(54) Title: CELL SPECIFIC PROMOTERS OF UNCOUPLING PROTEIN 3

(57) Abstract

The present invention is in the field of the gene promoters which mediate the transcription of proteins which play a part in the energy management of cells and in obesity. The invention relates to DNA molecules which contain recombinant cell-specific promoters of uncoupling protein 3 (UCP 3) or functional derivatives thereof. It also relates to cells which contain these DNA molecules. The invention also relates to uses of these DNA molecules and the cells according to the invention, and processes for finding substances which are capable of influencing transcription. These processes can also be carried out by the High Throughput Screening method. Processes are also disclosed which can be used to find substances or factors which bind to the DNA molecules according to the invention.
FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<table>
<thead>
<tr>
<th>Code</th>
<th>Country</th>
<th>Code</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>Albania</td>
<td>LS</td>
<td>Lesotho</td>
</tr>
<tr>
<td>AM</td>
<td>Armenia</td>
<td>LT</td>
<td>Lithuania</td>
</tr>
<tr>
<td>AT</td>
<td>Austria</td>
<td>LU</td>
<td>Luxembourg</td>
</tr>
<tr>
<td>AU</td>
<td>Australia</td>
<td>LV</td>
<td>Latvia</td>
</tr>
<tr>
<td>AZ</td>
<td>Azerbaijan</td>
<td>MC</td>
<td>Monaco</td>
</tr>
<tr>
<td>BA</td>
<td>Bosnia and Herzegovina</td>
<td>MD</td>
<td>Republic of Moldova</td>
</tr>
<tr>
<td>BB</td>
<td>Barbados</td>
<td>MG</td>
<td>Madagascar</td>
</tr>
<tr>
<td>BE</td>
<td>Belgium</td>
<td>MK</td>
<td>The former Yugoslav</td>
</tr>
<tr>
<td>BF</td>
<td>Burkina Faso</td>
<td>ML</td>
<td>Mali</td>
</tr>
<tr>
<td>BG</td>
<td>Bulgaria</td>
<td>MN</td>
<td>Mongolia</td>
</tr>
<tr>
<td>BJ</td>
<td>Benin</td>
<td>MR</td>
<td>Mauritania</td>
</tr>
<tr>
<td>BR</td>
<td>Brazil</td>
<td>MW</td>
<td>Malawi</td>
</tr>
<tr>
<td>BY</td>
<td>Belarus</td>
<td>MX</td>
<td>Mexico</td>
</tr>
<tr>
<td>CA</td>
<td>Canada</td>
<td>NE</td>
<td>Niger</td>
</tr>
<tr>
<td>CF</td>
<td>Central African Republic</td>
<td>NL</td>
<td>Netherlands</td>
</tr>
<tr>
<td>CG</td>
<td>Congo</td>
<td>NO</td>
<td>Norway</td>
</tr>
<tr>
<td>CH</td>
<td>Switzerland</td>
<td>NZ</td>
<td>New Zealand</td>
</tr>
<tr>
<td>CI</td>
<td>Côte d'Ivoire</td>
<td>PL</td>
<td>Poland</td>
</tr>
<tr>
<td>CM</td>
<td>Cameroon</td>
<td>PT</td>
<td>Portugal</td>
</tr>
<tr>
<td>CN</td>
<td>China</td>
<td>RO</td>
<td>Romania</td>
</tr>
<tr>
<td>CU</td>
<td>Cuba</td>
<td>RU</td>
<td>Russian Federation</td>
</tr>
<tr>
<td>CZ</td>
<td>Czech Republic</td>
<td>SD</td>
<td>Sudan</td>
</tr>
<tr>
<td>DE</td>
<td>Germany</td>
<td>SE</td>
<td>Sweden</td>
</tr>
<tr>
<td>DK</td>
<td>Denmark</td>
<td>SG</td>
<td>Singapore</td>
</tr>
<tr>
<td>EE</td>
<td>Estonia</td>
<td>SI</td>
<td>Slovenia</td>
</tr>
<tr>
<td>SK</td>
<td>Slovakia</td>
<td>SN</td>
<td>Senegal</td>
</tr>
<tr>
<td>SZ</td>
<td>Swaziland</td>
<td>TD</td>
<td>Chad</td>
</tr>
<tr>
<td>TG</td>
<td>Togo</td>
<td>TJ</td>
<td>Tajikistan</td>
</tr>
<tr>
<td>TR</td>
<td>Turkey</td>
<td>TM</td>
<td>Turkmenistan</td>
</tr>
<tr>
<td>TT</td>
<td>Trinidad and Tobago</td>
<td>UA</td>
<td>Ukraine</td>
</tr>
<tr>
<td>UG</td>
<td>Uganda</td>
<td>US</td>
<td>United States of America</td>
</tr>
<tr>
<td>UZ</td>
<td>Uzbekistan</td>
<td>VN</td>
<td>Viet Nam</td>
</tr>
<tr>
<td>YU</td>
<td>Yugoslavia</td>
<td>YU</td>
<td>Yugoslavia</td>
</tr>
<tr>
<td>ZW</td>
<td>Zimbabwe</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Cell specific promoters of uncoupling protein 3

The present invention is in the field of the gene promoters of proteins which play a part in the energy management of cells and in obesity. The subject matter of the invention is recombinant DNA molecules (DNA: deoxyribonucleic acid), which contain cell specific promoters of uncoupling protein 3 (UCP 3) or functional derivatives thereof. The invention also relates to cells which contain these recombinant DNA molecules, the uses of these DNA molecules and the cells according to the invention and process for finding substances which influence the transcription of the promoters according to the invention or which bind to the DNA molecules according to the invention.

Obesity is a disease in which the fatty tissue increases as a result of a positive energy balance sheet. This can be the result of an excessive food intake or a symptom of a metabolic disorder (ROCHE Lexicon of Medicine, 3rd Edition, Urban & Schwarzenberg). Obesity is a problem which is frequently caused by food in the industrial western countries and plays an important part as a cause of illness and death (McGinnis and Foege, 1993; Manson et al., 1995). Since obesity is caused by a constant imbalance between food intake and energy use, a chronic reduction in energy use might be a risk factor for this condition. In fact, a lower energy use at rest is a risk factor for obesity (Ravussin et al., 1988; Griffiths et al., 1990). Genetic factors contribute significantly to the level of basic energy use (Bogardus et al., 1986; Bouchard et al., 1989). Brown Adipose Tissue (BAT) is specialised for thermogenesis, one of the main factors of energy conversion (Himms-Hagen, 1989). The so-called uncoupling proteins 1, 2 and 3 (UCPs for short) have central importance in the
thermogenic function of BAT (Klaus et al., 1991; Boss et al., 1997; Fleury et al. 1997; Gimeno et al., 1997; Gong et al., 1997; Vidal-Puig et al., 1997). These proteins stimulate the production of heat by uncoupling substrate oxidation from ATP synthesis (Ricquier et al. 1991). Whereas the importance of BAT has been shown as a regulator of the body fat store for rodents (Lowell et al., 1993; Kopecky et al., 1995), its role in obesity in humans is unclear as adults have very little BAT (Krief et al. 1993).

Regardless of this restriction, a significantly lower UCP1-mRNA content (RNA: ribonucleic acid; mRNA: messenger ribonucleic acid) was found in the intraperitoneal fatty tissue of obese individuals than in slim control subjects (Oberkofer et al., 1997). A major part of the fluctuations in the quantity of UCP1-mRNA in obese individuals was explained by normal changes in sequence in the UCP1 gene locus (Esterbauer et al., in press). Therefore, it is possible that UCP1 and BAT contribute to the basic energy conversion.

UCP2 and UCP3, two other recently discovered members of the UCP family, are expressed in different tissues in humans and rodents. UCP2-mRNA can be found in white fatty tissue, BAT, lungs, liver, spleen and macrophages (Fleury et al., 1997; Gimeno et al., 1997), whereas large amounts of UCP3-mRNA are observed in the skeletal muscle and BAT (Boss et al., 1997; Gong et al., 1997; Vidal-Puig et al., 1997). Studies in mice deficient in the UCP1 gene (Enerback et al., 1997) and in slender rats which over-express leptin (Zhou et al., 1997) support a model in which UCP2 may act as a reserve system for thermogenesis if UCP1 is deficient and/or the quantity of leptin increases as the result of overweight. UCP3 could have profound effects on energy homeostasis, as muscle tissue contains a major proportion of catecholamine and is responsible for nutrition-induced
thermogenesis both in humans (Astrup et al., 1986) and in rats (Thurlby and Ellis, 1986).

