378
 425 Q   E   H   M   S   I   S   G   S   S   A   R   H   M   V   M   Q   K   L   L   444

1379 AGA AAA CAG GAG TCT ACA GTG ATG GTT CTC CGA AAC ATG GTG GAT CCC AAG GAC ATT GAT 1438
 445 R   K   Q   E   S   T   V   M   V   L   R   N   M   V   D   P   K   D   I   D   464

1439 GAT GAC CTG GAG GGA GAG GTG ACA GAG GAA TGT GGC AAA TTT GGT GCT GTG AAC CGG GTC 1498
 465 D   D   L   E   G   E   V   T   E   E   C   G   K   F   G   A   V   N   R   V   484

1499 ATT ATC TAC CAA GAA AAA CAG GGC GAG GAG GAA GAT GCA GAA ATC ATC GTC AAG ATT TTT 1558
 485 I   I   Y   Q   E   K   Q   G   E   E   E   D   A   E   I   I   V   K   I   F   504

1559 GTG GAA TTT TCC ATG GCT TCA GAG ACT CAC AAG GCC ATC CAG GCC CTC AAT GGG CGC TGG 1618
 505 V   E   F   S   M   A   S   E   T   H   K   A   I   Q   A   L   N   G   R   W   524

1619 TTT GGT GGT CGC AAG GTG GTG GCT GAA GTG TAT GAC CAG GAG CGT TTT GAT AAC AGC GAC 1678
 525 F   G   G   R   K   V   V   A   E   V   Y   D   Q   E   R   F   D   N   S   D   544

1679 CTC TCT GCG TGA CTTTGGCCTCATTTCCTGGACTTGCACTTGCCCCTTCCTTCCTTTGGGTTTTTGTAAAGTGTTA 1754
 545 L   S   A   *                                                                   547

1755 AAAGTGGTGTGTCCCTGGGACCAGGCACTTCTGCCCAGCCCACCCTGTAGTGTACATAAAGATGCAGACAGATACTGCTG 1834
1835 GGAAAAAAAAAAAAAAAAAAAAAAAAA                                                      1861
```

Fig. 2 Partial nucleotide and deduced amino-acid sequences of the clone AF165892, isolated in the three-hybrid screen. (A) The clone homologous to human Siah-BP and PUF60 (accession number AF165892). The nucleotide and amino-acid sequences are numbered on both sides. The indicated ORF is open at the 5′ end. The stop codon is indicated by an asterisk. The RRM domains are underlined. (B) Comparison of the predicted amino-acid sequences for the clone AF165892and its closest known homologue, human Siah-BP.

which in man is the Nab50/CUG-BP gene known also as NAPOR-1 (Fig. 1A). NAPOR-1 (neuroblastoma apoptosis-related RNA-binding protein) was isolated previously by differential display from apoptotic neuroblastoma cells [31]. The Nab50/CUG-BP protein was identified first as an interacting partner for yeast hnRNP Nab2p in a two-hybrid screen, and for the second time as a protein that binds RNA in a sequence-specific manner [32,33]. Its binding activity was specific for the CUG repeat, but not for single-stranded or double-stranded DNA CTG repeats. It was suggested that the

B

```
AF165892    1 MATATIALQVNGQQGGGSEPAAAAAAAAAAVVAAGDKWKPPQGTESIKMENGQSTGTKLGL  61
hSiahBP     1 MARATIALQVNGQQGGGSEPAAAA-----AVVAAGDKWKPPQGTDSIKMENGQSTAAKLGL  56

AF165892   62 PPLTPEQQEALQKAKKYAMEQSIKSVLVKQTIAHQQQQLTNLQMAAQRQRALAIMCRVYV  121
hSiahBP    57 PPLTPEQQEALQKAKKYAMEQSIKSVLVKQTIAHQQQQLTNLQMAAQRQGALAIMCRVYV  116

AF165892  122 GSIYYELGEDTIRQAFAPFGPIKSIDMSWDSVTMKHKGFAFVEYEVPEAAQLALEQMNSV  181
hSiahBP   117 GSIYYELGEDTIRQAFAPFGPIKSIDMSWDSVTMKHKGFAFVEYEVPEAAQLALEQMNSV  176

