

**(12) PATENT**  
**(19) AUSTRALIAN PATENT OFFICE**

**(11) Application No. AU 199648356 B2**  
**(10) Patent No. 710697**

(54) Title  
**Amplifier sequences derived from the desmin gene, vectors comprising same and uses thereof for producing proteins**

(51)<sup>6</sup> International Patent Classification(s)  
**C12N 015/85 C12N 015/11**  
**A61K 031/70 C12N 015/12**  
**C12N 001/21 C12N 015/67**

(21) Application No: **199648356** (22) Application Date: **1996.02.16**

(87) WIPO No: **WO96/26284**

(30) Priority Data

(31) Number	(32) Date	(33) Country
<b>95 01937</b>	<b>1995.02.20</b>	<b>FR</b>

(43) Publication Date : **1996.09.11**

(43) Publication Journal Date : **1996.10.31**

(44) Accepted Journal Date : **1999.09.30**

(71) Applicant(s)  
**Institut Pasteur; Universite Paris 7**

(72) Inventor(s)  
**Denise Paulin; Zhen Lin Li**

(74) Agent/Attorney  
**WATERMARK PATENT and TRADEMARK ATTORNEYS, Locked Bag 5, HAWTHORN VIC 3122**

OPI DATE 11/09/96 APPLN. ID 48356/96  
 AOJP DATE 31/10/96 PCT NUMBER PCT/FR96/00261



DEM

AU9648356

(PCT)

(51) Classification internationale des brevets <sup>o</sup> :  
 C12N 15/85, 15/67, A61K 31/70, C12N  
 1/21 // 15/11, 15/12, (C12N 1/21, C12R  
 1:19) A1 (11) Numéro de publication internationale: WO 96/26284  
 (43) Date de publication internationale: 29 août 1996 (29.08.96)

(21) Numéro de la demande internationale: PCT/FR96/00261  
 (22) Date de dépôt international: 16 février 1996 (16.02.96)

(81) Etats désignés: AU, CA, JP, MX, US, brevet européen (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

(30) Données relatives à la priorité:  
 95/01937 20 février 1995 (20.02.95) FR

Publiée  
 Avec rapport de recherche internationale.  
 Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si de telles modifications sont reçues.

(71) Déposants (pour tous les Etats désignés sauf US): INSTITUT PASTEUR [FR/FR]; 28, rue du Docteur-Roux, F-75724 Paris Cédex 15 (FR). UNIVERSITE PARIS 7 [FR/FR]; 2, place Jussieu, F-75251 Paris Cédex 05 (FR).

(72) Inventeurs; et  
 (75) Inventeurs/Déposants (US seulement): PAULIN, Denise [FR/FR]; 7, rue de Montreuil, F-94300 Vincennes (FR). LI, Zhen, Lin [CN/FR]; 33, rue Mazarine, F-75005 Paris (FR).

(74) Mandataire: MICHELET, Alain; Cabinet Harle & Phelip, 21, rue de la Rochefoucauld, F-75009 Paris (FR).

(54) Title: AMPLIFIER SEQUENCES DERIVED FROM THE DESMIN GENE, VECTORS COMPRISING SAME AND USES THEREOF FOR PRODUCING PROTEINS

(54) Titre: SEQUENCES AMPLIFICATRICES DERIVEES DU GENE DE LA DESMINE, VECTEURS PORTANT CES SEQUENCES ET LEURS UTILISATIONS POUR LA PRODUCTION DE PROTEINES

	MEF2	MT	MyoD1	ACTIN	ACTIN
3	AAGCTTCTCTCTATAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
4	AAGCTTCTCTCTAGAGAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
5	AAGCTTCTCTCTAGAGAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
6	AAGCTTCTCTCTATAAATAGGAGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
7	AAGCTTCTCTCTATAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
8	AAGCTTCTCTCTATAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
9	AAGCTTCTCTCTATAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
10	AAGCTTCTCTCTATAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
11	AAGCTTCTCTCTATAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
12	AAGCTTCTCTCTATAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
13	AAGCTTCTCTCTAGAGAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT
14	AAGCTTCTCTCTAGAGAAATACCCGCTCTGGATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT	TGGTATTTGGGGT

(57) Abstract

An amplifier sequence at least 90 % homologous with SEQ ID No. 1: TCTATAAATA X<sub>1</sub>X<sub>2</sub>X<sub>3</sub>GC Y<sub>1</sub>Y<sub>2</sub>Y<sub>3</sub>GG TATTGGGGT TGGCAGCTGT T, wherein X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub>, which are the same or different, may be independently C or G, C or G and C or A respectively; and Y<sub>1</sub>, Y<sub>2</sub> and Y<sub>3</sub>, which are the same or different, may be independently T or C, C or G and T or C respectively, is disclosed. Said sequence may be included in an expression vector or sequence further containing a sequence coding for a protein and a promoter. Such sequences enable very strong amplification of the transcription of the gene coding for the protein.



