What can be learned from intermediate filament gene regulation in the mouse embryo

PHILIPPE DUPREY* and DENISE PAULIN

Laboratoire de Biologie Moléculaire de la Différenciation, Université Paris VII and Institut Pasteur, Paris, France

CONTENTS

Introduction .......................................................................................................................... 444
Sequential expression of intermediate filament during development .................................. 444
Glial Fibrillary Acidic Protein (GFAP): tissue-specific and dynamic expression patterns ........ 445
GFAP is the last type III IF to be expressed during development ........................................ 445
Alteration in GFAP synthesis in pathological situations ....................................................... 445
Both 5' upstream and intragenic elements control GFAP gene transcription ....................... 445
GFAP transgenes suggest further heterogeneity among astroglial cells .............................. 446
Peripherin: an additional neuronal filament ...................................................................... 447
Peripherin is expressed in neurons of different embryological origins ............................... 447
A tissue-specific repressor controls peripherin gene transcription ...................................... 447
Peripherin upstream and intragenic sequences are required to target neuronal expression in transgenic mice ................................................................. 448
Desmin: a muscle-specific IF protein ................................................................................. 449
Desmin is expressed in skeletal, cardiac and smooth muscles ............................................ 449
Two independent enhancers control desmin gene expression in skeletal muscle ............... 449
Desmin transgenic mice provide evidence for distinct cardiac, smooth and skeletal muscle gene programs .......................................................... 451
Vimentin: a potentially ubiquitous IF ............................................................................. 451
Vimentin is the first type III IF to be expressed during development ................................. 451
Multiple enhancer and silencer elements control vimentin gene expression ....................... 451
Vimentin transgene recapitulates the complex and dynamic expression pattern .................. 452
Discussion ......................................................................................................................... 452
Conclusion ......................................................................................................................... 453
Summary and key words .................................................................................................... 454
References ......................................................................................................................... 454

*Address for reprints: Laboratoire de Biologie Moléculaire de la Différenciation, UFR de Biochimie, tour 42, Université Paris VII, 2 Place Jussieu, 75005 Paris, France. FAX: 33.1.44277907.

This paper is dedicated to Hubert Condamine.
Introduction

Embryonic development involves generation of numerous cell types, each of which is characterized by the expression of specific sets of genes. Analysis of the mechanisms controlling cell-specific transcription opens the route towards the identification of regulators' networks. For this purpose, genes encoding intermediate filament proteins (IFs) provide an attractive model system (Franke et al., 1978; Lazarides, 1980, 1982; Geisler and Weber, 1981, 1982, 1983; Osborn and Weber, 1982; Geisler et al., 1983, 1984; Traub, 1985; Weber and Geisler, 1985; Steiner and Roop, 1988; Zehner, 1991; Stewart, 1993).

IFs constitute a multigenic family whose members are expressed in a cell-specific manner. Among IF genes, sequence homologies are concentrated in the coding part, whereas regulatory sequences are considerably more divergent.

According to sequence homologies and exon-intron organization, the different IF genes have been classified into five classes (Table 1). The members of four classes are expressed in the cytoplasm. Lamins, which are found in the nucleus, constitute an independent class (type V) (Georgatos and Blobel, 1987). However a number of additional IF proteins have recently been identified in various vertebrate and invertebrate organisms (Weber et al., 1991; Brunkener and Georgatos, 1992; Hemmati-Brivanlou, 1992; Riemer et al., 1992; Gounari et al., 1993; Liem, 1993; Remington, 1993).

Type I (acidic) and type II (basic) cytokeratins found as heteropolymers in epithelial cells constitute the first two classes. The third class (type III) contains four members with different expression profiles: while desmin is found in muscle, GFAP in astrocytes, and peripherin in some neurons, vimentin is observed in many different cells. The members of the fourth class, (type IV), α-internexin, nestin and the neurofilament triplet, are all expressed in neural cells. In addition, nestin is also expressed in skeletal muscle precursors. Recently, two other proteins, filensin and tanabin, which share structural organization and sequence homology with intermediate filaments, have been characterized (Brunkener and Georgatos, 1992; Hemmati-Brivanlou, 1992; Gounari et al., 1993; Remington, 1993). Expressed in lens, filensin is associated with phakinin (Meredes et al., 1993) while tanabin can be found in the neuronal growth cones.

Sequential expression of intermediate filaments during development

The first intermediate filaments are detectable in oocytes and are composed of type 5, 6, 8 and 16 cytokeratins (Chisholm and Houliston, 1987; Lehtonen et al., 1987; Gallicano et al., 1994). Low expression of types 8 and 18 cytokeratins can be detected in early cleavage and morula stage mouse embryos (Duprey et al., 1985; Chisholm and Houliston, 1987; Lehtonen et al., 1987). The expression of types 8 and 18 cytokeratins increases at the blastocyst stage, where they are mostly expressed in the trophoderm cells (Bruilet et al., 1980; Jackson et al., 1980; Paulin et al., 1980; Oshima 1982; Duprey et al., 1985). Pairs of the different keratins will be expressed in the various epithelia later during development (Moli et al., 1982; Powell and Rogers, 1990; Fuchs, 1991; Vassar et al., 1991; Coulombe, 1993). Although most epithelial cells express only cytokeratins, transient expression of vimentin has been described in migrating parietal endodermal cells from 8.5 days p.c. mouse embryos (Lane et al., 1983; Lehtonen et al., 1983). In mesenchymal cells the first appearance of vimentin filaments was reported in primitive mesoderm cells at day 8.5 between the ectoderm and proximal endoderm layers (Jackson et al., 1981; Franke et al., 1982).

Vimentin is characterized by a transient expression in some precursor cells during early development. Vimentin is expressed in neural precursors and mesenchymal cells prior to the accumulation of the tissue-specific filaments (Hotzler et al., 1982; Lazarides, 1982; Cochrard and Paulin, 1984). In few cases, coexpression of two types of intermediate filament is maintained. This was shown to be the case for some glial cells, vascular smooth muscle cells and in the plexiform layer of the retina (Shaw et al., 1981; Dräger, 1983). In most cases, tissue specific filaments progressively replace vimentin, thus leading to the transient coexistence of several different intermediate filament proteins. Another intermediate filament protein, nestin, is transiently expressed in both myogenic precursors from the somites and neuroepithelial precursors of the central nervous system (Lendahl et al., 1990; Sejersen and Lendahl, 1993; Zimmerman et al., 1994). α-internexin is found in the neural tube and neural crest-derived neuroblasts before the neurofilament proteins appear (Liem, 1993). Post-mitotic neuroblasts, which are first generated at day 9 p.c., express the neurofilament proteins. A subset of these postmitotic neuronal precursors is characterized by the presence of an additional intermediate filament protein, peripherin. The last filament system to appear in the neural system is GFAP, which is expressed in astrocytes and Schwann cells at the end of gestation (E18). In striking contrast to the diversity of intermediate filaments in the nervous system, only one muscle-specific filament has been found in adult mouse cardiac, skeletal and smooth muscle. In the heart, desmin expression can be initially detected at day 8 p.c. while in the somitic skeletal muscle precursors, it appears shortly before day 9 p.c. On the other hand, transient expression of cytokeratins has been observed in developing human fetal heart (Kuruc and Franke, 1988) and the expression of cytokeratins has been reported in human smooth muscle (Azumi et al., 1988).