AF165892  182 MLGGRNIKVGRPSNIGQAQPIIDQLAEEARAFNRIYVASVHQDLSDDDIKSVFEAFGKIK  241
hSiahBP   177 MLGGRNIKVGRPSNIGQAQPIIDQLAEEARAFNRIYVASVHQDLSDDDIKSVFEAFGKIK  236

AF165892  242 SCTLARDPTTGKHKGYGFIEYEKAQSSQDAVSSMNLFDLGGQYLRVGKAVTPPMPLLTPA  301
hSiahBP   237 SCTLARDPTTGKHKGYGFIEYEKAQSSQDAVSSMNLFDLGGQYLRVGKAVTPPMPLLTPA  296

AF165892  302 TPGGLPPAAAVAAAAATAKITAQEAVAGAAVLGTLATPGLVSPALTLAQPLGALPQAVMA  361
hSiahBP   297 TPGGLPPAAAVAAAAATAKITAQEAVAGAAVLGTLGTPGLVSPALTLAQPLGTLPQAVMA  356

AF165892  362 AQAPGVITGVTPARPPIPVTIPSVGVVNPILASPPTLGLLEPKKEKEEEELFPESERPEM  421
hSiahBP   357 AQAPGVITGVTPARPPIPVTIPSVGVVNPILASPPTLGLLEPKKEKEEEELFPESERPEM  416

AF165892  422 LSEQEHMSISGSSARHMVMQKLLRKQESTVMVLRNMVDPKDIDDDLEGEVTEECGKFGAV  481
hSiahBP   417 LSEQEHMSISGSSARHMVMQKLLRKQESTVMVLRNMVDPKDIDDDLEGEVTEECGKFGAV  476

AF165892  482 NRVIIYQEKQGEEEDAEIIVKIFVEFSMASETHKAIQALNGRWFGGRKVVAEVYDQERFD  541
hSiahBP   477 NRVIIYQEKQGEEEDAEIIVKIFVEFSIASETHKAIQALNGRWFAGRKVVAEVYDQERFD  536

