The EMBO Journal vol.11 no.3 pp.1095-1098, 1992

Splicing with inverted order of exons occurs proximal to large introns

Claude Cocquerelle, Pierre Daubersies, Marie-Ange Majérus, Jean-Pierre Kerckaert and Bernard Bailleul[1]

Unité 124 INSERM, Institut de Recherches sur le Cancer, Place de Verdun, 59045 Lille Cedex, France

[1]Corresponding author

Communicated by E.Brody

Following our studies which showed that the α and β exons of the chicken c-*ets*-1 gene are not conserved in the human homologue, we succeeded in identifying a novel human c-*ets*-1 transcript in which the normal order of exons is scrambled. By PCR and RNase protection assays, we demonstrated that while the order of exons is different from that in genomic DNA, splicing of these exons in aberrant order occurs in pairs and at the same conserved consensus splice sites used in the normally spliced transcript. The scrambled transcript is non-polyadenylated and is expressed at much lower levels than the normal transcript. It is not the consequence of genomic rearrangement at the *ets*-1 locus nor is it due to the transcription of any *ets*-1 pseudogene. These results confirm previous observations of scrambled splicing.

Key words: alternative splicing/intron/oncogene/c-*ets*-1

Introduction

The c-*ets*-1 proto-oncogene encodes a transcription factor that cooperates with AP-1 in transcriptional activation (Wasylyk *et al.*, 1990). The chicken c-*ets*-1 locus encodes two alternatively spliced transcripts which encode $p54^{ets\text{-}1}$ and $p68^{ets\text{-}1}$ respectively (Duterque-Coquillaud *et al.*, 1988; Gegonne *et al.*, 1987). $p54^{ets\text{-}1}$ is homologous to the human $p51^{ets\text{-}1}$ (Watson *et al.*, 1988). $p68^{ets\text{-}1}$ contains two alternatively spliced exons, α and β (Leprince *et al.*, 1988), which to date have not been detected in human. It is the $p68^{ets\text{-}1}$ transcript which has been transduced by the E26 retrovirus (Leprince *et al.*, 1983; Nunn *et al.*, 1983). PCR experiments were performed in order to determine whether these exons exist in human. We failed to detect exons α and β in the human genome but in the process of our studies, we detected abnormally spliced products. In these transcripts, exon d or c splice donor sites were joined at the exon a1 splice acceptor site.

Splicing is mediated by a RNA–protein complex, the spliceosome, which catalyses splicing between exons at consensus splice sites (Steitz *et al.*, 1987; Brody and Abelson, 1985). As many pre-mRNAs contain more than one intron, a mechanism has been postulated which ensures that cognate 5′ and 3′ splice sites are appropriately used such that aberrant splicing does not occur (Aebi and Weissmann, 1987).

In this paper we present evidence from PCR and RNase protection experiments that pairs of exons of the human *ETS*-1 gene are joined at consensus splice sites, but in a different order from that present in genomic DNA. Similar observations have recently been described for the *DCC* gene (deleted in colorectal carcinomas) in which the normal order of exons is scrambled (Nigro *et al.*, 1991). It is noteworthy that abnormal splicing occurs only with *ETS*-1 exons c, d and a1. These exons are close to the largest introns (Kerckaert *et al.*, 1990; Jorcyk *et al.*, 1991) and the exon d skips during pre-mRNA maturation, at relatively high frequency (Jorcyk *et al.*, 1991; Reddy and Rao, 1988).

Results and discussion

We performed double PCR amplifications with anti-sense primers to exon a1 and with a sense primer to exon I of human *ETS*-1 (Reddy and Rao, 1988) or with sense primers homologous to the additional 5′ exons, α and β, that were identified in the chicken gene (Leprince *et al.*, 1988) (Figure 1a and b). PCR products of the expected size (93 bp) amplified from the *ETS*-1 6.8 kb mRNA (Reddy and Rao, 1988; Watson *et al.*, 1985) were found in different cell lines, including CEM, HeLa and A431 (Figure 1c). Using the oligomer O27 (27mer homologous to chicken exon β) as sense primer, double PCR amplification of cDNA of different cell lines yielded an amplification product of 158 bp (Figure 1d). We verified that the products obtained from different cell lines were identical by Southern blotting hybridization (Figure 1d). To do this, we used a probe derived from a PCR-amplified product and showed that it hybridized on pulsed field electrophoresis blots, to a specific 400 kb *Nru*I restriction fragment containing the *ETS*-1 locus (data not shown). We concluded from these data that we have amplified an identical product in the different cell lines which is very probably an *ETS*-1 transcript.