Amplifying sequences, vectors comprising these sequences and their uses in compositions for the expression of nucleotide sequences in transfected cells, therapeutic and vaccine applications.

The present invention relates to amplifying sequences and vectors comprising these sequences, to the use of these vectors in compositions for the expression of nucleotide sequences in these transfected cells, and to the use of these products for therapeutic and vaccine applications.

The desmine gene, one of the first muscle proteins to be detected in the developing embryo of mammals, was sequenced in 1989 by Li et al. (Gene, 78, 243-254).

Subsequently, an amplifying sequence of 280 base pairs, situated between the nucleotides -693 and -973 upstream from the transcription initiation site, was described (Li and Paulin, 1991, J. Biol. Chem., 266, 6562-6570).

This amplifying sequence was identified by producing a series of plasmid constructions carrying the bacterial gene coding for chloramphenicol acetyl transferase (CAT), and by introducing them into myogenic or non-myogenic mice cells. The authors show that this amplifying region of 280 base pairs can activate both homologous and heterologous promoters regardless of their orientation, position or distance in relation to this sequence.

The study of this sequence was continued (Li et al., 1993, J. Biol. Chem., 268, 10 403-10.415) and showed that this sequence contained two regions : one active in myotubes and the other active in myoblasts.

The part that is active in myotubes lies between positions -973 and -848, that is to say in a sequence of 125 nucleotides.

Within this sequence, a region lying between positions -910 and -870 is protected from the action of a DNase. This region contains binding sites of MEF2 and MyoD1 factors. No specific activity is indicated for this region of 40 nucleotides. In particular, no experiment has been carried out with constructions containing this isolated region. Mention has only been made that MyoD1 and MEF2 sites are necessary to



obtain full amplifying activity in the myotubes.

The amplifying effect of the sequence upstream from the desmine gene has also been tested in a transgenic mouse (Li et al., 1993, Development, 117,947-959). A fragment of 1kb  
5 containing the regulatory sequence upstream from the desmine gene is bound to a reporting gene coding for Escherichia coli  $\beta$ -galactosidase. The authors show that the amplifying activity of the desmine promoter is very considerable as the activity of  $\beta$ -galactosidase can be easily detected in tissue sections.

10 It arises from the prior art analysed above, that the amplifying properties of certain parts upstream from the human desmine gene are known. On the other hand, the regions responsible for this amplifying activity have not been identified.

15 They could therefore not be successfully modified with a view to improving their performance for the expression of genes or nucleotide sequences corresponding in whole or in part to these c-DNAs (complementary DNAs) coding for proteins of interest or for fragments of these proteins conveying  
20 epitopes able to induce protective antibodies or antibodies which recognise said epitopes. The preferable length of c-DNA or genomic DNA fragments is between 30 base pairs and 2 kb. They can be produced by chemical synthesis or by restriction enzyme cutting of DNA extracted from cells or micro-organisms  
25 selected by methods known to men of the art.

It is known that by introducing DNA into these eucaryote cells using the described methods, in particular by patent WO-90 11 092, it is not possible to obtain an extended or a high quality expression of said DNAs introduced in  
30 plasmid form.

The present invention brings an improvement in the expression level of genetic material introduced into transfected cells by means of constructions which use all or part of the activating sequences identified for the desmine



gene and which can be used for the expression of any other gene or genetic material (genomic DNA fragments, c-DNA fragments etc..) in muscle tissue.

In particular, it is possible to use said sequences to  
 5 express nucleotide sequences having an immunomodulatory function, such as cytokines, or immunogenic properties, such as the VPI protein of the polio virus or HBs protein of the hepatitis-B virus.

Another application lies in the area of gene therapy.  
 10 In particular and in general, techniques which require the introduction of DNA into tissues necessitate the use of nucleotide sequences or vectors containing such, or liposomes strongly expressing the genes or cDNAs carried by these sequences. To date, this problem has not been solved in  
 15 satisfactory manner.

The applicant therefore sought to determine which parts of the desmine gene regulatory sequences were responsible for the amplifying activity.

In surprising manner the applicant identified the role  
 20 played by a sequence lying between the MEF2 and MyoD1 sites in the amplifying activity.