The human UCP2 and UCP3 genes are located in chromosome region 11q13 (Fleury et al., 1997; Solanes et al., 1997; Boss et al., 1998). Investigation of the family trees of the Quebec family study produced very powerful indications that this chromosomal region is associated with energy conversion in the resting state, the body mass index (BMI) and the fatty mass in adults (Bouchard et al., 1997).

Moreover, the syntenic region on chromosome 7 in the mouse is responsible for obesity and non-insulin-dependent Diabetes mellitus (Hashimoto et al., 1994; Warden et al., 1995).

Compared with the slim control subjects, the quantity of UCP2-mRNA in the intraperitoneal fatty tissue of morbidly obese people was reduced. In accordance with the role in the pathophysiology of obesity, the quantity of UCP2-mRNA in patients who have overcome obesity is low, both before and after weight reduction (Oberkofler et al., in press). By contrast, the quantity of UCP3-mRNA in muscle tissue did not vary between obese and slender individuals (Millet et al., 1997), but these results must be interpreted cautiously. Firstly, tissue-specific differences were found in the regulation of UCP3 expression in rodents (Gong et al., 1997; Larkin et al., 1997; Boss et al., 1998). Secondly, two different isoforms (UCP3L and UCP3S) of the human UCP3 gene were found by cloning the DNA complementary to mRNA (complementary DNA; cDNA). A polyadenylation signal in Intron 6 is used in a significant proportion of the UCP3 transcripts for forming UCP3S, whilst a second polyadenylation signal in exon 7 is used for forming UCP3L. Because of a C-terminal truncation, the UCP3S form lacks the sixth predicted transmembrane domain and the purine nucleotide binding domain which has been deemed responsible
for inhibiting UCP activity by nucleotides (Jezek et al., 1994). The truncation of the UCP3 molecule could increase the UCP3 activity in this way, but defective membrane insertion could affect the stability and function of UCP3.

It is desirable to find better methods of treating obesity. Substances which can be administered to the patient in as painfree a manner as possible and with as few side effects as possible are advantageous. One possibility of simplifying the search for such substances is to identify factors native to the body which play an important role in the pattern of this illness. These might be a target molecule for suitable substances, interaction with which might influence the activity or properties of the native factor in a manner which is positive for the patient. Suitable target molecules could also be defined if known factors native to the body can be shown to participate in the events which precipitate the disease. After a suitable target molecule has been identified, these substances might be found with greater prospects of success by measuring the influence of as many compounds as possible on the activity or properties of the target molecule. These compounds might be derived from a library of natural or other substances, and combinatorial chemistry could make a valuable contribution. To speed up this process the analyses are carried out by high throughput screening (HTS). Naturally, only a few substances are found but these could be used as the basis for chemical derivatisation and optimisation and subsequent pharmacological characterisation.

For the UCP3 gene, the amino acid sequence derived from cDNA, the cDNA sequence (Gene Bank Accession number U84763), the genomic DNA sequence including the position of the exons and introns (Boss et al., 1998) and part of the 5'-non-coding region are already known in the prior art
(Gene Bank Accession number AF032871). Moreover the position of a promoter and a TATA signal were known (Entry AF032871 in the Gene Bank databank). The genomic DNA sequence of the UCP3 gene including the introns and exons and the 5'-region is enclosed for inspection with this application (see SEQ ID NO: 17).

Surprisingly, it has now been found within the scope of this invention that there is an additional promoter in the region upstream of TATA signal which is preferentially active in fat cells (see Example 1). The UCP3 gene is otherwise only strongly expressed in the skeletal muscle but is expressed there using the known promoter, as has surprisingly been found within the scope of this invention (see Example 1). It is advantageous to influence the promoter which is preferentially active in fatty cells or in muscle cells, since this could serve as a treatment for obesity. This is particularly advantageous as it can be used to modulate a factor which is active only in certain parts of the body. DNA constructs can thus be produced which make it possible to investigate the fat cell-specific or muscle cell-specific transcription of the UCP3 gene and find corresponding substances for modulating them. For this purpose, only the corresponding part of the UCP3 promoter which is active in the other type of cell need be removed.

The objective of preparing a new suitable target molecule can thus be achieved with the present invention as set forth in the specification and claims. According to the invention, recombinant DNA molecules with cell specific promoters of the UCP3 gene are prepared. In connection with this, recombinant DNA molecules which contain functional derivatives of the promoters according to the invention or comprise certain parts of the 5'-sequence of the UCP-3 promoter are also disclosed by the present
invention. The present invention also includes cells which contain DNA molecules according to the invention. In another embodiment, the use of these DNA molecules or the cells according to the invention for transcribing a gene or for discovering substances which influence transcription is also disclosed. The invention further includes processes which make it possible to discover substances which influence the transcription rate. The process according to the invention in the high throughput screening format is particularly preferred. Processes are also disclosed which make it possible to find substances or factors which bind to the DNA molecules according to the invention.

This invention discloses a recombinant DNA molecule which contains the UCP3 promoter preferentially active in fat cells but not the known UCP3 promoter which is active in muscle cells. This recombinant DNA molecule contains the sequences SEQ ID NO: 1 TATATTAAA and SEQ ID NO: 2 CACCTC, but not SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC.

However, the sequences which are important for the promoter which is active in muscle cells can also be altered by point mutation or deletions or combinations thereof so that they lose their function in transcription. Between the sequences SEQ ID NO: 1 and SEQ ID NO: 2 there may be 30 to 50 base pairs, preferably 40 to 45 base pairs, most preferably 42 base pairs. In another embodiment the DNA molecule described is further characterised in that upstream of SEQ ID NO: 1 there is additionally an RXR/PPAR element (Schoonjans et al., 1996) which preferably includes the sequence SEQ ID NO: 5 TGACCTTTGGACT, whilst there are preferably 65 to 75 base pairs between the sequences SEQ ID NO: 1 and SEQ ID NO: 5. A summary of essential promoter elements is provided by Locker and Buzard (1990). Another embodiment of the invention is characterised in that, upstream of SEQ ID NO: 1, there is additionally an Alu sequence, whilst there may be 255 to 265 base pairs between
the sequence SEQ ID NO: 1 and the Alu sequence. For the purposes of the invention an Alu sequence means a section of about 300 base pairs which occurs only in the human genome, which belongs to the highly repetitive DNA sequences and carries a cutting site for the restriction enzyme Alu I more or less in the middle. The DNA molecule described may be further characterised in that, upstream of SEQ ID NO: 1, there is additionally an E2A element (Locker and Buzard, 1990; Aronheim et al., 1991; Leshkowitz et al., 1992; Park and Walker, 1992) which corresponds to the sequence SEQ ID NO: 6 CAGATG, whilst there may be 440 to 450 base pairs between the sequences SEQ ID NO: 1 and SEQ ID NO: 6. The DNA molecule described may be further characterised in that there is additionally an E box (Aronheim et al., 1991; Leshkowitz et al., 1992; Park and Walker, 1992) upstream of SEQ ID NO: 1, which preferably includes the sequence SEQ ID NO: 7 CACTTG, whilst between the sequences SEQ ID NO: 1 and SEQ ID NO: 7 there may be 450 to 460 base pairs. The DNA molecule described may be further characterised in that there is additionally another E box (Aronheim et al., 1991; Leshkowitz et al., 1992; Park and Walker, 1992) upstream of SEQ ID NO: 1, which preferably includes the sequence SEQ ID NO: 8 CATT TG, whilst there may be 460 to 470 base pairs between the sequences SEQ ID NO: 1 and SEQ ID NO: 8. The DNA molecule according to the invention may be further characterised in that, upstream of SEQ ID NO: 1, there may additionally be an octamer sequence (Locker and Buzard, 1990) which preferably includes the sequence SEQ ID NO: 9 ATGAAATA GT, whilst there may be 515 to 525 base pairs between the sequences SEQ ID NO: 1 and SEQ ID NO: 9. The DNA molecule according to the invention may be further characterised in that, upstream of SEQ ID NO: 1, there may additionally be another CAAT box (Locker and Buzard, 1990) which preferably includes the sequence SEQ ID NO: 10 CCAAT, whilst there may be 560 to 570 base pairs between the sequences SEQ ID NO: 1
and SEQ ID NO: 10. The CAAT box (Locker and Buzard, 1990) is part of a base sequence which is conserved in many eukaryotic genes, observed in the 5'-flanking section of the coding region. The DNA molecule according to the invention may be further characterised in that, upstream of SEQ ID NO: 1, there may additionally be another CAAT box (Locker and Buzard, 1990) which preferably contains the sequence SEQ ID NO: 11 ATGG, whilst there may be 670 to 680 base pairs between the sequences SEQ ID NO: 1 and SEQ ID NO: 11. The DNA molecule according to the invention may be further characterised in that a CAAT box (Locker and Buzard, 1990) may additionally be provided upstream of SEQ ID NO: 1, preferably comprising the sequence SEQ ID NO: 12 ATGG, whilst there may be 730 to 740 base pairs between the sequences SEQ ID NO: 1 and SEQ ID NO: 12. The DNA molecule according to the invention may be further characterised in that, upstream of SEQ ID NO: 1, there is additionally a binding site for an upstream binding stimulating factor (USF; Locker and Buzard, 1992) which preferably includes the sequence SEQ ID NO: 13 CCACGTGC, whilst there may be 845 to 855 base pairs between the sequences SEQ ID NO: 1 and SEQ ID NO: 13.