The sequence of the 158 bp amplification product revealed that the 5′ exon amplified from the exon a1 was not the 5′ alternative exon as expected but to our surprise, the sequence was identical to exon d of *ETS*-1 (Reddy and Rao, 1988) which is 10 kb downstream of exon a1 (Kerckaert *et al.*, 1990; Jorcyk *et al.*, 1991). To explain this result, sequence comparison between primer O27 and exon d was performed. The result demonstrated high homology between O27 and exon d and this homology probably allowed for PCR amplification in the low stringency conditions used. We confirmed that no rearrangement of the *ETS*-1 locus had occurred in any of the cell lines used to prepare cDNA; indeed, probes from exon a1 and d show a single band in Southern blot hybridization experiments (data not shown). Similarly we failed to identify any potential pseudogene which may explain this transcript.

We designed PCR experiments using high stringency conditions with a new set of primers derived from sequences within exon d and a1 to determine whether the same products were obtained (Figure 2a). The PCR product of the predicted

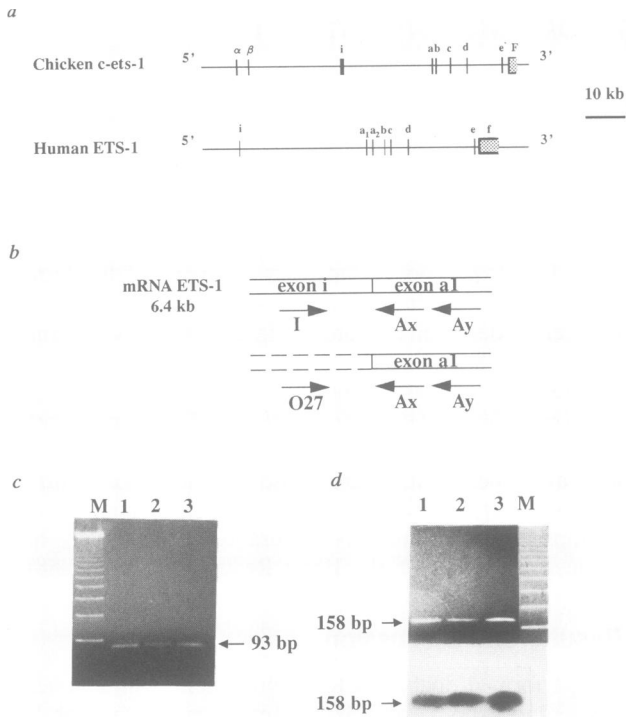

Fig. 1. Two stage PCR experiments to detect 5′ alternative exons to exon a1. (**a**) Comparison of genomic organization of chicken and human *ets*-1 loci. (**b**) The diagram shows the set of primers used in the two-step amplification strategy from total RNA, for control amplifications and for studying 5′ alternative exons. The primers I and O27 are homologous to exon i of the human *ETS*-1 and to exon β of the chicken *ets*-1, respectively. Primers to exon a1 of *ETS*-1 (Ax and Az) are common to both reactions. (**c**) Amplification results with the primers I; an amplified product of 93 bp is observed. (**d**) Amplification results with the primers O27. RNAs show an amplified product of 158 bp. (Bottom) Southern blot of the agarose gel above, using the 158 bp amplification product of A431 cell line as a probe (1, CEM; 2, A431. 3, HeLa. M; multimer of 123 bp).

