Genes encoding the $\beta 1$ subunit of voltage-dependent Na^+ channel in rat, mouse and human contain conserved introns

Sulayman D. Dib-Hajj, Stephen G. Waxman*

Department of Neurology, Yale University School of Medicine, New Haven, CT 06510, and Neuroscience Research Center, VA Medical Center, West Haven, CT 06516, USA

Received 9 November 1995

Abstract We provide evidence in this study that the 86-bp insert in the $\beta_{1.2}$ mRNA isoform of the voltage gated sodium channel is an intron. Transcripts still retaining this intron were detected in all tissues where the $\beta 1$ gene expression was investigated. We also show that the exon/intron boundaries of the last two introns are conserved among rat, mouse and human $\beta 1$ gene. Unlike the highly conserved cDNAs, introns in only the rat and mouse genes are highly related. The last intron is very short (86–90 bp) and is located in the 3′ untranslated sequence, both uncommon properties of mammalian pre-mRNA introns.

Key words: Voltage-gated sodium channel; $\beta 1$ subunit; Gene structure; Intron retention

1. Introduction

Voltage-gated Na^+ channels from rat brain are composed of three subunits: α (260 kDa), $\beta 1$ (36 kDa) and $\beta 2$ (33 kDa) [1,2]. A number of α subunits have been identified and are encoded by members of an expanding gene family [2–5]. Specific α subunits are expressed in a tissue-specific and developmentally specific manner [6,7]. Expression of the α subunit alone in *Xenopus* oocytes is sufficient to produce functional channels; however, the co-expression of the $\beta 1$ subunit results in a Na^+ current with properties closer to those of native channels [1]. Clones of the $\beta 1$ cDNAs have been obtained from rat and human [8–10] and were shown to be highly conserved at the nucleotide and amino acid levels. The sequence of the $\beta 2$ subunit has not yet been published and its function is yet to be clarified.

In an earlier study, restriction enzyme analysis, cloning and sequencing of $\beta 1$ cDNA from rat sciatic nerve, which does not contain neurons and thus is presumably derived from Schwann cells, revealed the presence of two isoforms $\beta_{1.1}$ and $\beta_{1.2}$ [11]. Amplification of similar DNA fragments is also reported from embryonic rat tissues [12]. In the present study we used RT-PCR and genomic PCR to investigate the presence of the two $\beta 1$ isoforms and to determine the source of $\beta_{1.2}$. In an attempt to determine if the $\beta 1$ subunit is encoded by members of a gene family, we analysed cDNA fragments that were amplified by RT-PCR from various adult rat tissue RNAs. We also amplified and compared portions of the $\beta 1$ gene from rat, mouse and human genomic DNA. We present here evidence that $\beta_{1.2}$ is an incompletely spliced RNA and confirm one of the polymorphic sites that were previously reported [11]. We also report on the relatedness of the last two introns in this gene among the three species.

2. Materials and methods

2.1. Cell culture: B104

The B104 neuroblastoma cell line was obtained from D. Schubert (Salk Institute, La Jolla, CA). Cells were cultured in Dulbecco's Modified Eagle's Medium supplemented with 20% fetal calf serum [13]. Penicillin and streptomycin were added at 500 U/ml each. Cells were grown in 5-ml flasks (source) at 37°C in a 5% CO_2/95% air atmosphere.

2.2. Nucleic acid isolation

Total cellular RNA was isolated by the single step Guanidinum isothiocyanate-acid phenol procedure [14] The quality of the RNA was assessed by electrophoresis in a 1% agarose-2.2 M formaldehyde gel [15]. High molecular weight DNA was isolated from liver tissue of adult Sprague-Dawley rats as described [15]. Mouse and human high molecular weight DNA were purchased from Clontech (Palo Alto, CA).

2.3. Reverse transcription

First-strand cDNA was reverse transcribed in a 50 μl final volume using 5 μg total RNA, 1 μM random hexamer (Boehringer Mannheim) and 500 U SuperScript II reverse transcriptase (Life Technologies) in the presence of 100 U of RNase Inhibitor (Boehringer Mannheim). The reaction buffer consisted of 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT and 5 mM dNTP. The reaction was allowed to proceed at 37°C for 90 min, 42°C for 30 min then terminated by heating to 65°C for 10 min. A control reverse transcription reaction contained all components except for reverse transcriptase.

