

(11) Número de Publicação: **PT 1259606 E**

(51) Classificação Internacional:

**C12N 15/12** (2007.10) **C12N 5/10** (2007.10)  
**C07K 14/47** (2007.10) **C07K 16/18** (2007.10)  
**G01N 33/50** (2007.10) **A61K 38/17** (2007.10)

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **2001.02.15**

(30) Prioridade(s): **2000.02.17 FR 0001980**  
**2000.04.18 US 198489 P**

(43) Data de publicação do pedido: **2002.11.27**

(45) Data e BPI da concessão: **2010.04.28**  
**147/2010**

(73) Titular(es):

**AVENTIS PHARMA S.A.**  
**20, AVENUE RAYMOND ARON 92160 ANTONY** FR  
**INSTITUT NATIONAL DE LA SANTE ET DE LA**  
**RECHERCHE MEDICALE (INSERM)** FR

(72) Inventor(es):

**ALAIN FOURNIER** FR  
**ALEXIS BRICE** FR  
**LAURENT PRADIER** FR  
**HANA KOUTNIKOVA** FR  
**CATHERINE PRADES** FR

(74) Mandatário:

**ELSA MARIA MARTINS BARREIROS AMARAL CANHÃO**  
**RUA DO PATROCÍNIO 94 1399-019 LISBOA** PT

(54) Epígrafe: **COMPOSIÇÕES UTILIZÁVEIS PARA REGULAR A ACTIVIDADE DA PARKINA**

(57) Resumo:

## RESUMO

### **"COMPOSIÇÕES UTILIZÁVEIS PARA REGULAR A ACTIVIDADE DA PARKINA"**

A presente invenção refere-se a novos compostos e às suas utilizações, nomeadamente farmacêuticas, de diagnóstico ou como alvos farmacológicos. Mais particularmente, a presente invenção refere-se a uma nova proteína, designada PAP1, assim como a novos péptidos e compostos capazes de modular, pelo menos parcialmente, a actividade da Parkina.

## DESCRIÇÃO

### “COMPOSIÇÕES UTILIZÁVEIS PARA REGULAR A ACTIVIDADE DA PARKINA”

A presente invenção refere-se a composições e a métodos utilizáveis para regular a actividade da parkina. Esta refere-se nomeadamente, a uma nova proteína, designada PAP1, parceira da parkina, assim como aos péptidos ou polipéptidos derivados ou homólogos desta proteína. Esta refere-se igualmente a compostos capazes de modular, pelo menos parcialmente, a actividade da Parkina, nomeadamente de interferir com a interacção entre a parkina e a PAP1. A presente invenção é utilizável nos domínios terapêuticos, diagnóstico ou para a constituição de alvos farmacológicos permitindo o desenvolvimento de novos medicamentos.

O gene da Parkina está mutado em determinadas formas familiares (juvenis autossómicas recessivas) da doença de Parkinson (Kitada *et al.*, 1998). A doença de Parkinson (Lewy, 1912) é uma das doenças neurodegenerativas mais comuns, afectando mais de 1% da população com mais de 55 anos. Os doentes que sofrem desta doença apresentam distúrbios neurológicos agrupados sob o termo de síndrome de Parkinson, caracterizados por rigidez, bradicinesia e tremor em repouso. Estes sintomas são a consequência de uma degenerescência dos neurónios dopaminérgicos da substância negra do cérebro.

A maior parte dos casos de doença de Parkinson não tem historial familiar. No entanto, existem casos familiares em que

alguns correspondem a uma forma monogénica da doença. Presentemente, foram apenas identificados três genes diferentes em algumas formas hereditárias raras. A primeira forma corresponde a uma forma autossómica dominante, na qual o gene responsável codifica para a Sinucleína alfa (Polymeropoulos *et al.*, 1997). Esta proteína é um constituinte abundante das inclusões intracitoplasmáticas, designadas corpos de Lewy, que servem como marcadores da doença de Parkinson (Lewy, 1912). A segunda forma, igualmente autossómica dominante, está relacionada com uma mutação num gene codificante para uma hidrolase designada hidrolase do terminal carboxilo da ubiquitina L1 (Leroy *et al.*, 1998). Esta enzima é suposta hidrolisar os polímeros ou os conjugados de ubiquitinas em monómeros de ubiquitina. A terceira forma distingue-se das anteriores por uma transmissão autossómica recessiva e um aparecimento frequentemente antes dos 40 anos assim como por uma ausência de corpos de Lewy. Estes doentes respondem mais favoravelmente à levodopa, um precursor da dopamina utilizado como tratamento da doença de Parkinson. O gene relacionado com esta forma codifica para uma nova proteína designada Parkina (Kitada *et al.*, 1998).