This invention discloses another recombinant DNA molecule which does not contain the UCP3 promoter preferentially active in fat cells but rather contains the promoter which is active in muscle cells. This recombinant DNA molecule contains the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC, but not the sequences SEQ ID NO: 1 TATATTAAA and SEQ ID NO: 2 CACCTC. However, the sequences which are important for the promoter active in fat cells may be altered by point mutation or deletions or combinations thereof so that they lose their function in transcription. There may be 45 to 70 base pairs, preferably 52 to 58 base pairs, most preferably 55 base pairs between sequences SEQ ID NO: 3 and SEQ ID NO: 4. In another embodiment the DNA
molecule described is further characterised in that, upstream of SEQ ID NO: 3, there is additionally an RXR/PPAR element which preferably includes the sequence SEQ ID NO: 5 TGACCTTTTGGACT. A further embodiment of the invention is characterised in that upstream of SEQ ID NO: 3 there is additionally an Alu sequence. The DNA molecule described may be further characterised in that upstream of SEQ ID NO: 3 there is additionally an E2A element which corresponds to the sequence SEQ ID NO: 6 CAGATG. The DNA molecule described may be further characterised in that, upstream of SEQ ID NO: 3, there is additionally another E box which preferably contains the sequence SEQ ID NO: 8 CATT TG. The DNA molecule according to the invention may be further characterised in that, upstream of SEQ ID NO: 3, there may additionally be an octamer sequence which preferably includes the sequence SEQ ID NO: 9 ATGAAATGT.

The DNA molecule according to the invention may be further characterised in that upstream of SEQ ID NO: 3 there may additionally be another CAAT box which preferably includes the sequence SEQ ID NO: 10 CCAAT. The CAAT box is part of a base sequence which is conserved in many eukaryotic genes and observed in the 5'-flanking section of the coding region. The DNA molecule according to the invention may be further characterised in that, upstream of SEQ ID NO: 3, there may additionally be another CAAT box which preferably includes the sequence SEQ ID NO: 11 ATTGG. The DNA molecule according to the invention may be further characterised in that, upstream of SEQ ID NO: 3, there is additionally a binding
site for an upstream binding stimulation factor which preferably includes the sequence SEQ ID NO: 13 CCACGTGC.

The invention also include DNA molecules which contain the sequence SEQ ID NO 14 or a sequence which hybridises to SEQ ID NO 14 under stringent conditions. By stringent conditions is meant conditions which select for more than 85 %, preferably more than 90 % homology (Sambrook et al., 1989). The hybridisations are carried out in 6x SSC/5x Denhardt's solution/0.1 % SDS (SDS: sodium dodecylsulphate) at 65°C. The degree of stringency is determined in the washing step. Thus, for selection for DNA sequences with approximately 85 % homology or more, the conditions of 0.2 x SSC/0.01 % SDS/65°C are suitable and for selection for DNA sequences with about 90 % homology or more, the conditions of 0.1x SSC/0.01 % SDS/65°C are suitable. The composition of the reagents is described in Sambrook et al. (1989).

The invention further comprises functional derivatives of the DNA molecules according to the invention with the cell specific UCP3-promoters. Under the terms of the invention, the term functional derivatives includes all DNA sequences which have come into being by point mutations or deletions or combinations of point mutations or deletions from the original sequence. These functional derivatives have according to the invention, the common ability of mediating transcription of a gene fused at the 3'-end of the functional derivative. It is within the capabilities of the average skilled person to produce these functional derivatives by targeted or random mutagenesis or by inserting deletions from the sequences according to the invention. The precise methods are described by Sambrook et al. (1989). For example, oligonucleotide-directed or random mutagenesis methods may be used. For the latter process, polymerase chain reaction may also be used, which
operates imprecisely if it is carried out under suboptimal conditions (Lin-Goerke et al., 1997). The deletions may be inserted, for example, by using oligonucleotides with methods of targeted mutagenesis or preferably by amplification of desired areas by polymerase chain reaction. The average skilled person can find a functional derivative by determining the activity of a reporter gene such as luciferase fused at the 3'-end of the functional derivative. It is particularly preferred to determine the activity of a functional derivative by the chloramphenicol transferase test ("CAT assay") carried out according to Sambrook et al. (1989). The average skilled person will regard a sequence obtained by point mutation or deletion or a combination of both changes, which produces a signal above the background in one of the two tests described above, as a functional derivative.

In a preferred embodiment of the invention the DNA molecules according to the invention are present in the form of vectors, particularly expression vectors. In a particularly preferred embodiment the DNA molecule according to the invention is a plasmid. In another embodiment the DNA molecule according to the invention is a viral vector.

In another preferred embodiment, DNA molecules are disclosed which in addition to the DNA sequence of the promoters according to the invention contain other genes which are functionally linked to the promoters according to the invention. For the purposes of this invention the term functional means that the linking is carried out in a way which makes it possible either to transcribe the gene or to transcribe the gene and obtain its translation product. The genes may be the cDNA or the genomic sequence of the uncoupling protein 3 or they may be reporter genes. The translation products of the reporter genes are preferably
proteins the quantity of which can easily be determined by measuring the activity, absorption, luminescence or fluorescence, e.g. green fluorescent protein (GFP), chloramphenicol-acetyl-transferase, β-galactosidase, secreted alkaline phosphatase or luciferase. For the purposes of this invention green fluorescent protein also includes all the variants thereof which fluoresce in other wavelengths but are derived from the amino acid sequence of green fluorescent protein. Most preferably, the proteins coded by the reporter gene are enzymes which catalyse a reaction the end product of which is easy to determine quantitatively. Thus, for example, the enzyme luciferase catalyses a chemical reaction in which light is emitted by luminescence which can be measured in a luminometer.

Another aspect of the invention is a DNA molecule which contains the sequence SEQ ID NO 15 or a sequence which hybridises with SEQ ID NO 15 under stringent conditions. By stringent conditions, the skilled person means conditions which select for more than 85 %, preferably more than 90 % homology (Sambrook et al., 1989). The hybridisations are carried out in 6x SSC/5x Denhardt's solution/0.1% SDS at 65°C. The degree of stringency is determined in the washing step. Thus, for selection for DNA sequences with about 85 % homology or more, the suitable conditions include 0.2 x SSC/0.01% SDS/65°C and for selection for DNA sequences with about 90 % homology or more, the suitable conditions include 0.1x SSC/0.01% SDS/65°C. The composition of the reagents described is also described in Sambrook et al. (1989).

A further aspect of the invention is a host cell into which a DNA molecule according to the invention has been introduced. This may be a eukaryotic host cell, preferably a yeast or mammalian cell. Particularly preferred host cells are adipocytes in primary culture, preferably from
humans or from mice, adipocyte cell lines such as, for example, NIH-3T3L1 (a murine adipocyte cell line), muscle cells in primary culture, muscle cell lines such as for example C2C12 (a muscle cell line from the mouse) or cell lines such as HEK293 or HeLa. The average skilled person knows how to introduce suitable DNA molecules into certain cells using standard methods. These methods are described in Sambrook et al. (1989). For eukaryotic cells it is preferable to use the calcium precipitation method, lipofection or electroporation. It is also within the capabilities of the average skilled person to take parameters from the literature and develop them without too much effort so as to use the methods for introducing DNA molecules into these cells successfully.

A further aspect of the invention is the use of a DNA molecule according to the invention with above-mentioned cell-specific promoters for the transcription of a gene. In order to produce suitable DNA molecules, the skilled man will use molecular biological methods as described by Sambrook et al. (1989) which he will select with a knowledge of the sequence of the DNA molecules according to the invention, the sequence of the gene to be transcribed and other elements. The transcripts or the translation product is or are preferably detected by standard methods, e.g. detection of RNA by sequence-specific hybridisation (Northern blot), detection of the translation product by means of antibodies (e.g. Western blot) or by its activity in suitable detection systems. The translation product may also have an enzymatic activity. The methods involved are described by Sambrook et al. (1989).

