1018–1022 Nucleic Acids Research, 1994, Vol. 22, No. 6

# A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon

M.Caputi, G.Casari+, S.Guenzi§, R.Tagliabue, A.Sidoli§, C.A.Melo and F.E.Baralle*
International Centre for Genetic Engineering and Biotechnology – UNIDO, Padriciano 99, Area Science Park, 34012 Trieste, Italy

Received December 17, 1993; Accepted February 16, 1994

## ABSTRACT

**EDA is a facultative type III homology of human fibronectin encoded by an alternative spliced exon. The EDA+ and EDA– mRNA forms show a cell type specific distribution with their relative proportion varying during development, aging and oncogenic transformation. We have previously demonstrated that an 81 bp nucleotide sequence within the exon itself is essential for differential RNA processing. Fine mapping of cis acting elements within this region has been carried out to identify possible target sites for the modulation of alternative splicing. There are at least two short nucleotide sequences involved. Element A (GAAGAAGA) is a positive modulator for the recognition of the exon, its deletion results in constitutive exclusion of the EDA exon. Element B (CAAGG) is a negative modulator for exon recognition, its deletion results in constitutive inclusion of the EDA exon. This bipartite structure of the splicing enhancer is a novel feature of the mammalian exons.**

## INTRODUCTION

Alternative splicing of a single primary gene transcript is a widespread mechanism for the regulation of gene expression. Differential RNA processing can generate variant proteins from a single gene which may have different functions and may be expressed in a particular developmental stage, age and tissue.

Alternative RNA processing of human (1–4) and rat (5) fibronectin pre mRNA occurs in three regions of the primary transcript EDA, EDB and III CS. Structural and functional analysis of EDA and EDB alternative splicing shows suggestive similarities, as well as important differences. In both areas, alternative splicing results in the facultative insertion of the ED exons. In most adult tissues, ED+ and ED– mRNAs are found in various ratios. In contrast, in liver a distinct pattern of splicing is observed in which the ED exons are always omitted (1,3,5).

Extensive data are now available about the features and steps of the generic splicing reaction (6–8), however little is known about the mechanisms responsible for alternative splicing of fibronectin and other pre-mRNAs, and even less about those underlying the tissue specificity of the process. The cis acting factors that have been implicated in pre mRNA splicing are: 5′ and 3′ splice site sequences (9,10), branchpoint sequences and their locations (11,12), intron sequences (11,13) and exon sequences (14–19). These sequences are target for interactions with trans acting factors such as the small nuclear ribonucleoproteins U1, U2, U4, U5 and U6 (20), and multiple auxiliary proteins (9,21). Recently protein factors directly involved in the alternative splicing process have been identified. The better known are the Drosophila Sxl protein (22) and several members of the SR protein family, like the Tra and Tra 2 in Drosophila (23) and SC35 and SF2 in mammalian cells (24,25).

We have previously shown that both ED-A (26) and ED-B (27) alternative splicing can be accurately reproduced in HeLa cells by transient expression of a hybrid minigene containing a 3 kb fragment of the human fibronectin gene inserted into the third exon of the human $\alpha$1-globin gene and under the control of the promoter of this latter gene. This system was used to study the cis and trans acting factors involved in the tissue specificity of this process (4). Particularly interesting was the study of the sequences involved in the recognition of the EDA exon by the splicing machinery, which were found to be located in the centre of the EDA exon (15). In fact, deletion or replacement of 81 nucleotides located internally in the EDA exon abolished alternative splicing and resulted in the omission of the EDA exon transcript from the mature fibronectin mRNA.

More recently (28) it has been demonstrated that purine rich sequences within exons were essential for a precise recognition of the exon by the splicing machinery. Two of these exon recognition sequences (ERS) have been identified (28) in the 81 bp region of EDA mentioned above. They are likely targets for non specific splicing factors that identify exon sequences and promote their inclusion in the mature mRNA.