O gene da Parkina é constituído por 12 exões abrangendo uma região genómica de mais de 500000 pares de bases no cromossoma 6 (6q25.2-q27). Presentemente, são conhecidos dois grandes tipos de mutação deste gene, na origem da doença, quer eliminações de tamanho variável na região que abrange os exões 2 a 9, quer mutações pontuais que produzem o aparecimento prematuro de um codão de terminação ou a modificação de um aminoácido (Kitada *et al.*, 1998; Abbas *et al.*, 1999; Lucking *et al.*, 1998; Hattori *et al.*, 1998). A natureza destas mutações e o modo de

transmissão autossômica recessiva sugerem uma perda da função da Parkina conduzindo a doença de Parkinson.

Este gene é expresso num grande número de tecidos e nomeadamente na substância negra. Existem vários transcritos correspondentes a este gene que provêm de processamentos alternativos diferentes (Kitada *et al.*, 1998; Sunada *et al.*, 1998). No cérebro, encontram-se dois tipos de ARN mensageiros em que um é desprovido da parte correspondente ao exão 5. Foram identificados, nos leucócitos, ARN mensageiros da Parkina não contendo a região codificante para os exões 3, 4 e 5. Os ARN mensageiros da Parkina mais longos, presentes no cérebro, contêm 2960 bases e codificam para uma proteína com 465 aminoácidos.

Esta proteína tem uma homologia baixa na sua parte do terminal N com a ubiquitina. A sua parte do terminal C contém dois motivos "ring finger" separados por um domínio IBR (In Between Ring) correspondente a uma região rica em cisteína que se pode ligar a metais como os domínios "ring finger" (Morett, 1999). Foi demonstrado por imunocitoquímica que a parkina está localizada no citoplasma e no aparelho de Golgi de neurónios da substância negra que contêm a melanina (Shimura *et al.*, 1999). Além disso, esta proteína está presente em determinados corpos de Lewy de doentes com Parkinson. A função celular da Parkina não está ainda demonstrada, mas esta pode desempenhar um papel de transportador nas vesículas sinápticas, na maturação ou na degradação de proteínas e no controlo do crescimento, da diferenciação ou do desenvolvimento celular. A Parkina está ausente nas formas autossômicas recessivas juvenis, confirmando-se, assim, que a perda desta função é responsável pela doença.

A elucidação do papel exacto da proteína Parkina no processo da degenerescência dos neurónios dopaminérgicos constitui, conseqüentemente, uma vantagem significativa para a compreensão e a abordagem terapêutica da doença de Parkinson e mais geralmente das doenças do sistema nervoso central.

A presente invenção baseia-se na identificação de um parceiro da parkina que interaja com esta proteína em condições fisiológicas. Este parceiro representa um novo alvo farmacológico para a produção ou a investigação de compostos capazes de modular a actividade da parkina, nomeadamente a sua actividade na degenerescência dos neurónios dopaminérgicos e/ou no desenvolvimento de patologias nervosas. Esta proteína, os anticorpos, os ácidos nucleicos correspondentes, assim como as sondas ou os iniciadores específicos, são igualmente utilizáveis na detecção ou no doseamento de proteínas em amostras biológicas, nomeadamente em amostras de tecido nervoso. Estas proteínas ou ácidos nucleicos são também utilizáveis em abordagens terapêuticas, para modular a actividade da parkina, assim como qualquer composto de acordo com a invenção, capaz de modular a interacção entre a parkina e os polipéptidos da invenção.

A presente invenção resulta mais particularmente, da identificação pela requerente de uma nova proteína humana, designada PAP1 (Parkine Associated Protein 1) ou LY111, que interage com a parkina. A proteína PAP1 (sequência SEQ ID N°: 1 ou 2) apresenta alguma homologia com as sinaptotagminas e é capaz de interagir, mais particularmente, com a região central da parkina (representada na sequência SEQ ID N°: 3 ou 4). Foi igualmente clonada, sequenciada e caracterizada a proteína PAP1 a partir de diferentes tecidos de origem humana, nomeadamente do

pulmão (SEQ ID N°: 12, 13) e do cérebro (SEQ ID N°: 42, 43), assim como formas curtas, correspondente às variantes de processamento (SEQ ID N°: 14, 15, 44, 45).