A particularly preferred embodiment of the invention is the use of the DNA molecules according to the invention for discovering substances which influence the transcription of the UCP3 promoters according to the invention. For this,
DNA molecules are prepared which include the sequence of the promoters according to the invention, the sequence SEQ ID NO 14 or SEQ ID NO 15. These may also contain functional derivatives thereof. Preferably, these DNA molecules are linked to other DNA molecules which contain the sequences for genes which allow easy detection of the gene product (reporter genes). These reporter genes may be, for example, the luciferase gene or green fluorescent protein (GFP for short). The RNA transcribed by the promoter according to the invention can also be detected and quantified. In another embodiment of the invention, cells are used here which have been transformed with the DNA molecules according to the invention.

A preferred aspect of the invention is a process for discovering substances which are capable of influencing transcription of the UCP3-promoter which is preferentially active in fat cells or preferentially active in muscle cells and in which the DNA molecules according to the invention can be used. The process according to the invention may be carried out in a cell-free or cell-based system.

One embodiment of such a process consists of a cell-free in vitro transcription system which contains as components at least cell extract, ribonucleotides and a DNA molecule according to the invention. A preferred embodiment consists in measuring the transcription rate of the promoter used in the present of a test substance and comparing it with the transcription rate in the absence of the test substance. If the quantity of RNA transcribed per unit of time in the presence of the test substance is lower than the comparison value (i.e. in the absence of test substance), the test substance inhibits transcription in this test. If the quantity of RNA transcribed per unit of time in the presence of the test substance is higher than
the comparison value (i.e. in the absence of the test substance), the test substance increases transcription in this test.

Another embodiment of the process for finding substances capable of influencing the transcription of the promoter according to the invention consists in measuring the change in activity of a reporter gene fused to the 3'-end of a DNA molecule according to the invention, e.g. the luciferase gene or green fluorescent protein. Such a process can be carried out in a cell-free system. One embodiment of the process according to the invention comprises converting the RNA obtained after in vitro transcription in the presence of a test substance into proteins by in vitro translation, then determining the quantity, activity, absorption, luminescence or fluorescence of said proteins. The RNA obtained is converted into proteins which catalyse an enzymatic reaction the product of which is easily determined. It is most particularly preferred to use luciferase or green fluorescent protein.

A particularly preferred embodiment is a cell based process according to the invention wherein cells which contain a DNA molecule according to the invention grow in the presence of a test substance and under conditions under which transcription is carried out by the UCP3-promoters according to the invention. The principle of the process is based on using the UCP3-promoter according to the invention to express a reporter gene in order to investigate the effects of added test substances. In this way, the expression of the reporter gene simulates the expression of the UCP3 gene. The promoter-reporter constructs according to the invention can be used in tests after being introduced into the cell or after a cell line has been produced which has these constructs stably integrated in the genome. Preferably, at a given time, the
quantity, activity, luminescence or fluorescence of the translation product of the reporter gene fused to the 3'-end of the cell-specific UCP3-promoter according to the invention will be determined, the cells possibly having been lysed beforehand. A preferred embodiment of such a process comprises measuring the results in the presence of a test substance and comparing them with results in the absence of the test substance. If the quantity, activity, absorption, luminescence or fluorescence of the reporter protein produced per unit of time in the presence of the test substance is lower than the comparison value (i.e. in the absence of test substance), the test substance inhibits the transcription in this test. If the quantity, activity, absorption or fluorescence of the reporter protein produced per unit of time in the presence of the test substance is higher than the comparison value (i.e. in the absence of test substance), the test substance increases transcription in this test. The signal level is a measurement of the activity of the UCP3-promoter.

In practice, the promoters according to the invention could be cloned into plasmids which harbour reporter genes such as luciferase, β-galactosidase, chloramphenicol-acetyltransferase or secreted alkaline phosphatase. Plasmids which contain a reporter gene include, for example, the commercially obtainable pGL2basic or pGL3basic (Promega, Madison, Wisconsin, USA) and pSEAP2basic or pbgal-basic (Clontech, Heidelberg, Germany). In the same way it is possible to use cells which have the above-mentioned DNA constructs stably integrated in their genome. Cells which may be used for the screening test include, for example, primary cell cultures e.g. of adipocytes or muscle tissue, or cell lines such as HEK293, HeLa, C2C12 or NIH-3T3. If primary cell cultures are cultivated for longer periods, the UCP3-expression is lost. No UCP3-expression can be detected in some cell lines. However, these cell lines are
suitable for finding activators of UCP3 transcription since they also have an activating effect in cells which express UCP-3. Activators of UCP-3-transcription in the muscle tissue are particularly suitable for the preparation of pharmaceutical compositions for treating obesity. However, activators of UCP-3-transcription in the fatty tissue can also be used for this purpose.

The average skilled person can perform a test of this kind (see Example 2), by using standard methods and ready-made reagents such as those sold ready-made by Promega (Madison, Wisconsin, USA), for example (Luciferase Assay System). He will also obtain instructions as to how to carry out the test with the commercially available reagents. The skilled person produces suitable constructs using standard methods and transfekt them into suitable cell cultures (Sambrook et al., 1989). After growth in the presence of the test substance, the cell cultures are optionally washed and the cells are then lysed. After centrifugation to remove the cell debris the supernatant is mixed with the test reagent containing luciferin and the light emitted is measured in a luminometer.

In a preferred embodiment a process according to the invention is carried out by high throughput screening (HTS). A large number of test substances are tested simultaneously in a suitable arrangement. An HTS method of this kind may advantageously be fully or partially automated and evaluation may be carried out by electronic data processing. Advantageously, a test substance which exhibited an increasing effect in the cell-free system is additionally tested in a cell-based system. The process described in the previous paragraph can readily be adapted by anyone skilled in the art to the requirements of the high throughput screening without any inventive step.
Another embodiment of the invention comprises using the DNA molecules according to the invention to identify other factors or substances which bind to them. Preferably, factors which mediate, lower or raise the transcription of a gene fused at the 3'-end (transcription factors) are sought. For this purpose, the DNA molecules according to the invention may be bound to insoluble carriers and brought into contact with a mixture of different factors or substances, e.g. a core extract or a library of natural substances, under suitable conditions. After one or more washing steps, and possibly repetition of the binding process, the bound factors or substances are identified by suitable methods, such as amino acid sequencing, mass spectroscopy, gel electrophoresis.

The DNA molecules for the purposes of this invention are all molecules which consist of naturally occurring deoxyribonucleotides. However, according to the invention, they also include molecules which have retained the basic chemical nature of the DNA after chemical derivatisation. The possible types of derivatisation include for example derivatisation with fluorescent dyes, changes in the basic structure such as the replacement of oxygen atoms for sulphur atoms in order to stabilise the molecule or convert the DNA molecule into peptide nucleic acids (PNA for short).
Examples

Example 1 - Identification of two transcription start sites in the 5'-region of the UCP3 gene

Test subjects, fatty tissue samples from the skeletal muscle and abdominal cavity

A total of 63 unrelated patients, 38 morbidly obese people, 10 people who have overcome obesity and 15 non-obese people are examined. Tissue samples are taken from the Musculus rectus abdominis and in the abdominal cavity of the fatty tissue of the omentum of morbidly obese people in whom the stomach had been surgically bandaged for weight reduction. Control subjects and subjects who had overcome obesity were subjected to certain surgical interventions such as removal of the gall bladder, inguinal hernia operations and regulation or removal of the band around the stomach.

After some instruction, the test subjects had consented to the study, which had also been authorised by the Ethical Commission of the Institute. The subjects were anaesthetised on an empty stomach with fast acting barbiturates and kept unconscious with alfentanil-hydrochloride ((INN); IUPAC: N-\{1-\{2-(4-ethyl-5-oxo-2-tetrazolin-1-yl)ethyl\}-4-methoxy-methyl-4-piperidyl\}-propionanilide). Tissue samples were taken at the start of surgical intervention, divided into aliquots and frozen at -70°C. The ratio of the weight to the size was determined after measurements of the weight and size (Body mass index; BMI).

