The EMBO Journal vol.8 no.6 pp.1719 – 1726, 1989

Multiple mRNAs encode peripherin, a neuronal intermediate filament protein

Françoise Landon, Marguerite Lemonnier[1], Richard Benarous[2], Claude Huc, Marc Fiszman[1], François Gros and Marie-Madeleine Portier

Biochimie Cellulaire, Collège de France, 11 place Marcellin Berthelot, 75231 Paris Cedex 05, [1]Unité de Biochimie, Département de Biologie Moléculaire, Institut Pasteur, 75724 Paris Cedex 15, and [2]Unité Inserm U15, CHU Cochin, 75674 Paris Cedex 14, France

Communicated by F.Gros

Three cDNA clones of 1.6 (3u), 1.2 (5g) and 0.6 (5b) kbp, specific for peripherin, a neuronal intermediate filament protein (IFP), have been isolated from a murine neuroblastoma cell λgt11 library by immunoscreening using peripherin antiserum. Antibodies eluted from the fusion proteins produced by clones 3u and 5g recognize the peripherin spots on immunoblots. Where they overlap the three cDNAs have identical sequences. cDNA 5g exhibits the closest homology to type III IFP cDNAs. cDNA 3u is identical to the corresponding region of cDNA 5g, except for the insertion of a 96 bp fragment at a position corresponding to the junction of exons 4 and 5 in type III IFP cDNAs. cDNA 5b is also identical to the corresponding region of cDNA 5g, except for the deletion of a 62 bp fragment at the junction of exons 8 and 9 in type III IFP cDNAs. S1 mapping experiments performed with probes covering the 3′ end of the two unexpected regions show that three distinct mRNAs correspond to the three cDNAs. Moreover, three peripherin products, two minor 61 and 56 kd products in addition to the major 58 kd peripherin, are observed when poly(A)$^+$ RNA is *in vitro* translated, the 61 kd peripherin being translated from the 3u-selected RNA. The three RNAs originate from alternative splicing of a unique peripherin gene, thus generating polymorphism of peripherin.

Key words: alternative splicing/intermediate filament protein/neuroblastoma/neuronal/peripherin

Introduction

Intermediate filament proteins (IFP) are assembled into 7 – 10 nm elements in most vertebrate cells and their tissue specificity is well established (Lazarides, 1980). Biochemical studies have revealed conserved structural features of all IFPs. These include a central α-helical rod domain of some 310 amino acids, forming coiled-coil segments 1a, 1b and 2, together with flanking non-α-helical ends (Geisler and Weber, 1986). The amino-terminal head and carboxy-terminal tail regions are variable in their length and sequence. Regarding the sequence homologies, IFPs can be divided into several classes: types I and II for acidic and basic keratins, respectively; type III for vimentin, desmin and glial fibrillary acidic protein (GFAP); type IV for neurofilament proteins (NFPs) (Steinert and Roop, 1988). Referring to this classification, the IFP-like nuclear lamins A and C belong to a type V (Steinert and Roop, 1988). IFPs are encoded by a large multigene family of single copy genes each generating only one protein (Geisler and Weber, 1982; Steinert *et al.*, 1983; Marchuk *et al.*, 1984; for a review, see Osborn and Weber, 1986). A possible exception concerns the lamin A and C transcripts which are supposed to be generated through the alternative splicing of a unique gene (McKeon *et al.*, 1986; Fisher *et al.*, 1986). Interestingly, the location of introns appears to be conserved in each of the I – IV classes of IFP genes.

Another member of the IFP family, peripherin, has been described as being specific to certain neuronal populations (Portier *et al.*, 1984a,b; Escurat *et al.*, 1988). Peripherin exhibits the biochemical characteristics of IFPs and its rate of synthesis responds to differentiating agents such as 1-methylcyclohexane carboxylic acid (CCA) for neuroblastoma cells or nerve growth factor (NGF) for PC12 cells. Recently, several authors reported the presence of a protein in PC12 cells whose characteristics were similar to those of peripherin, with respect, for example, to the stimulation of its synthesis in NGF-treated cells (Franke *et al.*, 1986; Parysek and Goldman, 1987; Aletta *et al.*, 1988). While we had undertaken the cloning of mouse peripherin cDNA, Leonard *et al.* (1988) described a clone 73 cDNA obtained from NGF-treated PC12 cells and showed that this cDNA encoded a protein with significant homology to type III IFPs.

Here, we report the cloning of three mouse neuroblastoma peripherin cDNAs in an expression vector, the nucleotide sequence of these different cDNAs and the deduced sequence of their encoded peptides. Where they overlap the three cDNAs have strictly identical sequence. Homologies between one of our three murine cDNAs and the rat clone 73 indicate that they both encode peripherin. The other two cDNAs display in their coding sequences unique regions located at the junctions of exons 4 and 5, on the one hand, of exons 8 and 9, on the other hand, when compared with the type III IFP gene sequences (Quax *et al.*, 1983, 1985; Balcarek and Cowan, 1985). Furthermore, multiple translations products are found in relation to the multiple cDNAs. Taken together these results show that multiple peripherin mRNAs and isoforms are derived from a unique peripherin gene by differential splicing.

Results

Identification of peripherin cDNA clones by epitope selection

The immunoscreening of ~2.5 × 10^5 plaques of the library resulted in 54 positive clones. When tested by dot blot analysis, 15 clones exhibited strong reactivity, 18 medium reactivity and 21 weak reactivity with the antiserum

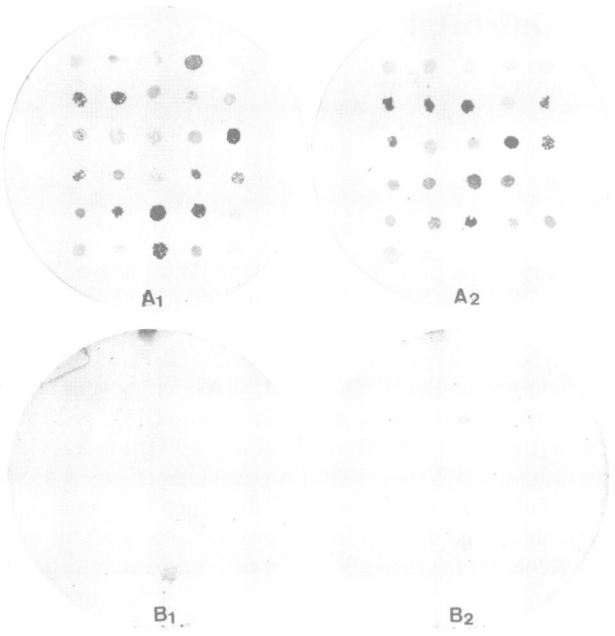

Fig. 1. Dot blot analysis of the 54 positive clones. λgt11 recombinant clones were plated in drops of 6×10^2 p.f.u. on *E.coli* Y 1090. $\mathbf{A_1}$ and $\mathbf{A_2}$: probing with antiserum specific to peripherin. $\mathbf{B_1}$ and $\mathbf{B_2}$: probing with anti-IFA. Non-recombinant λgt11 phage produced the plaque signal in the extreme upper left and right and lower right of each array.