We have now extended the analysis of the fibronectin EDA exon splicing and dissected the specific nucleotide sequences involved in the alternative processes. These sequences seem to be a novel bipartite splicing enhancer.

*To whom correspondence should be addressed

Present addresses: +Prassis-Sigma Tau Research Institute, Via Forlanini 3, 20019 Settimo Milanese, Milan and §Department of Sciences and Biomedical Technologies, S. Raffaele Hospital, Via Olgettina 60, 20132 Milan, Italy

## MATERIALS AND METHODS

### Plasmids construction

The two complementary synthetic oligonucleotides: (5′ TCGACGTACAGGGTGACCTACTCGAGCCCTGAGG-ATGGAATCCATGAGCTATTCCCTGCACCTGATGGTG-AAGAAGACACTGCAGAGCTGCAAGGCC -3′) and (5′ GATCCAGGCCTTGCAGCTCTGCAGTGTCTTCTTCACC-ATCAGGTGCAGGGAATAGCTCATGGATTCCATCCTC-AGGGCTCGAGTAGGTCACCCTGTACG -3′) were annealed generating a synthetic fragment of 99 bp, with the sequence of the central part of the EDA exon and SalI/BamHI sticky ends. The fragment was then cloned into SalI/BamHI digested M13 mp18 obtaining the construct mp ED-A-99. The fragment −1/ED-A (1235 bp) comprising the last 7 bases of the exon−1, the 5′ end of the exon EDA and the −1 intron was obtained after digestion of the vector pSVEDα1W/FN (15) with PstI and BstEII. The synthetic fragment E/−1 (77 bp) carrying rest of the 3′ part of the exon −1 and two sticky ends complementary to digested SalI and PstI sites was obtained after annealing the two synthetic oligonucleotides (5′ TCGACAGTGGAGTATGTGGTTAGTGT-CTATGCTCAGAATCCAAGCGAGAGAGTCAGCCTC-TGGTTCAGACTGCA -3′) and 5′ GTCTGAACCAGAGG-CTGACTCTCTCCGCTTGGATTCTGAGCATAGACA-CTAACCACATACTCCACT/G -3′). The −1/ED-A and the E/−1 fragments were then contemporary cloned into the SalI/BstEII mp ED-A-99 digested construct, originating the mp-1 ED-A.

Finally the ED-A/+1 fragment (1676 bp) that contain the 3′ end of the EDA exon, intron +1 and the 5′ end of exon +1 was obtained from the StuI/BstEII pSVα1W/FN digested vector, blunt ended and cloned into the StuI site of mp-1 ED-A obtaining the final construct M13 mp 18 EDA Tot. From this construct the EDA Tot (3078 bp) fragment was excised with SalI and BamHI, blunt ended and cloned into the BstEII site in the third exon of the α1 globin of the pSVα1W vector (29) obtaining the final construct pSVED-A Tot (Fig. 1).

The pSVED-A Tot Sac vector was obtained cloning a mutated ED-A fragment carrying the SacI site. The construct pSVED-A ΔXho−Stu was obtained after SalI/XhoI and SalI/StuI digestion of the pSVED-A Tot Sac, blunt ending of the XhoI site and ligation of the SalI/XhoI and the SalI/StuI fragments. The vector pSVED-A ΔXho−Sac was obtained excising the XhoI/SacI fragment from the pSVED-A Tot Sac vector and religating it after blunt ending. The pSVED-A ΔSac−Stu construct was obtained after SalI/SacI and SalI/StuI digestion of the pSVED-A Tot Sac, blunt ending of the SacI site and ligation of the two fragments. The constructs pSVED-A Δ1, Δ2c, Δ2e, Δ3, Δ4 were obtained cloning the SalI/SacI fragment, obtained from the vector pSVED-A Tot Sac, into the pTZ18R vector obtaining the pTZ-SS3.9 construct. After SacI/StuI digestion of pTZ-SS3.9 the synthetic fragments Δ1, Δ2c, Δ2e, Δ3, Δ4, carrying the deletions shown in Fig. 2, were ligated obtaining the final constructs were the SalI/SacI fragment was excised and ligated into the previously SalI/SacI pSVED-A Tot Sac digested vector.