He also found that by coupling such amplifying sequence to a promoter, it is possible to obtain the expression of genes or cDNAs coding for peptides, polypeptides  
 25 or proteins of therapeutic interest through the direct introduction of the DNA.

The object of the present invention is an amplifying sequence showing at least 90% homology with the following SEQ ID N°1 sequence :

30 TCTATAAATA X<sub>1</sub>X<sub>2</sub>X<sub>3</sub>GC Y<sub>1</sub>Y<sub>1</sub>Y<sub>1</sub>GG TATTTGGGGT TGGCAGCTGT T

in which :

- X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may independently represent respectively C or G, C or G, and C or A.

- Y<sub>1</sub>, Y<sub>2</sub> and Y<sub>3</sub> may independently represent



respectively T or C, C or G, and T or C.

Such sequence can advantageously hybridize under stringent conditions with the sequence SEQ ID N°1.

Such conditions are for example the following :

5 Hybridization at 65°C overnight :

buffer :

0.2% polyvinyl pyrrolidone

0.2% ficoli 400

0.2% bovine serum albumin (BVS)

10 2XSSC

100 µg/ml salmon sperm DNA

100 µg/ml yeast RNA

Washings:

2xSSC 10 minutes at 65°C once

15 2xSSC 30 minutes at 55°C 4 times.

Another object of the present invention is an amplifying sequence showing at least 90% homology with the following SEQ ID N°2 sequence :

ZGGTATTT

20 in which Z may be T or C.

Such sequence advantageously hybridizes under stringent conditions with sequence SEQ ID N°2.

It will be noted that the present invention covers all amplifying sequences differing in one or several nucleotides  
25 in relation to sequences SEQ ID N°1 or SEQ ID N°2 whose functioning is not appreciably altered as well as their complementary sequences.

Under the present application, by amplifying sequence is meant a sequence which very strongly amplifies the  
30 transcription of one or more gene situated in cis.

The sequences of the invention include also sequence N°3 or sequence N°6, or sequence N°7 or sequence N°11, as defined with reference to the figure 1, either alone, in the form of polymers with two or eight identical fragments or in



association with one of the four sequences.

The term "vector" is used to designate a nucleic acid molecule into which it is possible to insert fragments of foreign nucleic acid, for subsequence introduction into and maintaining within a host cell.

By "liposome" is meant any lipid composition able to fix RNA or DNA in covalent or non-covalent manner.

The present invention also relates to a sequence or an expression vector including said sequence and at least:

- a sequence of type SEQ ID No 1 or SEQ ID No 2 such as described above, in one or more copies, with the exception of sequences in this X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub> and Y<sub>2</sub> represent a cytosine (C), and Y<sub>1</sub> and Y<sub>3</sub> represent a thymidine (T) or Z represents (T), in which case the sequences SEQ ID No 1 or SEQ ID No 2 are at least two copies thereof, and

- a sequence coding for a protein or corresponding to the genomic sequence coding for said protein or part thereof.

Advantageously, such sequence, or said vector is completed by one or more promoter(s) enabling the expression of said protein in cells transfected by said DNA.

The protein may be at least part of a bacterial antigen, for example H. pylori urease, so-called HSP (A) an abbreviation for "Heat-shock protein A", or the HSP (B) protein, or of a viral antigen, for example the surface antigen of the hepatitis virus, in particular hepatitis-B, and preferably the HBs protein in one of its S forms, S-preS2 or S-preS2-preS1. It may also be a protein of a hepatitis-A virus, or of a non-A, non-B hepatitis virus such as hepatitis C, E or delta.

The sequences of the genes or proteins of the bacteria or viruses of these hepatites are described or can be deduced from the following documents : patent FR-79 21 811 (published under N° 2 464 269); FR-80 09 039 (published under N° 2 480



779), EP-81 400 634 (published under N° 0 038 765), FR-84 03 564 (published under N° 2 560 890), EP-91 830 479 (published under N° 0 485 347) and FR-88 13 135 (publication N° 2 637 612), PCT application published under N° WC 94/26 901 and  
5 article by Najarian et al. (Proc. Natl. Acad. Sci. USA, 1985, 82, 2627-2631).

The protein may also be a protein or part of a protein of the HIV-1 virus or the HIV-2 virus or the HTLV-1 virus, in particular the ENV protein (corresponding to the viral  
10 envelope) or the gag protein. The sequence of the part of the protein or peptide is chosen in relation to the function it is wished to express. The minimal nucleotide sequences lie advantageously between 20 and 50 base pairs.