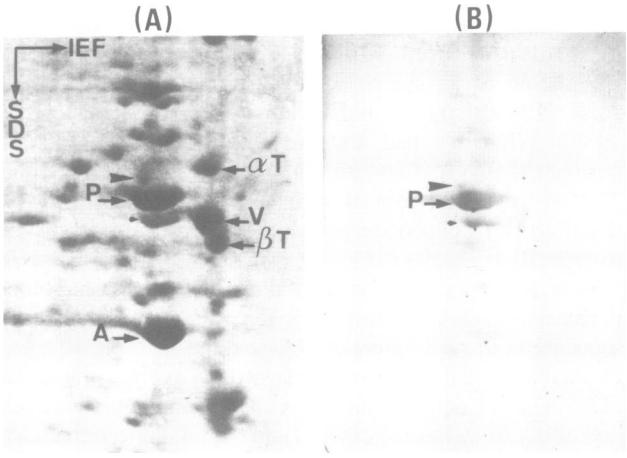

Fig. 2. Two-dimensional gel electrophoresis of the Triton-insoluble proteins from the mouse neuroblastoma cells. (**A**) Coomassie blue stained gel. IEF: Direction of the first dimension separation, using isoelectric focusing. SDS, direction of the second dimension separation by SDS−PAGE. P, peripherin; V, vimentin; A, actin; α- and β-T, α- and β-tubulins. Large and small arrowheads: two proteins recognized by antiperipherin. (**B**) Western blot probed with antibodies selected by the fusion protein produced by clone 3u. Location of proteins named in (A) was indicated by a pencil point after transient staining of the nitrocellulose replica with Ponceau red.

specific to peripherin (Figure 1A). When tested similarly for their ability to produce antigens recognized by IFA antibody, a monoclonal antibody specific to a common epitope of IFPs, the reactivity was always found to be weaker (Figure 1B) than with antiperipherin and no correlation existed among the clones between the levels of reactivity with each antibody. Reactivity with anti-IFA should depend not only on the presence of the epitope in the fusion protein

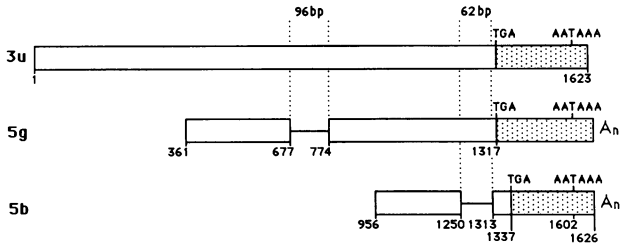

Fig. 3. Schematic diagram of cDNAs 3u, 5b, and 5g, derived from sequence analysis. The clear boxes specify amino acid coding sequences and the shaded boxes, 3′ untranslated sequences. TGA represents the position of the termination codon and AATAAA, the position of consensus polyadenylation signal. Bars between boxes represent missing sequences. Numbers refer to the nucleotide numbering of cDNA 3u.

but also on its accessibility to the antibody. Some fusion proteins may thus exhibit such a conformation as they do not bind IFA. No reactivity was observed with antibodies specific to vimentin, α- and β-tubulins and actin (not shown).

Subsequent analyses were performed with three clones 3u, 5b and 5g: 3u and 5g reacted strongly with antiperipherin and 5b reacted moderately; 5g and 5b were positive with anti-IFA and 3u was negative. The epitope selection technique was used to confirm the identity of the clone 3u: the antibody, affinity-purified from the fusion protein, detected peripherin on a Western blot of a neuroblastoma cell extract (Figure 2A and B). A similar result was obtained with clone 5g. Epitope-selected antibodies from wild type λgt11 showed no reactivity.

Sequencing of cDNA clones

Three cDNAs, 3u (1623 bp), 5g (1182 bp) and 5b (621 bp), were sequenced entirely on both strands. Their schematic diagram is shown in Figure 3. The three cDNAs overlap with short missing sequences in the coding regions of 5g and 5b, as compared with the cDNA 3u. All cDNAs display identical sequences in their overlapping regions. They all possess a stop codon, a consensus polyadenylation signal and, in the case of clones 5b and 5g, at least part of the poly(A) tail, 25 nucleotides down the polyadenylation signal, indicating that they contain the 3′-terminus of the transcripts.

The sequence of the longest cDNA, 3u, is shown in Figure 4 in parallel with that of clone 73 from rat PC12 cells (Leonard *et al.*, 1988). It begins within the codon 55 of the rat PC12 clone 73 and ends 127 nucleotides downstream of the 3′-cloning site of clone 73. It is presented in Figure 4 with the last three nucleotides and the poly(A) tail of the cDNAs 5b and 5g. With the exception of the sequence 678−773, which is lacking in clone 73, homology of cDNA 3u with clone 73 is evident. In the overlapping coding region, only 63 nucleotides are different from rat clone 73: 49 do not change the deduced amino acid, four contribute to three conservative amino acid changes (residue 416: Met/Val; residue 425: Lys/Arg; residue 452: Asp/Glu), 10 contribute to five non-conservative amino acid changes (residue 151: Glu/Ser; residue 152: Arg/Ala; residue 153: Asp/Tyr; residue 167: Ala/Gly; residue 419: Leu/Pro). After completion of our work, Thompson and Ziff (1989) reported the sequence of the rat peripherin gene: the corrected residues 151−153 are identical to those of the mouse peripherin. The estimated identity between the known part of mouse