A presente invenção resulta, igualmente, da identificação e da caracterização de regiões particulares da proteína PAP1, envolvidas na modulação da função da parkina. A detecção da existência desta proteína e de regiões envolvidas na sua função permite nomeadamente preparar novos compostos e/ou composições utilizáveis como agentes farmacêuticos e desenvolver métodos industriais de rastreio desses compostos.

Um primeiro objecto da invenção refere-se a um polipéptido que interage com a proteína apresentando, em todo o seu comprimento, uma identidade de, pelo menos, 95% com a sequência SEQ ID N°: 2.

Um outro aspecto da invenção reside num ácido nucleico codificando a proteína PAP1, aos seus fragmentos, derivados ou homólogos, assim como qualquer vector compreendendo esse ácido nucleico, qualquer célula recombinante contendo esse ácido nucleico ou vector e qualquer mamífero não humano compreendendo, nas suas células, esse ácido nucleico.

A invenção refere-se também a anticorpos capazes de se ligarem à proteína PAP1, aos seus fragmentos, derivados e homólogos, nomeadamente anticorpos policlonais ou monoclonais, de um modo mais preferido, anticorpos capazes de se ligarem à proteína PAP1 e de inibirem, pelo menos parcialmente, a sua interacção com a parkina.

Um outro aspecto da invenção refere-se a sondas ou a iniciadores nucleotídicos, específicos da PAP1, utilizáveis para detectar ou amplificar o gene de PAP1 ou uma região deste em qualquer amostra biológica.

A invenção refere-se ainda a composições farmacêuticas, a métodos de detecção de anomalias genéticas, a métodos de produção de polipéptidos, tal como aqui anteriormente definidos, assim como a métodos de separação ou de caracterização de compostos activos.

Dentro do âmbito da presente invenção, a denominação proteína PAP1 designa a proteína, assim como todas as suas formas homólogas. Por forma homóloga pretende-se designar todas as proteínas equivalentes à proteína considerada, de origem celular diversa e, nomeadamente, proveniente de células de origem humana ou de outros organismos e possuindo uma actividade do mesmo tipo. Esses homólogos compreendem igualmente as variantes naturais da proteína PAP1 com a sequência SEQ ID N°: 2, nomeadamente as variantes polimórficas ou de processamento. Esses homólogos podem ser obtidos por experiências de hibridação entre ácidos nucleicos codificantes (nomeadamente o ácido nucleico da sequência SEQ ID N°: 1). No âmbito da invenção, é suficiente que uma sequência deste tipo apresente uma percentagem significativa de identidade para originar um comportamento fisiológico comparável ao da proteína PAP1, tal como reivindicada. No que a isto respeita, são descritas variantes e/ou homólogos da sequência SEQ ID N°: 2 nas sequências SEQ ID N°: 13, 15, 43 e 45, identificadas a partir de tecidos de origem humana. A denominação PAP1 abrange, deste modo, igualmente estes polipéptidos. A “percentagem de identidade” entre duas sequências de nucleótidos ou de

aminoácidos, no âmbito da presente invenção, pode ser determinada comparando duas sequências alinhadas de uma forma óptima, através de uma janela de comparação.

A parte da sequência nucleotídica ou polipeptídica na janela de comparação pode, assim, compreender adições ou eliminações (por exemplo "espaços") em comparação com a sequência da referência (que não compreende estas adições ou estas eliminações), de forma a obter um alinhamento óptimo das duas sequências.

A percentagem é calculada determinando o número de posições em que é observada uma base nucleica ou um resíduo de aminoácido idêntico para as duas sequências comparadas (nucleica ou peptídica), dividindo, depois, o número de posições em que se verifica identidade entre as duas bases ou resíduos de aminoácido pelo número total de posições na janela da comparação, multiplicando depois o resultado por 100 para obter a percentagem da identidade de sequência.

O alinhamento óptimo das sequências para a comparação pode ser realizado por via informática com a ajuda de algoritmos conhecidos contidos no software da Empresa WISCONSIN GENETICS SOFTWARE PACKAGE, GENETICS COMPUTER GROUP (GCG), 575 Science Doctor, Madison, WISCONSIN.