In respect of the HTLV-1 virus, the proteins or  
15 sequences coding for these proteins may be those described in the following documents, or may be deduced from these documents : applications PCT/WO 93/05 843, PCT/WO 90 15 820, EP-O 352 060 and EP-O 269 445, and articles by Gray et al. (1990, Virology, 177, 391-395), Tanaka et al. (1991,  
20 J.Immunol., 147,354-360) and Nakamura et al. (1987, Int.J. Cancer, 40,403-407).

Generally, said protein may be one or more proteins of therapeutic or vaccine interest, such as proteins of interest for immunotherapy, for example the Interleukins, the growth  
25 factors, for example fibroblast growth factors (FGF) or the nerve growth factor (NGF) or the maltase gene, proteins able to induce an immune response, whether humoral, cytotoxic or cell immunity, or proteins which allow complementary gene activity in a gene normally expressed in the individual to be  
30 treated but which is no longer expressed either through mutation or deletion of its sequence.

Said protein may be a hybrid protein, made up for example of at least one part of the desmine and at least one part of a protein whose deficiency it is sought to



counterbalance. In this case, the sequence or the vector may include an activating or amplifying sequence of the desmine gene, a sequence coding for at least one part of the desmine and a sequence coding for the protein whose deficiency it is sought to counterbalance.

It will be noted that the amplifying sequences may be positioned indifferently either upstream or downstream from the sequence coding for a protein, or a fragment thereof, or for a peptide lying between 10 and 50 aminoacids, and its promoter, and may lie in the two possible orientations.

In addition to a sequence coding for a protein or a fragment thereof, and an amplifying sequence, said sequence or vector may also include other regulatory sequences of Krox type, such as the Krox 24 sequence (LEMAIRE, P. et al. 1990, Mol. Cell. Biol. ,10, 3456-3467).

The promoter may be any promoter either alone or in association with another promoter allowing the expression of the protein as defined above. It may therefore be a promoter internal to the gene to be expressed or to the cDNA coding for the protein or fragment thereof to be expressed. It may also be a promoter that is homologous to the host, such as the promoter of a gene in a protein of the cytoskeleton, in particular that of desmine as described by BOLMONT et al. (Journal of submicroscopic cytology and pathology, 1990, 22, 117-122) or LI et al. (Gene, 1989, 78, 243-254).

It may also be a heterologous promoter, such as that of a virus, for example that of thymidine kinase in the herpes virus (HSV).

In addition to the promoter, positioned upstream from the sequence coding for the protein or any fragment thereof, the sequence may also include a terminal transcription sequence positioned downstream from the sequence coding for the protein.

Finally, such sequence, or such vector, may include



sequences allowing homologous recombination in the treated organism, specific to the gene to be replaced, said sequences being positioned upstream and downstream from the sequence coding for a protein, and possibly from the natural activating sequences and/or from the promoter or promoters. Owing to the presence of such sequences, the undesired gene existing in the treated organism will be replaced by the gene carried by the vector or the sequence, and which it is sought to have expressed in the organism.

Said homologous recombination method may be of the type described by Le Mouellic et al. (1990, Proc. Natl. Acad. Sci. USA, 87, 4712-4716) or in application PCT WO 91/05 667.

A vector such as defined above may also include a replicating origin that is active in a micro-organism, such as a bacterium.

Said vector may also include a gene allowing its selection in said micro-organism, such as an antibiotic resistant gene. It may be plasmid.

Advantageously, such micro-organism is *Escherichia coli*.

The expression vectors containing the sequences to be expressed and the transfectable sequences, that are the subject of the present invention, may be obtained by methods known to men of the art, in particular by chemical synthesis, such as the method marketed by Applied Biosystem, or using genetic engineering methods. Such methods are those described in particular by Maniatis T. et al. 1982 - Molecular Cloning, A Laboratory Manual, Cold Spring Harbor - Ed. New York, or one of its recent re-editions.

The invention also relates to two plasmids whose identification references are respectively EO3D CAT and DMT tk CAT (see Examples 1 and 2 below), which were filed with the CNCM collection of Institut Pasteur on December 16th 1994 under registration numbers I-1508 and I-1509 respectively.



A further object of the present invention is a pharmaceutical composition including a pharmacologically effective quantity of one of the expression vectors of one of the nucleotide sequences including at least one of the  
 5 activating sequences of the invention described above which can be expressed without a vector in the cells to be transfected.