It is noteworthy that each structural class of intermediate filaments (see Table 1) corresponds to major expression lineages.
such as cytokeratins in epithelia or neurofilaments in neurons, except for type III filaments, vimentin, desmin, GFAP and peripherin, which are expressed in different lineages. Recent reviews on intermediate filament regulation focused on epithelia (Oshima, 1992) or neurons (Liem, 1993) and thus discussed the regulation of highly related genes within a given cell lineage.

Type III IF genes encode highly related proteins, and the organisation of these genes is very similar. Indeed they are encoded by rather small genes (not exceeding 10 kbp) all made of 9 exons and 8 introns. By contrast the control sequences are completely divergent. These similarities and differences reinforce the interest of the use of type III IF as a model to analyze cell-specific gene regulation.

In this review we will focus on the regulation of type III IF gene expression.

**Glial Fibrillary Acidic Protein (GFAP): tissue-specific and dynamic expression pattern**

**GFAP is the last type III IF sequence to be expressed during development**

Glial Fibrillary Acidic Protein is a 54 kDa (430 aminoacids) type III intermediate filament protein which is the major component of astroglial intermediate filaments (Bignami et al., 1972; Maunoury et al., 1976; Lewis et al., 1984; Tardy et al., 1988). GFAP is expressed in most, but not all, astroglial cells of adult mammals. GFAP expression has been detected in tanyocytes (De Vity et al., 1981), cerebellar golgi epithelial cells, retinal Müller cells, fibrous astrocytes and to a lesser degree, in protoplasmic astrocytes. During embryonic development GFAP is not expressed by radial glial cells, the exclusive astroglial cells of the embryonic central nervous system (CNS), which express vimentin (Dahl et al., 1981). In the mouse, GFAP expression was first detected at day 16 p.c. in the 4th ventricle (Galou, unpublished result) and at day 17 p.c. in the retina (Boloventa et al., 1987). However in the rat immunochemical studies allow us to detect GFAP as early as at day 14 p.c. in the brain and day 16 p.c. in the spinal cord (Noetzl and Agrawal, 1985). Later on, immunocytochemical methods revealed the presence of GFAP in the rat embryo CNS at day 17 p.c. in the dentate gyrus (Rickmann et al., 1987), while other studies described GFAP staining at day 18 p.c. in rat embryo brain and spinal cord (Raju et al., 1981; Valentino et al., 1983).

GFAP expression further increases after birth, concomitantly with the gradual disappearance of vimentin intermediate filaments in most, but not all, astroglial cells (Eng, 1980, 1985; Lewis and Cowan 1985; Tardy et al., 1988; Mokuno et al., 1989; Sarthy et al., 1991). Indeed, it was observed that, while many astroglial cells still co-express GFAP and vimentin in the adult, a subset of these cells does not express GFAP. Expression of GFAP or of a highly similar protein was also detected in some non-astroglial cells such as Schwann cells, peripheral glial cells and perisinusoidal cells from rat liver.

While GFAP expression is mostly detected in cells of neural origin, some extra-neural GFAP expression has been reported. In the mouse, GFAP expression is detected in epithelial cells of the lens. This lens expression of GFAP is only detected in some mouse species. For instance, GFAP expression is detected in *Mus musculus* but not in *Mus spretus* mice (Boyer et al., 1991).

**Alteration in GFAP synthesis in pathological situations**

In the adult, increased GFAP expression is observed in several CNS pathologies such as Alzheimer's disease (Delacourte, 1990), epilepsy and multiple sclerosis (Eng et al., 1971; Eng, 1980). Theses pathologies are characterized by astrocytic hyperplasia and hypertrophy correlated with a marked increase in GFAP expression which can also be detected in cells surrounding glioma as well as upon trauma and ageing (Beach et al., 1989). Developmental and pathological expression of GFAP in astroglial cells has led to the hypothesis of the involvement of this protein in some of the functions of these cells which include neurotransmitter processing, control of extracellular channel activity, vascular-brain barrier and immune reactions. However the recent disruption of the mouse gene encoding GFAP through homologous recombination in embryonic stem cells argues against this hypothesis. Indeed mice devoid of GFAP develop normally. No obvious anatomical abnormalities have been observed and the astrocytic response against injury occurs normally when the mutant mice are infected with prions (Gomi et al., 1995).

**Both 5' upstream and intragenic elements control GFAP gene transcription**

So far, the GFAP gene has been isolated and characterized in rat, mouse and human. The human gene was found to be located on chromosome 17, while the mouse gene is located on chromosome 11 (Boyer et al., 1991). Regulatory sequences of the mouse gene have been extensively analyzed both in vitro in astroglial cell lines and in vivo in transgenic mice.

In the case of the murine GFAP gene, whole promoter sequence and transcriptional start sites have been determined (Miura et al., 1990, 1991). In those studies, the luciferase gene was set under the control of various deletion mutants of the murine GFAP gene promoter and the resulting luciferase activity was measured in GFAP-positive (rat C6 glioma) and negative (neuroblastoma, and NIH-3T3 fibroblastic) cells. It was observed that sequences required for tissue-specific expression were located within 256 bp from the transcription start point. Using DNase footprinting, three binding sites for transcription factors designed as GFI (-104 to -82), GFI (-124 to -104) and GFI3 (-163 to -140) were defined in this region (Fig. 1).