2.4. PCR

We used nested primers for both RT-PCR and genomic PCR to reduce non-specific amplification. The same primer set was used in the primary PCR of the two applications. The 5′ oligonucleotide primer (A) (5′ GATATGGCATCCATCGTGTC 3′) corresponds to nucleotides 676–685 (rat) and 554–573 (human) while the 3′ oligonucleotide primer (B) (5′ CCAGGTCCAGCCGGAGGAAGC 3′) corresponds to nucleotides 1325–1345 (rat) and 1239–1259 (human) [8,9]; this sequence is completely conserved between rat and human and is expected to also be conserved in the mouse. The 5′ primer has a mismatch (C to T) at position 3 compared with the human sequence. This difference is not expected to affect the priming efficiency under the PCR conditions used in this study.

For RT-PCR, the primer set of the secondary PCR is expected to amplify sequences between positions 810 and 1255 of the cDNA (numbers according to [8]). The 5′ primer (C) (5′ CTCGGAATACCTGGCCAT 3′) and the 3′ primer (D) (5′ CCCTCTTCACCCCATCAAG 3′) correspond to nucleotides 810–827 and 1271–1255, respectively [8]. The amplified fragment is predicted to be 445 bp in length.

For secondary genomic PCR, the 5′ oligonucleotide primer (A) is the same primer used for primary PCR. The 3′ oligonucleotide primer (E_R) (5′ ACTGCAGAACTGTGAGGCT 3′) correspond to nucleotides 1000–1018 (rat) and (E_H) (5′ CAGGACTCTGAGGCTTTCT 3′) correspond to nucleotides 852–870 (human). The 3′ ends of these primers are complementary to different sequences in rat and human and, therefore, do not efficiently prime the template under the stringent amplification

*Corresponding author. Department of Neurology, LCI 707, Yale School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA. Fax: (1) (203) 785 7826.

SSDI 0014-5793(95)01400-4

conditions utilized in this study (data not shown). Mouse β1 sequence are not available but we assumed that they will be highly related to those of the rat sequence.

Rat β-actin sequences were amplified from the same cDNA pool, using commercial primers (Clontech) to determine the presence of contaminating genomic DNA in the RNA preparations. PCR from a cDNA template results in a 764-bp fragment while a genomic template results in a 1440-bp fragment. A single DNA fragment, comigrating with the 800-bp molecular weight marker, is detected in control PCR reactions which indicates the absence of appreciable genomic DNA contamination in the RNA samples (data not shown).

Amplification was typically performed in 60 μl volume using 1 μl of the first-strand cDNA or 100 ng genomic DNA template, 0.8 μM of each primer and 1.75 U of Expand Long Template DNA polymerase enzyme mixture (Boehringer Mannheim). Control PCR reactions in which the template was substituted by water or an aliquot from a control reverse transcription reaction lacking reverse transcriptase produced no amplification products (data not shown). The use of the Expand Long Template enzyme mixture increased the yield of the PCR products without an increase in non-specific amplification [16,17]. The PCR reaction buffer consisted of 50 mM Tris-HCl (pH 9.2), 16 mM $(NH_4)_2SO_4$, 2.25 mM $MgCl_2$, 2% (v/v) DMSO and 0.1% Tween-20. Amplification was carried out in two stages using a programmable thermal cycler (PTC-100, MJ Research, Cambridge, MA). First, a denaturation step at 94°C for 4 min, an annealing step at 58°C for 2 min and an elongation step at 72°C for 90 s. Second, a denaturation step at 94°C for 1 min, an annealing step at 58°C for 1 min and an elongation step at 72°C for 90 s. The second stage was repeated 33 times for a total of 35 cycles with the elongation step in the last cycle extended to 10 min. Secondary PCR was performed essentially as described above except that the template was a 1 μl of a 1:500 (genomic PCR) and 1:1000 (RT-PCR) dilution of the primary PCR product and amplification was carried out for 25 cycles.

2.5. Cloning and sequence analysis

Rat genomic PCR products were cloned into the *SmaI* site of pBluescript II (KS^-) vector (Stratagene). Mouse and human genomic PCR products were cloned into the *SrfI* site in the vector pCR-Script AMP (KS^+) (Stratagene). Inserts of three independent isolates of rat and mouse and one isolate of human PCR products were sequenced using an automated sequencer at the Keck Biotechnology Laboratory at Yale University. The sequences were analysed using software from the Genetics Computer Group.