A further object of the present invention is the use of a pharmacologically effective quantity of one of the  
 10 expression vectors or one of the nucleotide sequences including at least one of the activity sequences of the invention described above which can be expressed without a vector in the cells to be transfected in the manufacture of a pharmaceutical composition when used for gene therapy.

15 Such composition may also include pharmaceutically compatible excipients.

The present invention also relates to a medicinal product or vaccine containing these sequences and vectors.

Although this invention is not limited to a specific  
 20 treatment, the vectors and sequences are preferably intended for use with a vaccine purpose, and particular for injection into muscle tissues. For this method of use of the invention the chosen promoter is advantageously that of a gene or a cDNA coding for a muscle protein or fragment thereof.

25 The present invention is illustrated but not limited by the following examples.

The figure 1 illustrates the mutation effect in the amplifying sequence of the DMT tk CAT vector on the activity of the gene coding for the CAT it carrier.

30 The figure 2 is an autoradiogram of a gel of nuclear extracts of cellular lines C 2,7 (well C2) and CCL 136 (well Rd) incubated with labelled MT oligonucleotide. The position of the MT-protein complex is indicated by the number 1.

The figure 3 is an autoradiogram of a filter obtained by  
 35 Southwestern blotting of cellular extra acts of cellular lines C 2,7 and CCL 136 with labelled MT oligonucleotide. Molecular weight markers (Sigma) have been migrated in well MW.

**EXAMPLE 1 :**

**PREPARATION OF THE EXPRESSION VECTOR D MT k CAT:**

40 Using two primers:



(1) DesHindIII 5' -GAAAGCTTCTCCTCTATAAATACC-3'

or (2) DesXbaI 5' -CCTCTAGAGTCAACCCAACCTCT-3'

a fragment of 84 base pairs containing a fixation site (GGCAGCTGTT) for the MyoD1 transactivator factor, a fixation site (CTATAAATACC) for the MEF2 factor and a site (GGTATTT) called MT was amplified from the promoter of the human desmine gene. This 84 bp fragment was inserted into the HindIII-XbaI restriction sites of the pBLCAT2 plasmid (Lucknow and Schütz, 1987, Nucl. Acid. Res., 15, 5490) which contains the basal promoter of the thymidine kinase gene of the herpes simplex virus and the region encoding the bacterial gene chloramphenicol acetyltransferase (CAT).

Owing to the co-operation between the MT site and sites MyoDI and MEF2, this fragment of 84 bp increases the expression of the CAT indicator gene 60 times in the differentiated muscle cells, the myotubes.

EXAMPLE 2 :

Preparation of the expression vector EO3 D CAT :

A fragment of 280 base pairs (-973 to -693) containing the amplifier of the human desmine gene was placed before the 228 bp region (Li and Paulin 1991, 1993, previously quoted) which contains the basal promoter of the human desmine gene and the region coding for the bacterial gene chloramphenicol acetyltransferase (CAT).

A fragment of 280 base pairs (-973 to -693) containing the amplifier of the human desmine gene was tailed by the polymerase DNA (Klenow) and inserted into the Hind III restriction site which was also free-ended with the polymerase DNA (Klenow).

This amplifier not only operates in differentiated muscle cells, but also in mononucleated muscle cells, the myoblasts, This fragment activates the expression of the gene in the myotubes 100 times and in the myoblasts 10 times.

EXAMPLE 3 :



Mutation effects in the amplifying region of the desmine gene on the activity of a reporting gene in the myotubes.

5 The DMT tk CAT vector whose preparation was described in Example 1, was mutated using the guided mutagenesis method (Kunkel et al. 1985, PNAS, 82, p.488-492). These mutated vectors were transfected in culture myoblasts and the CAT activity was evidenced in the manner described by Li et al. (1991 and 1993 previously quoted).

10 The results of these experiments are summarized in the figure 1. Eleven constructions numbered from 4 to 14 showing various underlined mutations were tested for their effect on CAT activity. The non-muted amplifying sequence is the one represented at position 3.

15 It is very clearly apparent that the mutations of constructions 6 and 7 give rise to overamplification of CAT activity in relation to construction N'3.

20 The two expression vectors in Examples 1 and 2 may be used to construct vectors for the expression of the gene of therapeutic interest in muscle cells through replacing the CAT gene by the chosen gene. The EcoRI and BamH I sites may be used to remove the CAT gene.