### Cell culture, DNA transfection and RNA analysis

HeLa cells were maintained in 60 mm Petri dishes in Dulbecco's MEM containing 10% FCS. Subconfluent cultures were transfected with 2 μg of a plasmid DNA construct and DEAE dextran as already described (30). Cells were cotransfected with an equal amount of pβ5'sVBgl II (31) which contains the large

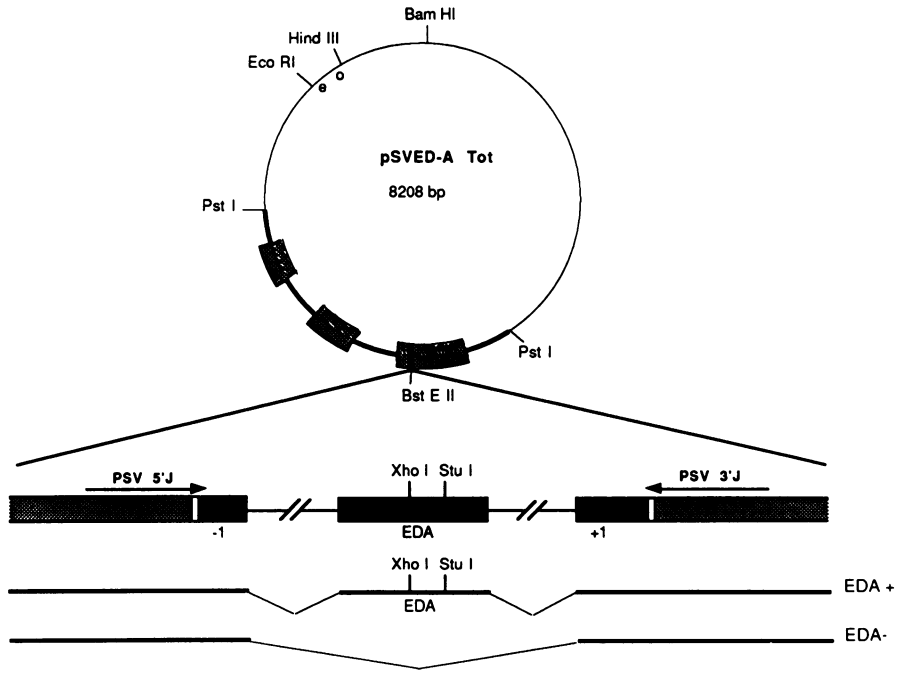


**Figure 1.** Schematic representation of the pSVED-A Tot construct. The EDA human fibronectin exon with its −1 and +1 introns, plus 81 bp at the 3′ end of the −1 exon and 112 bp of the 5′ end of the +1 exon have been inserted in the BstEII site of the third α globin exon of the pSVα1W vector. White boxes indicate polylinker sequences. The position of the primers used in the PCR analysis is shown together with the alternative splicing pattern of the EDA exon. The SV40 enhancer and origin of replication (e, o) are also indicated.

T antigen gene, whose expression is necessary for SV40 replication. The cells were harvested after 24 hours and the total RNA was extracted with guanidinium thiocianate (32). Control mock transfections were carried out in parallel with the vector plasmid DNA.