AF165892  542 NSDLSA 547
hSiahBP   537 NSDLSA 542
```

Fig. 2. continued.

CUG repeat found in the 3′ UTR of the myotonin protein kinase mRNA is a target for CUG-BP. There are two known isoforms of human CUG-BP, which are distributed in both cytoplasmic and nuclear fractions. The CUG-BP1 cDNA contains an ORF that encodes a 482-amino-acid protein with three RNA-recognition motif (RRM) domains. The AF169013 clone isolated in our screen is a partial cDNA, coding for the C-terminal part of the corresponding protein and large 3′ UTR. In a region coding for 3d C-terminal RRM, this clone shows high homology to ETR-R3b as well as to human CUG-BP (Fig. 1B). 3′ UTR resembles that of the mouse NAPOR-3 gene (not shown).

The second clone, AF165892, is a rat counterpart of the human Siah-BP/PUF60 gene (Fig. 2). Siah-BP (GenBank accession number U51586; Hu, Holloway, and Bowtell, personal communication) was initially isolated as an interacting partner of sina ring-finger protein, using the two-hybrid system. Rat Siah-BP shows predominant cytoplasmic localization as a GFP fusion protein (data not shown). This result contradicts the report describing the PFU60gene, which seems to be an isoform of Siah-BP, as splicing factor [34].

The isolated clone is 1861 nt long, contains an ATG at the 5′ end and ends with a TGA stop codon (nucleotides 1688–1690), which is followed by a short 3′ UTR (Fig. 1B). The cDNA contains three regions with trinucleotide repeats (nucleotides 107–137, 303–362 and 1153–1369) found in a wide variety of unrelated genes. The predicted product of the

Table 1. Specific interactions of clones isolated in the three-hybrid screen with different baits. Both RNA-expressing plasmid (containing sequence of natural or artificial splicing enhancer derived from agrin, clathrin light chain B, amyloid precursor protein genes or CUG repeat, respectively) and clone isolated in the three-hybrid screen (partial cDNA of CUG-BP-related gene, AF169013, or rat Siah-BP, AF165892, respectively) were introduced into yeast by the mating protocol, and plated on SD-agar without leucine, uracil, histidine and with 5 mM 3-aminotriazole. Growth was estimated after 3 days of incubation at 30 °C. (+) Weak interaction in the three-hybrid assay; (+ + +) strong interaction; (–) no interaction. IRE, iron-regulatory protein responsive element.

	AF169013	AF165892
Agrin	–	–
Anti-agrin	–	–
EN CLB	–	–
Anti-(EN CLB)	–	–
APP	+ + +	+ + +
Anti-APP	–	–
Anti-(CUG)x8	–	–
(CUG)x8	+	+
IRE	–	–

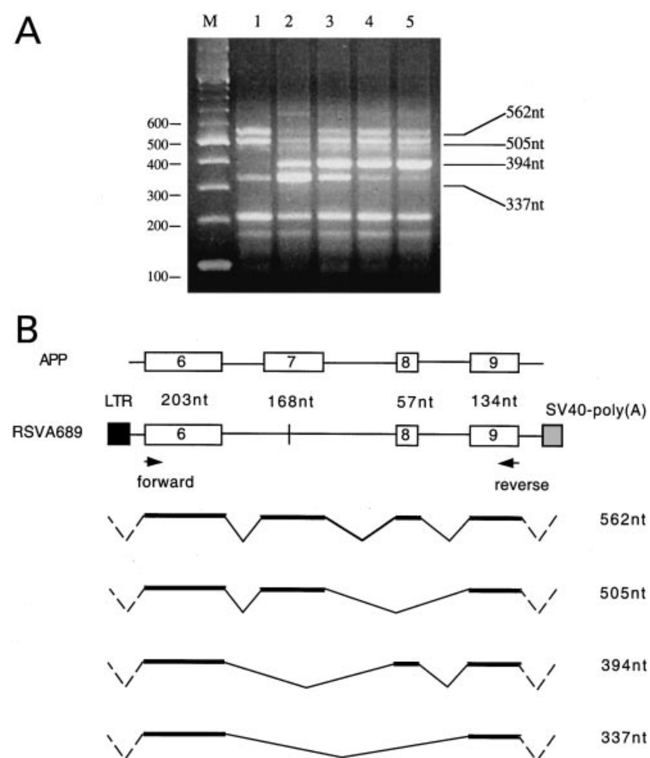

Fig. 3. Inclusion of exon 8 of the APP minigene is strongly inhibited by CUG-BP and its deletion mutant in a splicing assay. (A) Human CUG-BP, deletion mutant of rat CUG-BP and Siah-BP were cotransfected in HEK293 cells together with RSVA689, as indicated in Experimental procedures. RT-PCR experiments was performed with total cellular RNA with reverse and forward primers to amplify exons 6–9 from processed transcripts. M, 100 bp DNA marker; lane 1, no transfection; lane 2, minigene cotransfected with human CUG-BP expression vector, with the deletion mutant of rat CUG-BP-like clone (lane 3), Siah-BP expression vector (lane 4), or alone (lane 5). Predicted size for PCR fragments consisting of exons 6–9 is 562 bp, exons 6, 7, and 9 is 505 bp, exons 6, 8 and 9 is 394 bp, and exons 6 and 9 is 337 bp. (B) Diagram of the splicing pattern, showing the endogenous gene (APP), the minigene (RSVA689), the expected pre-mRNA and size of PCR products, amplified from these using the indicated primers (forward and reverse).

ORF is a protein with three RRM domains, which are highly conserved between the rat and human proteins (Fig. 2B). PUF60 and Siah-BP are identical except for a 17-amino-acid insertion in PUF60 (position 102/103 of Siah-BP) and one difference in protein sequence (G at position 105 in Siah-BP, R in position 123 in PUF60).

No homology outside of the RRM domains exists between CUG-BP and Siah-BP, but these proteins show similar compositions. Both contain three RRMs, two in the N-terminus and one in the C-terminus. The RRM domains of these proteins show about 30% amino-acid identity and 50% similarity at the amino-acid level to the RRM domains of polyadenylate-binding protein, spliceosome-associated protein 49, splicing factor U2AF 65KD subunit, hnRNP A3, Hu proteins, and weaker homologues with the sex-lethal protein of *Drosophila*.