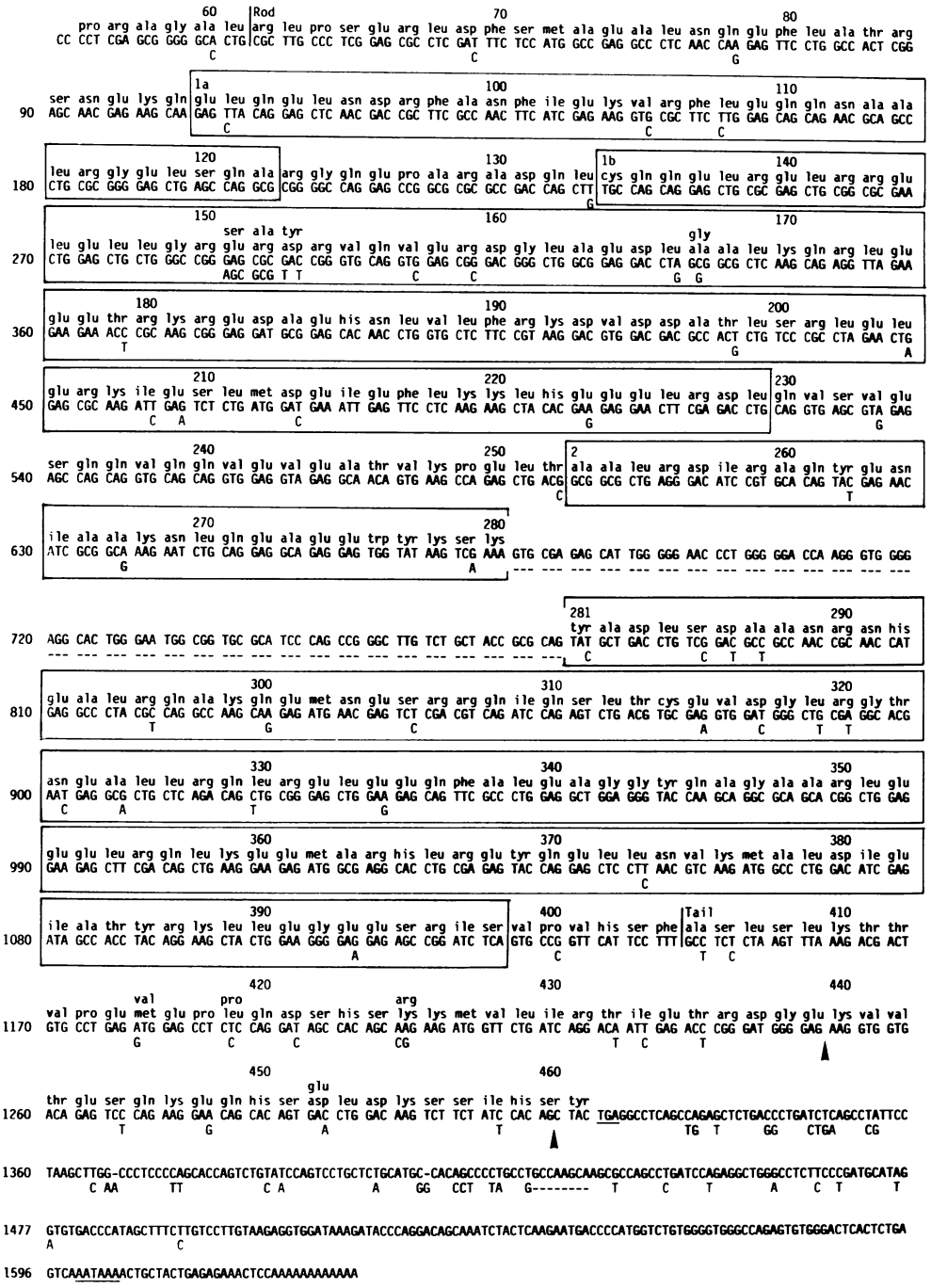

Fig. 4. Nucleotide and deduced amino acid sequence of mouse neuroblastoma peripherin cDNA. The peripherin sequence is aligned with the rat clone 73 up to nucleotide 1496 (3′ end of clone 73) (Leonard *et al.*, 1988): only nucleotides (lower row) and amino acids (upper row) that are different in clone 73 are indicated. The recently published sequence of the rat peripherin gene (Thompson and Ziff, 1989) shows that the codons 151–153 are identical to those of mouse peripherin cDNA. Dashes indicate the absence of the nucleotide and allow for a better alignment in the 3′ untranslated region. Polyadenylation signal and the termination codon are underlined. Numbering of the nucleotides, indicated on the left, refers to Figure 3. Numbering of the amino acids refers to clone 73, so that the extra sequence from nucleotide 678 to nucleotide 773 is not translated here. Vertical bars indicate rod and tail domains. Coil 1a, 1b and 2 of the rod domain are enclosed in boxes. Arrowheads indicate the beginning (1251) and the end (1312) of the cDNA 5b deletion.

peripherin and the corresponding one of rat peripherin is 95 and 99% at the nucleotide and amino acid levels, respectively. In the aligned 3′ untranslated ends, many differences are observed together with regions of conservation giving a figure of 75% identity.

The additional 96 bp fragment (678–773) which characterizes the cDNA 3u can be translated using the same open reading frame (ORF) into a 32-amino acid sequence which is particularly rich in tryptophan and proline residues (Figure 5A). This fragment with its initial G and its 3′ CAG end, fits the consensus ends of an exon (Mount, 1982). The nucleotides downstream of nucleotide 678 and upstream from nucleotide 773 also fit the consensus of the 5′ and 3′ ends of intron sequences (Figure 5A).

The cDNA 5g exhibits the closest homology with clone 73 due to the absence of the nucleotide sequence 678–773.

The cDNA 5b represents another form of peripherin cDNA, identical to cDNAs 5g and 3u where they overlap

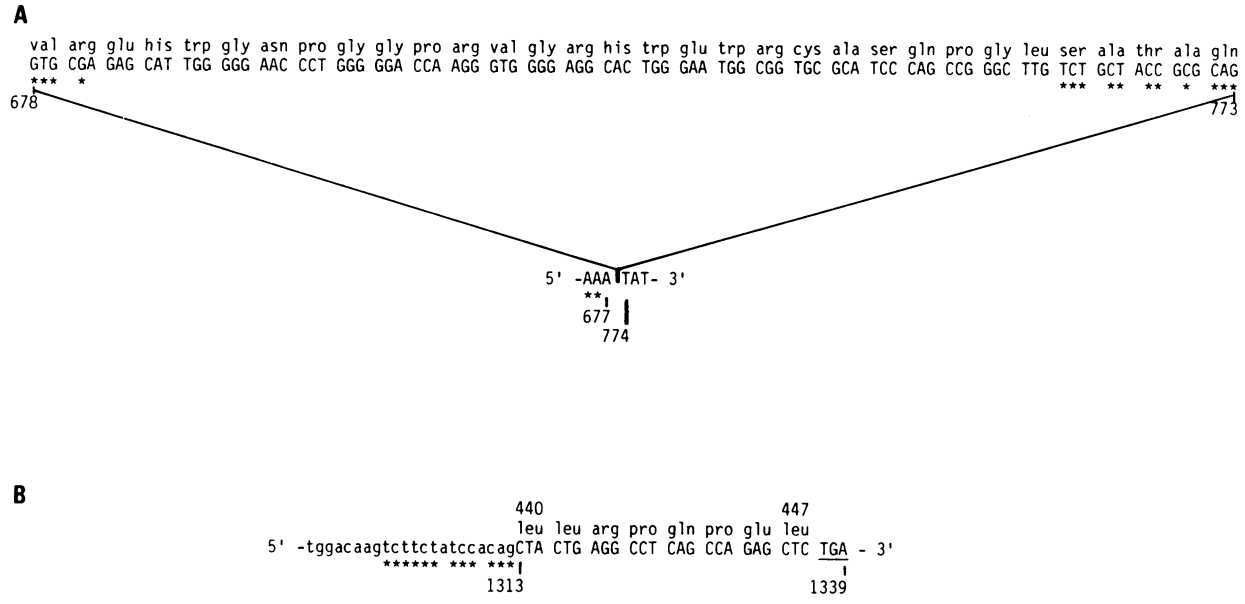

Fig. 5. Specific sequence of the cDNAs 3u and 5b. (**A**) Sequence of the 96 bp fragment found in cDNA 3u between nucleotides 677 and 774 and its deduced amino acid sequence inserted in the sequence of cDNA 5g where a splice site is expected. (**B**) Nucleotide sequence surrounding nucleotide 1313 and deduced amino acid sequence of cDNA 5b translated downstream from nucleotide 1313 using an ORF shifted one nucleotide relative to that used in cDNA 3u and 5g. The 3′ end of the sequence deleted in cDNA 5b is shown in lower case letters. Asterisks indicate the nucleotides homologous to the consensus sequence of exon−intron junctions (Mount, 1982). Numbering of the nucleotides refers to that used in Figure 3.

except for a deletion of 62 bp between positions 1250 and 1313. The 3′ end of the deleted sequence shares homology with the 3′ end of an intron (Figure 5B). Due to this deletion, a new ORF is found with the use of another stop codon located 20 bp downstream from the one used in cDNAs 3u and 5g (Figure 5B). The derived amino acid sequence is thus 14 amino acids shorter than the one translated from the cDNA 5g.