The GFI binding site is homologous to the binding site for the AP2 factor, while GFI1 and GFI3 binding sites correspond to NFI

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Types</th>
<th>Sub-units</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Acidic Cytokeratins</td>
</tr>
<tr>
<td>II</td>
<td>Basic Cytokeratins</td>
</tr>
<tr>
<td>III</td>
<td>Desmin, GFAP, Peripherin, Vimentin</td>
</tr>
<tr>
<td>IV</td>
<td>Neurofilaments, IL light, (M) medium, (H) Heavy</td>
</tr>
<tr>
<td>V</td>
<td>α-interneurinin, Nestin</td>
</tr>
<tr>
<td>Other non classified</td>
<td>Lamin A, B, C</td>
</tr>
</tbody>
</table>

---

**Intermediate filament gene expression in the mouse embryo** 445
These studies, exploiting promoter-reporter gene fusions, suggested that the sequences required for the cell-specific expression of the murine GFAP gene, were located close to the transcriptional start site. However, other investigators, using the entire murine GFAP gene found evidence for additional important elements located in the intragenic regions. Using such an approach, it was found that an upstream region located between nucleotides -1631 and -1479 is a strong glial-specific positive element, while a negative regulatory element located in the first intron prevents transcription in non-glial cells such as Hela cells (Sarkar and Cowan, 1991). In this study, the glial-specific proximal element observed by Miura et al. (1990) was not observed. This might suggest that a negative interaction occurs between intragenic negative regulatory sequences and the previously described proximal cell-specific element. The significance of this potential interaction deserves further experimentation.

Other investigations using transfection methods demonstrated the presence of two cell-specific negative regulatory elements within the related rat GFAP gene (Kaneko and Sueoka, 1993). GDR1, a 2.7 kb region extending from the first intron to the fifth exon, prevents the activation of a 10.8 kb promoter in non-neural tissue. Addition of GDR2, a 1.7 kb region located downstream of the polyadenylation site, further suppresses expression in neuronal cells. These negative regulatory elements function in a position-independent manner, and both of them were found to be required for the proper expression of the GFAP gene (Fig. 1).

Analysis of the human GFAP gene has led to the discovery of an additional initiator site, located downstream from the transcriptional start site (nucleotides +10 to +40). These sequences are perfectly conserved in the murine GFAP gene (Nakatani et al., 1990; Besnard et al., 1991).

While much attention has been devoted to the tissue-specific expression of the GFAP gene, it is well established that GFAP expression can also be modulated by hormonal stimuli. Indeed both sex steroids, growth factors and cytokines can affect GFAP expression (Laping et al., 1994). Comparative analysis of the 5' upstream sequences of the rat, mouse and human genes has shown the occurrence of positive response elements, which are conserved between the three genes (Laping et al., 1994). However, no functional analysis of these possible control sequences has been reported yet.

Altogether, these investigations using cell lines suggest that multiple elements are required for the proper expression of the GFAP gene. Indeed, control of GFAP gene transcription is conferred by the combination of ubiquitously active positive regulatory elements and tissue-specific negative control regions (Fig. 1). However, the ultimate demonstration of the importance of these multiple regulatory elements requires in vivo experiments using a transgenic mouse approach.

**GFAP transgenes suggest further heterogeneity among astroglial cells**

Two different constructs derived from the murine GFAP gene have been used to generate transgenic mice. In a first study, the lac-Z gene was inserted in frame with the coding sequence of the GFAP gene. The resulting construct contained all introns, 2 kb of 5' upstream sequences, and 2.5 kb of the 3' flanking sequences fused to the lac-Z gene. During insertion of the lac-Z gene, possible additional murine initiator elements (located within a 50
Intermediate filament gene expression in the mouse embryo

Fig. 2. Expression of GFAP-lac-Z in the embryo and in the adult mouse brain.
Upstream sequences (-1913 to +92) from the murine GFAP gene were linked to the lac-Z reporter gene and the resulting construct was used to generate transgenic mice. (A) Initial expression of GFAP-lac-Z transgene in 16-day p.c. embryo: lateral view. (B) Brain section was stained for beta-galactosidase activity (nuclear blue staining), the same section is treated for immunodetection of GFAP revealed by peroxidase-conjugated antibodies (brown color). Note that the transgene is expressed in most but not all GFAP-positive cells, as revealed by the antibody staining. (Unpublished data from P. Dupouey, M. Galou and L. Lossouarn).

bp fragment of the first intron of the GFAP gene) were deleted (Mucke et al., 1991). Another construct fused upstream sequences between -1913 and +92 to the lac-Z gene. (Galou et al., 1994). In both cases the transgene was expressed in astroglial cells. However a large fraction of astroglial cells expressing endogenous GFAP did not express either transgene, suggesting the occurrence of further heterogeneity among astroglial cells (Fig. 2). On the other hand, expression of both transgenes could be detected in some neuronal cells. Thus it is clear that other mechanisms than those defined using cell lines must be involved in the control of the proper expression of the GFAP gene.

Although these transgene expression patterns slightly differ from the endogenous GFAP pattern, it was observed that both transgenes began to be expressed at embryonic day 16 (Fig. 2). In adult mice, induction of experimental gliosis led to an increased expression of the transgenes and to the induction of their expression in previously lac-Z negative astroglial cells. Thus these two types of transgenes will be very useful for deciphering the mechanisms involved in the control of the GFAP gene during normal ontogeny and in pathological situations.

Peripherin: an additional neuronal filament

Peripherin is expressed in neurons of different embryological origin

Peripherin is a 57 kDa (451 amino acids) type III IF protein whose expression is restricted to a subset of neuronal cells. Its expression was first described in neuroblastoma cells derived from the peripheral nervous system and in PC12 pheochromocytoma cells where expression of peripherin mRNA is enhanced by nerve growth factor (NGF) (Portier et al., 1984; Leonard et al., 1988). In fact, its expression is also observed in neuronal cells not derived from the peripheral nervous system (Escurat et al., 1990; Troy et al., 1991).

In adult rodents (rat and mouse), peripherin is found in all sensory neurons, even though these cells are of various embryonic origins (neural crest: spinal ganglionic neurons; placodes: acoustic and olfactory neurons; neural tube: retinal ganglionic neurons). Peripherin expression is also detected in some neural tube-derived motor neurons in retinal preganglionic neurons and in neural crest-derived spinal ganglia and sympathetic neurons. All these neurons have different embryological origins. Peripherin expression in these neurons correlates with the fact that they all extend their neurites outside the brain-spinal cord axis. Additional peripherin-expressing cell types include medullosurrenal cells and pancreatic endocrine cells, which raises the question of the embryonic origin of these cells.

In the mouse, initial expression of peripherin is first detected at day 9 p.c. while rat embryos start to express this protein at day 11 p.c. In both cases initial peripherin expression parallels or immediately follows expression of the 70 kDa neurofilament protein. As the embryo develops, the peripherin expression domain will gradually extend to all neuronal cell types which will express it in the adult. (Escurat et al., 1990; Gorham et al., 1990; Troy et al., 1991).