CUG-BP and Siah-BP specifically bind sequences derived from the APP splicing enhancer

Both clones isolated in the three-hybrid screen specifically interact with the bait used for screening, as confirmed in the three-hybrid mating tests of specificity (Table 1). Whereas interaction with the APP splicing enhancer element was strong, only weak interactions were observed with bait containing the CUG repeat, and no interactions were observed between either prey and a number of different RNA baits, including other potential splicing-enhancer sequences. Interestingly, the CUG-BP-like clone, containing only one of the three RRM domains, is sufficient to interact with RNA bait in a sequence-specific manner.

CUG-BP is a potential regulator of alternative splicing of the APP gene

To examine whether the factors isolated in our three-hybrid screen could influence alternative splicing *in vivo*, cotransfection experiments with an APP minigene were performed. Construct RSVA689 [13], consisting of exons 6–9 where exon 7 is deleted, was transfected into the HEK293 cell line alone or together with human CUG-BP, deletion mutant AF169013 or the rat Siah-BP homologue AF165892 expression vectors. In the HEK293 cell line, in which the endogenous APP gene is expressed, exon 8 is included in only a fraction of mature RNAs (Fig. 3, lane 1). After introduction of the RSVA689 construct into cells, exon 8 is expressed strongly from the minigene, and is abundantly present in the processed RNA (Fig. 3, lane 5). On coexpression of the reporter gene with human CUG-BP, exon 8 is strongly excluded (Fig. 3, lane 2). A weaker effect on the exclusion of exon 8 was seen with the portion of rat CUG-BP isolated in the screen (Fig. 3, lane 3). No effects on splicing were observed with Siah-BP (Fig. 3, lane 4).

Interaction of CUG-BP and Siah-BP with other RNA-binding proteins

Factors that regulate differential splicing are expected to bind

Table 2. Interaction of CUG-BP and Siah-BP with certain RNA-binding proteins in the two-hybrid assay. The full coding regions of human CUG-BP or rat Siah-BP were subcloned in GAD424 vector (Stratagene); other cDNAs were cloned in GBT9 vector. Constructs were introduced into yeast by the mating protocol, and diploid cells were plated on SD-agar without leucine, tryptophan, histidine, and with 5 mM 3-aminotriazole, and incubated for 3 days at 30 °C. CUG-BP, human CUG-BP1; Siah-BP, clone AF165892 homologues to human Siah-BP; ELAV, rat ELAV-like gene [46]; SF2, human SF2/ASF [47]; U2AF65, human gene [48]; U2AF35, human gene [49]; SC35, human gene [50]; X16, mouse SRp20 [51]; tra2b1, human tra2b1 [52]; SAF-B, rat SAF-B [53]; EWS, rat gene (obtained from Oliver Stoss, MPI for Neurobiology, Munich, Germany).

	CUG-BP	Siah-BP
EWS	+ +	–
SC35	–	–
U2AF35	–	–
X16	–	–
tra2b1	–	–
U2AF65	–	+ +
SAF-B	–	–
SF2	–	+
ELAV	–	–
CUG-BP	–	–
Siah-BP	–	+ +

not only to their RNA target, but perhaps also to the essential splicing machinery, and may act in a complex with other RNA-binding proteins. To investigate these possible interactions, cDNAs coding for proteins involved in RNA metabolism, including essential splicing factors, were tested for protein–protein interactions with CUG-BP and Siah-BP in the two-hybrid system (Table 2). Human CUG-BP interacted only with EWS, an RGG-type RNA-binding protein associated with Ewing sarcoma. Siah-BP interacts with U2AF65, the large subunit of the essential splicing factor U2AF, as well as with itself.

DISCUSSION