### RNA analysis

a) RT-PCR: cDNAs were obtained using the 'First Strand cDNA synthesis Kit' supplied from Pharmacia. For each reaction 1 μg of total RNA was used and 0.1 μg of specific primer pSVcDNA (5'-GGTATTTGGAGGTCAGCA-3') was added to the reaction mixture immediately after short denaturation of the RNA sample. The reaction mixture was then incubated for 45 minutes at 37°C. Each cDNA was then analysed by PCR using the primers PSV5'J (5'-CACTGCCTGCTGGTGACTCGA-3') and PSV3'J (5'-GCGGCCAGGGGTCACGAT-3'). Both primers can discriminate the construct transcripts from host globin and fibronectin since the 3' bases match the residual linker sequences and is hence specific for the mRNA transcribed from the construct. The amplification reactions were carried out in a final volume of 50 μl with 1.5 units of TaqPol and 5 μl of 10× Taq buffer supplied from Boheringer Mannheim, dNTPs (Pharmacia) up to a final concentration of 0.2 mM, 1.5 μl of DMSO (Fluka) and the above primers at a final concentration of 1 μM. The amplifications were performed on a Perkin Elmer Cetus DNA Thermo Cycler. Each amplification cycle was carried out as follows: denaturation step 45 seconds at 93°C, annealing step 1 minute at 60°C, extension step 40 seconds at 72°C. The number of rounds of amplification was 28.

10 μl of each amplification reaction were then analysed by electrophoresis on a 1.2% agarose gel, which was stained with ethidium bromide and subject to densitometric analysis. The optical density assigned to each band was then normalised on the basis of the length of the product.

Six independent transfection experiments and their PCR analysis were carried out to ensure the reproducibility of the method.

b) Northern blots: 15 μg of each RNA sample was subject to electrophoresis, blotting and hybridization as previously described (15).

## RESULTS

A hybrid minigene α1globin/fibronectin was constructed as described in methods, preserving the fibronectin gene structure in this region. A synthetic approach was chosen to insert the −1, EDA and +1 exons, including a minimum of linker sequences (Fig. 1 and Materials and Methods). Site directed mutagenesis was carried out to obtain a new SacI site that allowed the manipulation of the central region of the EDA exon. The sequence of the region extending between the XhoI and StuI sites in the centre of the EDA exon is shown in Fig. 2. In order to obtain the relative amount of ED-A+ and ED-A− isoforms by transfection of different constructs into a small number of cells ($10^6$) a PCR-based method has been set up. Specific fibronectin cDNA was synthesized from total RNA extracted from transfectd cells and then amplified using the primers described in Material and Methods (Fig. 1). These primers match specifically the sequences derived from the transfected DNA and are thus able to amplify only the cDNAs synthesized from messengers originated from the constructs and not from the endogenous genes. The amplification gives two different bands, one of 658 bp that corresponds to the EDA+ (EDA included) form, and one of 388 bp that corresponds to the EDA− (EDA excluded) form (Fig. 3). Two versions of the complete sequence, pSVEDA Tot and pSVEDA Tot Sac, produced identical relative proportions of EDA+ and EDA− mRNAs when transfected into Hela cells (lanes 11 and 10—Fig. 3). This demonstrates that the A to C