A título de ilustração, a percentagem da identidade de sequência pode ser realizada com a ajuda do software BLAST (versões BLAST 1.4.9 de Março de 1996. BLAST 2.0.4 de Fevereiro de 1998 e BLAST 2.0.6 de Setembro de 1998), utilizando exclusivamente os parâmetros por omissão (Altschul *et al.*, *J. Mol. Biol.*, (1990) **215**: 403-410; Altschul *et al.*, *Nucleic Acids*

Res. (1997) **25**: 3389-3402). O Blast pesquisa sequências homólogas a uma sequência "inquirida" de referência, utilizando o algoritmo de Altschul *et al.* (*Supra*). A sequência inquirida e as bases de dados utilizadas podem ser peptídicas ou nucleicas, sendo possível qualquer combinação.

No âmbito da presente invenção, o termo derivado indica qualquer sequência diferente da sequência considerada devido a uma degenerescência do código genético, obtida por uma ou várias modificações de natureza genética e/ou química, assim como qualquer péptido codificado por uma sequência hibridando com a sequência nucleica SEQ ID N°: 1 ou um seu fragmento, por exemplo com a sequência nucleica SEQ ID N°: 12, 14, 42 ou 44 ou um seu fragmento e tendo a capacidade para interferir no nível da interação entre a proteína PAP1, ou um dos seus homólogos, e a Parkina. Por modificação de natureza genética e/ou química, pode-se entender qualquer mutação, substituição, eliminação, adição e/ou modificação de um ou mais resíduos. O termo derivado compreende, igualmente, as sequências homólogas à sequência considerada, proveniente de outras fontes celulares e nomeadamente de células de origem humana ou de outros organismos e possuindo uma actividade do mesmo tipo. Essas sequências homólogas podem ser obtidas por experiências de hibridação. As hibridações podem ser realizadas a partir de bibliotecas de ácidos nucleicos, utilizando como sonda a sequência nativa ou um seu fragmento, em condições variáveis de hibridação (Maniatis *et al.*, 1989). Além disso, o termo "fragmento" ou "parte" designa qualquer porção da molécula considerada, compreendendo, pelo menos, 5 resíduos consecutivos, de um modo preferido, pelo menos, 9 resíduos consecutivos, de um modo ainda mais preferido, pelo menos, 15 resíduos consecutivos. Os fragmentos típicos podem compreender, pelo menos, 25 resíduos consecutivos.

Esses derivados ou fragmentos podem ser produzidos com objectivos diferentes, tais como nomeadamente o de aumentar a sua eficácia terapêutica ou de reduzir os seus efeitos secundários ou de lhes conferir novas propriedades farmacocinéticas e/ou biológicas.

Um objectivo específico desta invenção refere-se à proteína PAP1 de sequência SEQ ID N°: 13 ou 43.

Um outro objecto da invenção reside em anticorpos ou fragmentos ou derivados de anticorpos policlonais ou monoclonais dirigidos contra um polipéptido tal como aqui anteriormente definido. Esses anticorpos podem ser produzidos por métodos conhecidos pelo especialista na técnica. Em particular estes anticorpos podem ser preparados por imunização de um animal contra um composto peptídico da invenção (nomeadamente um polipéptido ou um péptido compreendendo a totalidade ou a parte da sequência SEQ ID N°: 2), recolha de sangue e isolamento de anticorpos. Estes anticorpos podem ser igualmente produzidos por preparação de hibridomas de acordo com as técnicas conhecidas pelo especialista na técnica.

De um modo mais preferido, os anticorpos ou os fragmentos dos anticorpos da invenção apresentam a capacidade para modular, pelo menos parcialmente, a interacção dos péptidos reivindicados com a Parkina.

Além disso, estes anticorpos podem ser igualmente utilizados para detectar e/ou dosear a expressão da PAP1 em amostras biológicas e por, este motivo, informar sobre o seu estado de activação.

Os fragmentos ou derivados de anticorpos são, por exemplo, fragmentos Fab, Fab'2, anticorpos de cadeia simples (ScFv), etc. Trata-se, nomeadamente, de qualquer fragmento ou derivado conservando a especificidade antigénica dos anticorpos dos quais estes provêm.