Splicing enhancers and repressors have been described as short sequences, without loops or structural conformations, that influence differential splicing. These elements are diverse in sequence and location relative to the regulated splice site, and function as tissue-specific and/or stage-specific regulatory elements for alternative splicing. Therefore, such sequences can be more correctly defined as *cis*-acting elements regulating alternative splicing. It has recently been shown that multiple enhancer elements additively increase the efficiency of splicing [35]. The effect of a single *cis*-acting element depends on the natural sequence and environmental context, for example the presence of tissue-specific *trans*-acting factors. The involvement of such factors make differential splicing more complex but at the same time more precise, allowing fine-tuning of tissue-specific or developmentally regulated expression. However, many natural sequences that match the simple motifs for splicing enhancers could fail to function in artificial systems unless they are placed in an appropriate context. At present, there is too little information about how splicing enhancers function: which proteins bind these sequences; what causes the tissue-specific and stage-specific assembly of these proteins into a splicing-enhancer complex; how such a complex interacts with obligatory splicing components, etc.

It has been postulated, mainly on the basis of studies on *Drosophila*, that SR proteins may be *trans*-acting factors that regulate alternative splicing, and whose targets are splicing enhancers (reviewed in [2,36]). Experiments performed with different splicing factors in mammalian systems show that they act non-specifically and can bind several different sequences strongly [19]. Therefore the SELEX procedure and statistical methods used to identify targets of RNA-binding proteins are useful in identification of functional splicing enhancers but not their natural targets [37].

A different approach, utilizing the biochemical purification of RNA-binding proteins on an RNA-affinity column, has been used successfully to isolate the protein complexes associated with a splicing enhancer. Surprisingly, KSRP protein cloned as a component of a protein complex assembled on c-*src* splicing enhancer cannot individually bind that sequence (D. Black, UCLA, USA, personal communication). The same effect has been observed in the three-hybrid assay: KSRP was not able to interact with any RNA bait tested in our laboratory (data not shown). Furthermore, no other splicing factors or RNA-binding proteins tested so far in our laboratory interacts with RNA baits in the three-hybrid assay. This includes the target sequences described for WT1 and SC35, U2AF35, human tra2b, mouse X16, hnRNP A1, SF2, U170K, hnRNPG and some other factors.

An increasing number of proteins appear to be multifunctional, participating in different post-transcriptional events. WT1, at first thought to be a typical transcription factor, may also be involved in splicing, as well as in regulation of mRNA [38]. The +KTS isoform of WT1 is found preferentially in nuclear speckles, and colocalizes with snRNPs and splicing factors, such as the Sm proteins, U170 and U2AF65. Additional evidence indicates involvement of WT1 in post-transcriptional events, as WT1 also binds to the IGF-2 transcript. CUG-BP was initially reported to bind to the CUG repeat found in the 3′ UTR of myotonin protein kinase mRNA, regulating its cellular localization and/or transport/stability as well as being associated with poly(A)-rich RNA [32]. Our data indicate that this factor could also be involved in pre-mRNA processing, and could be a regulator of differential splicing in the APP gene.

Recently, a group of neuronal cytoplasmic Elav-like proteins have been described. These proteins, like CUG-BP and Siah-BP, contain three RRMs, and it has been shown that each RRM plays a distinct role in RNA-binding activity. The first RRM of mHuC is required for AU-rich element binding, and this binding is greatly enhanced when the first RRM is connected to the second RRM. The third RRM has specific affinity for the poly(A)-rich sequence [39]. The third RRM of the CUG-BP homologue described in this paper alone possesses sequence-specific RNA-binding activity in the three-hybrid system. However, it is not efficient in driving splicing from the APP minigene in transfection experiments, indicating that the activities of the three RRM domains of CUG-BP are not redundant.