3. Results and discussion

Recently, we described a novel β1 subunit RNA isoform ($\beta_{1.2}$) from rat brain, optic nerve, sciatic nerve and skeletal muscles that contains several nucleotide substitutions and an 86 nt insert in the 3′ untranslated [11]. PCR amplification of cDNAs from brain, spinal cord, dorsal root ganglia (DRG) and the neuroblastoma cell line B104 (Fig. 1) reveals the presence of the two β1 isoform RNAs (Fig. 1). Lanes 2–6 contain fragments from the secondary RT-PCR reaction using primer pair C/D and the following templates: forebrain (lane 2), hindbrain (lane 3), spinal cord (lane 4), DRG (lane 5) and B104 cells (lane 6). Lane 7 contains PCR product using water as a template. The size of the DNA fragments in these lanes is in agreement with the expected sizes of 445 bp and 531 bp of $\beta_{1.1}$ and $\beta_{1.2}$, respectively. The presence of the two transcripts in forebrain, hindbrain, spinal cord and DRG (each of which contain neuronal cell bodies and glial cells), and B104 cells (a neuroblastoma cell line) as well as rat optic and sciatic nerves (which contain glial and Schwann cells but no neuronal cell bodies) argues against the possibility that this is a glia-specific phenomenon. It appears that the two forms coexist in all of the tissues where the β1 gene expression is tested. However, this data does not demonstrate the co-existence of the two transcripts in the same cell.

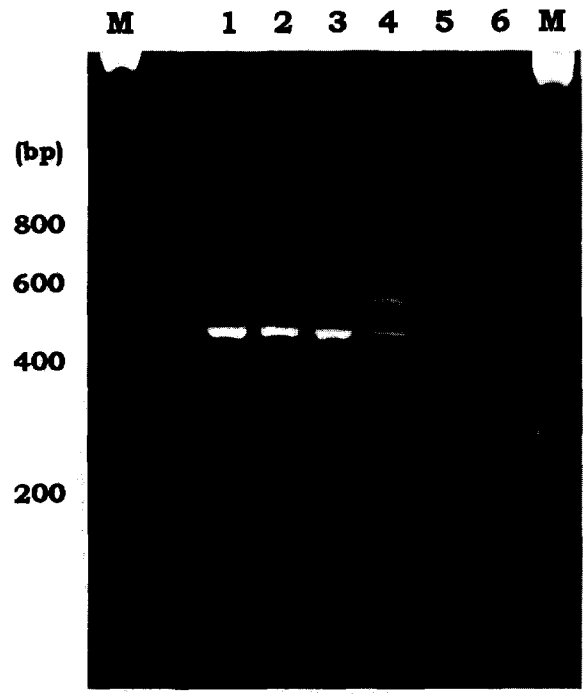

Fig. 1. RT-PCR amplification of Na channel β1 subunit sequences. Secondary PCR products from forebrain (lane 1), hindbrain (lane 2), spinal cord (lane 3), DRG (lane 4) and B104 cells (lane 5). Lane 6 contains amplification products using water as a template. The lower band represents amplification products from the mature, fully spliced β1 mRNA. The top band is that expected from an amplification product of a template retaining intron 5. The reason for the appearance of two top bands is not clear to us at this time.

Table 1
Percentage nucleotide similarity in the exons and introns of the amplified β1 gene fragments from rat, mouse and human

	Exon 4	Intron 4	Exon 5	Intron 5	Exon 6
rat Vs mouse	95.5	87	95.8	93	97
rat Vs human	95.5*	46.5[a]	88.9	46.5	55
mouse Vs human	91.5[§]	61[a]	86	41	54.5

*Deduced from the published sequences [10] and this study.
[a]Based on the partial sequence available for the human β1 intron 4 from this study.
[§]Although the overall similarity is the same, the different positions are not the same compared with the rat sequence.

Based on its sequence and transcription pattern, the insert in the $\beta_{1.2}$ isoform is a pre-mRNA intron. Eukaryotic pre-mRNA introns are characterized by GT as the 5′ dinucleotide, AG as the 3′ dinucleotide preceded by a polypyrimidine-rich tract of about 14 bp and a branch site that has the loose motif C/TTA/GA*C/T (the A marked by an asterisk is the branch nucleotide) located typically 20–30 bp upstream of the 3′ splice site ([18] and references therein). The sequence of the 86 nucleotide insert in $\beta_{1.2}$ (Fig. 2, [11]) contains all of these hallmarks. Consistent with its splicing, transcripts lacking this insert account for the majority of the β1 mRNA ([11,12] and this study). Additionally, the human β1 gene structure has already been published and an intron of 90 bp (intron 5) is reported to be present at precisely the same location as the insert in $\beta_{1.2}$, but the complete sequence of this intron was not reported [19].