Os anticorpos de acordo com a invenção são, de um modo mais preferido, capazes de se ligar às proteínas PAP1 compreendendo a sequência da SEQ ID N°: 2, 13, 15, 43 ou 45, nomeadamente, a região desta proteína envolvida na interacção com a parkina. Estes anticorpos (ou fragmentos ou derivado) são, de um modo mais preferido, capazes de se ligarem a um epitopo presente na sequência compreendida entre os resíduos 1 e 344 da sequência SEQ ID N°: 2.

A invenção refere-se igualmente a compostos não peptídicos ou não exclusivamente peptídicos utilizáveis como agente farmacêuticos. É de facto possível, a partir dos motivos proteicos activos descritos no presente pedido, preparar moléculas moduladoras da actividade da PAP1 não exclusivamente peptídicas e compatíveis com uma utilização farmacêutica, em particular, duplicando os motivos activos dos péptidos com uma estrutura não peptídica ou de natureza não exclusivamente peptídica.

A presente invenção tem igualmente como objecto qualquer ácido nucleico codificando para um composto peptídico de acordo com a invenção. Pode-se tratar, em particular, de um ácido nucleico compreendendo a totalidade ou parte da sequência SEQ ID N° 1, 12, 14, 42 ou 44 ou um dos seus derivados. Por sequência derivada, entende-se, no âmbito da presente invenção,

qualquer sequência hibridando com a sequência apresentada na SEQ ID N° 1 ou com um seu fragmento e codificando para um composto peptídico de acordo com a invenção, assim como as sequências resultantes destes últimos pela degenerescência do código genético. Os ácidos nucleicos de acordo com a invenção compreendem, por exemplo, a totalidade ou parte da sequência nucleica SEQ ID N°: 12, 14, 42 ou 44.

A presente invenção refere-se, além disso, a sequências apresentando uma percentagem significativa de identidade com a sequência apresentada na SEQ ID N° 1 ou com um seu fragmento e codificantes para um composto peptídico apresentando um comportamento fisiológico comparável com o da proteína PAP1. Entende-se por percentagem de identidade significativa uma percentagem de, pelo menos, 60%, de um modo preferido, 80%, de um modo mais preferido, 90% e, de um modo ainda mais preferido, 95%.

As diferentes sequências nucleotídicas da invenção podem ser de origem artificial ou não. Podem-se tratar de sequências genómicas, ADNc, ARN, sequências híbridas ou sequências sintéticas ou semi-sintéticas. Estas sequências podem ser obtidas por rastreio de bibliotecas de ADN (biblioteca de ADNc, biblioteca de ADN genómico) ou por síntese química ou por métodos variados incluindo a modificação química ou enzimática de sequências obtidas por rastreio de bibliotecas quer por pesquisa de homologia em bases de dados nucleicos ou proteicos. A hibridação referida acima é, de um modo preferido, realizada nas condições descritas por Sambrook *et al.* (1989, páginas 9.52 - 9.55).

É vantajosamente realizada em condições da hibridação de estringência. Por “condições da hibridação de estringência” no âmbito da presente invenção, entender-se-ão as condições seguintes:

1- Competição das membranas e PRÉ-HIBRIDAÇÃO:

- Misturar: 40 µL ADN de esperma de salmão (10 mg/mL) + 40 µL de ADN placentário humano (10 mg/mL)
- Desnaturar durante 5 min a 96 °C, depois, mergulhar a mistura em gelo.
- Descartar o tampão SSC 2 X e verter 4 mL de mistura de formamida no tubo da hibridação contendo as membranas.
- Adicionar a mistura dos dois ADN desnaturados.
- Incubação a 42 °C, durante 5 a 6 horas, com rotação.

2- Competição da sonda marcada:

- Adicionar 10 a 50 µL de ADN Cot I à sonda marcada e purificada, de acordo com a quantidade de hibridações não específicas.
- Desnaturar durante 7 a 10 min a 95 °C.
- Incubar a 65 °C, durante 2 a 5 horas.

### 3- Hibridação:

- Descartar a mistura de pré-hibridação.
- Misturar 40 µL de ADN de esperma de salmão + 40 µL de ADN placentário humano; desnaturar durante 5 min a 96 °C, depois, mergulhar em gelo.
- Adicionar 4 mL da mistura de formamida, a mistura dos dois ADN e a sonda marcada/ADN desnaturado com Cot I ao tubo de hibridação.
- Incubar durante 15 a 20 horas a 42 °C, com rotação.

