Reprod Nutr Dev (1990) 30, 423-429

Original article

Regulatory elements of the human vimentin gene: activation during proliferation

D Paulin [1,2*], A Lilienbaum [1], P Duprey [2], Z Li [1], P Vicart [1,2]

[1] *Institut Pasteur, 25, rue du Docteur Roux, 75015 Paris;*
[2] *Université Paris 7, Laboratoire de Biologie Moléculaire de la Différenciation, Paris, France*

(15th meeting of the INRA development group, Paris, 24–26 May 1989)

Summary — We investigated the constitution of the vimentin regulatory region through the use of cloned deletion mutants and the nucleotide sequence analysis in order to determine the elements which are implicated in the various physiological stimulations. We report that the vimentin promoter is constituted of a juxtaposition of at least 20 different putative regulatory elements illustrating the molecular tinkering theory. Fifty-eight motifs were found, representing 20 different sequences. Each of these mini-elements displays a consensus sequence homologous to or closely related to that found in regulatory regions of different genes correlated with processes of cell activation and proliferation.

activation / binding motif / development / homeotic gene / intermediate filament

Résumé — Le gène vimentine : bricolage moléculaire et activation durant la prolifération cellulaire. *Le gène vimentine est activé en réponse à de multiples facteurs liés aux conditions de croissance : sérum, facteurs de croissance, ester de phorbol, infections virales. L'expression peut se faire* in vitro *dans tous les types cellulaires perdant ainsi la régulation spécifique observée* in vivo. *Nos résultats montrent que cette réponse à la multiplicité d'inducteurs est due à la constitution de la région régulatrice qui comprend une succession de mini-promoteurs empruntés aux gènes viraux et cellulaires.*

activation / motif de liaison / développement / gène homéotique / filament intermédiaire

INTRODUCTION

The vimentin gene is a member of the intermediate filament family which is developmentally regulated (Cochard and Paulin, 1984) and tissue–specific (Osborn and Weber, 1982). Of all the IF genes, only the vimentin gene deviates from the established pattern of tissue expression and can be expressed *in vitro* in all cell types (Franke *et al*, 1979). Its inducibility with a variety of agents ranging from mitogens, hormones, virus or specific differentiation agents, suggests complex and versatile regulatory mechanisms. Because genes coding for different proteins becoming part of the same developmental program are expressed in a given time at a certain stage and must therefore be co-ordinately regulated, we assume that they display homologous regulatory motifs able to bind the same transacting protein factors.

* Correspondence and reprints.

RESULTS AND DISCUSSION

Nucleotide sequences of 2 050 base pairs (1 710 bp upstream the initiation site) were compared to the regulatory motifs present in other genes responsive to the same inducers or activated concomitantly. A comparison of these sequences has allowed us to deduce putative *cis*-regulatory elements.

Our results show evidence that the vimentin regulatory region is constituted of modules of binding motifs running through the 1 700 bp length upstream the cap site. The stretches of 6–12 base pairs found in other different genes, are linked to each other in the vimentin gene to produce an integrative system: enhansons (Dynan, 1989). Fifty-eight motifs were found, representing 20 different sequences repeated 2–6 times.

Some individual sequences differ from others in a family as a result of some base changes. We presume that binding of factor to these receptor sequences will occur for a degree of perfect sequence homology. However, at some degree of divergence, binding would be lost and the genes would fail to be activated as part of its previous battery. The possibility of increasing sequence divergences suggests an evolutionary mechanism.

From -529 to -63 : multiple GC boxes

Some motifs, such as the TATA box, are shared by many promoters, whereas others, such as SP1 binding sites and the CCAAT box, are found in selected promoters. The SP1 binding sites occur in many other viral and cellular promoters: SV40, HSV, tk, AIDS retrovirus, mouse dihydrofolate reductase, human metallothionein IIA, IA, hamster hydroxmethyl-glutamyl coenzyme A reductase, mouse hypoxanthine phosphoribosyl transferase, mouse adenine phosphoribosyl transferase, human adenosine deaminase and rat type II procollagen. Some of the putative binding sites have been tested for their ability to bind SP1, but in most cases, binding is merely predicted from the DNA sequence.

The presence of multiple GC boxes located at positions -1 038, -714, -555, -496, -389, -275, -139, -63 is unique to the vimentin promoter region, since they have not been found in other IFP genes. However, a single GC box has been found in 3 IFP genes (keratins and desmin).