S1 protection analysis of alternate forms

In order to determine the copy number of peripherin in the genome, labelled 3u cDNA was hybridized under stringent conditions to a Southern blot of mouse genomic DNA cut with restriction endonucleases. Number, size and intensity (not shown) of the fragments (*Bam*HI: one 12 kb fragment, no site in the probe; *Hin*dIII: two fragments of 13 and 1 kb, one site in the probe; *Sac*I: three fragments of 1.8, 1.1 and 0.7 kb, three sites in the probe; *Eco*RI: three fragments of 14.5, 10 and 2 kb, no site in the probe) are compatible with those arising from the cDNA probe and also with those expected from the location of introns in the type III IFP genes (Quax *et al.*, 1983, 1985; Balcarek and Cowan, 1985).

The size of the peripherin RNA was investigated by Northern blot analysis using the 3u cDNA as a probe (Figure 6). A broad signal was obtained at 2 kb. To determine if the different cDNA sequences were representative of different peripherin mRNAs and not due to reverse transcriptase artifacts, S1 mapping experiments were performed.

In order to probe the 3u-specific sequence, a 69mer oligonucleotide was designed to span the putative splice site 773−774 such that exclusion of the sequence 678−773 would result in protection of 50 nucleotides (nt) and inclusion in protection of 60 nt of the labelled probe (Figure 7A). Annealing was performed at two temperatures and was found to be more efficient at 50°C than at 42°C. Two protected

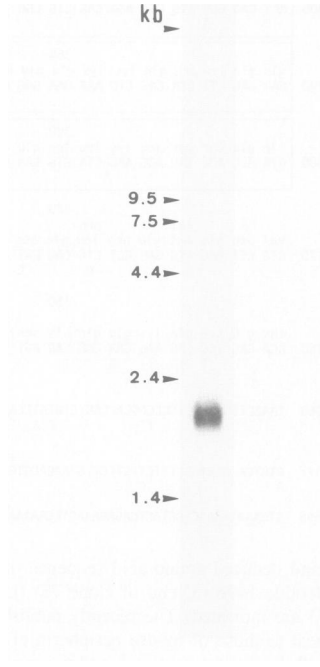

Fig. 6. Size of peripherin mRNA estimated by Northern analysis. 0.9 μg of poly(A)$^+$ RNA was loaded onto the gel and hybridized with cDNA 3u probe after transfer to nitrocellulose. An RNA ladder (BRL) was run on a parallel lane.

fragments were obtained, each migrating as a doublet: the major one 49−50 nt in length corresponded to the mRNA 5g species and the minor (10−15%) one 59−60 nt in length corresponded to the mRNA 3u species. The doublets probably result from hybrid instability and consequent nuclease nibbling due to the presence of AT nucleotides.

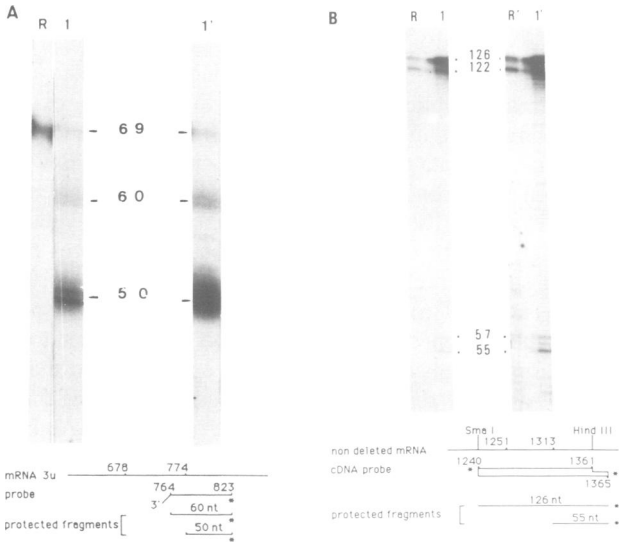

Fig. 7. S1 mapping of peripherin transcripts containing the 3u-specific 678−773 sequence (**panel A**) or deleted in the 1251−1312 sequence, i.e. corresponding to the cDNA 5b (**panel B**). Panel A: A 69mer oligonucleotide probe was designed with 60 nt annealing to the RNA 3u and with an additional segment 5′-AGCTTTGCC-3′ added in 3′, as illustrated in the drawing. R, probe treated at 50°C in the presence of *E.coli* tRNA before S1 treatment; 1,1′, probe annealed in the presence of neuroblastoma poly(A)$^+$ RNA at 50°C. Panel B, a *Sma*I−*Hin*dIII cDNA fragment was used as a probe, as illustrated in the drawing; R and R′, probe treated at 42°C in the presence of *E.coli* tRNA before S1 treatment; 1,1′, probe annealed in the presence of neuroblastoma poly(A)$^+$ RNA at 42°C. Exposure time: overnight (1,R) or 48 h (1′,R′). Size markers were provided by samples used for sequence analysis. nt, nucleotides.

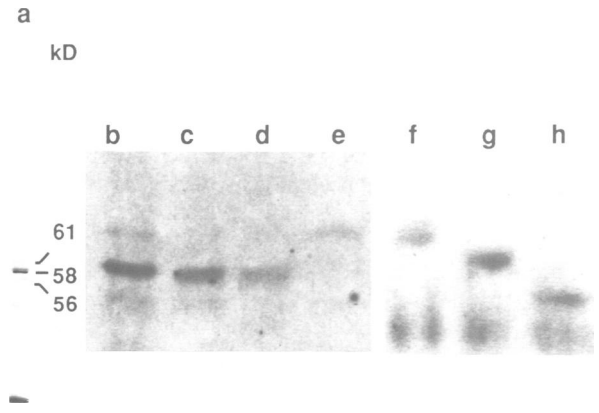

Fig. 8. SDS−PAGE analysis of the peripherin proteins translated from mRNAs and immunoprecipitated with the peripherin antiserum. (**a**)−(**d**) Autoradiograph of the gel showing the immunoprecipitated peripherin products translated from poly(A)$^+$ RNA (a and b) and from RNAs selected by hybridization to the cDNA clones 3u (c) and 5g (d) and to the 3u-specific oligonucleotide (e). In (a), the gel is shown from the origin to the front of migration, while in (b), only the enlarged part of the 58 kd region is shown. Molecular weight markers (94, 67 and 43 kd) were run on the same 8% acrylamide gel. (**f**)−(**h**) Coomassie blue stained gel of the proteins designated by P (g), large (f) and small (h) arrowheads in Figure 2 and co-electrophoresed with the products of (a). The diffuse bands on the lower part of lanes (f), (g) and (h) correspond to the light chains of immunoglobulins used in immunoprecipitation assays.