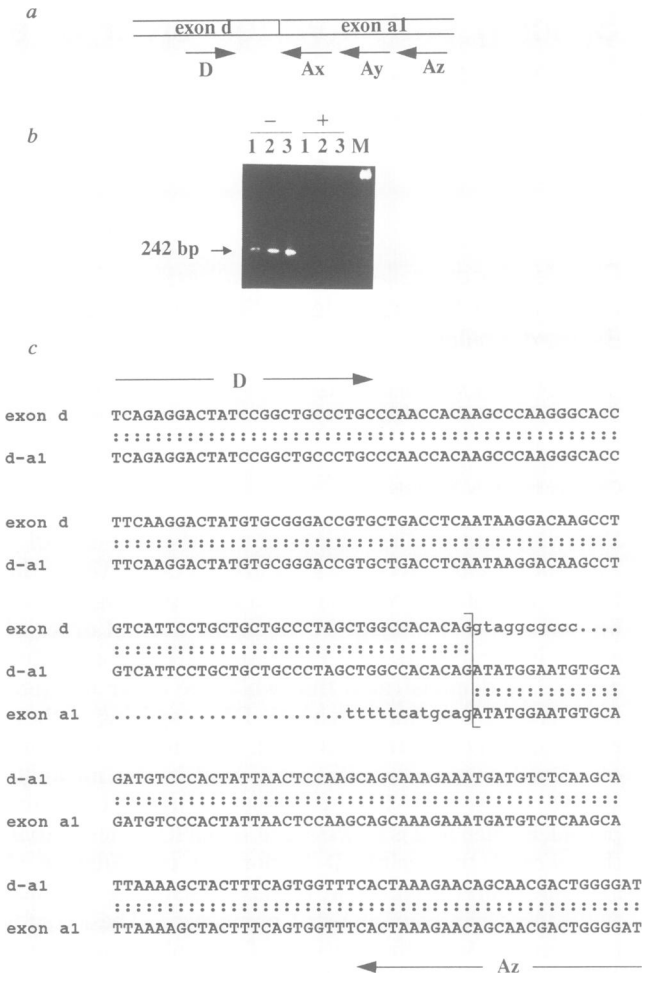

Fig. 2. PCR amplification and sequence comparison of scrambled splicing product with genomic DNA. (**a**) Diagram shows the new primers D and Az used to amplify the scrambled splicing product. Both primers were totally equivalent to their respective exons. (**b**) PCR analysis using one-step amplification with primers D and Az, onto poly(A)$^+$ (+) and poly(A)$^-$ (−) RNA. 1, CEM; 2, A431; 3, HeLa; M, marker. (**c**) Sequence comparison of the amplification product and genomic sequences. Exon−exon and exon−intron junctions are indicated by a vertical line. Homology is indicated by dots. Intron sequences are in small letters. Arrows correspond to primers.

size (242 bp) was found after 30 cycles from the poly(A)$^-$ RNA and was not observed with poly(A)$^+$ RNA from the three cell lines (Figure 2b). Sequence comparison of this 242 bp amplified product, with genomic clones of exons a1 and d was carried out and showed 100% homology between the PCR product and the genomic exons a1 or d which both end correctly at the exon−intron junction (Figure 2c). This is in agreement with a splicing reaction occurring between splice donor site of exon d and splice acceptor site of exon a1 (i.e. no mutation or base addition was observed, and both cleavage positions of the donor and acceptor splicing sites are correct). We concluded from these data that this product is the result of an abnormal splicing reaction, in which the predicted order of exons is scrambled relative to the genomic order of exons. Additionally, this product is not polyadenylated and consensus splice sites are used.

To rule out the possibility that amplified and cloned sequences were artefacts from PCR, we performed RNase protection assays which allowed quantification of the scrambled transcript with regard to the normal transcript. An antisense RNA probe, that included the scrambled transcript containing exon a1 and d (Figure 3a), was hybridized to total RNA. The RNase-resistant products were analysed by gel electrophoresis. Protected fragments corresponding to the abnormally spliced product (242 bp) and normally spliced products (130 and 122 bp for exon d and a1 respectively), were observed (Figure 3b). This result is the first observation of scrambled exons by an RNase protection experiment. The level of scrambled transcript was estimated to be close to 0.01 the level of the normally spliced transcript.