Like the human gene [10], rat and mouse β1 subunit is likely

to be encoded by a single gene. In support of this view, the sequences of three independent isolates from the amplification product, including the introns, were identical for the rat and mouse, respectively. Intron sequence variation has been reported even in the relatively recent gene duplicates α1 and α2 of the human α-globin genes [20] and the MCP-1 and MCP-2 rabbit defensin genes [21]. The presence of introns 4 and 5 in all three isolates strongly suggests the absence of a completely processed pseudogene in the rat and mouse genomes. The additional two cytosine residues at position 912 of the cDNA that was reported earlier [11] are also present in these genomic isolates. The difference between this sequence and original published sequence may identify a polymorphic site in this gene.

We found that the position of introns 4 and 5 in the human gene is conserved in the rat and mouse genes. Intron 4 from rat and mouse genes show 87% identity which is lower than that of intron 5 and the flanking exon sequences (Table 1). Intron 4 in the human gene apparently contains a recognition site for the restriction endonuclease *Srf*I because we did not recover sequences from exon 4, which contains the sequence of the upstream PCR primer, and a portion of this intron. Based on the comparable size of the amplified fragment from the three genomic DNAs (data not shown) we estimate that we are missing around 100 bp of this intron. Intron 4 from the human gene is only 46.5% and 61% identical to intron 4 from the rat and mouse genes, respectively (Table 1). In contrast, exon 4 from the human gene is 91.5% identical, at the nucleotide sequence level, compared with the rat and mouse exon 4 (Table 1). The reduction in similarity relative to the flanking exon sequences is not surprising since intron primary sequences tolerate changes except at the conserved splicing signals [22]. The similarity of the introns between rat and mouse genomes is also not surprising given that they are known to have nearly equal mutation rates [23].

The sequence of intron 5 in the three species is shown in Fig. 2. Analysis of this data shows that mouse β1 intron 5 is 87 bp

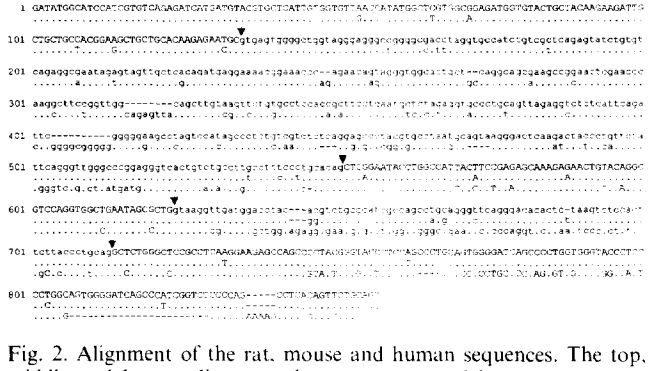

Fig. 2. Alignment of the rat, mouse and human sequences. The top, middle and bottom lines are the rat, mouse and human sequences, respectively. We only show the portion of intron 4 of the human gene that could be reliably aligned with the rat and mouse sequences: the remaining available sequence is deposited in the GenBank data base. Exon and intron sequences are in upper and lower case type, respectively. This alignment was generated by the GAP program of GCG. Dots and dashes represent identities and deletions, respectively. Exon/intron boundaries are identified by the arrow heads and were determined for the human and rat sequences by comparison to published cDNA and genomic sequences, and for the mouse by homology to the rat sequences. These sequences are deposited in the GenBank database [accession numbers L48687 (mouse), L48688 (rat) and L48689 (human)].