25 The CAT gene may therefore be replaced for the immunisation or vaccination of mammals with bare DNA or genomic DNA corresponding to the genes coding for the envelope and gag protein of the HTLVI virus (Cary et al., 1990; Virology 177, 391-395; Tanaka et al., 1991, J. Immunology 147; 354-360; Nakamura et al. 1987; Int. J. Cancer 40: 403-407); the gene coding for all or part of the envelope of the hepatitis B virus or of human papillomae such as HPV16, HPV33 etc.; the  
30 gene coding for the FIV envelope, or for gene therapy, the genes coding for the interleukins and growth factors (FGF, NGF, MG-CSF etc...) and for gene therapy of myopathy due to maltase deficiency.



EXAMPLE 4: Identification of a protein which specifically binds the MT sequence (GGTATTT).

Materials and methods

1. Method of gel delay.

5 Nuclear extracts obtained from muscular cell lines C 2;7 ( Li and Paulin, 1991, J. Biol. Chem., 266, 6562-6570) and Rhabdomyosarcoma ( accessible to ATCC under N° CCL 136) are incubated with radiolabelled oligonucleotide MT, at room temperature during 10 minutes. The reaction mixtures are then  
10 resolved in 5% acrylamide gel. After 3 hours migration, the gel is dried and autoradiographed.

The mixtures have been performed in 20  $\mu$ l, and contain 750 ng of sonicated salmon sperm DNA, 3  $\mu$ g of nuclear extract, 15 mM HEPES pH 8, 0.1 mM EDTA, 1 mM DTT, 2 mM  $MgCl_2$ , 70 mM  
15 NaCl, 15% glycerol.

2) Southwestern blotting

Nuclear extracts are prepared from muscular cell lines C 2,7 and Rhabdomyosarcoma. 40  $\mu$ g of proteines are resolved on a 10% SDS-PAGE gel. The proteins are then transferred on a  
20 nitrocellulose filter. The filter is incubated during 1 hour at 37 °C in a solution containing 5% of skimmed milk , 50 mM Tris pH 7,5; 50 mM NaCl; 0,1 mM EDTA, 1 mM DTT and then in a buffer containing the labelled probe MT. The filter is washed three times in the same buffer and autoradiographed.

25 Results

The gel-delay experimentations ( figure 2) show that the MT sequence (GGTATTT) specifically binds a nuclear factor from murin ( C 2,7) and human (Rhabdomyosarcoma) muscular cell lines.

30 The autoradiogram obtained by Southwestern blotting (figure 3) shows that the nuclear factor which binds the MT sequence is a 35 kDa protein.

This protein is found in the muscular cells such as C 2,7 and Rhabdomyosarcoma cells.



The quantity of this protein increases during the muscular differentiation.



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

(A) NAME: INSTITUT PASTEUR  
(B) STREET: 28 RUE DU DOCTEUR ROUX  
(C) CITY: PARIS  
(E) COUNTRY: FRANCE  
(F) POSTAL CODE (ZIP): 75724

(A) NAME: UNIVERSITE PARIS 7  
(B) STREET: 2 PLACE JUSSIEU  
(C) CITY: PARIS  
(E) COUNTRY: FRANCE  
(F) POSTAL CODE (ZIP): 75251

(ii) TITLE OF INVENTION: AMPLIFYING SEQUENCES, VECTORS COMPRISING THESE SEQUENCES AND THEIR USES IN COMPOSITIONS FOR THE EXPRESSION OF NUCLEOTIDE SEQUENCES IN TRANSFECTED CELLS, THERAPEUTIC AND VACCINE APPLICATIONS

(iii) NUMBER OF SEQUENCES: 2

## (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

## (v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: WO PCT/WO96

## (vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: FR 95 01 937  
(B) FILING DATE: 20-FEB-1995

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens



(ix) FEATURE:  
 (A) NAME/KEY: allele  
 (B) LOCATION:replace(11, "g")

(ix) FEATURE:  
 (A) NAME/KEY: allele  
 (B) LOCATION:replace(12, "g")

(ix) FEATURE:  
 (A) NAME/KEY: allele  
 (B) LOCATION:replace(13, "a")

(ix) FEATURE:  
 (A) NAME/KEY: allele  
 (B) LOCATION:replace(16, "c")

(ix) FEATURE:  
 (A) NAME/KEY: allele  
 (B) LOCATION:replace(17, "g")

(ix) FEATURE:  
 (A) NAME/KEY: allele  
 (B) LOCATION:replace(18, "c")

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TCTATAAATA CCCGCTCTGG TATTTGGGGT TGGCAGCTGT T

41

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 8 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: Homo sapiens

(ix) FEATURE:  
 (A) NAME/KEY: allele  
 (B) LOCATION:replace(1, "c")

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TGGTATT

8



THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. Amplifying sequence showing a homology of at least 90% with the following SEQ ID N'1 sequence :  
TCTATAAATA X<sub>1</sub>X<sub>2</sub>X<sub>3</sub>GC Y<sub>1</sub>Y<sub>2</sub>Y<sub>3</sub>GG TATTTGGGGT TGGCAGCTGT T

5 in which :

- X<sub>1</sub>, X<sub>2</sub> and X<sub>3</sub> may independently represent respectively C or G, C or G, and C or A,

- Y<sub>1</sub>, Y<sub>2</sub> and Y<sub>3</sub> may independently represent respectively T or C, C or G, and T or C.