A tissue-specific repressor controls peripherin gene transcription

The peripherin gene has been cloned in rat and mouse. In the mouse, the peripherin locus is located in the E-F region of chromosome 15, while the human gene has been assigned to chromosome 12 q12-q13 (Moncla et al., 1992). The regulation of peripherin gene expression was analyzed mostly in cultured neuroblastoma and PC12 pheochromocytoma cells (Desmarais et...
In both rat (PC12) and mouse (NI8TG2) peripherin-expressing cell lines, the use of a series of peripherin promoter deletion mutant constructs led to the conclusion that the peripherin upstream control region contains alternating positive and negative elements. It was observed that only 98 bp of the mouse peripherin gene promoter suffice to confer cell specific expression to the reporter lac-Z gene, while no activity was observed in negative cell lines, rat and mouse glioma, mouse olfactory bulb and mouse lung carcinoma.

This 98 bp proximal cell-specific region contains three areas, PER1 (TATAAAGCCGCCCCGCATCGGTCT), PER2 (CCCCCACCACC) and PER3 (TGGGAGGAGC), which are protected by nuclear factors (Fig. 1) (Desmarais et al., 1992), while the PER2 and PER3 elements interact with widely distributed factors, the PER1 sequence binds factors that are only found in peripherin-expressing cells. The PER1 and PER3 elements are perfectly conserved in rat and mouse, while the PER2 motif is not found in the rat gene. However the corresponding rat sequence is also C-rich. The PER3 element is found in the promoter regions of all type III IF genes and bears homology to the binding sequence for the E. coli MalT helix-turn-helix factor (Richet et al., 1991). The PER1 sequence, which includes the TATA box of the gene, is distinct from the known TATA box binding factor consensus sequences and has not been described previously. Furthermore, it was shown that the factor that interacts with the PER1 motif bears no affinity to TA-rich sequences. Therefore, it seems likely that the PER1 factor is not related to any TATA box-binding factors. Mutational analyses were conducted to determine the relative contributions of the three elements to the activity of the peripherin promoter. From these studies it appears that PER2 and PER3 are important for determining the level of expression, while PER1 was important for cell type specificity. Therefore, it seems that peripherin promoter activity is controlled by the combination of non-specific positive elements containing the PER2 and PER3 motifs and a negative cell-specific element corresponding to the PER1 sequence (Desmarais et al., 1992). In this respect, the mechanisms involved in the control of the activity of the peripherin promoter are very similar to the ones controlling cell-specific expression of the GFAP gene.

An additional level of regulation of peripherin promoter activity is observed upon addition of NGF to PC12 cells, which leads to further activation of peripherin gene transcription (Thompson et al., 1992). This is correlated with an increase in promoter activity due both to the release of a negative control mechanism involving sequences located between nucleotides -111 and -179 and to the activation of an upstream positive regulatory element located between nucleotides -2680 and 2290 (Fig. 1) (Thompson et al., 1992).

**Peripherin upstream and intragenic sequences are required to target neuronal expression in transgenic mice**

While most studies dealing with peripherin regulation have been performed using cell lines, peripherin gene-lac-Z fusion constructs have been recently used to generate transgenic mice (Fig. 3). Up to 5.8 kb of peripherin upstream sequences have been used. Analysis of the transgenic mice, which is in progress, shows expression in peripherin-expressing cell types with some additional ectopic expression sites which vary according to the transgenic mouse line.

These embryos present an expression in the peripheral nervous system (Fig. 3) (trigeminal ganglia and maxillary branches; cranial nerves; dorsal root ganglia; sympathetic chain) but ectopic expression was found in mesoderm (cartilages of limbs) and in the central nervous system. These results suggest that the 5.8 kb upstream region of the mouse peripherin gene is not able to drive the expression of lacZ gene in a tissue-specific manner. This result is comparable to that obtained by Begemann et al. (1990) for promoters of other genes (insulin and neurofilament).

Using another construct with 5.1 kb of additional intragenic sequences added, the spatial and temporal patterns of transgene expression are consistent with the reported developmental regulation of the endogenous peripherin gene. In the absence of
intragenic sequences, ectopic expression was obtained. This result illustrates the role of intragenic sequences in the regulation of the peripherin gene.

Desmin: a muscle-specific IF protein

**Desmin is expressed in skeletal, cardiac and smooth muscles**

Desmin is a 55 kDa muscle cytoskeletal protein (468 amino acids) whose gene belongs to the class III of intermediate filaments (Lazarides and Hubbard, 1976; Small and Sobieszak, 1977; Geisler and Weber, 1982). Desmin is one of the first muscle-specific proteins to be detected in the mammalian embryo (Table 2) since it is expressed before titin, skeletal muscle actin, myosin heavy chains and nebulin (Hill et al., 1986; Fürst et al., 1989; Babai et al., 1990). During mouse embryonic development, desmin has been first detected at 8.25 days p.c. in the ectoderm where it was transiently coexpressed with keratin and vimentin (Schaart et al., 1989). Desmin has been detected in the heart rudiment at 8.5 days p.c. and its expression increased in the myocardial cells during subsequent cardiogenesis (Schaart et al., 1989). From 9 days p.c. onwards, desmin expression in newly formed myotomes follows the somitic maturation rostro-caudal gradient (Table 2), (Mayo et al., 1992). By day 11 p.c., clusters of desmin positive cells can be detected in cephalic muscles. In limb buds, by day 12 p.c., primary fibers display desmin expression (Table 2). Subsequently, the levels of desmin expression in skeletal, cardiac and smooth muscles remain high throughout embryogenesis and in early postnatal life (Li et al., 1993). Finally, both smooth muscles from visceral organs and from blood vessels express desmin (Table 2).

**Two independent enhancers control desmin gene expression in skeletal muscle**

Desmin is derived from a fully characterized single copy gene (Quax et al., 1985; Li et al., 1989; Li and Capetanaki, 1993) which has been mapped to band q35 of the long arm of human chromosome 2 (Viegas-Péquignot et al., 1989) and band C3 of mouse chromosome 1 (Li et al., 1990).

The expression of the desmin gene differs from that of most other genes sharing the characteristic of being repressed in proliferating undifferentiated myoblasts, and whose expression is activated concomitantly with myoblast fusion. In contrast, desmin expression is initiated in replicating myoblast and accumulates to a high level as muscle cells differentiate (Pieper et al., 1987; Kaufman and Foster, 1988; Li and Paulin, 1991). This pattern is recapitulated in the mouse C2 myogenic cell line, where desmin is expressed at low levels in myoblasts. To date, all the muscle-specific enhancers, such as the rat and mouse muscle creatine kinase enhancer (Jaynes et al., 1988; Sternberg et al., 1988; Horlick and Benfield, 1989), rat myosin light chain (MLC1/3) enhancer (Donoghue et al., 1988), quail tropomyosin I enhancer (Yutzey et al., 1989), chicken and mouse acetyl choline receptor β-subunit enhancer (Baldwin and Burden, 1989; Wang et al., 1990) and chicken acetyl choline receptor α-subunit enhancer (Wang et al., 1988) function in differentiated cells but not in their precursors.