Three of the GC boxes contain an A at position 5 exactly like that in the vimentin promoter of the hamster. This sequence was also found in the AIDS retrovirus (Kadonaga *et al*, 1986). Comparison with the 5' flanking sequence of the chicken gene shows 5 GC boxes exhibiting at least 9 out of 10 nucleotides from the consensus sequences. Results of Sax *et al* (1988), indicate that SP1 enriched protein fractions are capable of specifically interacting with at least the 3 most proximal GC boxes found in chicken vimentin promoter.

Functional domains of SV40 were defined in 3 modules A, B, C. The combinational and modular nature of the SV40 early control region is typical of many transcriptional control elements found in the vimentin gene.

From -241 to +1 : binding motifs conferring serum inducibility and trans-activation by the HTLV1 virus

It was shown that the HTLV1 tax (p40x) protein is able to transactivate the *IL2R* gene, the *IL2* gene and the GM-CSF gene, all implicated in the mitogenic response (Gazzolo and Duc-Dodon, 1987). Vimentin

Table I. Sequence of putative DNA binding motifs found in the 5' regulatory region of the vimentin gene.

POSITION	TYPE		SEQUENCE	CONSENSUS
-1618	Octamer element		ATTTaCAT	ATTTGCAT
-1562	Hox1.3	*	ACTAATAttG	CPyPyNATTAT/GPy
-1528	Octamer element		ATTTcCcT	ATTTGCAT
-1527	HSP70 element	*	TTTCCC	GGGAAA
-1443	HSP70 element	*	TaTCCC	GGGAAA
-1429	c-fos basal element	*	AAcGTCA	TGACGTA/T
-1387	AP5	*	CATTtCcCAG	CTGTGGAATG
-1385	HSP70 element	*	TTTCCC	GGGAAA
-1355	AP5		CTGTGcAtTG	CTGTGGAATG
-1288	Hox1.3	*	cATAATTGcc	CPyPyNATTAT/GPy
-1268	Hox1.3		aCTCATTAaT	CPyPyNATTAT/GPy
-1264	Hox1.3	*	AtTAATCAAa	CPyPyNATTAT/GPy
-1235	Hox1.3		CCTAATTAaT	CPyPyNATTAT/GPy
-1197	ATF	*	TGACtTC	T/GT/ACGTCA
-1144	ATF	*	gGACGTC	T/GT/ACGTCA
-1143	ATF		GACGTCG	T/GT/ACGTCA
-1132	Lymphoid factor element		TCCTCTTT	TCCTCTTT
-1087	GTIA		GGGTGTGG	GGGTGTGG
-1047	c-fos SCM element		GAggGACGGGggC	GATTGACGGGAAC
-1038	SP1		GGGCGG	GGGCGG
-965	Sph		AAGgcTGCA	AAGT/CATGCA
-937	Sph		AAGTccGCA	AAGT/CATGCA
-929	PEB2		ACATTCgA	ACATTCCA
-859	PEA2		GACCGCA	GACCGCA
-834	GTIIB	*	cCAGCTG	CAGCTGT
-833	GTIIB		CAGCTGT	CAGCTGT
-756	c-fos-SCM element		tTTCCCcTaAAcC	GTTCCCGTCAATC
-756	HSP70	*	TTTCCC	GGGAAA
-739	c-fos-SCM element	*	GgagCCcTCAATC	GTTCCCGTCAATC
-714	SP1		GGGCGc	GGGCGG
-707	AP1		TGAGTCA	TT/GAGTCA
-692	ATF	*	TGACtAA	T/GT/ACGTCA
-693	AP1	*	TGACTAA	TT/GAGTCA
-650	PEB1		CACcGCCgTCT	CACTGCCCTCT
-555	SP1		GGGCgG	GGGCGG
-496	SP1	*	CCaCCC	GGGCGG
-450	ATF		GACtTCA	T/GT/ACGTCA
-389	SP1	*	CCCGCC	GGGCGG
-382	PEA2		GACCGCA	GACCGCA
-375	AP2		GcCCCgAGGC	GTCCCCAGGC
-338	NF1		TGGCgtggGCCA	TGGC(N5)GCCA
-275	SP1	*	CCCGCC	GGGCGG
-227	NF-KB	*	GGAAAGCCCCC	A/GGGGA/GA/CTT/CT/CCC
-197	ATF	*	TGAaGTA	T/GT/ACGTCA
-197	c-fos basal element		TGAaGTA	TGACGTA/T
-197	CLE2		TgAaGTA	TCAGGTA
-195	c-fos-SCM element		aAgTaACGGGAcC	GATTGACGGGAAC
-177	16 mer element		GTAAcGGGACcaTGcC	GTAAGGGGACTGTGTC
-160	CLE2		gCAGGaA	TCAGGTA
-158	PEA3		AGGAAG	AGGAAG
-139	SP1	*	CCaCCC	GGGCGG
-98	c-fos-SCM element		GATTGgCtGGggC	GATTGACGGGAAC
-75	PEB1		gGAtGGCAGTG	AGAGGGCAGTG
-63	SP1	*	GGGaGG	GGGCGG
-55	18 mer elemen t		GGATGGCAGTGGGAGGGG	GGATGGCAGTGGGAGGG
+19	AP2		GTCCCCgcGC	GTCCCCAGGC
+55	GTIA	*	CCACACCC	GGGTGTGG
+192	AP2		GCCcGGGcAC	GCCTGGGGAC
+230	ATF		TACGTgA	T/GT/ACGTCA
+234	ATF		TGACtAC	T/GT/ACGTCA
+238	ATF		TACGTCc	T/GT/ACGTCA
+307	AP2		GTCCCCgGGC	GTCCCCAGGC
+348	AP2		GCCTGcGGAg	GCCTGGGGAC