UCP-3 Gene structure

A λ-EMBL3-SP6/T7 genomic library from the human placenta (produced by: Clontech, Palo Alto, California, USA) was screened by plaque hybridisation using a probe which corresponded almost to the total UCP3-cDNA with a length of 1049 base pairs and included nucleotides +159 to +1207 (Gene Bank Accession Number U84763). Three overlapping clones with a size of about 16.5 kbp and with the complete
human UCP3 gene were isolated and subcloned into the ZERO-Background™ cloning system (Invitrogen, Carlsbad, California, USA). Sequencing of the plasmid DNA was carried out with ready assembled, commercially obtainable reagents (PRISM™ Ready Reaction dRhodamine Terminator Kit, manufactured by Perkin Elmer-Applied Biosystems, Foster City, California, USA) using dRhodamine Terminators and an ABI PRISM™ 310 DNA sequencing machine (Perkin Elmer-Applied Biosystems, Foster City, California, USA), synthesising successive suitable oligonucleotides in order to perform the sequencing reactions ("Migrating on the DNA", "Primer walking").

RNA isolation, experiments on rapid amplification of the 5' and 3' cDNA ends (Rapid amplification of cDNA ends; abbreviated to: RACE)

Total RNA was isolated from 0.5 g of human skeletal muscle tissue using the method of Chumczynski and Sacchi (1987). The integrity of the RNA samples was confirmed by analysing their electrophoretic flow properties in formaldehyde gels (Sambrook et al., 1989). RNA concentrations were determined by absorption measurements at 260 nm. 5' and 3' RACE were carried out as described by Frohman et al. (1988). For the 5'-RACE, 3 μg of total RNA from human skeletal muscle was reverse transcribed with Superscript™ II Reverse Transcriptase (GibcoBRL Life Technologies, Paisley, Great Britain) and a UCP3 gene-specific oligonucleotide from the intron 1 (5'-TACACCTGCT TGACGGAG-3'). After digestion with ribonuclease H (produced by Boehringer Mannheim Corp., Indianapolis, Indiana, USA) and lengthening of the polyA-tail (polyA-tailing) with terminal transferase (made by Boehringer Mannheim Corp., Indianapolis, Indiana, USA) the first strand cDNA was subjected to polymerase chain reaction (PCR) with 5'-GCTGTGTCCA GTGGAAAGGT AACGAGGTCA GCAA-3' as the gene specific oligonucleotide and 5'-GAGGACTCGA
GCTCAAGCT\textsubscript{(20)}-3' as the adapter oligonucleotide. The PCR-products were amplified again using 5'-GAGGACTCGA
GCTCAAGC-3' as the anchor oligonucleotide and 5'-TGGAGGACGTCTGAAG-3' as the internal (nested) gene
specific oligonucleotide. (For the 3' RACE, 3 μg of human skeletal muscle RNA were reverse transcribed using the
above-mentioned polyA-specific adapter-oligonucleotide.
The anchor oligonucleotide and two nested oligonucleotides
in the coding strand 5'-CCTCGACTGT ATGATAAAGA TG- 3' (+921
to +942) and 5'-CCTCCTGGGC CACCACCTT-3' (+ 937 to +955,
Gene Bank Accession Number U84763) were used for the
amplification). 5' and 3' RACE-PCR products were subcloned
into a common vector using the standard methods known to
those skilled in the art (Sambrook et al., 1989) and
sequenced after being gel-purified.

Test for protection of the UCP3-5'-transcript region from
degradation by ribonuclease (RNAse Protection Assay)
A probe covering the 425 bp long intron 1 (+6816 to +9074;
sequence SEQ ID NO 17) was prepared by reverse
transcription-PCR (RT-PCR) with the oligonucleotides 5'-
CCTCACCAGC CAGCCTCTTG TC-3' (+ 6816 to +6837) and 5'-
GCTGTGTCCA GTGAAAGGTA-3' (+9054 to +9074), in the coding
and non-coding regions. The PCR products were cloned into
the pZERC\textsuperscript{TM} plasmid vector and the sequence was verified by
sequencing with dye-coupled terminators (Dye-Terminator
cycle sequencing).

\(^{32}\text{P}\)-labelled counterstrand RNA was obtained with a combined
system which uses the SP6 and the T7 promoter (Riboprobe
Combination System SP6/ T7; manufactured by Promega Corp.,
Madison, Wisconsin, USA), and \(^{32}\text{P}\)-dUTP (3000 Ci/mmol;
manufactured by Amersham Life Science, Buckinghamshire,
Great Britain) in accordance with the manufacturer's
instructions. RNA transcribed in vitro was gel-purified
and the radioactivity incorporated was determined by liquid
scintillation counting (manufacturer: Wallac 1450 Microbeta PLUS, EG&G Berthold, Bad Wildbach, Germany). After denaturing at 95°C for 5 minutes, aliquots of the $^{32}$P-labelled RNA with an activity of $8 \times 10^4$ cpm (counts per minute), were hybridised overnight at 43°C with 5 μg of total RNA from human muscle or 15 μg of total RNA from fatty tissue. Unprotected RNA was digested at 37°C for 30 minutes in a total volume of 220 μl with 0.5 units of ribonuclease A and 20 units of ribonuclease T1 (Ambion RPAII Kit; Ambion Inc., Austin, Texas, USA). The ribonuclease inactivation/precipitation mixture supplied by the manufacturer was added in order to precipitate the $^{32}$P-labelled RNA-RNA hybrids. The precipitates were washed with 70% ethanol and protected fragments were separated by electrophoresis in 4% polyacrylamide-urea gels (Sambrook et al., 1989).

Identification of two transcription starting sites
The 5'-end of the UCP3-mRNA from the skeletal muscle was determined by 5'-RACE experiments using a counterstrand oligonucleotide from intron 1. Two specific PCR-fragments were obtained and characterised by sequencing. The products went 184 bp and 331 bp beyond the transcription start in the 5'-direction, which can be explained by two transcriptional start sites (Fig. 1A). The use of the most remote initiation site was demonstrated by RT-PCR (data not shown).

Ribonuclease protection tests (RNAse protection assays) showed that the majority of the transcripts were started in the muscle tissue of numerous obese and slender people at the downstream site at -184. The proportion of UCP3-mRNA which originates from the upstream site at -331 was less than 5%. However, the upstream site was predominantly, although not exclusively, used in the fatty tissue (Fig. 1B). Similar results were found in the fatty tissue of 10
obese and 4 slender individuals (data not shown). These experiments show that certain promoter regions can be used in the fatty tissue and in the skeletal muscle. Computer analysis of the sequences adjacent to the alternative starting sites showed that both regions contain eukaryotic consensus sequences for initiating transcription. Two TATA-like sequences (Locker and Buzard, 1990) are found 29 to 36 bp upstream and two conserved sites, on which a modified 7-methylguanosine group can be hooked onto the transcribed RNA (cap site; Locker and Buzard, 1990), are found 16 to 33 bp downstream from the corresponding starting site. By sequence comparison, fundamental promoter elements such as CAAT boxes at -749, -861 and -922 were found, separated from the transcription starting site by an Alu sequence, relative to the transcription starting site located downstream, an octamer motif was found at -713 and a conserved binding site for an upstream activating factor (USF) was found at -1038 (Fig. 1A). In addition, three successive E-box elements were found at -653, -640 and -631, a putative peroxisome-proliferator-activated receptor/retinoid-X-receptor affected element (PPAR/RXR) was found at -265 and a putative thyroid response element (TRE) (Locker and Buzard, 1990) was found at position -3340, located in an Alu-sequence from -3396 to -3094.

**Example 2 - Screening test for finding substances which modulate the transcription of the UCP3-promoter**

The average skilled person carries out such a test by using standard methods and ready-assembled reagents such as those sold by Promega (Madison, Wisconsin, USA) as the `Luciferase Assay System`. With this system he also obtains instructions as to how the test should be carried out with the commercially obtainable reagents. However, he can also refer to the the teaching of Example 1 of US Patent 5,283,179 in which the procedure of the test is described.
By a combination of suitable cloning techniques and PCR-techniques the sequence SEQ ID NO 16 is prepared (Sambrook et al., 1989), which no longer contains the muscle cell-specific UCP3-promoter. This DNA sequence is inserted by standard methods through suitable restriction cutting sites into the vector pGL2basic or pGL3basic (Promega, Madison, Wisconsin, USA). This vector is transfected by standard methods (Sambrook et al., 1989) into primary human adipocyte cell cultures obtained from obese persons as described in Example 1 who had undergone surgical treatment. This cell culture is grown under standard conditions in standard cell culture medium in a cell culture dish 100 mm in diameter. The cells are then cultivated in the presence of an aqueous solution or a solution in dimethylsulphoxide of a test substance for a certain length of time. At ambient temperature, 1 ml of buffer which lyses the cells (25 mM Tris-phosphate pH 7.8, 2 mM dithiothreitol, 10 % glycerol, 1 % Triton® X-100, 1 mg/ml bovine serum albumin, 2 mM cyclohexylene diamine tetraacetate (CDTA)), is added for about 2 to 5 minutes and the cell debris is eliminated by brief centrifugation. Optionally, the growth medium can previously be removed from the cultivated cells and the cells may be washed once or twice with PBS buffer (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.4 mM KH₂PO₄, pH 7.3; Mg²⁺- and Ca²⁺-free). Part of the supernatant is mixed with five times its volume of test reagent (20 mM tricine (N-tris-(hydroxymethyl)-methylglycine) pH 7.8, 33.3 mM dithiothreitol (DTT), 8 mM Mg²⁺, 0.13 mM ethylenediamine tetraacetic acid (EDTA), 0.53 mM adenosine triphosphate (ATP), 0.47 mM luciferin (a polyheterocyclic organic acid: D-(-)-2-(6′-hydroxy-2′-benzothiazolyl)+A²-thiazolin-4-carboxylic acid) and 0.27 mM of coenzyme A (CoA)) and the light emitted is measured in a luminometer. The volume of the test reagent may be up to 25 times the volume of the solution to be tested. Then the
signal is compared with the signal from a control (aqueous solution or dimethylsulphoxide solution without the test substance).