Recently, similar abnormally spliced transcripts were identified for a candidate tumour suppressor gene, the *DCC* gene, by PCR experiments (Nigro *et al.*, 1991) and the authors also demonstrated that these products are not polyadenylated. Both results clearly establish that scrambling of exons may occur during the splicing process *in vivo*. In both cases, the scrambled exons were joined at the same splice sites, with the same precision as normal splicing (Mount, 1982), except that the exon order was changed. The first step in assembly of the spliceosome involves splice site recognition (Robberson *et al.*, 1990) and is carried out normally, but during scrambled splicing, the subsequent step, which pairs sequential exons appropriately, is executed aberrantly. Aberrant annealing of homologous sequences or

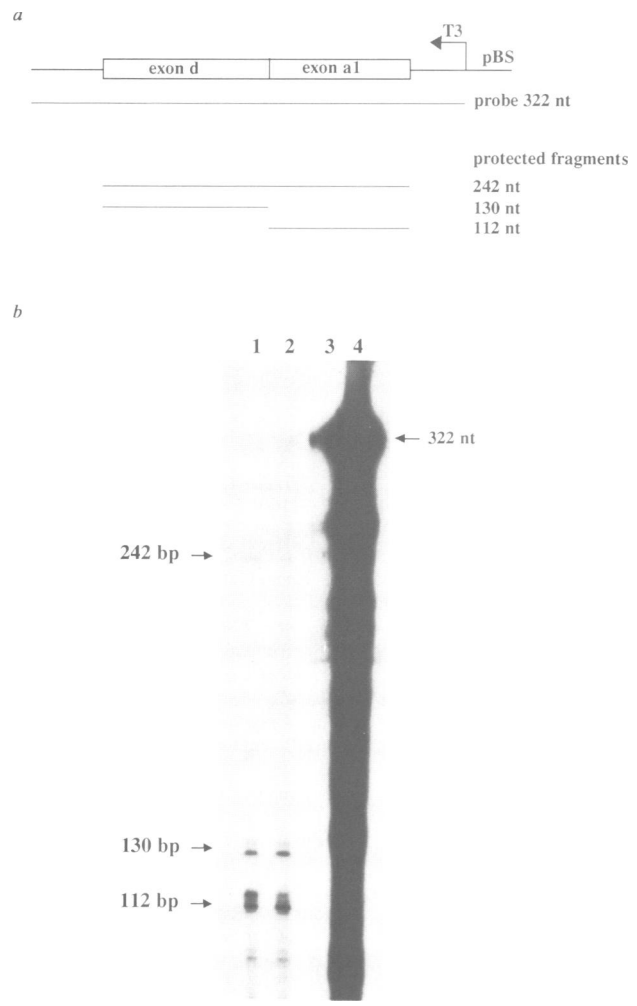

Fig. 3. RNase protection analysis of scrambled splicing product. (**a**) The top line shows the construct used to generate specific probe. Open boxes indicate exon structure in the scrambled splicing product and the lines correspond to the polylinker of the pBluescript II SK−vector. The second line shows the RNase protection probe of 322 nucleotides. The expected protected fragments over the scrambled splicing product, exon a1 and exon d are drawn in the third, fourth and fifth lines, respectively, and the sizes are indicated. (**b**) Autoradiogram of RNase protection analysis products of 20 μg total RNA sample from CEM cell line (1−2), 20 μg tRNA (3) and undigested probe (4). Arrows indicate protected fragments which correspond, from top to bottom, to the scrambled product and the normally spliced exons d and a1.