The presence of RNA 5b was assayed with a 126 bp *Sma*I−*Hin*dIII cDNA fragment spanning the whole deletion 1251−1312 such that three fragments were expected: a 126 nt fragment protected by the undeleted mRNAs 3u and

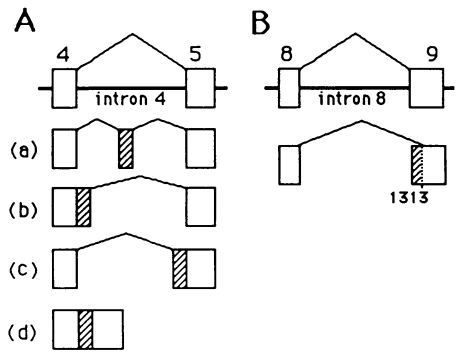

Fig. 9. Hypothetic patterns of alternative splicing from a unique peripherin gene compared to the established pattern of splicing generating the type III IFP mRNAs (Quax *et al.*, 1983, 1985; Balcarek and Cowan, 1985). (**A**) Four possibilities for the splicing of the 3u-specific sequence 678−773 (striped) between exons 4 and 5: a = a new exon; b and c = an extension of an existing exon by the use of a cryptic splice site in intron 4, 96 bp downstream from exon 4 (b) or 96 bp upstream from exon 5 (c); d = a retained intron. (**B**) mRNA 5b probably generated through the use of an internal acceptor site in exon 9 at nucleotide 1313. Constitutive exons (white), alternative sequences (striped), and introns (thick line) are spliced according to different pathways (thin line). The splicing generating the type III IFP mRNAs (upper drawings) is supposed to occur in producing mRNA 5g with the 3′ end of exon 4 at nucleotide 677, the 5′ end of exon 5 at nucleotide 774, the 3′ end of exon 8 at nucleotide 1250 and the 5′ end of exon 9 at nucleotide 1251.

5g, a 122 nt sense strand fragment protected by reannealing of the two strands of the probe and a 55 nt rather than a 53 nt fragment protected by the deleted mRNA 5b, since nucleotides 1311−1312 are identical to nucleotides 1249−1250 (see Figure 4). Annealing was also performed at two temperatures and was found to be more efficient at 42°C than at 50°C. Figure 7B (lane R and R′) shows the 122 and 126 nt fragments of the probe reannealed in the absence of specific mRNA. When annealing was performed in the presence of specific mRNA (Figure 7B, lanes 1 and 1′), the intensity of the signal at 126 nt relative to that at 122 nt represented the proportion of the undeleted mRNA species and the intensity of the signal at 55 nt, the amount of the deleted mRNA species. Clearly, the deleted mRNA 5b is a minor (1−5%) species relative to the undeleted mRNAs. Minor additional protected fragments of 56 and 57 nt may result from non-Watson−Crick type base-pairings, since GG and GT interactions can be involved in stabilizing duplexes of short oligonucleotides (Wilson *et al.*, 1988).

Products translated from hybrid selected mRNAs

When translated *in vitro*, poly(A)$^+$ RNA gave three products precipitable by the antiserum specific to peripherin (Figure 8a and b). The major 58 kd product is the previously described peripherin (Portier *et al.*, 1984a,b), the other two peripherin products are characterized by apparent M_rs of 61 and 56 kd. mRNAs selected by hybridization to the cDNA clones 3u and 5g were translated into the major 58 kd plus the other two barely discernible peripherin isoforms (Figure 8c and d). The ability of the complementary 3u-specific oligonucleotide in selecting the mRNA encoding the only 61 kd peripherin is shown in Figure 8e. The 61 and 56 kd peripherin isoforms were compared to the proteins (designated by large and small arrowheads in Figure 2) recognized by the antiserum specific to peripherin on a

Western blot of the Triton-insoluble extract of neuroblastoma cells. Co-electrophoresis showed that the 61 and 56 kd products co-migrate with the proteins designated by large and small arrowheads, respectively (compare lanes f and h with lane a in Figure 8).

Discussion

The present study describes the characterization of three murine neuroblastoma cDNAs, 3u, 5g and 5b, identified as encoding peripherin. The comparison of the three cDNAs (except for the 96 bp sequence specific to cDNA 3u) with the rat PC12 clone 73, encoding an IFP whose tissue distribution is quite similar to that of peripherin (Leonard *et al.*, 1988), shows that their coding regions are highly conserved at both the nucleotide and amino acid levels, so that the mouse and rat cDNAs encode the same protein: peripherin. The few amino acid differences are found in the heptad repeats of coil 1b or in the tail region, at positions which are poorly conserved among other IFPs (Geisler and Weber, 1986). No differences occur in the two consensus-type sequences at the amino-terminal end of coil 1a and at the carboxy-terminal end of coil 2 of all IFPs (Osborn and Weber, 1986), or in the short region of homology in the tail domain of type III IFPs (Leonard *et al.*, 1988). Hence, given the already reported homology of the clone 73 protein with the type III IFPs, it follows that mouse peripherin is more closely related to type III IFPs than to the other types. Such a conclusion is reinforced by recent data from Parysek *et al.* (1988) who described a rat PC12 cDNA, 199E, shorter than clone 73, of identical sequence and encoding a type III IFP.

An important conclusion emerging from the present study pertains to the existence of different peripherin mRNAs generated by a mechanism of alternative splicing. The three peripherin cDNAs, 3u, 5g and 5b, exhibit overlapping regions of identical sequences, including wobble positions and the 3′ untranslated ends. In combination with evidence from Southern analysis, these data suggest strongly that the three cDNAs arise from the same gene. The occurrence of unique regions in the coding moiety, the 3u-additional and the 5b-missing sequences, is confirmed by the S1 data. Moreover, the comparison of the three cDNA sequences with the sequences of the type III IFP genes (Quax *et al.*, 1983, 1985; Balcarek and Cowan, 1985) clearly shows that the 3u-additional and the 5b-deleted fragments are located at the junctions of exons 4 and 5, and of exons 8 and 9, respectively. Taken together these findings provide evidence of alternative splicing of a unique peripherin primary transcript. Matches of the additional or deleted sequences with exon−intron consensus boundaries allow various hypotheses which comply with some of the known patterns of alternative splicing (Breitbart *et al.*, 1987).