### 4- Lavagens:

- Uma lavagem à temperatura ambiente em SSC 2 X, para enxaguar.
- 2 vezes 5 minutos em SSC 2 X e SDS a 0,1% à temperatura ambiente.
- 2 vezes 15 minutos SSC 0,1 X e SDS a 0,1% a 65 °C.

Enrolar as membranas em película aderente e expor.

As condições de hibridação descritas acima são adaptadas à hibridação em condições de severidade forte, de uma molécula nucleico de ácido comum comprimento variável de 20 nucleótidos a várias centenas de nucleótidos.

Será evidente que as condições da hibridação descritas acima podem ser adaptadas de acordo com o comprimento do ácido nucleico cuja hibridação é necessária ou do tipo de marcação seleccionado, de acordo com as técnicas conhecidas pelo especialista na técnica.

As condições de hibridação adequadas podem, por exemplo, ser adaptadas de acordo com os ensinamentos contidos no trabalho de HAMES e HIGGINS (1985) (*Nucleic acid Hybridization iapractical Approach*, Hames e Higgins Ed., IRL Press, Oxford) ou no trabalho de F. AUSUBEL *et al.* (1999) (*Currents Protocols in Molecular Biology*, Green Publishing Associates e Wiley Interscience, N.I).

Um ácido nucleico particular no âmbito da invenção codifica para um polipéptido compreendendo a sequência SEQ ID N°: 2 ou um seu fragmento ou derivado, nomeadamente para a proteína humana PAP1. Trata-se vantajosamente de um ácido nucleico compreendendo a sequência nucleica SEQ ID N°: 1, 12, 14, 42 ou 44.

Esses ácidos nucleicos podem ser utilizados para a produção dos compostos peptídico da invenção. O presente pedido refere-se deste modo a um método de preparação desses compostos peptídicos de acordo com o qual se cultiva uma célula contendo um ácido nucleico de acordo com a invenção, em condições de expressão do referido ácido nucleico e se recupera o composto peptídico produzido. Neste caso, a parte codificante para o referido composto peptídico é, geralmente, colocada sob o controlo de sinais, permitindo a sua expressão num hospedeiro celular. A selecção destes sinais (promotores, terminadores, sequência "líder" da secreção, etc.) pode variar de acordo com o

hospedeiro celular utilizado. Além disso, os ácidos nucleicos da invenção podem fazer parte de um vector que pode ser de replicação autónoma ou integrativa. Mais particularmente, podem ser preparados vectores com replicação autónoma utilizando sequências com replicação autónoma no hospedeiro seleccionado. Tratando-se de vectores integrativos, estes podem ser preparados por exemplo utilizando sequências homólogas a determinadas regiões do genoma do hospedeiro, permitindo a integração do vector, por recombinação homóloga. Pode-se tratar de um vector do tipo plasmídico, epissómico, cromossómico, viral, etc.

Os hospedeiros celulares utilizáveis para a produção dos compostos peptídico da invenção por via recombinante são tanto hospedeiros eucariotas como procariotas. De entre os hospedeiros eucariotas convenientes, podem-se referir células animais, leveduras ou fungos. Nomeadamente, tratando-se de leveduras, podem-se referir as leveduras do género *Saccharomyces*, *Kluyveromyces*, *Pichia*, *Schwanniomyces* ou *Hansenula*. Tratando-se células animais, podem-se referir as células COS, CHO, C127, PC12, etc. De entre os fungos, podem-se referir mais particularmente *Aspergillus* ssp. ou *Trichoderma* ssp. Como hospedeiros procariotas, é preferido utilizar as bactérias seguintes *E. coli*, *Bacillus* ou *Streptomyces*.

A presente invenção tem, além disso, como objecto mamíferos não-humanos compreendendo nas suas células um ácido nucleico ou um vector de acordo com a invenção.