Siah-BP interacts with the sequence derived from the APP gene as well as with CUG repeat, and binds RNA in a sequence-specific manner. On the other hand, this factor does not interact with the APP minigene in transfection experiments. Cytoplasmic localization of the GFP–rat Siah-BP fusion protein is at variance with studies on PFU60 (as briefly discussed above). This also contradicts the fact that Siah-BP interacts with U2AF65 in the two-hybrid system, as well as the results of Page-McCaw *et al*., describing PUF60 as a splicing factor [34]. This is the first publication describing the Siah-BP/PFU60 gene, and its biochemical and genetic features need to be further investigated. It should be mentioned that a lot of splicing factors do not have unequivocal cellular localization and function, and shuttle between the nucleus and cytoplasm [40–42]. CUG-BP/Nab50 has been shown to be present in both cytoplasmic and nuclear fractions, and its intranuclear distribution is highly variable between different cell types [32]. If the HEK293 cell line is not an appropriate test system for analysing Siah-BP function, then all experiments should be repeated in other cell lines. The interaction of Siah-BP with itself in the two-hybrid system is confirmed by functional analysis of the PFU60 protein. The third RRM domain has been shown to mediate protein–protein interaction to form stable dimers of PFU60 [34]. If Siah-BP binds RNA as a homodimer or heterodimer, it could be an interesting phenomenon for further investigation.

Preliminary data on proteins that interact with the isolated clones indicate possible links to other cell pathways. EWS, an RNA-binding protein [43] that interacts in the two-hybrid assay with CUG-BP, also copurifies with another RNA-binding protein and protein kinase C substrate, TLS/FUS [44] as well as with the splicing factor PSF [45]. EWS also interacts with ESAF (O. Stoss, MPI for Neurobiology, Munich, Germany, personal communication), described as an important factor involved in pre-mRNA metabolism.

In this report, we have cloned nuclear RNA-binding protein using an RNA target described previously as a splicing enhancer in the APP gene. No such factors have previously been isolated by the three-hybrid approach. Two other

sequences regarded as potential splicing enhancers were not effective in the isolation of RNA-binding proteins. This may indicate that the baits used for the cloning of proteins that bind to splicing enhancers/silencers must possess other features as for cloning of the cytoplasmic proteins that bind mRNA. If multiprotein complexes bind preferentially to such sequences rather than individual factors, biochemical purification of these complexes and subsequent cloning of their components would be more appropriate. If enhancer-binding proteins require other proteins to stabilize RNA binding, the baits should be modified to combine splicing-enhancer sequences with consensus 5′ or 3′ sequences, polypyrimidine tract, etc. and tested in the three-hybrid screens.

ACKNOWLEDGEMENTS

We thank Marvin Wickens and Stanley Fields for providing the yeast strains and plasmids for the three-hybrid system, Professor Yoshiyuki Sakaki for the APP-minigene construct, and Douglas Black for KRSP cDNA. This work was supported by a DFG grant to S. S.

REFERENCES

1. Blencowe, B.J. (2000) Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends Biochem. Sci.* **25**, 106–110.
2. Hertel, K.J., Lynch, K.W. & Maniatis, T. (1997) Common themes in the function of transcription and splicing enhancers. *Curr. Opin. Cell Biol.* **9**, 350–357.
3. Mattox, W. & Backer, B.S. (1991) Autoregulation of the splicing of transcripts from the transformer-2 gene of *Drosophila*. *Genes Dev.* **5**, 786–796.
4. Tian, M. & Maniatis, T. (1993) A splicing enhancer complex controls alternative splicing of doublesex pre-mRNA. *Cell* **74**, 105–114.
5. Siebel, C.W., Admon, A. & Rio, D.C. (1995) Soma-specific expression and cloning of PSI, a negative regulator of P element pre-mRNA splicing. *Genes Dev.* **9**, 269–283.
6. Koushika, S.P., Lisbin, M.J. & White, K. (1996) Elav, a *Drosophila* neuron-specific protein, mediates the generation of an alternetively spliced neural protein isoform. *Curr. Biol.* **6**, 1634–1641.