(i) With regard to the 3u-specific sequence, four possibilities arise (Figure 9A). The recent report of the sequence of the rat peripherin gene (Thompson and Ziff, 1989) supports the hypothesis of a retained intron (Figure 9A.d), since intron 4 has the same location as the 3u-specific sequence, an identical length of 96 bp and a nucleotide identity of 80%. This intron would be very short, nevertheless introns as short as 91−97 bp exist in the hamster vimentin (intron 3) (Quax *et al.*, 1983), hamster desmin (intron 2) (Quax *et al.*, 1985), human vimentin (intron 3) (Perreau *et al.*, 1988) and chicken vimentin (intron 6) (Zehner *et al.*, 1987) genes.

(ii) The form 5b could arise through the choice of a downstream cryptic acceptor site for the intron 8−exon 9 junction resulting in exclusion of part of exon 9 (Figure 9B). The sequence data of Thompson and Ziff (1989) show that the exon 8−intron 8 junction of the rat peripherin gene is located at the beginning of the 5b-specific deletion. Homologies with exon−intron consensus sequences have been observed in cDNAs of lamins (McKeon *et al.*, 1986; Fisher *et al.*, 1986) just downstream of the divergence point where the sequence of the lamin C cDNA is identical to the consensus sequence of the 5′ end of an intron and that of lamin A cDNA represents a consensus sequence of the 5′ end of a spliced exon. These lamin A and C RNAs have been tentatively ascribed to a mechanism of alternative splicing: they share identical sequences of 566 codons up to a point of divergence followed by unrelated short sequences, and the Southern analysis suggests that they are transcribed from a single gene (McKeon *et al.*, 1986; Fisher *et al.*, 1986).

Three mRNA species are expected from the three cDNAs. An mRNA species of ~2 kb is shown when cDNA 3u is hybridized to RNA from neuroblastoma cells. If the unknown 5′ ends of the three RNA species 3u, 5g and 5b are identical, their total length should be very close to each other. As far as we know from S1 mapping experiments, the mRNAs 5g, 3u and 5b contribute to the peripherin RNA population for 80−90, 10−15 and 1−5%, respectively. The homologous rat PC12 RNA appears also as a ~2 kb species (Leonard *et al.*, 1987), and an additional 3.2 kb message occurs as a minor species in spinal cord and brain of the rat (Parysek *et al.*, 1988). In the case of NF-L, two RNAs (2.5 and 4 kb) have similarly been reported (Julien *et al.*, 1985; Lewis and Cowan, 1985). It is not known whether the size differences between these mRNAs are uniquely in non-coding regions.

The three mRNAs must be translated into peripherin isoforms of different sizes: the major species 5g would encode the peripherin described previously (Portier *et al.*, 1984a) and the minor ones 3u and 5b would be translated into proteins 3.8 kd larger or 1.5 kd smaller, respectively, assuming that their unknown 5′ coding regions are identical. *In vitro* translation of poly(A)$^+$ RNA indicates that there are indeed three peripherin products precipitated by an antiserum specific to peripherin, the major previously described 58 kd peripherin being translated from the major mRNA 5g and the minor 61 kd peripherin being the product translated from the selected RNA 3u. The 61 kd peripherin 3u can be related to the protein of 61 kd present in Triton-insoluble neuroblastoma cell extracts (designated by a large arrowhead in Figure 2), on the basis of their comigration and of their more basic character than the major peripherin (+4 Arg/2 Glu for peripherin 3u; location on the basic side of P in Figure 2). The minor 56 kd peripherin is possibly translated from mRNA 5b. A relation between the peripherin 5b and the protein of 56 kd present in Triton-insoluble neuroblastoma cell extracts (designated by a small arrowhead in Figure 2) is based on the correspondence with the expected size. Furthermore, the C-terminal moiety generated by *N*-chlorosuccinimide treatment of this Triton-insoluble 56 kd species is shorter than the corresponding moiety of the major 58 kd species, while the N-terminal moieties exhibit similar

length (unpublished results). Since a shorter C-terminal fragment is what is expected from the peripherin isoform 5b, the Triton-insoluble 56 kd protein is a good candidate as being the translation product of the RNA 5b.

The peripherin isoforms 3u and 5b would exhibit properties different from those of the major 5g peripherin, due to the extra 96 bp sequence of cDNA 3u and to the 62 bp deletion of cDNA 5b.

(i) The 3u-specific sequence (32 amino acids) is characterized by an overall basicity (4 Arg/2 Glu), the presence of helix-breaking residues (3 Pro and 3 Trp) and location within helix 2 of the rod domain, five amino acids beyond the centrally located unique tryptophan of most IFPs. It is noteworthy that the 32-amino acid sequence predictable from the ORF of the rat peripherin intron 4 (Thompson and Ziff, 1989) exhibits similar location and characteristics (5 Arg/3 Glu and 2 Trp + 1 Pro). The conservative character (69% identity at the amino acid level) suggests that the intron 4 of peripherin genes is under a selective constraint originating in a sequence-dependent function of the peripherin translated from the mRNA retaining the intron 4. The presence of three proline and three tryptophan residues together with the interruption of the heptad repeat in the coil 2 of peripherin 3u appears incompatible with the preservation of the α-helical structure in that region. This situation is reminiscent of that of the lamin type V IFPs in which an additional 42-residue fragment is inserted in coil 1b (McKeon *et al.*, 1986; Fisher *et al.*, 1986). However, in lamins, the insertion continues the heptad repeat of the coiled-coil arrangement, whereas in the peripherin isoform 3u, the 32-residue sequence would rather correspond to a spacer joining two subsegments inside coil 2. A putative spacer dividing coil 2 of IFPs has been localized either prior to the centrally located tryptophan or 12 residues past that tryptophan (Geisler and Weber, 1986; Steinert and Roop, 1988), each location close to that of the 3u-specific sequence. Since this spacer does not exceed 8–10 residues in all IFPs, the 32 residues of the isoform 3u might rather indicate a particular loop or hinge region and might prevent the filament formation.

(ii) The isoform 5b of peripherin would be 14 amino acids shorter than the isoform 5g and the eight C-terminal amino acids would be different, due to a shift in the coding phase. Such a frame-shifted overlap of RNAs resulting in differing carboxy-termini for two proteins encoded by the same DNA segment has been observed with other proteins such as the polyoma middle and large T antigens (Ziff, 1980). The tail regions of all IFPs are variable in length and sequence even within one type. At least their more carboxy-terminal halves are located outside the filament wall, keeping individual filaments apart and possibly interacting with other molecules within the cytoplasm (Geisler and Weber, 1986; Steinert and Roop, 1988). The carboxy-terminal end of the isoform 5b in which three lysine and four glutamic–aspartic residues have been exchanged for one arginine and one glutamic residues may therefore associate with proteins which differ from those interacting with the carboxy-termini of the isoform 5g.

The major questions which we are now investigating are whether there exist other minor mRNA species and what is the relative importance of all the different species in the neuronal cell populations where peripherin is expressed. Correlatively, we are investigating the relevant polymorphism of peripherin and the involvement of the different isoforms in filament formation and in the interactions with the surrounding proteins.

Materials and methods

Cell culture

Culture of the mouse neuroblastoma clonal line NIE 115 was carried out as previously described (Croizat *et al.*, 1979). When cultured in growth medium supplemented with 1% (w/v) 1-methylcyclohexane carboxylic acid, the cells acquire a neuronal phenotype.