In the human desmin gene, a positive regulatory element that is important for high-level expression, is located in the region between -973 and -693 bp relative to the start site of transcription and shows characteristics of a muscle-specific enhancer (Fig. 1). This enhancer can function not only in differentiated myotubes, but also in undifferentiated myoblasts (Li and Paulin, 1993). Downstream of this enhancer, a negative region located between nt -693 and -228 has been found (Fig. 1). Between nt -228 and +75, another positive region containing a classical TATA box and transcriptional initiation sites was defined as the desmin promoter and can confer low level muscle-specific expression to a CAT reporter gene (Fig. 1) (Li and Paulin, 1991). The mouse desmin promoter was recently characterized and found to have a similar functional organization (Li and Capetanaki, 1993). In the case of the human gene, the enhancer was further characterized and found to be composed of two independent enhancer elements (Li and Paulin, 1993). One of these is active only in myotubes, while the activity of the second region has been detected only in myoblasts (Fig. 1). The myoblast-specific region contains one MyoD1 site (TTGGCAGCTTGTG) and one MEF2 site (TCTATAATCC). These two sites are necessary for full enhancer activity in myotubes (Fig. 1). MyoD1 is the prototype of a small family of helix-loop-helix (HLH) transcription factors that play a key role in the control of skeletal myogenesis (for reviews see

---

**TABLE 2**

<table>
<thead>
<tr>
<th>Vimentin</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>primitive mesoderm</td>
<td>6</td>
</tr>
<tr>
<td>notochord</td>
<td>8.5</td>
</tr>
<tr>
<td>cephalic mesenchyme</td>
<td>8.5</td>
</tr>
<tr>
<td>heart epimyocardia</td>
<td>8.5</td>
</tr>
<tr>
<td>blood vessels</td>
<td>8.5</td>
</tr>
<tr>
<td>somatopleurala</td>
<td>8.5</td>
</tr>
<tr>
<td>splanchnopleurala</td>
<td>8.5</td>
</tr>
<tr>
<td>otic arch</td>
<td>9</td>
</tr>
<tr>
<td>spinal gaaligia</td>
<td>9</td>
</tr>
<tr>
<td>cardiac trabeculae</td>
<td>10.5</td>
</tr>
<tr>
<td>intestine mesenchyme</td>
<td>10.5</td>
</tr>
<tr>
<td>liver mesenchyme</td>
<td>10.5</td>
</tr>
<tr>
<td>kidney mesenchyme</td>
<td>10.5</td>
</tr>
<tr>
<td>sympathetic ganglia</td>
<td>14</td>
</tr>
<tr>
<td>spinal cord</td>
<td>14</td>
</tr>
</tbody>
</table>

Desmin

| heart | 8 |
| somites | 1-14 |
| diaphragm | 11 |
| intercostal muscles | 11 |
| forelimb | 10 |
| hind limb | 11 |
| mandibular muscles | 11 |
| lingual muscle | 11 |
| tongue | 11 |
| palate | 11 |
| vascular smooth muscles | 10 |

---

Data from: Gabbiani et al., 1981; Jackson et al., 1981; Franke et al., 1982; Cochard and Paulin, 1984; Fürst et al., 1989; Schart et al., 1989; Babai et al., 1990; Mayo et al., 1992; Li et al., 1993; and unpublished observations from Z. L. Li, A. Lilenbaum and D. Paulin.
Weintraub, 1993; Duprey and Lesens, 1994; Olson and Klein, 1994). Myocyte-specific MEF2 factors bind to an A+T rich site that is highly conserved in many muscle genes (Gossett et al., 1989). The myotube-specific enhancer region does not function in myoblasts, and deletion or mutation of the MyoD1 or the MEF2 site does not influence the myoblast-specific enhancer activity of the desmin gene. This correlates with the fact that some myogenic HLH proteins are present in myoblasts but unable to activate muscle-specific gene expression while MEF2 is not expressed in proliferating myoblasts.

As mentioned above, a major difference between the regulation of the desmin gene and the regulation of other contractile protein genes is that the desmin enhancer is active in myoblasts. This activity is due to the presence of a 150 bp myoblast-specific enhancer region. Footprinting experiments performed with nuclear extracts from myogenic cells revealed the presence of four protected regions within the 150 bp myoblast-specific fragment. These four regions contain several GC-rich sequences; three of them contain a Krox-like sequence which has 8 bp homology to the 9 bp consensus binding site for Krox factors (Christy and Nathans, 1989; Lemaire et al., 1990). Two Krox-like sequences, K1 (GTGTGTGGCGTG) and K3 (GCGGCGGTGGAG), tested in gel mobility shift assays, are capable of binding the same factors that bind to the Krox site (Fig. 1). Deletion or mutation of any of them reduced CAT activity in myoblasts by at least 50% but did not influence the enhancer activity in myotubes. The Mb sequence (GGGCCAGGAGCCAC) is also a GC-rich sequence (Fig. 1). Although it shares 13 bp out of the 14 bp MyoD1 binding site, the Mb sequence does not interact with factors of the MyoD family. The Mb sequence binds the same (or very closely related) factors as the K1 and K3 sequences. The Krox-family factors containing three zinc-finger motifs, are encoded by immediate-response genes which are induced by diverse signals such as growth factors and mitogens. It is interesting to note that one of the Krox members, Krox 24 (also known as egr-1, Zif268, TIS8 or NGF1-A), is present in the C2 myoblasts. Egr-1 has been shown to be present in developing muscles in vivo, and seems to participate in the regulation of the rat cardiac α-myosin heavy chain gene in cardiac muscle cells (McMahon et al., 1990; Gupta et al., 1991).
Desmin transgenic mice provide evidence for distinct cardiac, smooth and skeletal muscle gene programs

Even though cardiac, skeletal and smooth muscles express desmin, it is of considerable interest to define whether desmin expression would be regulated with identical or different cis-elements in these three tissues. The use of in vitro models to address this question however, has inherent limitations. On the other hand, the transgenic methodology permits the analysis of both temporal and spatial regulation of the transgene (Fig. 4). This was achieved by generating transgenic mice bearing transgenes in which the previously characterized 2.5 or 1 kb 5’ regulatory sequences [Li and Paulin, 1991] of the desmin gene were linked to a reporter lac-Z gene. Both transgenes were expressed in the committed mononucleate myogenic cells as well as in myotubes (Fig. 4) (Li et al., 1993 and unpublished results). In addition, transgenes were expressed in the mononucleate cells leaving the somites and invading the limb buds, indicating that the cells migrating from the somites are predetermined for myogenesis (Fig. 4).