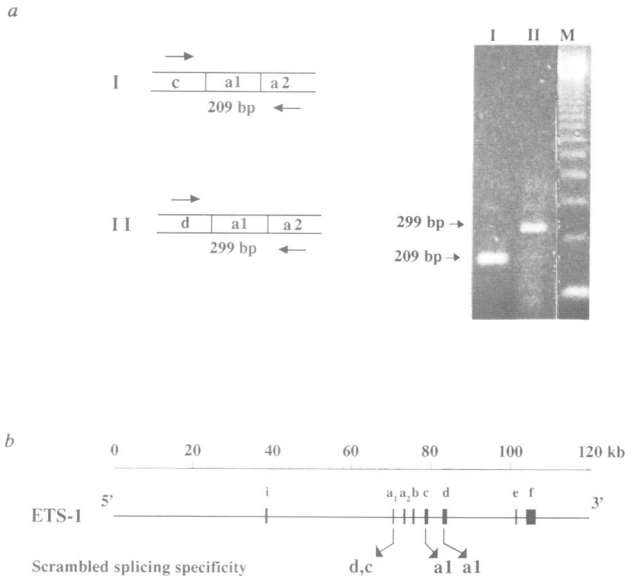

Fig. 4. Specificity of scrambling splicing. (**a**) Only two scrambled products (I and II) were observed by PCR from CEM total RNA, using a combination of primers which could allow visualization of scrambled products between donor sites of the exons c, d and e and the acceptor splice site of the exons a1, a2, b and c. These products correspond to splicing of exon d or exon c with exon a1 which normally splices 3′ with exon a2. M, 123 bp mutimer. (**b**) Genomic organization of the human *ETS*-1 locus and location of the splice sites used by scrambled splicing (see arrows).

fortuitous protein interactions within introns leading to a particular secondary structure of the pre-mRNA, might facilitate illegitimate interactions of snRNPs to incorrectly paired exons. A possible mechanism for intramolecular splicing generating abnormally spliced product has been proposed (Nigro *et al.*, 1991) which would generate circular RNAs, but an intermolecular reaction is also possible. *Trans*-splicing mechanisms described in lower eukaryotes (Agabian, 1990) bear no comparison to the scrambled splicing observed here.

Using a combination of sense and antisense primers to the different exons of *ETS*-1, which could allow visualization of scrambled products between donor splice sites of exons c, d, e and the acceptor site of the exons a1, a2, b and c, we performed PCR to determine whether scrambling may be observed with other exons of this locus. We observed only an additional unexpected product of splicing which involved the donor splice site of exon c and, as before, the acceptor splice site of exon a1 (Figure 4a, product I). Both scrambled transcripts I and II were detected using primers to exons a2 (Figure 4a) and in these experiments scrambled products between exon c or d and exon a2 were never observed. Other PCR experiments to detect scrambled splicing between the other exons were unsuccessful. Thus, only the acceptor splice site of exon a1 and the donor splice sites of exons c and d appear to participate in the scrambled splicing mechanism. The specificity observed in the scrambled splicing is represented in Figure 4b. It is noteworthy that the intervening sequences downstream of exon d and upstream of exon a1 are very large: 15 kb and 30 kb respectively (Kerckaert *et al.*, 1990; Jorcyk *et al.*, 1991). In addition, alternative splicing of exon d has been demonstrated in different cell lines (Jorcyk *et al.*, 1991; Reddy and Rao, 1988; Koizumi *et al.*, 1990) and this frequently occurs at a significant level with regard to the normal transcript (J.P.Kerckaert, unpublished data). It is interesting to note that the *DCC* gene also contains large intervening sequences (Fearon *et al.*, 1990). This gene shares sequence homology with the *NCAM* gene (Fearon *et al.*, 1990) which expresses a combination of alternatively spliced forms (Reyes *et al.*, 1991) and it is possible that some exons of *DCC* gene are alternatively spliced. A relationship between scrambled splicing and alternative splicing may exist.

By PCR amplification of aliquots from sucrose gradient separation of total RNA, we observed that the scrambled product (242 bp amplified product as in Figure 2b) is only detected in one or two fractions corresponding to ~2 kb size (data not shown). Although the *ETS*-1 pre-mRNA spans over 70 kb (Kerckaert *et al.*, 1990; Jorcyk *et al.*, 1991),

it is unlikely that the signal observed results from degradation of a larger product which should be detected in more than two fractions of the sucrose gradient separation. In addition, the amplification of scrambled products I and II (Figure 4a) demonstrated that normal splicing occurs downstream of the scrambled exons. It may be concluded that the 2 kb molecule containing scrambled exons (d−a1) which came from large pre-mRNAs, has been processed by splicing to eliminate large intervening sequences. To date, we do not know the 5′ and 3′ limits of the aberrant transcripts. The open reading frame is preserved by joining exon c or d with exon a1 but it is unlikely that this has any biological relevance since in the case of scrambled exons of the *DCC* gene, two out of three products lose the open reading frame (Nigro *et al.*, 1991).