Antibodies

An antiserum was raised against peripherin by injecting a rabbit with mouse neuroblastoma peripherin isolated from two-dimensional gel electrophoresis. Specificity was checked by immunoblotting of various cellular extracts separated on two-dimensional gels; cross-reactivity with neuroblastoma peripherin was observed with 1:2000 to 1:5000 dilutions. No cross-reactivity was observed with vimentin or with NF-L. This antiserum was also used concomitantly in an immunohistochemical study of the expression of peripherin during rat ontogenesis (Escurat *et al.*, 1988). The murine monoclonal IFA antibody (Pruss *et al.*, 1981) was a kind gift of Dr Brian Anderton; it was used as hybridoma supernatant at dilution 1:5.

Construction and immunoscreening of the cDNA library

The library was constructed in the λgt11 vector (Huynh *et al.*, 1985). Total RNA was isolated by the LiCl–urea extraction method (Auffray and Rougeon, 1980) from a 1-day-CCA-treated NIE 115 culture. Poly(A)$^+$ RNA was selected by one cycle of binding to oligo(dT)–cellulose (Collaborative Research) and was ~3% of the RNA loaded onto the column. The presence of peripherin messenger was ascertained by *in vitro* translation and analysis of the translation products by two-dimensional gel electrophoresis, as previously described (Portier *et al.*, 1984a). The amount of [^{35}S]methionine incorporated into peripherin was 0.15% of the protein radioactivity. Double strand blunt-end cDNA was synthesized using a cDNA synthesis kit (Amersham), methylated with *Eco*RI methylase (Appligene) and ligated to phosphorylated *Eco*RI linkers (Pharmacia). The cDNA was digested with *Eco*RI, purified by chromatography on Sepharose 4B (Pharmacia) and ligated to *Eco*RI-digested phosphatase-treated λgt11 arms (Promega-Biotech). Recombinant DNA was packaged using an *in vitro* packaging kit (Gigapack gold, Stratagene) and resulting phages were plated on *Escherichia coli* Y1090 R$^-$M$^+$. The library contained 10^6 clones, half of which were amplified in the same *E.coli* strain. The amplified library was plated at a density of 0.5×10^5 p.f.u. per 15 cm plate. Nitrocellulose filters were left in contact for 3.5 h with the plates and then treated as described by Huynh *et al.* (1985) with a polyclonal antibody against mouse neuroblastoma peripherin at a 1:1000 dilution. The antibody retained on the filters was reacted with alkaline phosphatase-conjugated anti-rabbit IgG (Fc) (Promega-Biotech) and revealed by the phosphatase coloured reaction (Tuan and Fitzpatrick, 1986). Positive plaques were picked and purified by two to four subsequent rounds of screening on 10 cm plates. Dot blots of the purified clones were performed as described by Young and Davis (1983), using various antibodies. The identity of the cDNAs was confirmed by epitope selection (Weinberger *et al.*, 1985): immobilized fusion proteins from individual clones were used to select epitope-specific antibody which was subsequently eluted and identified by binding to protein blots of cellular extracts.

Subcloning and DNA sequencing

cDNA inserts were isolated from plate lysates (Maniatis *et al.*, 1982) of three clones, 3u, 5b and 5g, and subcloned into pUC18 and M13 bacteriophage. Single-stranded DNA templates were prepared from M13 cultures and sequenced by the dideoxynucleotide chain-termination method (Sanger *et al.*, 1977), using [^{35}S]dATP and sequenase enzyme (USB) (Tabor and Richardson, 1987). Products were analysed on denaturing acrylamide gels. Both strands were sequenced and synthetic 17mer oligonucleotides were used as primers. Some GC-rich regions were sequenced using 7-deaza dGTP instead of dGTP, with the Klenow large fragment of DNA polymerase (5 U/μl) (Boehringer).

RNA blotting and hybridization

Poly(A)$^+$ RNA was denatured and size fractionated on a 1% agarose gel in the presence of 0.8% formamide, and blotted onto a nitrocellulose filter (Maniatis *et al.*, 1982). The filter was baked at 80°C for 2 h, prehybridized in 50% formamide, 5 × SSPE, 1 × Denhardt's solution, 0.5% SDS, 250 μg/ml salmon sperm DNA, for 2 h at 42°C. Hybridization was performed

in the same buffer with 3×10^7 c.p.m./ml of ^{32}P-labelled restriction fragment for 15 h at 42°C. Probe was obtained from clone 3u and then ^{32}P-labelled by nick translation to a specific activity of $\geq 10^7$ c.p.m./μg. Following hybridization, the filter was washed three times at room temperature with 2 × SSC−0.1% SDS, for 15 min each, then once with 0.1 × SSC−0.1% SDS at 65°C for 15 min and exposed overnight to a Kodak X-Omat AR film at −70°C with an intensifier screen.

S1 mapping

The procedure was a modification of the method described by Berk and Sharp (1977). An oligonucleotide and a cDNA fragment were used as probes. The synthetic oligonucleotide (1.2 pmol) was 5′ end-labelled with 10 units of T4 polynucleotide kinase (Biolabs) using [γ-^{32}P]ATP (3000 Ci/mol). The cDNA fragment (2.5 pmol) was treated with 0.05 U of alkaline phosphatase, then heated at 80°C, phenol-extracted and ethanol-precipitated prior to 5′ end-labelling as reported above. Labelled probes (40 000 c.p.m./labelled end) were precipitated with neuroblastoma poly(A)$^+$ RNA (1.5 μg) or with *E.coli* tRNA (1 μg), then boiled 10 min at 90°C and annealed to RNA in 80% formamide, 0.4 M NaCl, 0.04 M Pipes, pH 6.4, overnight at 42 or 50°C. S1 nuclease cleavage was for 90 min at 37°C with 200 U of S1 nuclease (Amersham) and the protected products were analysed on a sequencing gel with a sequence sample as size marker. Time of exposure of Kodak X-Omat AR film is indicated in the legend to Figure 7. After development, signal intensity was measured with a Biorad videodensitometer and a Shimadzu Chromatopac CR3A integrator.

Hybrid selection, translation and immunoprecipitation

Selection of mRNA was performed using 20 μg of plasmid 3u or 5g, or 2 μg of a 96mer oligonucleotide complementary to the 3u-specific sequence, immobilized on 0.4 × 0.4 cm squares of nitrocellulose after NaOH denaturation and neutralization. The filters were baked, then prehybridized in 65% formamide, 20 mM Pipes, pH 6.4, 0.4 mM NaCl, 0.2% SDS and 0.1 mg/ml *E.coli* tRNA for 3 h at 50°C. For hybridization, poly(A)$^+$ RNA was prepared at 10 μg/ml in 0.13 ml of the same solution and heated at 70°C for 10 min. The filters were incubated in the solution for 3 h at 50°C, then removed and washed extensively at 65°C 10 times with 10 mM Tris−HCl, pH 7.6, 0.15 M NaCl, 1 mM EDTA, 0.5% SDS and twice with the same solution without SDS. In the case of the selection by the 96mer oligonucleotide, poly(A)$^+$ RNA was 100 μg/ml and hybridization was at 42°C in 65% formamide. The RNA was eluted from the hybrid by boiling each filter for 1 min in 0.3 ml of 0.1 mg/ml *E.coli* tRNA, then quick-frozen in a methanol−dry ice bath and thawed. The RNA was extracted with phenol−chloroform and precipitated with ethanol overnight at −20°C. The entire precipitate was used in a translation reaction. *In vitro* translation was performed as previously described (Portier *et al.*, 1984a), using the reticulocyte lysate system with [^{35}S]methionine as label. Translated products were immunoprecipitated (Jagus, 1987) using the antiperipherin serum and analysed on one-dimensional SDS−PAGE with subsequent fluorography.