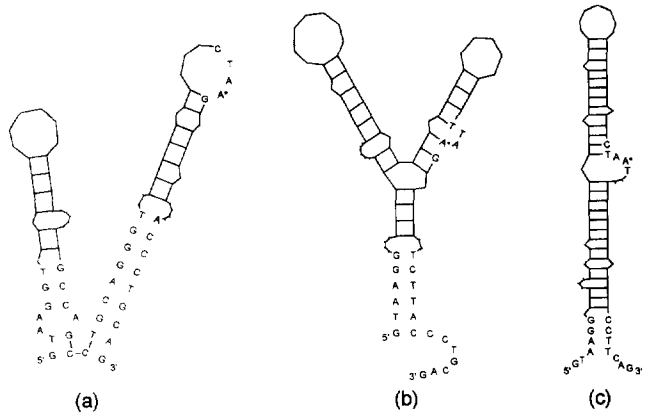

Fig. 3. Predicted secondary structure of intron 5 from rat, mouse and human β1 gene. This figure shows the predicted secondary structure of intron 5 from rat (a), mouse (b) and human (c). These structures were determined by computer analysis using the FOLD program of GCG. The stability of the structure was calculated at −24.7 kcal/mol, −16.8 kcal/mol and −37.1 kcal/mol for the rat, mouse and human sequence, respectively. Only the nucleotides at the 5′and 3′ splice sites and the putative branch site are shown. The second A in the branch site motif (denoted by an *) is the branch residue.

long and is highly conserved relative to its rat cognate. Intron 5 from the rat and mouse genes share 93% sequence identity while they share only 46.5% and 41% identity with the human cognate (90 bp long), respectively (Table 1). The sequence of the human intron 5 from our study is identical to the corresponding published sequence [19]. The divergence between the human and rat sequences is in stark contrast with the conservation of the coding sequences (90%) and the 3′ untranslated sequences (67%) previously reported [9,10]. Rat and human exon 6 sequences that were obtained in this study were 55% identical to each other (Table 1). This reduced identity compared with the 67% figure cited above is due to the selection of a more divergent region of the 3′ untranslated sequences to enhance the selectivity of the amplification process in our study.

The location of intron 5 is uncommon among vertebrate genes which rarely contain introns in their 3′ untranslated sequences [24,25]. Also, the length of intron 5 (86–90 bp) is more typical of internal introns than of introns occurring in the 3′ untranslated sequences [24]. Intron 5 is the shortest intron in the human β1 gene [19] and, interestingly, the only intron to be retained in the $\beta_{1.2}$ mRNA isoform ([11,12] and this study). The presence of $\beta_{1.2}$ RNA is consistent with recent published reports which demonstrate the accumulation of unspliced, polyadenylated RNA as an intermediate during mRNA maturation [26]; the presence of a functional terminal intron is essential for the selection of a polyadenylation site and proper 3′ end formation [27].

Several cases of intron retention has been reported in the literature [28–31]. Retained introns could maintain the open reading frame and result in an alternate protein [29,30]. In the case of the bovine growth hormone, intron retention is dependent on suboptimal 5′ and 3′ splice sites [29,32]. Intron 5 of β1 has optimal splice sites, a polypyrimidine-rich tract and a branch site. Therefore, its retention is not likely to be caused by an inefficient spliceosome complex assembly. Intron 5 in the three genes has the potential to fold into secondary structures that involve the 5′ and 3′ boundaries paired with each other

(mouse and human) or to other intronic sequences (rat) (Fig. 3). The sequestration of one or more of the splicing signals may prevent the recruiting of splicing factors and could therefore block splicing of this intron. It is conceivable that self-folding and the assembly of a splicing complex compete at this intron site, resulting in the co-existence of the two transcripts in the same cell. Formation of secondary structures has been proposed to explain the use of cryptic donor or acceptor sites or for alternative splicing (reviewed in [33]) but more importantly, secondary structures that sequester the 5′ splice site in introns from yeast and plants block splicing [34,35].

Alternatively, splicing of this intron may require a specific splicing factor that is cell-specific or developmentally regulated such that, at any one time, only a subset of cells will be able to produce a fully functional transcript of the $\beta 1$ subunit. The presence of both transcripts in the B104 cell line is intriguing. B104 cells express a variety of Na currents with significant cell-to-cell heterogeneity, and it is possible that a fraction of those cells are unable to splice intron 5. Indeed only a subpopulation of B104 is capable of producing an action potential (Gu et al., unpubl.). Whether the two $\beta 1$ transcripts co-exist in the same cell, and whether intron 5 effects the translation of $\beta_{1.2}$ remain to be determined.

Acknowledgements: We thank S. Baserga for helpful discussion. We also acknowledge the technical assistance of K. Zur and P. Khungar. This work was supported in part by grants from the National Multiple Sclerosis Society and the Medical Research Service, Department of Veterans Affairs. S.D.D.-H. was supported in part by a Spinal Cord Research Fellowship from the EPVA.