The expression of alternatively spliced forms is regulated according to the cell environment or the stage of development (Breitbart *et al.*, 1987) but relatively little is known about the mechanisms involved. The discovery and the study of scrambled exons may be a means to understanding the mechanisms responsible for alternate process and order in splicing.

Materials and methods

PCR amplification

Total RNA was prepared from different cell lines using the guanidium thiocyanate−caesium chloride method as previously described (Glisin *et al.*, 1974). Poly(A)$^+$ and poly(A)$^-$ RNA was isolated from total RNA using double passage over oligo(dT)−cellulose chromatography. Poly(A)$^+$ RNA constituted 7.5% of the total RNA. 1.5 μg RNA was reverse transcribed in 10 μl containing 200 ng Ay primer, 1 μg BSA, 1.5 mM each dNTP, 50 mM Tris−HCl pH 8.3, 75 mM KCl, 10 mM DTT, 3 mM $MgCl_2$ and 100 U MuMLV reverse transcriptase (Gibco-BRL) for 1 h at 37°C. Reverse transcription mixture was included in a final volume (100 μl) of PCR reaction containing 100 mM Tris−HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 1 U *Taq* polymerase (Stratagene) and 200 ng of 5′ primer (I or O27). For second stage of PCR, 1 μl of first step reaction was added to a second PCR mixture containing 200 ng of each primer (Ax and I or Ax and O27). Amplification was performed by 30 cycles: 92°C (1 min), 45 or 65°C (2 min) for the first and second stage of PCR respectively and 72°C (2 min) in a Techne PHC-2 thermocycler. PCR reactions shown in Figures 2 and 4 were performed as previously, but with a 65°C hybridization temperature. Sequences of the primers are: I, 5′-GATCTCAAGCCGACTCTCAC-CATCATC-3′; O27, 5′-TGCTGTGAGGACCCCTGGATGCCATGC-3′; Ax, 5′-TAGTGGGACATCTACACATTCCATATC-3′; Ay, 5′-TCTT-TGCTGCTTGGAGTTAATAGTGGG-3′; Az, 5′-ATCCCCAGTCG-TTGCTGTGTTCTTTAGTG-3′; D, 5′-TCAGAGGACTATCCGG-CTGCCCTGC-3′; exon a2 primer, 5′-GCCCACATCACCCAGTC-CCGAA-3′; and exon c primer, 5′-GTCACCCCAGACAACATGTG-3′.

Agarose gel analysis

A 10 μl aliquot of the PCR reaction was electrophoresed on a 2% agarose−TBE gel. The RNA was transferred onto a nylon membrane (Biotrace N) according to the protocol of the manufacturer and was then hybridized with 10^6 c.p.m. of 158 bp probe. The membrane was autoradiographed overnight using Hyperfilm^{TM} MP (Amersham). The 158 bp probe was prepared as follows: 1 μl amplification reaction obtained from A431 RNA (diluted 20 times) was added to a 20 μl PCR mixture as described, with 200 ng of primer Ax and O27, 200 μM dATP, dGTP, dTTP and 6 μl [^{32}P]dCTP (10 μCi/μl).

Cloning and sequencing

The 242 bp amplification product was cloned into the *Sma*I site of pBluescriptII SK-vector (Stratagene) and sequenced by the dideoxynucleotide method. For genomic sequences, subclones were obtained and sequenced by conventional methods from 10 overlapping recombinant clones which cover 120 kb of the *ETS*-1 locus (Kerckaert *et al.*, 1990 and our data).

RNase protection assay

The RNase protection was performed as previously described (Bailleul *et al.*, 1990). The amplification product of 242 bp (Figure 2c) was cloned into the *Sma*I site of pBluescript II SK−. The plasmid was linearized with *Hin*dIII and used as a template for T3 RNA polymerase (Stratagene). 20 μg of total RNA was hybridized to 1 × 10^5 c.p.m. of [^{32}P]uridine-labelled probe, overnight at 45°C. After RNase A and T1 digestion at 37°C for 30 min, samples were analysed by electrophoresis on a 6% acrylamide−8 M urea gel, and exposed to Cronex X-ray film in the presence of intensifying screens for three days.