No expression of the transgenes was observed in cardiac or in smooth muscles. Therefore, these results demonstrate that the upstream regulatory sequences are competent to promote specific expression of desmin in skeletal muscle, and suggest that other sequences are necessary for cardiac and smooth muscle expression.

The activity of the transgenes differs in the adult mice. Whereas the 2.5 kb regulatory sequences are active in fetal and adult skeletal muscles, the activity of the 1 kb proximal regulatory sequences decreases after birth and stops at eleven days post natal. At eleven days the activity of the muscle genes is regulated by the thyroid hormone. It seems that the region -1.3 kb to -2.5 kb includes a hormone responsive element (Li et al., in preparation).

Further experiments have shown that transgene including 2.5 kb of upstream sequences and intragenic sequences spanning the first four exons fused in frame to the lac-Z gene is expressed not only in skeletal muscle but also in heart muscle. Thus it seems that some cardiac-specific element(s) within the intragenic sequences are involved in the cardiac-specific expression of the desmin gene (Li et al., in preparation).

Vimentin: a potentially ubiquitous IF

Vimentin is the first IF III to be expressed during development

Vimentin is a 58 KDa (464 aminoacids) cytoskeletal protein, whose expression pattern is unique among type III IFs.

In contrast to that of the other IFs, vimentin expression is not related to a single lineage. Indeed, its expression can be detected in tissues originating from both the ectodermal and the mesodermal layers. A prominent vimentin expression is characteristic of precursor neural and mesenchymal cells from the mouse embryo. In these cells, vimentin expression is complex, due to rapid dynamic variations (Table 2).

Appearance of vimentin filaments is first detected at day 8.5 P.C. in mesoderm cells between the primitive streak and the proximal endoderm (Table 2) (Jackson et al., 1981; Franke et al., 1982). At day 8.5 p.c., vimentin expression is detected in some cephalic mesenchymal cells, in the notochord, the epimyocardium and blood vessels; somatopleura and splanchnopleura are also positive (Table 2) (Paulin, unpublished).

Expression in the neural tube first occurs at the cephalic level at day 9 p.c. At day 10.5 p.c., dorsal root ganglia begin to express vimentin (Table 2). Vimentin can be detected in sympathetic ganglia at day 14 p.c. (Table 2). In neuronal precursor cells, vimentin is progressively replaced by neurofilaments (Cochard and Paulin, 1984). At this stage, all blood vessels contain large amounts of vimentin, whereas no expression can be detected in the somites nor in the sclerotomes. Cartilaginous structures originating from the sclerotomes also lack vimentin, whereas dorsal mesoderm is positive (Table 2). The mesodermal layer from internal organs also expresses vimentin. In limb buds only blood vessels display vimentin IFs (Table 2). Later in the developing nervous system vimentin expression is observed in numerous glial cells. In most astroglial cells, vimentin is replaced by GFAP. However, in a subset of adult astroglial cells, both IFs can be found. Thus, it appears that vimentin is initially widely expressed in the embryo and becomes progressively restricted to fewer cell types. Although vimentin expression cannot be detected in most endodermal derivatives and in many cells derived from the ectoderm and the mesoderm in vivo, most vimentin-negative cells re-express this IF proteins upon in vitro culture.

Multiple enhancer and silencer elements control the vimentin gene

The vimentin gene has been cloned and extensively characterized in several vertebrate organisms such as human, mouse, hamster and chicken (Ferrari et al., 1986; Sax et al., 1988; Colucci et al., 1994). In human the vimentin gene has been allocated to chromosome 10p12 (Ferrari et al., 1987; Mattei et al., 1989; Baumgartner et al., 1991) and in the mouse to chromosome 2A.

Vimentin gene belongs to the family of immediate-early genes activated rapidly when cells are stimulated from quiescence to mitosis (Siebert and Fukuda, 1985; Ferrari et al., 1986; Sax et al., 1988; Vicart et al., 1994). Enhancer binding sites as well as negative elements have been characterized in human, hamster and chicken genes (Fig. 1) (Ritting et al., 1989; Sax et al., 1989; Zehner, 1991; Stover and Zehner, 1992; Salvetti et al., 1993). Two enhancers corresponding to AP1/c-jun and to NF-kB are involved in the control of the expression of the vimentin gene, and could explain one of the molecular mechanisms through which the vimentin gene is a mitogen-inducible gene (Fig. 1) (Ritting et al., 1989; Lilienbaum et al., 1990, 1993). The proximal NF-kB enhancer resides upstream of the promoter (Fig. 1). Further upstream silencer and “desilencer” elements may be important in limiting the expression of the gene. A (distal) enhancer element, which consists of a tandem of AP-1/Jun binding sites (TGAGTCA and TGACTAA), is located 5’ to the negative regulatory region. (Fig. 1). These two sites are involved in serum and TPA inducibility. The AP-1 sequences interact mostly with heterodimers made by the association of one member of the jun family with the fos factor, or some related gene products, through the leucine zipper dimerization motif.

Cellular factors, related to the NF-kB family are required for the induction of the vimentin gene by the tumor promoter PMA and by the human T-cell leukemia virus type I tax gene product. The NF-kB binding site is 5’ GGGGCTTTCC3’ in the antisense strand. Deletion of the NF-kB binding site from the human vimentin promoter (reduced to 210 nucleotides upstream the transcription
The possibility that homeodomain proteins play a role in the Hox family
experiments have shown that the Hox-A5 protein, a member of the Hox family of developmental control homeodomain proteins is able to bind specifically to these motifs. Ectopic expression of Hox gene product into embryonal carcinoma cells, which do not express the Hox gene products, led to a significant increase of the activity of a vimentin promoter deletion mutant containing the Hox binding sites. Such an increase was not observed in cells expressing the endogenous Hox genes. This observation raises the possibility that homeodomain proteins play a role in the regulation of the developmental expression of the vimentin gene.

Vimentin transgenesis in mouse recapitulates the complex and dynamic expression pattern

Constructs derived from the human and mouse vimentin genes have been used to generate transgenic mice. A first study used homologous recombination to insert the lac-Z gene inside of the endogenous gene in frame with the murine vimentin N-terminal coding sequences (Colucci-Guyon et al., 1994). Analysis of transgenic embryos showed an expression pattern strikingly similar to that of the endogenous vimentin gene (Fig. 4).