In order to find substances which influence the transcription of the muscle cell-specific UCP3-promoter, the skilled person can prepare a DNA molecule which no longer contains the fat cell-specific UCP3-promoter. To do this, he removes nucleotides 6739 to 6795 which contain the fat cell-specific part, from the DNA sequence SEQ ID NO: 14 using standard methods (Sambrook et al., 1989) and clones them into the vectors mentioned above. Moreover, he proceeds as described above except that he uses muscle cells in primary culture or suitable cell lines.

General activators of the UCP3-promoter are found by the same experimental procedure, and sequence SEQ ID NO: 14 which contains both promoters can be used. Here again, cells are used in which there is no or only very slight UCP3 expression, such as for example the cell line HEK293.

The skilled person can adapt a test of this kind to the particular circumstances of the microtitre plate format without any inventive step.
Figures:
Fig. 1: Characterisation of the UCP3 promoter region and evidence for various transcription initiation sites in the skeletal muscle and in the fatty tissue

A) Part of the nucleotide sequence of the human UCP3 gene; basic promoter elements such as e.g. CAAT box, octamer motifs and upstream binding sites for stimulation factor as well as putative transcription factor binding sites such as E boxes and a PPAR/RXR response element are emphasised by small frames. Half an Alu sequence is emphasised by a large frame. TATA-like sequences are printed in bold and underlined, conserved sites onto which a modified 7-methylguanosine group is attached to the transcribed RNA (known as cap sites), are printed in bold italics. The two transcription starting sites are indicated by arrows and the translation start codon is printed in bold. Part of the sequence of intron 1 is shown in lower-case letters.

B) Autoradiogram of the ribonuclease protection test (known as the RNAse protection assay) of RNA from human skeletal muscle and fatty tissue. The sizes of the protected fragments, calculated for transcripts initiated at the upstream or downstream site, are 425 and 316 nucleotides, respectively, as indicated on the right hand side.

Examining the gel:
M: end labelled size markers with the specified sizes

1: undigested $^{32}$P-UCP3 counterstrand probe of nucleotide 6816 to 9074 (sequence SEQ ID NO 17)

2: $^{32}$P-UCP3-counterstrand probe hybridised with yeast RNA and subjected to ribonuclease digestion

3-5: Ribonuclease protection assay of 5 µg of total RNA from the skeletal muscle of three individuals

6-7: Ribonuclease protection assay of 15 µg of total RNA from intraperitoneal fatty tissue from two individuals
Literature:
Claims:

1. Recombinant DNA molecule which contains the DNA sequence of the UCP-3 promoter active in fat cells but does not contain a functional sequence of the UCP-3 promoter active in muscle cells.

2. Recombinant DNA molecule according to claim 1, which contains the sequences SEQ ID NO: 1 TATATTTAAA and SEQ ID NO: 2 CACCTC but does not contain the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC.

3. Recombinant DNA molecule according to claim 1, which contains the sequences SEQ ID NO: 1 TATATTTAAA and SEQ ID NO: 2 CACCTC and derivatives of the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC formed by point mutations, as a result of which the sequences SEQ ID NO: 3 and SEQ ID NO: 4 lose their function in the transcription.

4. Recombinant DNA molecule according to claim 1, which contains the sequences SEQ ID NO: 1 TATATTTAAA and SEQ ID NO: 2 CACCTC and derivatives of the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC formed by deletions, as a result of which the sequences SEQ ID NO: 3 and SEQ ID NO: 4 lose their function in the transcription.

5. Recombinant DNA molecule according to claim 1, which contains the sequences SEQ ID NO: 1 TATATTTAAA and SEQ ID NO: 2 CACCTC and derivatives of the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC formed by combinations of point mutations and deletions, as a result of which the sequences SEQ ID NO: 3 and SEQ ID NO: 4 lose their function in the transcription.

6. Recombinant DNA molecule according to one of claims 2 to 5, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 2 there are 30 to 50 base pairs, preferably 40 to 45 base pairs, most preferably 42 base pairs.
7. Recombinant DNA molecule according to one of claims 2 to 6, characterised in that upstream of SEQ ID NO: 1 there is additionally an RXR/PPAR element.

8. Recombinant DNA molecule according to claim 7, characterised in that the RXR/PPAR element includes the sequence SEQ ID NO: 5 TGACCTTTGGACT.

9. Recombinant DNA molecule according to claim 8, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 5 there are 65 to 75 base pairs.

10. Recombinant DNA molecule according to one of claims 2 to 9, characterised in that upstream of SEQ ID NO: 1 there is additionally an Alu sequence.

11. Recombinant DNA molecule according to claim 10, characterised in that between the sequence SEQ ID NO: 1 and the Alu sequence there are 255 to 265 base pairs.

12. Recombinant DNA molecule according to one of claims 2 to 11, characterised in that upstream of SEQ ID NO: 1 there is additionally an E2A element.

13. Recombinant DNA molecule according to claim 12, characterised in that the E2A element includes the sequence SEQ ID NO: 6 CAGATG.

14. Recombinant DNA molecule according to claim 13, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 6 there are 440 to 450 base pairs.

15. Recombinant DNA molecule according to one of claims 2 to 14, characterised in that upstream of SEQ ID NO: 1 there is additionally an E box.

16. Recombinant DNA molecule according to claim 15, characterised in that the E box includes the sequence SEQ ID NO: 7 CACTTG.

17. Recombinant DNA molecule according to claim 16, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 7 there are 450 to 460 base pairs.
18. Recombinant DNA molecule according to one of claims 2 to 17, characterised in that upstream of SEQ ID NO: 1 there is additionally an E box.

19. Recombinant DNA molecule according to claim 18, characterised in that the E box includes the sequence SEQ ID NO: 8 CATT TG.

20. Recombinant DNA molecule according to claim 19, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 8 there are 460 to 470 base pairs.

21. Recombinant DNA molecule according to one of claims 2 to 20, characterised in that upstream of SEQ ID NO: 1 there is additionally an octamer sequence.

22. Recombinant DNA molecule according to claim 21, characterised in that the octamer sequence includes the sequence SEQ ID NO: 9 ATGAAAATGT.

23. Recombinant DNA molecule according to claim 22, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 9 there are 515 to 525 base pairs.

24. Recombinant DNA molecule according to one of claims 2 to 23, characterised in that upstream of SEQ ID NO: 1 there is additionally a CAAT box.

25. Recombinant DNA molecule according to claim 24, characterised in that the CAAT box includes the sequence SEQ ID NO: 10 CCAAT.

26. Recombinant DNA molecule according to claim 25, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 10 there are 560 to 570 base pairs.

27. Recombinant DNA molecule according to one of claims 2 to 26, characterised in that upstream of SEQ ID NO: 1 there is additionally a CAAT box.

28. Recombinant DNA molecule according to claim 27, characterised in that the CAAT box includes the sequence SEQ ID NO: 11 ATTGG.

29. Recombinant DNA molecule according to claim 28, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 11 there are 670 to 680 base pairs.
30. Recombinant DNA molecule according to one of claims 2 to 29, characterised in that upstream of SEQ ID NO: 1 there is additionally a CAAT box.

31. Recombinant DNA molecule according to claim 30, characterised in that the CAAT box includes the sequence SEQ ID NO: 12 ATTTGG.

32. Recombinant DNA molecule according to claim 31, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 12 there are 730 to 740 base pairs.

33. Recombinant DNA molecule according to one of claims 2 to 32, characterised in that upstream of SEQ ID NO: 1 there is additionally a binding site for an upstream binding stimulating factor.