Acknowledgements

We thank M.H.Loucheux-Lefebvre for encouragement and support, M.Collyn-d'Hooghe for discussion, S.Galiègue-Zouitina for pulsed field experiments and K.MacLeod for help in manuscript preparation. C.C. was supported by Institut de Recherches sur le Cancer de Lille and Région Nord-Pas de Calais (Conseil Régional) fellowship. This work was supported by the Association pour la Recherche sur le Cancer (ARC).

References

Aebi,M. and Weissmann,C. (1987) *Trends Genet.*, **3**, 102−107.
Agabian,N. (1990) *Cell*, **61**, 1157−1160.
Bailleul,B., Surani,M.A., White,S., Barton,S.C., Brown,K., Blessing,M., Jorcano,J. and Balmain,A. (1990) *Cell*, **62**, 697−708.
Breitbart,R.E., Andraedis,A. and Nadal-Ginard,B. (1987) *Annu. Rev. Biochem.*, **56**, 467−495.
Brody,E. and Abelson,J. (1985) *Science*, **228**, 963−967.
Duterque-Coquillaud,M., Leprince,D., Flourens,A., Henry,C., Ghysdael,J., Debuire,B. and Stehelin,D. (1988) *Oncogene Res.*, **2**, 335−344.
Fearon,E.R. *et al.* (1990) *Science*, **247**, 49−56.
Gegonne,A., Leprince,D., Duterque-Coquillaud,M., Vandenbunder,B., Flourens,A., Ghysdael,J., Debuire,B. and Stehelin,D. (1987) *Mol. Cell. Biol.*, **70**, 806−812.
Glisin,V., Crkvenjakov,R. and Byus,C. (1974) *Biochemistry*, **13**, 2633−2637.
Jorcyk,C.L., Watson,D.K., Mavrothalassitis,G.J. and Papas,T. (1991) *Oncogene*, **6**, 523−532.
Kerckaert,J.P., Duterque-Coquillaud,M., Collyn-d'Hooghe,M., Morel,P., Majérus,M.A., Laï,J.L., Fenaux,P., Bauters,F., Debuire,B. and Loucheux-Lefebvre,M.H. (1990) *Leukemia*, **4**, 16−19.
Koizumi,S., Fisher,R.J., Fujiwara,S., Jorcyk,C., Bhat,N.K., Seth,A. and Papas,T.S. (1990) *Oncogene*, **5**, 675−681.
Leprince,D., Gegonne,A., de Taisne,C., Schnuberger,A., Lagrou,C., and Stehelin,D. (1983) *Nature*, **306**, 395−397.
Leprince,D., Duterque-Coquillaud,M., Li,R.P., Henry,C., Flourens,A., Debuire,B. and Stehelin,D. (1988) *J. Virol.*, **52**, 3233−3241.
Lewin,B. (1980) *Cell*, **22**, 324−326.
Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459−472.
Nigro,J.M., Cho,K.R., Fearon,E.R., Kern,S.E., Ruppert,J.M., Oliner,J.D., Kinzler,K.W. and Vogelstein,B. (1991) *Cell*, **64**, 607−613.
Nunn,M.F., Seeburg,P.P.M., Moscovici,C. and Duesberg,P.H. (1983) *Nature*, **306**, 391−395.
Reddy,E.S.P. and Rao,V.N. (1988) *Oncogene Res.*, **3**, 239−246.
Reyes,A.A., Small,S.J. and Akeson,R. (1991) *Mol. Cell. Biol.*, **11**, 1654−1661.
Robberson,B.L., Cote,G.J. and Berget,S.M. (1990) *Mol. Cell. Biol.*, **10**, 84−94.
Sharp,P.A. (1987) *Science*, **235**, 766−771.
Steitz,J.A. (1987) In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer-Verlag, pp 71−99.
Wasylyk,B., Wasylyk,C., Flourens,P., Begue,A., Leprince,D. and Stehelin,D. (1990) *Nature*, **346**, 191−193.
Watson,D.K., McWilliams-Smith,M.J, Nunn,M.F., Duesberg,P.H., O'Brien,S.J. and Papas,T.S. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7294−7298.
Watson,D.K., McWilliams,M.J., Lapis,P., Lautenberger,J.A., Schweinfest,C.W. and Papas,T.S. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7862−7866.

Received on October 10, 1991; revised on November 25, 1991