7. Black, D.L. (1992) Activation of c-src neuron-specific splicing by an unusual RNA element *in vivo* and *in vitro*. *Cell* **69**, 795–807.
8. Min, H., Chan, R.C. & Black, D.L. (1995) The generally expressed hnRNP F is involved in a neural-specific pre-mRNA splicing event. *Genes Dev.* **9**, 2659–2671.
9. Min, H., Turck, C.W., Nikolic, J.M. & Black, D.L. (1997) A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer. *Genes Dev.* **11**, 1023–1036.
10. Jensen, K.B., Dredge, B.K., Stefani, G., Zhong, R., Buckanovich, R.J., Okano, H.J., Yang, Y.Y. & Darnell, R.B. (2000) Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. *Neuron* **25**, 359–371.
11. Tsukahara, T., Kunika, N., Momoi, T. & Arahata, K. (1995) Regulation of alternetive splicing in the amyloid precursor protein (APP) mRNA during neuronal and glial differentiation of P19 embryonal carcinoma cells. *Brain Res.* **679**, 178–183.
12. Rohan de Silva, H.A., Jen, A., Wickenden, C., Jen, L.S., Wilkinson, S.L. & Patel, A.J. (1997) Cell-specific expression of beta-amyloid precursor protein isoform mRNAs and proteins in neurons and astrocytes. *Brain Res. Mol. Brain Res.* **47**, 147–156.
13. Shibata, A., Hattori, M., Suda, H. & Sakaki, Y. (1996) Identification of cis-acting elements involved in an alternative splicing of the amyloid precursor protein (APP) gene. *Gene* **175**, 203–208.
14. Wei, N., Lin, C.Q., Modafferi, E.F., Gomes, W.A. & Black, D.L. (1997) A unique intronic splicing enhancer controls the inclusion of the agrin Y exon. *RNA* **3**, 1275–1288.
15. Stamm, S., Casper, D., Hanson, V. & Helfman, D.M. (1999) Regulation of the neuron-specific exon of clathrin light chain B. *Brain Res. Mol. Brain Res.* **64**, 108–118.
16. Smith, W.J., ed. (1998) *RNA–protein interactions. A Practical Approach.* Oxford University Press, Oxford, UK.
17. Coulter, L., Landree, M. & Cooper, T. (1997) Identification of a new class of exonic splicing enhancers by *in vivo* selection. *Mol. Cell. Biol.* **17**, 4096–4104.
18. Tian, H. & Kole, R. (1995) Selection of novel exon recognition elements from a pool of random sequences. *Mol. Cell. Biol.* **15**, 6291–6298.
19. Liu, H.-X., Zhang, M. & Krainer, A. (1998) Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes Dev.* **12**, 1998–2012.
20. Tuerk, C. & Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510.
21. Irvine, D., Tuerk, C. & Gold, L. (1991) SELEXION. Systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. *J. Mol. Biol.* **222**, 739–761.
22. Bachler, M., Schroeder, R. & von Ahsen, U. (1999) StreptoTag: a novel method for the isolation of RNA-binding proteins. *RNA* **5**, 1509–1516.
23. Paraskeva, E., Atzberger, A. & Hentze, M.W. (1998) A translational repression assay procedure (TRAP) for RNA–protein interactions *in vivo*. *Proc. Natl Acad. Sci. USA* **95**, 951–956.
24. SenGupta, D.J., Zhang, B., Kraemer, B., Pochart, P., Fields, S. & Wickens, M. (1996) A three-hybrid system to detect RNA–protein interactions *in vivo*. *Proc. Natl Acad. Sci. USA* **93**, 8496–8501.
25. Putz, U., Skehel, P. & Kuhl, D. (1996) A tri-hybrid system for the analysis and detection of RNA–protein interactions. *Nucleic Acids Res.* **24**, 4838–4840.
26. Wang, Z.F., Whitfield, M.L., Ingledue, T.C. 3rd, Dominski, Z. & Marzluff, W.F. (1996) The protein that binds the 3′ end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing. *Genes Dev.* **10**, 3028–3040.
27. Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J. & Wickens, M.P. (1997) A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line. *Nature (London)* **390**, 477–484.
28. Bacharach, E. & Goff, S.P. (1998) Binding of the human immunodeficiency virus type 1 Gag protein to the viral RNA encapsidation signal in the yeast three-hybrid system. *J. Virol.* **72**, 6944–6949.
29. January, E., Motzny, C.K., Graves, L.E. & Goodwin, E.B. (1999) The STAR protein, GLD-1, is a translational regulator of sexual identity in *Caenorhabditis elegans*. *EMBO J.* **18**, 258–269.
30. Fewell, S.W. & Woolford, J.L. Jr (1999) Ribosomal protein S14 of *Saccharomyces cerevisiae* regulates its expression by binding to RPS14B pre-mRNA and to 18S rRNA. *Mol. Cell. Biol.* **19**, 826–834.