34. Recombinant DNA molecule according to claim 33, characterised in that the binding site for an upstream binding stimulating factor includes the sequence SEQ ID NO: 13 CCACGTGC.

35. Recombinant DNA molecule according to claim 34, characterised in that between the sequences SEQ ID NO: 1 and SEQ ID NO: 13 there are 845 to 855 base pairs.

36. Recombinant DNA molecule which contains the DNA sequence of the UCP-3 promoter active in muscle cells but does not contain a functional sequence of the UCP-3 promoter active in fat cells.

37. Recombinant DNA molecule according to claim 36, which contains the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC but does not contain the sequences SEQ ID NO: 1 TATATTTAAA and SEQ ID NO: 2 CACCTC.

38. Recombinant DNA molecule according to claim 37, which contains the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC and derivatives of the sequences SEQ ID NO: 1 TATATTTAAA and SEQ ID NO: 2 CACCTC formed by point mutations, as a result of which the sequences SEQ ID NO: 1 and SEQ ID NO: 2 lose their function in the transcription.
39. Recombinant DNA molecule according to claim 37, which contains the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC and derivatives of the sequences SEQ ID NO: 1 TATATTTAAA and SEQ ID NO: 2 CACCTC formed by deletions, as a result of which the sequences SEQ ID NO: 1 and SEQ ID NO: 2 lose their function in the transcription.

40. Recombinant DNA molecule according to claim 37, which contains the sequences SEQ ID NO: 3 TATAAGA and SEQ ID NO: 4 CAATCC and derivatives of the sequences SEQ ID NO: 1 TATATTTAAA and SEQ ID NO: 2 CACCTC formed by combinations of point mutations and deletions, as a result of which the sequences SEQ ID NO: 1 and SEQ ID NO: 2 lose their function in the transcription.

41. Recombinant DNA molecule according to one of claims 37 to 40, characterised in that between the sequences SEQ ID NO: 3 and SEQ ID NO: 4 there are 45 to 70 base pairs, preferably 52 to 58 base pairs, most preferably 55 base pairs.

42. Recombinant DNA molecule according to one of claims 37 to 41, characterised in that upstream of SEQ ID NO: 3 there is additionally an RXR/PPAR element.

43. Recombinant DNA molecule according to claim 42, characterised in that the RXR/PPAR element includes the sequence SEQ ID NO: 5 TGACCTTTGGACT.

44. Recombinant DNA molecule according to one of claims 37 to 43, characterised in that upstream of SEQ ID NO: 3 there is additionally an Alu sequence.

45. Recombinant DNA molecule according to one of claims 37 to 44, characterised in that upstream of SEQ ID NO: 3 there is additionally an E2A element.

46. Recombinant DNA molecule according to claim 45, characterised in that the E2A element includes the sequence SEQ ID NO: 6 CAGATG.
47. Recombinant DNA molecule according to one of claims 37 to 46, characterised in that upstream of SEQ ID NO: 3 there is additionally an E box.

48. Recombinant DNA molecule according to claim 47, characterised in that the E box includes the sequence SEQ ID NO: 7 CACTTG.

49. Recombinant DNA molecule according to one of claims 37 to 48, characterised in that upstream of SEQ ID NO: 3 there is additionally an E box.

50. Recombinant DNA molecule according to claim 49, characterised in that the E box includes the sequence SEQ ID NO: 8 CATT TG.

51. Recombinant DNA molecule according to one of claims 37 to 50, characterised in that upstream of SEQ ID NO: 3 there is additionally an octamer sequence.

52. Recombinant DNA molecule according to claim 51, characterised in that the octamer sequence includes the sequence SEQ ID NO: 9 ATGAAAATGT.

53. Recombinant DNA molecule according to one of claims 37 to 52, characterised in that upstream of SEQ ID NO: 3 there is additionally a CAAT box.

54. Recombinant DNA molecule according to claim 53, characterised in that the CAAT box includes the sequence SEQ ID NO: 10 CCAAT.

55. Recombinant DNA molecule according to one of claims 37 to 54, characterised in that upstream of SEQ ID NO: 3 there is additionally a CAAT box.

56. Recombinant DNA molecule according to claim 55, characterised in that the CAAT box includes the sequence SEQ ID NO: 11 AT TGG.

57. Recombinant DNA molecule according to one of claims 37 to 56, characterised in that upstream of SEQ ID NO: 3 there is additionally a CAAT box.

58. Recombinant DNA molecule according to claim 57, characterised in that the CAAT box includes the sequence SEQ ID NO: 12 AT TGG.
59. Recombinant DNA molecule according to one of claims 37 to 58, characterised in that upstream of SEQ ID NO: 3 there is additionally a binding site for an upstream binding stimulating factor.

60. Recombinant DNA molecule according to claim 59, characterised in that the binding site for an upstream binding stimulating factor includes the sequence SEQ ID NO: 13 CCACGTGC.

61. Recombinant DNA molecule which contains the sequence SEQ ID NO: 14.

62. Recombinant DNA molecule which contains a sequence which hybridises with the sequence SEQ ID NO 14 under stringent conditions.

63. Recombinant DNA molecule which contains a functional derivative of a DNA molecule according to one of claims 1 to 62, characterised in that these functional derivatives have the ability to mediate the transcription of a gene.

64. Recombinant DNA molecule according to claim 63, characterised in that the functional derivatives have been produced by deletions.

65. Recombinant DNA molecule according to claim 63, characterised in that the functional derivatives have been produced by point mutations.

66. Recombinant DNA molecule according to claim 63, characterised in that the functional derivatives have been produced by combinations of point mutations and deletions.

67. Recombinant DNA molecule according to one of claims 1 to 66, characterised in that the DNA molecule contains an additional gene which is functionally linked.

68. Recombinant DNA molecule according to claim 67, characterised in that the additional gene is the gene of the uncoupling protein 3.
69. Recombinant DNA molecule according to claim 67, characterised in that the additional gene is a reporter gene.

70. Recombinant DNA molecule according to claim 69, characterised in that the reporter gene is the luciferase gene or green fluorescent protein.

71. Recombinant DNA molecule which contains the sequence SEQ ID NO 15.

72. Recombinant DNA molecule which contains a sequence which hybridises with the sequence SEQ ID NO 15 under stringent conditions.

73. A cell which contains a recombinant DNA molecule according to one of claims 1 to 72.

74. Use of a recombinant DNA molecule according to one of claims 1 to 72 for the transcription of a gene.

75. Use of a recombinant DNA molecule according to one of claims 1 to 72 for finding substances which are able to influence transcription.

76. Use of a cell according to claim 73 for finding substances which are able to influence transcription.

77. Process for finding substances which are able to influence transcription, characterised in that a recombinant DNA molecule according to one of claims 1 to 72 or a cell according to claim 73 is used.

78. Process according to claim 77, characterised in that it is a cell-free or cell-based process.

79. Process according to claim 78, characterised in that the transcription rate is measured in the presence of a test substance.

80. Process according to claim 79, characterised in that the transcription rate is measured in the presence of a test substance in a cell-free in-vitro system which contains at least cell extract, ribonucleotides and a recombinant DNA molecule according to one of claims 1 to 72.
81. Process according to claim 79, characterised in that the quantity, activity, luminescence or fluorescence of a reporter protein produced in the presence of a test substance is measured.

82. Process according to claim 81, characterised in that in the presence of a test substance RNA is transcribed in vitro, then translated in vitro and the amount of reporter protein produced is determined.

83. Process according to claim 81, characterised in that the amount of a reporter protein produced by a cell according to claim 73 in the presence of a test substance is measured.

84. Process according to one of claims 81 to 83, characterised in that the reporter protein is luciferase or green fluorescent protein.

85. Process for finding a modulator of the UCP3 promoter, characterised in that
   a) a host cell according to claim 73 is cultivated in the presence of a test substance,
   b) the transcription rate of the UCP3 promoter is measured,
   c) the transcription rate thus obtained is compared with the transcription rate which was obtained in the absence of the test substance.

86. Process according to one of claims 77 to 85, characterised in that it is carried out in a High Throughput Screening (HTS) format.

87. Use of a recombinant DNA molecule according to one of claims 1 to 72 for finding factors or substances which bind to the DNA molecule.