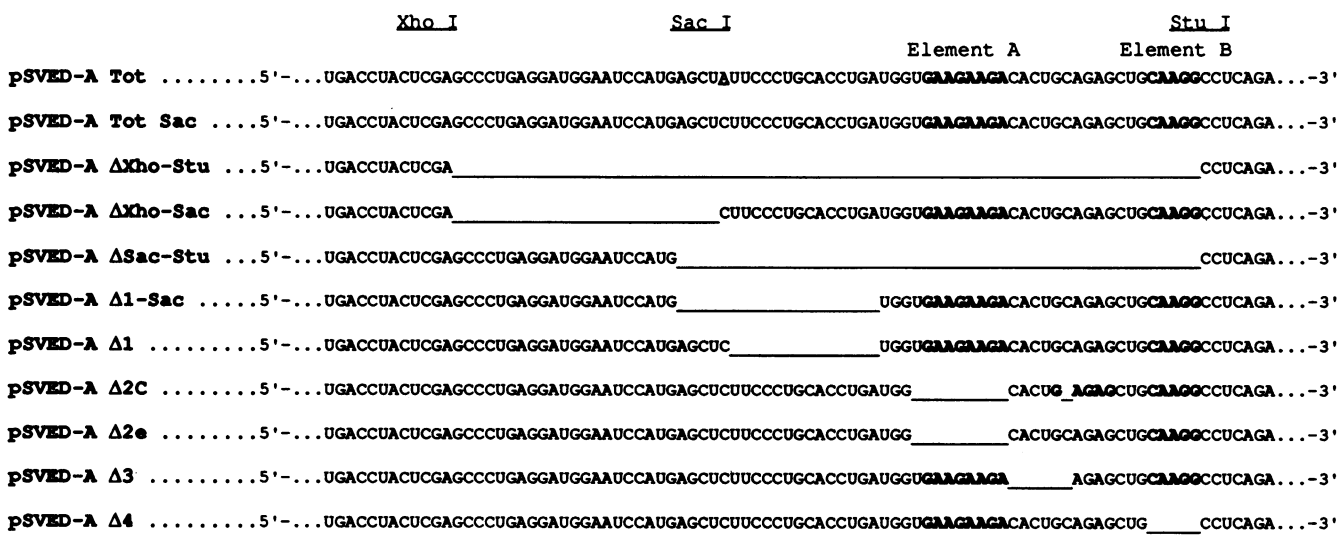


**Figure 2.** XhoI−StuI region in the pSVED-A constructs. The extent of the deletions carried out on the different pSVED-A constructs between the sites XhoI and StuI are reported. The sequences involved in the Δ2e and Δ4 deletions (elements A and B) are in bold characters. The A underlined in pSVED-A Tot has been mutated in C in the other constructs to create the SacI site, used for the construction of the Δ1 to Δ4 small deletions.

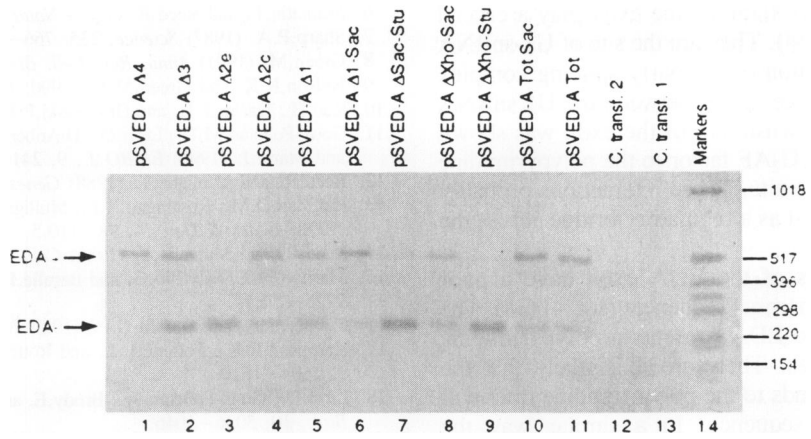


**Figure 3.** RT-PCR analysis of mRNA from HeLa cell transfected with the pSVED-A constructs. The band corresponding to the EDA+ (658 bp) and the EDA− (358 bp) products are indicated by arrows. Lane 13 shows a mock transfection control, lane 12 is the control of transfection with the pβ5'sVBglII plasmid alone. Lanes 1−11 show analysis of the transfection of the different pSVED-A constructs.