is also induced by the HTLV1 tax protein and the target is located between nucleotides -241 and -78. Two sequences located at -160 (gCAGGaA) and -197 bp (TgAaGTA) within the upstream region in the vimentin promoter are similar to a sequence (TCAGGTA) located in the GM-CSF gene promoter -88 to -94. This element is also found in the IL2 promoter at position -446 and -944 from the cap site and in the IL2 receptor gene at positions -127 and +441 from the TATA box. The NF-KB consensus site has also been implicated in the trans-activation of the *IL2-R* gene by the HTLVI tax protein and is present in the vimentin promoter. Therefore, it seems likely that trans-activation by tax gene products could be mediated by these common elements (Lilienbaum *et al*, 1990).

From -957 to -530 : Binding motifs conferring both serum and TPA inducibility

The upstream 5' region of the vimentin gene reveals the presence of 2 sequences of 7 nucleotides TT/GAGTCA identical to the consensus sequences found in other TPA inducing genes which are known to bind the AP1/*jun* protein (Bohmann *et al*, 1987). These elements are located at position -707 and -693 (Rittling and Baserga, 1987).

The promoter regions of several phorbol diester (TPA inducible genes collagenase stromelysin, hmT IIA, and SV40 share the conserved 7 bp motif. The consensus sequence was also found in the control region of IL2 (Fujita *et al*, 1986) and polyoma DNA (Zenke *et al*, 1986) and *c-fos* (Piette and Yaniv, 1987). The P motif which binds the transcription factor AP1 mediates transcriptional induction in response to phorbol ester tumor promoters and cAMP elevating agents.

There was approximately a 10-fold induction of vimentin biosynthesis following TPA treatment of human erythroleukemic cell line K562 (Siebert and Fukuda, 1985). This was found to be correlated with a rapid induction of vimentin mRNA as supported by Northern RNA gel hybridization.

Similarly, in response to TPA treatment, the human myeloid leukemic cell line HL–60 showed an increase at the level of vimentin mRNA abundance correlated with the macrophage monocyte-like shift.

Cyclic AMP, the major second messenger in eukaryotic cells, mediates its effects by activating cAMP dependent protein kinases that phosphorylate key cytoplasmic and nuclear molecules. The consensus sequence for cAMP response binds ATF factor is present in vimentin promoter 6 times and in *c-fos* and several adenovirus early promoters. It is interesting to note that the ATF binding sites also mediate trans-activation by the viral adenovirus E1a gene and the proto-oncogene *c-myc*.