88. Use according to claim 87 characterised in that the recombinant DNA molecule is bound to a carrier, contacted with a mixture of various substances or factors, subjected to one or more washing steps and the bound substances or factors are identified.
89. Process for identifying substances or factors which bind to a recombinant DNA molecule according to one of claims 1 to 72, characterised in that:
- the recombinant DNA molecules are bound to a carrier,
- contacted with a mixture of various substances,
- subjected to one or more washing steps,
- the bound factors are identified.
SEQUENCE LISTING

<110> Boehringer Ingelheim International GmbH
<120> Cell specific promoters of uncoupling protein 3
<130> 1-1061
<140> 
<141> 
<150> DE19838837.3
<151> 1998-08-27
<160> 17
<170> PatentIn Ver. 2.1

<210> 1
<211> 9
<212> DNA
<213> Homo sapiens

<220>
<221> TATA_signal
<222> (1)..<(9)

<400> 1
tatatta

<210> 2
<211> 6
<212> DNA
<213> Homo sapiens

<220>
<221> CAAT_signal
<222> (1)..<(6)

<400> 2
cacctc

<210> 3
<211> 7
<212> DNA
<213> Homo sapiens

<220>
<221> TATA_signal
<222> (1)..<(7)

<400> 3
tataga

<210> 4
<211> 6
<212> DNA
<213> Homo sapiens

<220>
<221> CAAT_signal
<222> (1)..<(6)

<400> 4
caatc
<210> 5
<211> 13
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(13)
<223> RXR/PPAR-Element

<400> 5
tgaccttttgtg act

<210> 6
<211> 6
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(6)
<223> Z2A-Element

<400> 6
cagatgt

<210> 7
<211> 6
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(6)
<223> E-Box

<400> 7
cactgtg

<210> 8
<211> 6
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(6)
<223> E-Box

<400> 8
catttg

<210> 9
<211> 10
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(10)
<223> Octamer
<400> 9
atgaaaatgt

<210> 10
<211> 5
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(5)
<223> CAAT-Box

<400> 10
ccaat

<210> 11
<211> 5
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(5)
<223> CAAT-Box

<400> 11
atgga

<210> 12
<211> 5
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(5)
<223> CAAT-Box

<400> 12
atgga

<210> 13
<211> 6
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)..<(6)
<223> USF-Element

<400> 13
ccaagtgc

<210> 14
<211> 6959
<212> DNA
<213> Homo sapiens

<220>
<221> promoter
<222> (1)..<(6959)
tttcctgtaag ttcaagtttt tgcctctggtt tgctgtgatt ggaagtgaatt aaccacacca 4020
gaaatcctcag agatctcctagt gacgcttcgct tttcaagatc ccaaacaagt tcaacccgaat 4080
aattctctctt tcattctactt tattttctttt ttcttttttt ttcttttttt ttttttttctag 4140
gagctgcctta gagaagacag cctcggcttg gacacagag aataaccccg ggcttgtagcc 4200
agctttagag agagaccaag gaaaggccct gacacagttg cacgctggaga tagagaccaag 4260
tctcactcgcc tttgcagcctc ttcgttcgctc tctgctaaaa agagatatac aacccaggtct 4320
catggagtttg tgcaggaagat ccaaattgag gacatgcttg aaaaagttct gcctcattttt 4380
ggcaagcccgt gagaagaggg tgaagagttc teeacgtcaaa cccccacccac ccctctctct 4440
dagcaagtctg gaggctcgac cttaccaagat cccctctctc ccccccccta cccaccccatc 4500
acggctcagaa gagctgctgtt gctgctcattt gcctgcctca ccctgctctc ccctgctctc 4560
cagagggagt cttcttctatc cccctcctaaa ccctctctct ccctctctct ccctctctct 4620
tccttcttctt tttctctcttt tttctctttt cccctctctt ctcttctctt ctcttctctt 4680
tctgagggagt cttcttcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 4740
ggtatgtgat gttggctttgt gctgctctttt gctgctctctt gctgctctctt gctgctctctt 4800
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 4860
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 4920
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 4980
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5040
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5100
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5160
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5220
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5280
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5340
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5400
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5460
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5520
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5580
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5640
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5700
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5760
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5820
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5880
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 5940
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6000
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6060
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6120
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6180
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6240
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6300
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6360
tgaggtgagtt gttggctttgt gctgctctttt gctgctctctt gctgctctctt gctgctctctt 6420
ggctctctct cccttctcttt cccttctcttt cccttctcttt cccttctcttt cccttctcttt 6480
gagctgctgtt gttggctttgt gctgctctttt gctgctctctt gctgctctctt gctgctctctt 6540
ttcacattact tatttcacatt tatttcacatt tatttcacatt tatttcacatt tatttcacatt 6600
tgttattacc cctttaaggtt ctcatcttttt ctctctctct ccctctctct ccctctctct 6660
ttcacattact tatttcacatt tatttcacatt tatttcacatt tatttcacatt tatttcacatt 6720
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6780
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6840
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6900
tccttctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt tttctctcttt 6960
<210> 15
<211> DNA
<212> Homo sapiens

<220>
<221> promoter
<222> (1..(3952)

<400> 15
aatctttggg ttttttgcgc tttttttttt tttttttttt tttttttttt ttttttttttt 60
tgcacccct caactccttg acttcttttt ttcttttttt ttcttttttt ttctttttttt 120
tcttgtgtc ccaactccttg aactacactcg gatctcctctc cctcggcttg gacacagag 180
ccaaatgtcg ggggttacga cggacgctctc cccacccgagac gggcttgtagcc 240
ttttctgtaa cagatgtgatt ggtcctgtaaa ccaacacatc gtcaccaggggt ggcctgtaaa 300
ttataacag agaatccgtgta tttttctcttt ttttttctttt ttttttttttt ttttttttttt 360
<221> promoter
<222> (1) ..(6795)

<400> 16
aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 6
tgacacacta cacccctcctta ctatgttgcgc caacagctgg ctccaaagctt ggggtttcaac 12
ttacagtgagt ccacagctgta acaacactct gcactgctagc acagctgcga 18
tttctagagt aagacatgaa tttccctttctttt ccactccatt ctgttggcttc aagtctgctg 24
ttttccaatg aagacatgaa tttccctttctttt ccactccatt ctgttggcttc aagtctgctg 30
ttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 36
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 42
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 48
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 54
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 60
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 66
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 72
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 78
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 84
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 90
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 96
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 102
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 108
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 114
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 120
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 126
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 132
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 138
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 144
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 150
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 156
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 162
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 168
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 174
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 180
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 186
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 192
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 198
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 204
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 210
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 216
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 222
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 228
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 234
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 240
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 246
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 252
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 258
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 264
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 270
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 276
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 282
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 288
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 294
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 300
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 306
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 312
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 318
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 324
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 330
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 336
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 342
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 348
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 354
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 360
cttcctcaacag aatcttcggg catcaacgcgt ctctccaacca cagccctctcg agtctctcggg accacagctg 366
aattcttggg cttcgaagct cacacccaca cagccttcctg tagtagctggg acccacaccgt 60
tgccacacta caccctctca atttttttttt tctcttttttt tttctgtcag caggtttcc 120
tctatggtgc cagctctcaata acccctcaag tgggcgggca cagcacttggtggtgagact 180
ttttgcataat accaatcactt ctccttttcct ctcctctctcct ccagagttgaca gggtttgagag 240
ttctattttt tacagaagagat tcagcactgc ctgccttcctac atcctgggtatc 300
tcttttttttt tcacccaccc cccgctcttcct ttggtggcatcg ccgctttct 360
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C12N15/63 C1201/68 C1201/66 C1201/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>DATABASE EMBL/GENBANK/DDBJ 'Online!' 11 March 1997 (1997-03-11) XP002126138 Acc. No. HS1159438</td>
<td>62,63, 72,73</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of box C.

* Special categories of cited documents:
  *A* document defining the general state of the art which is not considered to be of particular relevance
  *E* earlier document but published on or after the international filing date
  *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another document or to cite other special reason (as specified)
  *O* document relating to an oral disclosure, use, exhibition or other means
  *P* document published prior to the international filing date but later than the priority data claimed

Date of the actual completion of the international search: 21 December 1999

Date of mailing of the international search report: 11/01/2000

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tdl. (+31-70) 940-8040, Fax (+31-70) 940-8018

Authorized officer
Mata Vicente, T.
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>WO 98 31396 A (RICOUER DANIEL ;BOUILLAUD FREDERIC (FR); RECH SCIENT CENTRE DE RE) 23 July 1998 (1998-07-23) page 14, line 3 - line 9 page 18, line 1 - line 4 page 18, line 24 - page 19, line 26 example XI</td>
<td>1-89</td>
</tr>
</tbody>
</table>
## INTERNATIONAL SEARCH REPORT

**Information on patent family members**

<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AU 9798298 A</td>
<td>03-05-1999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>US 5976808 A</td>
<td>02-11-1999</td>
</tr>
</tbody>
</table>