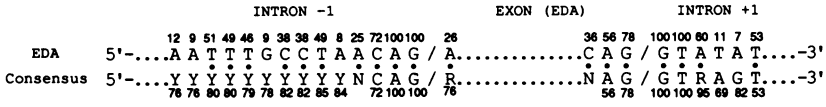


**Figure 4.** Intron−exon junction flanking the EDA exon. Homologies between the EDA and the consensus sequences are showed. Respective base frequencies are also indicated. N any nucleotide, Y pyrimidine, R purine.

mutation introduced to create the SacI site was irrelevant regarding the splicing specificity. Deletion of the Xho/Stu fragment (lane 9) abolished alternative splicing. The deleted sequence corresponds to the 81 bp original deletion described by Mardon *et al.* (15), and thus confirms the previous result. The deletion Xho/Sac (lane 8) affects the relative proportion of EDA+ and EDA− mRNA present, while the Sac/Stu deletion (lane 7) shows a pattern closer to the Xho−Stu deletion although not total abolition of the alternative splicing is observed. The analysis of the effect of deletions obtained with restriction enzymes cannot give a precise picture of the critical sequences. Hence the effect of short 5−10 nucleotide deletions obtained cloning suitable oligonucleotides between the SacI and StuI sites (Fig. 2) was analysed. Both the Δ1 Sac and Δ1 deletions (lanes 6 and 5), although not identical, result in a dominance of the EDA+ form, which is a wild type like pattern. Of particular interest is the Δ2e construct (lane 3) which has the same effect as the extensive Xho/Stu deletion (lane 9). The Δ2 deletion eliminates a polypurine sequence already pointed out (28) as a critical exon sequence for splicing. The Δ2c construct is identical to Δ2e but a C has been deleted, so forming again a continuous stretch of 5 purines that results in the restoration of a pattern close to the wild type (Fig. 3, lane 4). It is clear that a polypurine sequence at that position is essential for exon recognition, while a similar polypurine contained in the ΔXho/Sac deletion (GAGG-A) does not seem to have such a striking effect. The Δ3 deletion (lane 2) involves the region including the C deleted in Δ2c and produces a pattern with a predominance of the EDA− form.

A completely distinct pattern is observed with the Δ4 deletion (lane 1). With this sequence alternative splicing is abolished but, contrary to the effect of the Δ2e deletion, the EDA exon is always included in the messenger. The data obtained by PCR was confirmed by Northern blot analysis. Both methods gave consistently the same mRNA pattern, although quantitation of the bands was very difficult in Northern blots due to a high background (data not shown). Additionally a well represented pre mRNA was seen by Northern blot, particularly when both EDA isoforms are present, as previously reported (15).

## DISCUSSION

The results described here clearly show that exon elements regulate the relative proportion of EDA+ and EDA− fibronectin mRNA synthesized. Cis acting splice elements seem to belong to two types: element A (GAAGAAGA) is essential for exon recognition and its removal produces omission of the EDA exon, while removal of element B (CAAGG) results in 100% inclusion of the EDA exon (Figs. 2 and 3). It is interesting to note that eventual trans acting factor interactions may not only be a consequence of nucleotide sequence but a positional effect may also be operative. In fact, the sequence GAGGA (28), does not have the effect of the polypurine stretch of element A, as can be seen when comparing the patterns produced by the Xho/Sac and Sac/Stu deletion (Fig. 3 lane 8 and 7). Consistent with this interpretation, the Δ2c deletion reconstitutes a polypurine stretch with a different sequence but in the same position of element A and restores a transcription pattern similar to that of the wild type.

To understand the role of elements A and B in the splicing process, other factors involved should be considered. The transcription of an exon is the result of a complex network of interactions between protein factors, pre mRNA and snRNPs.

Both the 5′ and 3′ splice sites flanking the exon play a crucial role in these interactions (33,34). They are the site of U1-snRNA interaction during the formation of the early splicing complex (35–37). Furthermore, the interaction between the U1 snRNA bound at the 5′ splice site downstream of the exon was shown to enhance the binding of the $U_2AF$ factor to the polypyrimidine tract of the upstream 3′ splice site. These interactions probably involve other factors which act as a regulatory bridge across the exon (35).