From -1710 to -950: Five binding motifs for the product of the Hox 1-3 homeogene

The developmentally regulated mammalian embryo involved homeo gene products, located in the nuclei. These proteins could well function as regulators of transcription. The vimentin promoter displays 5 motifs whose sequence match with the consensus binding site of Hox 1.3 product, described by Odenwald *et al* (1989). The binding motif is shown at -1 258 and in reverse orientation at -1 549, -1 268, -1 255 and -1 226. Such sites could also be found in the Hox1.3 and SV40 promoters. During the early development of *Drosophila*, the transcripts of various homeogenes accumulate in a specific region of the embryo controlling its spatial organization.

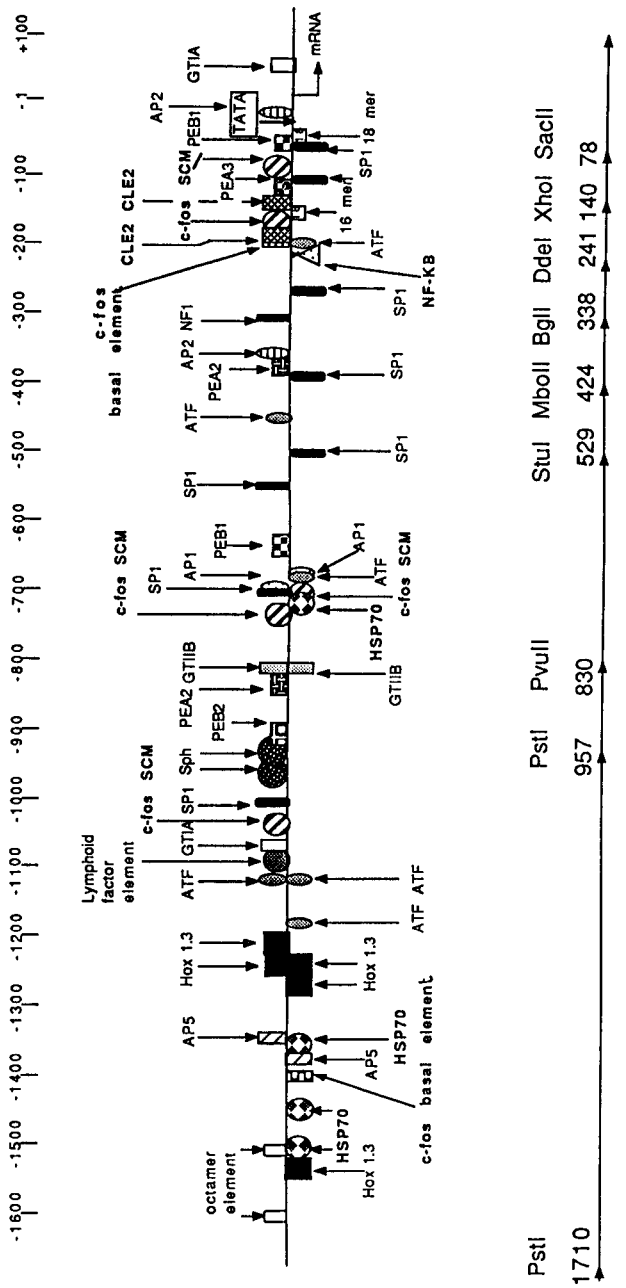

Fig 1. Schematic representation of the position of the putative DNA binding motifs. The thick line represents 1 700 nucleotides of the 5' regulatory region in front of the RNA starting site. Each DNA binding motif is deduced from the nucleotidic sequence which is presented in table I. Abbreviations are the same as in table I. Note the homologies between various species including common binding motifs: Human from +90 to −193; hamster from +88 to 188; chicken from −380 to −712. The position of the restriction sites used for subcloning were noted. Activity of the vimentin promoter – CAT fusion plasmids was measured after transfection into Hela cells; respectively 47%, 96%, 52% for the 1 710, 830 and 241 bp deletion mutants.

Study of the transcription of the murine homeobox containing genes has also revealed that they belong to genes whose transcription patterns seem to be regulated during ontogenesis and cell differentiation as judged by their enhanced expression at specific times in foetal life in a restricted number of adult tissues or following cell differentiation *in vitro*.

Odenwald *et al* (1987) have shown that the *Hox 1.3* gene encodes for a 35 kDa protein which is expressed during mouse embryogenesis in the nuclei of some mesodermal and neural cells. The *Hox 1-3* gene is also expressed in exponentially growing non-confluent fibroblasts. The protein is a nuclear sequence specific DNA binding protein and the consensus for the maximal binding site was determined from the protected region of the homeobox promoter (Odenwald *et al*, 1989).