Esses mamíferos (roedores, caninos, coelhos, etc.) são nomeadamente utilizáveis no estudo das propriedades da PAP1 e na identificação de compostos com perspectiva terapêutica. A modificação do genoma desse animal transgénico pode resultar de

uma alteração ou de uma modificação de um ou mais genes por "knock-in" ou "knock-out". Esta modificação pode ser realizada utilizando agentes alterantes ou mutagénicos normais ou, então, por mutagénese dirigida. A modificação do genoma pode igualmente resultar de uma inserção de gene(s) ou da substituição de gene(s) na sua forma selvagem ou mutada. As modificações do genoma são vantajosamente realizadas em células estaminais reprodutivas e, vantajosamente, nos pronúcleos. A transgénese pode ser realizada por microinjecção de uma cassete de expressão compreendendo os genes modificados nos dois pronúcleos fertilizados. Deste modo, um animal, de acordo com a presente invenção, pode ser obtido por injecção de uma cassete de expressão compreendendo um ácido nucleico. De um modo preferido, este ácido nucleico é um ADN que pode ser um ADN genómico (ADNg) ou um ADN complementar (ADNc).

A construção de animais transgénicos de acordo com a invenção pode ser realizada de acordo com as técnicas habituais bem conhecidas pelo especialista na técnica. O especialista na técnica será capaz de se referir nomeadamente à produção de animais transgénicos e, particularmente, à produção de murganhos transgénicos, como descrito nas patentes US 4873191, US 5464764 e US 5789215, sendo os conteúdos destes documentos, aqui incorporados por referência.

Em resumo, uma construção polinucleotídica compreendendo um ácido nucleico de acordo com a invenção é inserida numa linha de células estaminais do tipo ES. A inserção da construção polinucleotídica é realizada, de um modo preferido, por electroporação, como descrito por Thomas *et al.* (1987, *Cell*, Vol. **51**: 503-512).

As células que sofreram a etapa de electroporação são, então, rastreadas para a presença da construção polinucleotídica (por exemplo, por selecção utilizando marcadores ou, ainda, por PCR ou análise em gel de electroforese de ADN do tipo Southern) para seleccionar as células positivas tendo integrado a construção polinucleotídica exógena no seu genoma, se necessário, após um evento de recombinação homóloga. Essa técnica é, por exemplo, descrita por MANSOUR *et al.* (*Nature* (1988) **336**: 348-352).

Em seguida, as células seleccionadas positivamente são isoladas, clonadas e injectadas em blastocistos de murganhos com 3,5 dias, como é descrito por BRADLEY (1987, *Production and Analysis of Chimaeric mice*. Em: E. J. ROBERTSON (Ed., *teratocarcinomas and embryonic stem cells: A practical approach*. IRL Press, Oxford, página 113). Os blastocistos são, em seguida, introduzidos num animal hospedeiro do sexo feminino e o desenvolvimento do embrião continua até final.

De acordo com uma alternativa, as células do tipo ES seleccionadas positivamente, são colocadas em contacto com embriões com 2,5 dias numa estádio de 8-16 células (*morulae*) como descrito pela WOOD *et al.* (1993, *Proc. Natl. Acad. Sci. USA*, vol. **90**: 4582-4585) ou por NAGY *et al.* (1993, *Proc. Natl. Acad. Sci. USA*, vol. **90**: 8424-8428), estando as células ES internalizadas para colonizar extensivamente o blastocisto, incluindo as células que dão origem à linha germinal.

Os descendentes são, em seguida, testados para determinar os que integraram a construção polinucleotídica (o transgene).

Os ácidos nucleicos de acordo com a invenção podem ser igualmente utilizados para a preparação de oligonucleótidos anti-sentido ou anti-sentido genéticos, utilizáveis como agentes farmacêuticos. As sequências anti-sentido são oligonucleótidos de tamanho pequeno, complementares à cadeia codificante para um gene determinado e por este facto capazes de hibridar especificamente com o ARNm transcrito, inibindo a sua tradução em proteína. A invenção tem deste modo como objectivo as sequências anti-sentido capazes de inibir, pelo menos, parcialmente, a interacção das proteínas PAP1 com a Parkina. Essas sequências podem ser constituídas pela totalidade ou parte das sequências nucleicas aqui anteriormente definidas. Trata-se geralmente de sequências ou de fragmentos de sequências complementares a sequências codificantes para péptidos que interagem com a Parkina. Esses oligonucleótidos podem ser obtidos por fragmentação, etc, ou por síntese química.