The 5′ and 3′ splice sites of the EDA exon have a poor homology with the known consensus sequences (fig. 4) (38). The 3′ splice site upstream of the EDA presents a polypyrimidine stretch interrupted by purines. Therefore it is likely that the auxiliary factor $U_2AF$ that binds to the polypirimidine tract will have low affinity for this sequence. In a similar way the downstream 5′ splice site is not conserved at the positions +4 and +5. Particularly, the substitution of the consensus G in position +5 is known to produce exon skipping in several systems (33,36). Moreover, the 5′ splice junction downstream of the EDA exon has a structure similar to some disease-causing mutations that produce exon skipping in β thalassemia (39) and Lesch–Nyhan syndrome (40). The 5′ and 3′ splice sites flanking the EDA may then make exon definition dubious in the early splicesome complex. This is probably overcome by the presence of the splicing enhancer sequence GAAGAAGA, which promotes an accurate recognition and definition of the exon by the splicing machinery.

We have shown here that the EDA exon system has an added complexity. The exon recognition and definition is achieved not only by a positive mechanism but seems to involve a 'repressor', the element B sequence. This is not surprising, since the control of EDA alternative splicing and its specific tissue and developmental modulation would be difficult to achieve by a one way system. To balance exon recognition there seems to exist an interaction at the site of the Δ4 deletion (element B) that prevents recognition perhaps by disrupting the formation of the already weak early splicing complexes with one or both of the flanking splice sites. The element B seems to be recessive to the element A as simultaneous deletion of both (ΔSac–Stu) results in a pattern of EDA splicing similar although not identical to the one produced by deletion of element A alone (Δ2c). Models of negatively regulated splicing have already been described for the exon 7 of β rat tropomyosin (41) and the Drosophila P element (42), where it has been shown that a tissue specific protein binds to a target RNA sequence in the exon (P element) or in its neighbouring introns (tropomyosin) and causes exon skipping. The element B mapped in this work is the first mammalian example of exon sequences involved in negative regulation. A detailed analysis of splicesome formation, particularly of the splice factors' interactions with both the elements A and B of the bipartite enhancer is currently under way.