However, transgenes including up to 4 kb of vimentin upstream sequences and 8 kb coding region, previously used for homologous recombination, are not able to drive expression in vessels when integrated outside of the normal site. Thus it seems that these vimentin upstream and intragenic regulatory sequences do not carry all the necessary information required for the spatio-temporal expression in the embryo.

The lac-Z gene was also placed under the control of 1710 bp of human vimentin upstream regulatory sequences. Like the endogenous vimentin gene, the transgene was expressed in ganglia and superficial connective tissues (Gabbiani et al., 1981). Differences were found in vessels and neuroectoderm where the transgene was not expressed in contrast to the endogenous vimentin gene (Paulin, unpublished results; Lilienbaum et al., in preparation).

Strikingly, transgene containing 830 bp of upstream sequences displayed a very limited expression in the developing embryo (Lilienbaum et al., in preparation). Thus, in contrast to the situation observed in vitro, sequences located upstream of the 830 bp deletion mutant play a major role in the control of the developmental expression of the vimentin gene. It is not yet known whether the Hox-binding sites located in the upstream regulatory region of the vimentin gene, play a role in the expression of the 1710 bp construct in transgenic embryos. Breeding vimentin-lac-Z transgenic mice with mice carrying inactivated alleles of homeobox containing genes obtained through homologous recombination, will be required to answer this question.

Altogether, these data suggest the existence of multiple elements required for the proper expression of the vimentin gene. The elements responsible for the expression in vessels could be located either far upstream, in the 3' portion or even downstream from the gene. Thus, in spite of its complexity, the 5' upstream vimentin region is not able to reproduce the endogenous vimentin expression pattern in mouse embryos. To address the biological role of vimentin in the context of the living organism, a vimentin null mutation was introduced into the germ line of mice (Colucci-Guyon et al., 1994). Surprisingly, animals homozygous for this mutation developed and reproduced without any obvious phenotype. Immunoblotting and immunofluorescence analysis confirmed the absence of vimentin and of the corresponding filament network. Furthermore, no compensatory expression of another intermediate filament could be demonstrated. It was observed that in a subset of astrocytic cells, which do normally coexpress GFAP and vimentin, no GFAP filaments could be observed (Galou et al., in preparation). While these results leave open the question of the possible role of vimentin in unusual or pathological situations, they show that a conspicuous developmental and cell-specific structure, which is an integral part of the cytoskeleton, can be eliminated without apparent effect on mouse reproduction and development.

Discussion

In this review, we have focused on the regulation of type III intermediate filament genes, which constitute a subfamily of highly related genes. A very high degree of sequence homology suggests that these four genes have arisen recently through successive duplications of an ancestral gene. Therefore, we would like to suggest that the acquisition of different cell-specific control elements occurred relatively recently. This reinforces the interest of using type III IF genes as tools to analyze cell-specific gene regulation.

5' upstream sequences play a major role in the regulation of type III IF genes transcription

It was observed that the 5' upstream sequences of type III IF genes contain key regulatory elements (Fig. 1). While this observation does not exclude the occurrence of other control elements located downstream of the transcriptional initiation sites, as was found in the case of GFAP, desmin and peripherin, it does contrast with the situation observed for other classes of intermediate filaments (see below).

Analysis of the cis and trans-acting factors involved in the control of type III IF gene regulation is likely to provide some key information on the mechanisms involved in the control of cell-specific transcription. This assumption is supported by the data obtained in the case of desmin. Indeed a lot of information has been accumulated on skeletal muscle-specific transcription, which collectively point out to the key involvement of transcription factors belonging to the MyoD and MEF-2 families cooperating in the control of transcription in differentiated myotubes. The situation observed for the desmin myotube-specific enhancer was found to perfectly mirror the data obtained for the regulatory elements of other genes specifically expressed in differentiated myotubes.

In contrast, little is known about the mechanisms involved in the control of myoblast-specific gene expression. We would like to suggest that the mechanisms involved in the activity of the desmin myoblast-specific enhancer are likely to be similar to those involved in the control of the activity of other myoblast-specific regulatory elements. It is also likely that the mechanisms
controlling GFAP gene expression in astrocytes and peripherin gene expression in some neurons will also help decipher the control mechanisms involved in gene expression in these cell types, for which little information is currently available.

In the case of desmin, positive elements located in the proximal promoter and in the enhancer region appear to be responsible for cell-specific transcription, while a negative region flanked by these two positive control areas reduces transcriptional activity in an apparently ubiquitous manner.

In the case of GFAP and peripherin, ubiquitously active positive elements have been described. Cell-specific transcription of these two genes was found to be controlled by negative elements counteracting the activity of the positive elements in non-expressing cell types (Fig. 1).

Intragenic sequences play a major role in the regulation of other IF genes

Expression of the genes encoding for cytokeratins 8 and 18 was found to be controlled by enhancer elements located respectively 3' to the gene (CK8) and in the first intron (CK18) (Oshima et al., 1990; Takemoto et al., 1991). In contrast, the 5' upstream regions exhibited little or no activity. Additional mechanisms such as methylation and insulation have also been shown to be involved in the control of these genes (Oshima et al., 1990; Takemoto et al., 1991; Thorey et al., 1993). In the study by Thorey et al. (1993), insulation was found to be partially dependent on the transcription of a 5'-located alu sequence.

A prominent role of intragenic elements was also shown for the type IV nestin gene. Nestin expression in neuroepithelial precursors appears to be controlled by a neural-specific enhancer located in the second intron, while a muscle-specific enhancer located in the first intron is responsible for the expression in somitic myogenic precursors (Zimmerman et al., 1994).

Expression of the genes encoding the medium (140 kDa) and heavy (200 kDa) subunits of neurofilaments is controlled by intragenic cell-specific regulatory elements (Lee et al., 1992; Charron et al., 1993) while an intermediate situation is observed for the light (70 kDa) subunit. Indeed apparently redundant cell-specific regulatory elements in the NF-70 gene appear to be localized both in the 5' upstream region (between nucleotides -1.8 kb and -0.3 kb) and inside the gene (Julien et al., 1987; Beaudet et al., 1992; Ivanov and Brown 1992). It is worth noting that a similar situation was observed for the type III GFAP and peripherin genes, also expressed in neural tissue.

However, as noted previously, the 5' regions of the GFAP and peripherin genes play an important regulatory role.