10 2. Sequence in accordance with claim 1, characterized in that it hybridizes under stringent conditions with sequence SEQ ID N'1.

3. Amplifying sequence having a homology of at least 90% with the following SEQ ID N'2 sequence :

15 ZGGTATTT

in which Z may be T or C.

4. Sequence in accordance with claim 3, characterized in that it hybridizes under stringent conditions with sequence SEQ ID N'2.

20 5. Sequence including at least :

- one amplifying sequence in accordance with any one of claims 1 to 4, in one or more copies, with the exception of sequences in which X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub> and Y<sub>2</sub> represent a cytosine (C) and Y<sub>1</sub> and Y<sub>3</sub> represent a thymidine (T) or Z represents (T),  
25 in which case sequences SEQ ID No 1 or SEQ ID No 2 are at least two copies thereof, and

- a sequence coding for a protein or corresponding to the genomic sequence encoding said  
30 protein or part thereof.

6. Sequence which is complementary of a sequence according to one of claims 1 to 5.

7. Sequence in accordance with claim 5, characterised in that it includes, in addition, one or more promoter(s)  
35 allowing the expression of said protein in cells transfected by said DNA.



8. Sequence in accordance with claim 5, characterised in that it includes a promoter internal to the gene or to the cDNA.

9. Sequence in accordance with claim 5, characterised in that it includes a promoter that is heterologous to the gene.

10. Sequence in accordance with either of claims 5 and 7 characterised in that it includes sequences allowing the homologous recombination of a gene to be replaced in an organism, said sequences being placed upstream and downstream from the sequence coding for a protein, and possible from the promoter.

11. Sequence in accordance with claim 5, characterised in that it includes a Krox type sequence.

12. Sequence in accordance with either of claims 7 and 9, characterised in that said promoter is that of the desmine gene or thymidine kinase.

13. Sequence in accordance with any one or more of claims 5 and 7 to 12, characterised in that the encoded protein is a protein having therapeutic or vaccine properties.

14. Sequence in accordance with any one or more of claims 5 and 7 to 12, characterised in that the encoded protein is at least part of a bacterial antigen, for example urease, the protein of *H. pylori* so-called HSP (A), or the HSP (B) protein, or of a viral antigen, for example the surface antigen of a hepatitis virus, in particular hepatitis-B, and preferably protein HBs, of a hepatitis-A or of a non-A non-B hepatitis, such as hepatitis C, E or delta, or even part of a protein of the HIV-1 virus, of the HIV-2 virus or of the HTLV-1 virus, in particular the ENV protein (corresponding to the virus envelope) or the gag protein, and generally a protein having immunotherapy interest, for example the Interleukins,



fibroblast (FGF) or nerve (NGF) growth factors, proteins which are able to induce an immune, humeral or cytotoxic response, or proteins which allow complementary gene activity in a gene normally expressed in the individual to be treated but which is no longer expressed either through mutation or deletion of tis sequence such as the maltase gene.

15. Vector including a sequence in accordance with any one of claims 5 to 14.

16. vector in accordance with claim 15 characterised in that it includes an active replication origin in a micro-organism.

17. Vector in accordance with either of claims 15 and 16, characterised in that it includes an active replication origin in Escherichia coli.

18. Vector in accordance with any one or more of claims 15 to 17, characterised in that it includes a resistance gene expressing itself in a micro-organism.

19. Vector in accordance with any one or more of claims 15 to 18, characterised in that it is a plasmid.

20. Micro-organism carrying a sequence or a vector in accordance with any one or more of claims 15 to 18.

21. Micro-organism in accordance with claim 20, such as filed and registered with the CNCM under No. I-1508 or I-1509.

22. Pharmaceutical composition characterised in that it contains a pharmacologically effective quantity of a sequence or of a vector in accordance with any one or more of claims 5 to 18.

23. Composition in accordance with claim 22, characterised in that it contains pharmacologically compatible



excipients.