## REFERENCES


1. Kornblihtt,A.R., Vibe-Pedersen,K. and Baralle,F.E. (1984) *Nucleic Acid Res.*, 12, 5853–5868.
2. Kornblihtt,A.R., Umezawa,K., Vibe-Pedersen,K. and Baralle,F.E. (1985) *EMBO J.*, 4, 1755–1759.
3. Zardi,L., Carnemolla,B., Siri,A., Petersen,T.E., Paolella,G., Sebastio,G. and Baralle,F.E. (1987) *EMBO J.*, 6, 2337–2342.
4. Barone,M.V., Henchcliffe,C., Baralle,F.E. and Paolella,G. (1989) *EMBO J.*, 8, 1079–1085.
5. Schwarzbauer,J.E., Patel,R.S., Fonda,D. and Hynes,R.O. (1987) *EMBO J.*, 6, 2573–2581.
6. Maniatis,T. and Reed,R. (1987) *Nature*, 325, 673–678.
7. Sharp,P.A. (1987) *Science*, 235, 766–771.
8. Green,M. (1991) *Annu. Rev. Cell. Biol.*, 7, 559–599.
9. Nelson,K.K. and Green,M.R. (1990) *PNAS*, 87, 6253–6257.
10. Kuo,H., Nasim,F.H. and Grabowski,P.J. (1991) *Science*, 251, 1045–1050.
11. Goux-Pelletan,M.D., Libri,D., DiAubenton-Carafa, Fiszman,M., Brody,E. and Marie,J. (1990) *EMBO J.*, 9, 241–249.
12. Reed,R. and Maniatis,T. (1988) *Genes & Dev.*, 2, 1268–1276.
13. Helfman,D.M., Roschigno,R.F., Mulligan,G.J., Finn,L.A. and Weber,K.S. (1990) *Genes & Dev.*, 4, 98–110.5,
14. Reed,R. and Maniatis,T. (1986) *Cell*, 46, 681–690.
15. Mardon,H.J., Sebastio,G. and Baralle,F.E. (1987) *Nucleic Acid Res.*, 15, 7725–7733.
16. Cooper,T.A. and Ordahl,C.P., (1989) *Nucleic Acid Res.*, 17, 7905–7921.
17. Hampson,R.K., Follette,L.L. and Rottman,F.M. (1989) *Mol. Cell. Biol.*, 9, 1604–1610.
18. Libri,D., Goux-Pelletan,M., Brody,E. and Fiszman,M.J. (1990) *Mol. Cell. Biol.*, 10, 5036–5046.
19. Streuli,M. and Saito,H. (1989) *EMBO J.*, 8, 787–796.
20. Steitz,J.A., Black,D.L., Gerke,V., Parker,K.A., Kramer,A.,Frendwey,D. and Keller,W. (1988) In Birnstiel,M.L., (ed.), Structure and Function of Major and Minor Small Nuclear Ribonuclear Protein Particles. Springer Verlag, Berlin and Heidelberg.
21. Smith,C.W.J., Patton,J.G. and Nadal-Ginard,B. (1989) *Annu. Rev. Gen.*, 23, 537–577.
22. Valcàrel,J., Ravinder,S., Zamore,P.D. and Green,M. (1993) *Nature*, 362, 171–175.
23. Tiang,M. and Maniatis,T. (1993) *Cell*, 74, 105–114.
24. Zahler,A.M., Neugebauer, K.M., Lane,W.S. and Roth,M.B. (1993) *Science*, 260, 219–222.
25. Fu,X.D. (1993) *Nature*, 365, 82–85.
26. Vibe-Pedersen,K., Kornblihtt,A.R. and Baralle,F.E. (1984) *EMBO J.*, 3, 2511–2516.
27. Paolella,G.,Henchcliffe,C.,Sebastio,G. and Baralle F.E. (1988) *Nucleic Acid Res.*, 16, 3545–3557.
28. Watakabe,A. Kenji,T. and Shimura,Y. (1993) *Genes & Dev.*, 7, 407–418.
29. Higgs,D.R., Goodburn,S.E.Y., Lamb,J., Clegg,J.B., Weatherall,D.J. and Proudfoot,N.J., (1983) *Nature*, 306, 398–400.
30. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecolar Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
31. Grosveld,G.C., de Boer,E., Shewmaker,L.K. and Flavell,R.A. (1983) *Nature*, 295, 120–126.
32. Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, 162, 156–159.
33. Talerico,M. and Berget,S.M. (1990) *Mol. Cell. Biol.*, 10, 6299–6305.
34. Robberson,D.L. Cote,G.J. and Berget,S.M. (1990) *Mol. Cell. Biol.*, 10, 6299–6305.
35. Hoffman,B.E. and Grabowski,P.J. (1992) *Genes & Dev.*, 6, 2554–2568.
36. Nelson,K.K. and Green,R.M. (1990) *PNAS*, 87, 6253–6257.
37. Michaud,S. and Reed,R. (1993) *Genes & Dev.*, 7, 1008–1020.
38. Shaphiro,M.B. and Senapathy,P. (1987) *Nucleic Acid Res.*, 15, 7158–7174.
39. Krainer,A.R., Maniatis,T.,Ruskin,B. and Green,M.R. (1984) *Cell*, 36, 993–1005.
40. Gibbs,R.A., Nguyen,P., Edwards,A.E., Civitello,A.B. and Caskey,C.T. (1990) *Genomics*, 7, 235–244.
41. Guo,W., Mulligan,G.J., Wormsley,S. and Helfman,D. (1991) *Genes & Dev.*, 5, 2096–2107.
42. Siebel,C.W. and Rio,D.C. (1990) *Science*, 248, 1200–1207.