References

[1] Isom, L.L., De Jongh, K.S. and Catterall, W.A. (1994) Neuron 12, 1183–1194.
[2] Catterall, W.A. (1993) TINS 16, 500–506.
[3] Schaller, K.L., Krzemien, D.M., Yarowsky, P.J., Krueger, B.K. and Caldwell, J.H (1995) J. Neurosci. 15, 3231–3242.
[4] Klugbauer, N., Lacinova, L., Flockerzi, V. and Hofmann, F. (1995) EMBO J. 14, 1084–1090.
[5] Felipe, A., Knittle, T.J., Doyle, K.L and Tamkun, M.M. (1994) J. Biol. Chem. 269, 30125–30131.
[6] Beckh, S., Noda, M., Lubbert, H. and Numa, S. (1989) EMBO J. 12, 3611–3616.
[7] Mandel, G. (1992) J. Membrane Biol. 125, 193–205.
[8] Isom, L.L., De Jongh, K.S., Patton, D.E., Reber, B.F.X., Offord, J., Charbonneau, H., Walsh, K., Goldin, A.L. and Catterall, W.A. (1992) Science 256, 839–842.
[9] McClatchey, A.I., Cannon, S.C., Slaugenhaupt, S.A. and Gusella, J.F. (1993) Hum. Mol. Genet. 2, 745–749.
[10] Makita, N., Bennett, P.B. and George, A.L. (1994) J. Biol. Chem. 269, 7571–7578.
[11] Oh, Y. and Waxman, S.G. (1994) Proc. Natl. Acad. Sci. USA 91, 9985–9989.
[12] Lara, A., Dargent, B., Julien, F., Alcaraz, G., Tricaud, N., Couraud, F. and Jover, E. (1995) Brain Res., in press.
[13] Schubert, D., Heinemann, S., Carlisle, W., Tarikas, H., Kimes, B., Patrick, J., Steinbach, J.H., Culp, W. and Brandt, B.L. (1974) Nature (London) 249, 224–227.
[14] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[15] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) in Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Plainview, NY).
[16] Barnes, W.M. (1994) Proc. Natl. Acad. Sci. USA 91, 2216–2220.
[17] Cheng, S., Fockler, C, Barnes, W.M. and Higuchi, R. (1994) Proc. Natl. Acad. Sci. USA 91, 5695–5699.
[18] Senapathy, P., Shapiro, M.B. and Harris, N.L. (1990) Methods Enzymol. 183, 252–278.
[19] Makita, N., Sloan-Brown, K., Weghuis, D.O.Ropers, H.H. and George, A.L. (1994) Genomics 23, 628–634.
[20] Liebhaber, S.A., Goossens, M. and Kan, Y.Q. (1981) Nature (London) 290, 26–29.
[21] Ganz, T., Rayner, J.R., Valore, E.V., Tumolo, L., Talmadge, K. and Fuller, F.H. (1989) J. Immunol. 143, 1358–1365.
[22] Padget, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986). Annu. Rev. Biochem. 55, 1119–1150.
[23] Li, W. and Graur, D. (1991) in: Fundamentals of Molecular Evolution, p. 82, Sinauer Associates, Sunderland, MA.
[24] Hawkins, J.D. (1988) Nucl. Acid Res. 16, 9893–9908.
[25] Wilcox, A.S., Khan, A.S., Hopkins, J.A. and Sikela, J.M. (1991) Nucl. Acid Res. 19, 1837–1843.
[26] Niwa, M. and Berget, S.M. (1991) Gene Expr. 1, 5–14.
[27] Liu, X and Mertz, J.E. (1993) Nucl. Acid Res., 21, 5256–5263.
[28] Roche, S.E., Schiff, M. and Rio, D. (1995) Genes and Development, 9, 1278–1288.
[29] Dirksen, W.P., Sun, Q and Rottman, F.M. (1995) J. Biol. Chem. 270, 5346–5352.
[30] Altieri, D.C. (1994) Biochemistry, 33, 13848–13855.
[31] Weil, D., Brosset, S. and Dautry, F. (1990) Mol. Cell. Biol., 10, 5865–58–75.
[32] Dirksen, W.P., Hampson, R.K., Sun, Q and Rottman, F.M. (1994) J. Biol. Chem. 269, 6431–6436.
[33] Balvay, L., Libri, D and Fiszman, M.Y. (1993) BioEssays, 15, 165–169.
[34] Liu, H, Goodall, G.J., Kole, R. and Filipowicz (1995) The EMBO J., 14, 377–388.
[35] Goguel, V., Wang, Y. and Rosbash, M. (1993) Mol. Cell. Biol., 13, 6841–6848.