31. Choi, D.K., Ito, T., Mitsui, Y. & Sakaki, Y. (1998) Fluorescent differential display analysis of gene expression in apoptotic neuroblastoma cells. *Gene* **223**, 21–31.
32. Timchenko, L.T., Miller, J.W., Timchenko, N.A., DeVore, D.R., Datar, K.V., Lin, L., Roberts, R., Caskey, C.T. & Swanson, M.S. (1996) Identification of a (CUG) n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res.* **24**, 4407–4414.
33. Timchenko, L.T., Timchenko, N.A., Caskey, C.T. & Roberts, R. (1996) Novel proteins with binding specificity for DNA CTG repeats and RNA CUG repeats: implications for myotonic dystrophy. *Hum. Mol. Genet* **5**, 115–121.
34. Page-McCaw, P.S., Amonlirdviman, K. & Sharp, P.A. (1999) PUF60: a novel U2AF65-related splicing activity. *RNA* **5**, 1548–1560.
35. Hertel, & Maniatis (1998) The function of multisite splicing enhancers. *Mol. Cell* **1**, 449–455.
36. Valcárcel, J. & Green, M.R. (1996) The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends Biochem. Sci.* **21**, 296–301.
37. Tacke, R. & Manley, J.L. (1999) Determinants of SR protein specificity. *Curr. Opin. Cell Biol.* **11**, 358–362.
38. Caricasole, A., Duarte, A., Larsson, S.H., Hastie, N.D., Little, M.,

Holmes, G., Todorov, I. & Ward. A. (1996) RNA binding by the Wilms tumor suppressor zinc finger proteins. *Proc. Natl Acad. Sci. USA* **93**, 7562–7566.

39. Abe, R., Sakashita, E., Yamamoto, K. & Sakamoto, H. (1996) Two different RNA binding activities for the AU-rich element and the poly (A) sequence of the mouse neuronal protein mHuC. *Nucleic Acids Res.* **24**, 4895–4901.
40. Caceres, J.F., Screaton, G.R. & Krainer, A.R. (1998) A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes Dev.* **12**, 55–66.
41. Keene, J.D. (1999) Why is Hu where? Shuttling of early-response-gene messenger RNA subsets. *Proc. Natl Acad. Sci. USA* **96**, 5–7.
42. Weidhardt, F., Biamonti, G. & Riva, S. (1996) The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. *Bioessays* **18**, 747–756.
43. Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., Aurias, A. & Thomas, G. (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature (London)* **359**, 162–165.
44. Calvio, C., Neubauer, G., Mann, M. & Lamond, A.I. (1995) Identification of hnRNP P2 as TLS/FUS using electrosprey mass spectrometry. *RNA* **7**, 724–733.
45. Patton, J.G., Porro, E.B., Galceran, J., Tempst, P. & Nadal-Ginard, B. (1993) Cloning and characterization of PSF, a novel pre-mRNA splicing factor. *Genes Dev.* **7**, 393–406.
46. King, P.H., Levine, T.D., Fremeau, R.T. Jr & Keene, J.D. (1994) Mammalian homologs of *Drosophila* ELAV localized to a neuronal subset can bind in vitro to the 3′ UTR of mRNA encoding the Id transcriptional repressor. *J. Neurosci.* **14**, 1943–1952.
47. Ge, H., Zuo, P. & Manley, J.L. (1991) Primary structure of the human splicing factor ASF reveals similarities with *Drosophila* regulators. *Cell* **66**, 373–382.
48. Zamore, P.D., Patton, J.G. & Green, M.R. (1992) Cloning and domain structure of the mammalian splicing factor U2AF. *Nature (London)* **355**, 609–614.
49. Zhang, M., Zamore, P.D., Carmo-Fonseca, M., Lamond, A.I. & Green, M.R. (1992) Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc. Natl Acad. Sci. USA* **9**, 8769–8773.
50. Fu, X.D., Mayeda, A., Maniatis, T. & Krainer, A.R. (1992) General splicing factors SF2 and SC35 have equivalent activities *in vitro*, and both affect alternative 5′ and 3′ splice site selection. *Proc. Natl Acad. Sci. USA* **89**, 11224–11228.
51. Ayane, M., Preuss, U., Kohler, G. & Nielsen, P.J. (1991) A differentially expressed murine RNA encoding a protein with similarities to two types of nucleic acid binding motifs. *Nucleic Acids Res.* **19**, 1273–1278.
52. Beil, B., Screaton, G. & Stamm, S. (1997) Molecular cloning of htra2-beta-1 and htra2-beta-2, two human homologs of tra-2 generated by alternative splicing. *DNA Cell Biol.* **16**, 679–690.
53. Nayler, O., Straetling, W., Bourquin, J.-P., Stagljar, I., Lindemann, L., Jasper, H., Hartmann, A.M., Fackelmeyer, F.O., Ullrich, A. & Stamm, S. (1998) SAF-B couples transcription and pre-mRNA splicing to SAR/MAR elements. *Nucleic Acids Res.* **26**, 3542–3549.