G0/G1 transition is thought to be a critical regulatory period involving activation or repression of regulatory genes whose expression is essential in the control of cell proliferation and differentiation. This category includes immediate early genes which do not require *de novo* protein synthesis, *ie* proto-oncogenes *c-myc, c-fos,* genes encoding zinc-finger proteins, etc and other early genes whose kinetics and induction following serum stimulation show a delay in the response. The presence of common *c-fos*-SCM elements in vimentin and *c-fos* promoter suggest the same primary mechanism which could be modulated by interaction with different combinations of other regulatory elements. Obviously, similar developmental programs using disperse coding sequences through the genome might be coordinated. Genomic tinkering works on what already exists, combining several systems to produce a more complex one (Jacob, 1977).

ACKNOWLEDGMENTS

The authors wish to thank H Condamine for helpful discussions. This work was supported by grants from the Association Française contre les Myopathies (AFM), the Ministère de l'Enseignement Supérieur and the Université Paris 7.

REFERENCES

Bohmann D, Bos TJ, Admon A, Nishimura T, Vogt PK, Tjian R (1987) Human proto-oncogene *c-jun* encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science* 238, 1386-1392

Cochard P, Paulin D (1984) Initial expression of neurofilaments and vimentin in the central and peripheral nervous system of the mouse embryo *in vivo*. *J Neurosci* 4, 2080-2094

Dynan WS (1989) Modularity in promoters and enhancers. *Cell* 58, 1-4

Fujita T, Shibuya H, Ohashi T, Yamanishi K, Taniguchi T (1986) Regulation of human interleukin-2 gene: functional DNA sequences in the 5' flanking region for the gene expression in activated T lymphocytes. *Cell* 46, 401-407

Franke WW, Schmid E, Winter S, Osborn M, Weber K (1979) Widespread occurrence of intermediate-sized filaments of the vimentin type in cultured cells from diverse vertebrates. *Exp Cell Res* 123, 25-46

Gazzolo L, Duc-Dodon M (1987) Direct activation of resting T lymphocytes by human T-lymphotropic virus type I. *Nature* 326, 714-717

Jacob F (1977) Evolution and tinkering. *Science* 196, 1161-1166

Kadonaga JT, Jones KA, Tijian R (1986) Promoter–specific activation of RNA polymerase II transcription by Sp1. *Trends Biochem Sci* 11, 20-23

Lilienbaum A, Duc-Dodon M, Alexandre C, Gazzolo L, Paulin D (1990) Effect of human T cell leukemia virus, type I Tax protein on the activation of the human vimentin gene. *J Virol* 64, 256-263

Odenwald WF, Taylor CF, Palmer-Hill FJ, Friedrich V, Tani M, Lazzarini RA (1987) Expression of a homeo domain protein in noncontact-inhibited cultured cells and postmitotic neurons. *Genes Dev* 1, 482-496

Odenwald WF, Garbern J, Arnheiter H, Tournier-Lasserve E, Lazzarini RA (1989) The Hox 1.3 Homeobox Protein is a sequence specific DNA–binding phosphoprotein. *Development* 3, 157-163

Osborn M, Weber M (1982) Intermediate filaments: cell-type specific markers in differentiation and pathology. *Cell* 31, 303-306

Piette J, Yaniv M (1987) Two different factors bind to the domain of the polyoma virus enhancer, one of which also interacts with the SV40 and c-fos enhancers. *EMBO J* 6, 1331-1336

Rittling SR, Baserga R (1987) Functional analysis and growth factor regulation of the human vimentin promoter. *Mol Cell Biol* 7, 3908-3915

Sax CM, Farell FX, Tobian JA, Zehner ZE (1988) Multiple elements are required for expression of an intermediate filament gene. *Nucl Acids Res* 16, 8057-8076

Siebert D, Fukuda M (1985) Induction of cytoskeletal vimentin and actin gene expression by a tumor–promoting phorbol ester in the human leukemic cell line K562. *J Biol Chem* 260, 3868-3874

Zenke M, Grundström T, Matthes H, Wintzerith M, Schatz C, Wildeman A, Chambon P (1986) Multiple sequence motifs are involved in SV40 enhancer function. *EMBO J* 5, 387-397