The regulation of vimentin gene expression is unique among IF genes

Vimentin expression is not restricted to a single cell lineage. The control of its expression appears to require multiple regulatory elements some of which could be located far upstream, in the 3' portion or even downstream of the gene. The vimentin promoter, which is not able to confer the complete expression pattern to a reporter gene, is complex and its activity appears to be controlled by a combination of positive and negative elements. These are characterized by some degree of ubiquitous activity which can be further enhanced or lowered depending on the cell state. Therefore it seems that mechanisms controlling vimentin expression are somewhat intermediate between those controlling neural type III IF genes (GFAP and peripherin) and the mesodermal muscle-specific desmin gene. Interestingly, vimentin expression can be detected in precursor cells from both neural and mesodermal lineages. From an evolutionary point of view, vimentin appeared before the other type III IF genes which could be derived from a vimentin-like ancestor (Weber and Geisler, 1985; Weber et al., 1991). While vimentin expression is restricted to some cell types in the animal, it can be reinduced in most cell types upon in vitro culture and is thus potentially ubiquitous.

Why several type III IF genes?

The very high similarity between the various type III IF genes raises the question of the biological significance of the occurrence of four type III IF genes.

Embryonic development is characterized by a transient expression of vimentin in mesodermal and neural precursor cells from the early embryo. Subsequently, vimentin disappears from the cells which start to accumulate other type III IFs (desmin, GFAP, peripherin).

We would like to suggest that the various type III IF genes evolved from an ancestral vimentin gene, whose expression was ubiquitous. The increasing complexity of organisms called for subtle control mechanisms required to trigger the dynamic type III IF expression pattern. A single regulatory region might have become insufficient to ensure the control of such dynamic and complex IF gene transcription. The need for these dynamic variations might have led to gene duplication events therefore allowing type III IF gene expression in precursor cells (vimentin), disappearance of vimentin type III IF from these cells followed by accumulation of other type III IFs (desmin, GFAP, peripherin) in differentiating cells. Such hypothesis leads to the suggestion that type III IF function might become more apparent upon variation of the amount of these proteins in the cell rather than upon presence or absence of these proteins in the cell. Such hypothesis, which contrasts with observations made for cytokeratins (Oshima, 1992), can be tested by inducing a permanent expression of type III IF in the neural or mesodermal lineages. Such experimental strategy, together with targeted inactivation of the genes, might shed light on the still enigmatic function of type III IFs.

Conclusion

So far, gene disruption experiments, performed in the cases of vimentin and GFAP, have not allowed to assign a precise physiological role for the type III IF gene products whose biological function remains enigmatic. While the possibility remains open that the lack of obvious phenotypes could be ascribed to redundancy between very similar gene products, it is noteworthy that no evidence for compensation by the expression of another type III protein could be found in vimentin and GFAP null embryos.

The very high similarity between the different type III IF proteins further raises the question of the evolutionary significance of the gene duplications which led to the amplification of this small gene subfamily.

Whatever is the rationale for the occurrence of four type III IFs, if any, it is clear that analysis of the mechanisms which control their expression will further contribute to our knowledge on cell-specific transcription.
The combined use of both in vivo (transgenic mice) and in vitro (cell lines) experiments has already demonstrated that multiple cell-specific elements contribute to the establishment of the complete expression pattern of the different type III IF genes and the occurrence of both upstream and intragenic elements has been established.

Further work will be required to determine how the impact of these multiple cell-specific elements is integrated at the level of a single transcription unit. It is clear that the understanding of the dynamic interactions between multiple scattered regulatory elements will be a major challenge for the deciphering of the molecular basis of cell type specific type III IF gene expression.

Summary

In recent years, intermediate filaments (IFs) have attracted much interest, largely because their constitutive polypeptide units are specifically expressed in various cell types and thus represent excellent differentiation markers. Data obtained through biochemical studies and molecular cloning have allowed the classification of IFs into five types according to their protein structure. The expression of most IF types is characteristic of a classification of IFs into five types according to their protein structure. The expression of most IF types is characteristic of a classification of IFs into five types according to their protein structure. The expression of most IF types is characteristic of a classification of IFs into five types according to their protein structure. The expression of most IF types is characteristic of a classification of IFs into five types according to their protein structure. The expression of most IF types is characteristic of a classification of IFs into five types according to their protein structure. The expression of most IF types is characteristic of a classification of IFs into five types according to their protein structure.

On the other hand the four type III IFs are highly related proteins which are expressed in different cell types. Thus the study of type III IF gene regulation provides an excellent approach towards the analysis of cell-specific transcription. This review focuses on type III IF gene regulation during mouse embryogenesis and describes the latest data obtained through the combination of both in vitro (in cell lines) and in vivo (in transgenic mice) approaches. It appears that, while intragenic sequences play a major role in the regulation of the expression of the genes encoding other types of IFs, a major contribution to the transcriptional regulation of type III IF genes is brought by 5’ upstream sequences. However, recent evidence obtained through the use of transgenic mice indicate that upstream sequences must cooperate with intragenic elements to establish the complex and dynamic expression pattern characteristic of type III IF genes.

The very high similarity between the coding sequences of type III IF genes raises the question of the significance of the occurrence of four members of this class. We propose a model for the amplification of this small gene family based on the increasing complexity of expression patterns in higher organisms. This could have led first to the requirement for a highly sophisticated control region in an ancestral type III IF gene, followed by two successive gene duplications, thus leading to the appearance of four different regulatory regions directing the cell-specific transcription of nearly identical genes in different cell types.

KEY WORDS: intermediate filaments, transcriptional control, mouse embryo, gene duplication.

Acknowledgments

This paper is dedicated to the memory of Hubert Condamine. We wish to thank him again for many enlightening discussions. We are indebted to Z.L. Li and A. Benazzouz for sharing unpublished observations on desmin and vimentin control sequences. We would like to thank C. Babinet, E. Colluci, K. Djabali, P. Dupouey, M. Galou, Z.L. Li, A. Lilienbaum, L. Lossouarn, J.J. Panther, M.M. Porter, and J.L. Simonneau for providing the illustrations for Figures 2 to 4 and for sharing unpublished data. We are grateful to Dr. P. Coulombe and M.H. Kryszke for critical review of the manuscript. We are indebted to P. Ave and M. Huere for histology examinations. Part of the work was granted from Ministère de la Recherche, Association Française contre les Myopathies, Institut National de la Santé et de la Recherche Médicale (Contrat 920810) and Association pour la Recherche sur le Cancer (Contract 6518).

References


Intermediate filament gene expression in the mouse embryo


Intermediate filament gene expression in the mouse embryo


Acfset for publication: February 1995