Acknowledgements

We would like to thank Mrs Danièle Gros for her expert technical assistance. We are indebted to Dr Domenico Libri for his advice on S1 mapping, to Dr Benoît Robert for his generous gift of mouse genomic blots and to Dr Margaret Buckingham for helpful discussion and critical comments on the manuscript. This work was supported by grants of the Centre National de la Recherche Scientifique, the INSERM, the Ministère de l'Industrie et de la Recherche, the Fondation pour la Recherche Médicale and the Ligue Française contre le Cancer.

References

Aletta,J.M., Angeletti,R., Liem,R.K.H., Purcell,C., Shelanski,M.L. and Greene,L.A. (1983) *J. Neurochem.*, **51**, 1317−1320.
Auffray,C. and Rougeon,F. (1980) *Eur. J. Biochem.*, **107**, 303−314.
Balcarek,J.M. and Cowan,N.J. (1985) *Nucleic Acids Res.*, **13**, 5527−5543.
Berk,A.J. and Sharp,P.A. (1977) *Cell*, **12**, 721−732.
Breitbard,R.E., Andreadis,A. and Nadal-Ginard,B. (1987) *Annu. Rev. Biochem.*, **56**, 467−495.
Croizat,B., Berthelot,F., Ferrandes,B., Eymard,P., Sahuquillo,C. and Gros,F. (1979) *C.R. Acad. Sci., Ser. III*, **289**, 1283−1286.
Escurat,M., Gumpel,M., Lachapelle,F., Gros,F. and Portier,M.M. (1988) *C.R. Acad. Sci.*, **306**, 447−456.
Fisher,D.Z., Chaudhary,N. and Blobel,G. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 6450−6454.
Franke,W.W., Grund,C. and Achtstatter,T. (1986) *J. Cell Biol.*, **103**, 1933−1943.
Geisler,N. and Weber,K. (1982) *EMBO J.*, **1**, 1649−1656.
Geisler,N. and Weber,K. (1986) In Shay,J.W. (ed.), *Cell and Molecular Biology of the Cytoskeleton*. Plenum, NY, pp. 41−68.
Huynh,T.H., Young,R.A. and Davis,R.W. (1985) In Glover,D.M. (ed.), *DNA Cloning. A Practical Approach*. IRL Press, Oxford, Vol. 1, pp. 49−78.
Jagus,R. (1987) *Methods Enzymol.*, **152**, 296−304.
Julien,J.P., Ramachandran,K. and Grosveld,F. (1985) *Biochim. Biophys. Acta*, **825**, 398−404.
Lazarides,E. (1980) *Nature*, **283**, 249−256.
Leonard,D.G.B., Ziff,E.B. and Greene,L.A. (1987) *Mol. Cell. Biol.*, **7**, 3156−3167.
Leonard,D.G.B., Gorham,J.D., Cole,P., Greene,L.A. and Ziff,E.B. (1988) *J. Cell. Biol.*, **106**, 181−193.
Lewis,S.A. and Cowan,N.J. (1985) *J. Cell Biol.*, **100**, 843−850.
Maniatis,T.F., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
Marchuk,D., McCroham,S. and Fuchs,E. (1984) *Cell*, **39**, 491−498.
McKeon,F.D., Kirschner,M.W., and Caput,D. (1986) *Nature*, **319**, 463−468.
Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459−472.
Osborn,M. and Weber,K. (1986) *Trends Biochem. Sci.*, **11**, 469−472.
Parysek,L.M. and Goldman,R.D. (1987) *J. Neurosci.*, **7**, 781−791.
Parysek,L.M., Chisholm,R.L., Ley,C.A. and Goldman,R.D. (1988) *Neuron*, **1**, 395−401.
Perreau,J., Lilienbaum,A., Vasseur,M. and Paulin,D. (1988) *Gene*, **62**, 7−16.
Portier,M.-M., Brachet,P., Croizat,B. and Gros,F. (1984a) *Dev. Neurosci.*, **6**, 215−226.
Portier,M.-M., de Néchaud,B. and Gros,F. (1984b) *Dev. Neurosci.*, **6**, 335−344.
Pruss,R.M., Mirsky,R., Raff,M.C., Thorpe,R., Dowding,A.J. and Anderton,B.H. (1981) *Cell*, **27**, 419−428.
Quax,W., Egberts,W.V., Hendriks,W., Quax-Jeuken,Y. and Bloemendal,H. (1983) *Cell*, **35**, 215−223.
Quax,W., van den Broek,L., Egberts,W.V., Ramaekers,F. and Bloemendal,H. (1985) *Cell*, **43**, 327−338.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463−5467.
Steinert,P.M., Rice,R.H., Roop,D.R., Trus,B.L. and Steven,A.C. (1983) *Nature*, **302**, 794−800.
Steinert,P.M. and Roop,D.R. (1988) *Annu. Rev. Biochem.*, **57**, 593−625.
Tabor,S. and Richardson,C.C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4767−4771.
Thompson,M.A. and Ziff,E.B. (1989) *Neuron*, **2**, 1043−1053.
Tuan,R.S. and Fitzpatrick,D.F. (1986) *Anal. Biochem.*, **159**, 329−335.
Weinberger,C., Hollenberg,S.M., Ong,E.S., Harmon,J.M., Brower,S.T., Cidlowski,J., Thompson,E.B., Rosenfeld,M.G. and Evans,R.M. (1985) *Science*, **228**, 740−742.
Wilson,W.D., Dotrong,M.-H., Zuo,E.T. and Zon,G. (1988) *Nucleic Acids Res.*, **16**, 5137−5151.
Young,R.A. and Davis,R.W. (1983) *Science*, **222**, 778−781.
Zehner,Z.E. Li,Y., Roe,B.A., Paterson,B.M. and Sax,C.M. (1987) *J. Biol. Chem.*, **262**, 8112−8120.
Ziff,E.B. (1980) *Nature*, **287**, 491−499.

Received on November 15, 1988; revised on March 9, 1989

Note added in proof

After this manuscript was accepted, the paper on the alternative splicing of dystrophin mRNA [Feener,C.A., Koenig,M. and Kunkel,L.M. (1989), *Nature*, **338**, 509−511] was published which supplies an example for generating isoforms through a portion of a putative unspliced intron or through a modified ORF, similarly to the generation of the peripherin isoforms.