24. medicinal or vaccine product characterised in that it contains a vector or a sequence in accordance with any one or more of claims 5 to 18.

25. Application of the composition in accordance with either of claims 22 and 23, or of the medicinal product in accordance with claim 24 which express nucleotide sequences having an immunomodulatory function, for example cytokines, or immunogenic properties, for example the VPI protein of the poliovirus or HBS of hepatitis-B, for gene therapy through the introduction of DNA into tissues requiring the use of sequences or vectors strongly expressing the genes carried by these sequences.

26. The use of a composition in accordance with either of claims 22 and 23, or of the medicinal product in accordance with claim 24 which express nucleotide sequences having an immunomodulatory function, for example cytokines, or immunogenic properties, for example the VPI protein of the poliovirus or HBS of hepatitis-B, for the manufacture of a pharmaceutical product when used for the gene therapy through the introduction of DNA into tissues requiring the use of sequences or vectors strongly expression the genes carries by these sequences.

INSTITUT PASTEUR and UNIVERSITE PARIS 7

DATED this 26th day of February 1999

WATERMARK PATENT & TRADEMARK ATTORNEYS  
290 BURWOOD ROAD  
HAWTHORN VICTORIA 3122  
AUSTRALIA

VAX DOC025 AU4835696.WPC : IAS/KMH/RES



1/3

	MEF2	MT	MyoD1	CAT ACTIVITY
3	AAGCTTCTCCTCTATAAAATACCCGGCTCTGGIAITTTGGGGTTGGCAGCIGI			100
4	AAGCTTCTCCTCGAGAAATACCCGGCTCTGGTATTTGGGGTTGGCAGCIGI			8
5	AAGCTTCTCCTCTAGGCCCTACCCGGCTCTGGTATTTGGGGTTGGCAGCIGI			28
6	AAGCTTCTCCTCTATAAAATAGGAGCTCTGGTATTTGGGGTTGGCAGCIGI			120
7	AAGCTTCTCCTCTATAAAATACCCGGCGGGTATTTGGGGTTGGCAGCIGI			130
8	AAGCTTCTCCTCTATAAAATACCCGGCTCTGAGATCTGGGGTTGGCAGCIGI			27
9	AAGCTTCTCCTCTATAAAATACCCGGCTCTGGTATCTGGGGTTGGCAGCIGI			41
10	AAGCTTCTCCTCTATAAAATACCCGGCTCTGGATCTGGGGTTGGCAGCIGI			15
11	AAGCTTCTCCTCTATAAAATACCCGGCTCTGGTATTTAAAGTTGGCAGCIGI			108
12	AAGCTTCTCCTCTATAAAATACCCGGCTCTGGTATTTGGGGTTGGAAICTGTT			3
13	AAGCTTCTCCTCGAGAAATACCCGGCTCTGGGATCTGGGGTTGGCAGCIGI			5
14	AAGCTTCTCCTCTAGGCCCTACCCGGCTCTGGGATCTGGGGTTGGCAGCIGI			15

Figure 1

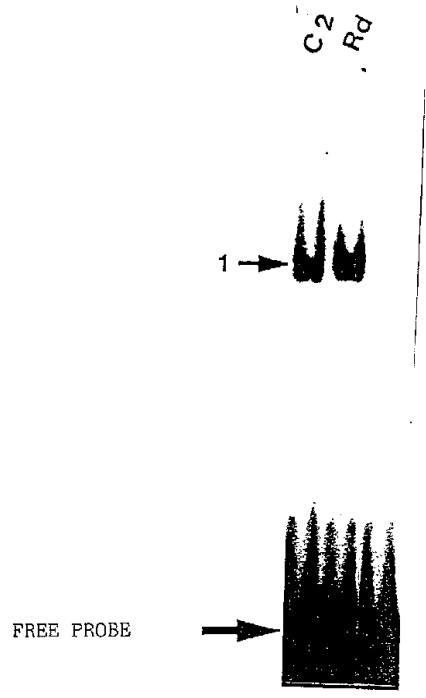


Figure 2

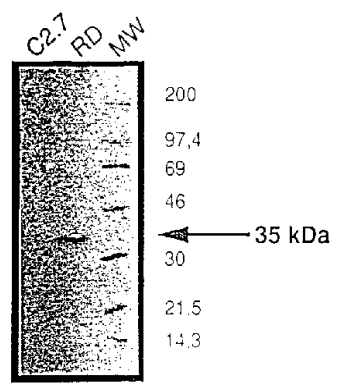


Figure 3