As sequências reivindicadas podem ser utilizadas no âmbito das terapias génicas, para transferência e expressão *in vivo* de sequências anti-sentido ou de péptidos capazes de modular a interacção da proteína PAP1 com a Parkina. Neste contexto, as sequências podem ser incorporadas em vectores virais ou não virais, permitindo a sua administração *in vivo* (Kahn *et al.*, 1991). A título de vectores virais de acordo com a invenção podem-se referir, particularmente, os vectores do tipo adenovírus, retrovírus, vírus associado ao adenovírus (AAV) ou vírus do herpes. O presente pedido tem igualmente como objecto, vírus recombinantes deficientes compreendendo um ácido nucleico codificante para um polipéptido de acordo com a invenção, nomeadamente um polipéptido ou um péptido compreendendo a totalidade ou parte da sequência SEQ ID N°: 2 ou de um seu

derivado, por exemplo, a totalidade ou parte da sequência SEQ ID N°: 12, 14, 42 ou 44 ou dos seus derivados.

A invenção permite igualmente a preparação de sondas nucleotídicas, sintéticas ou não, capazes de hibridar com as sequências nucleotídicas aqui definidas anteriormente ou com a sua cadeia complementar. Essas sondas podem ser utilizadas *in vitro* como ferramenta de diagnóstico, para a detecção da expressão ou da sobre-expressão da PAP1 ou ainda para a descrição de anomalias genéticas (processamento incorrecto, polimorfismo, modificações pontuais, etc.). Estas sondas podem ser igualmente utilizadas para a detecção e o isolamento de sequências de ácidos nucleicos homólogas codificando para péptidos, tal como anteriormente definidos, a partir de outras fontes celulares e de um modo preferido de células de origens humanas. As sondas da invenção compreendem geralmente, pelo menos, 10 bases e estas podem, por exemplo, compreender até a totalidade de uma das sequências anteriormente referidas ou a sua cadeia complementar. De um modo preferido, estas sondas são marcadas previamente à sua utilização. Consequentemente, podem ser utilizadas várias técnicas conhecidas pelo especialista na técnica (marcação radioactiva, fluorescente, enzimática, química, etc.).

A invenção refere-se igualmente a iniciadores ou a pares de iniciadores que permitem amplificar a totalidade ou parte de um ácido nucleico codificando uma PAP1, por exemplo, um iniciador com uma sequência seleccionada das SEQ ID N°: 16-41.

A invenção ainda tem como objectivo qualquer composição farmacêutica compreendendo como ingrediente activo, pelo menos,

um composto, tal como definido acima, nomeadamente um composto peptídico.

Esta tem nomeadamente como objecto qualquer composição farmacêutica compreendendo como ingrediente activo, pelo menos, um anticorpo e/ou um fragmento de anticorpo, tal como definido acima, assim como qualquer composição farmacêutica compreendendo como ingrediente activo, pelo menos, um ácido nucleico ou um vector tal como definido acima.

Tem igualmente como objecto qualquer composição farmacêutica compreendendo como ingrediente activo uma molécula química capaz de aumentar ou reduzir a interacção entre a proteína PAP1 e a Parkina.

Além disso, esta tem igualmente como objectivo as composições farmacêuticas, nas quais os péptidos, anticorpos, moléculas químicas e sequências nucleotídicas acima definidas estão associados entre si ou com outros ingredientes activos.

As composições farmacêuticas, de acordo com a invenção, podem ser utilizadas para modular a actividade da proteína Parkina e por este motivo manter a sobrevivência dos neurónios dopaminérgicos. Mais particularmente, estas composições farmacêuticas destinam-se a modular a interacção entre a proteína PAP1 e a Parkina. Trata-se de um modo mais preferido de composições farmacêuticas destinadas ao tratamento de doenças do sistema nervoso central, tal como, por exemplo, a doença de Parkinson.

A invenção tem, ainda, como objecto a utilização de moléculas aqui descritas anteriormente para modular a actividade

da Parkina ou a tipagem de doenças do sistema nervoso central. Em particular a invenção refere-se à utilização destas moléculas para modular, pelo menos parcialmente, a actividade da Parkina.

A invenção refere-se igualmente a um processo para o rastreio ou a caracterização de moléculas activas sobre a função da parkina, compreendendo a selecção de moléculas capazes de se ligar à sequência SEQ ID N°: 2 ou à sequência SEQ ID N°: 4 ou a um seu fragmento (ou derivado). O processo compreende, vantajosamente, colocar em contacto, *in vitro*, da ou das moléculas teste com um polipéptido compreendendo a sequência SEQ ID N°: 2 ou a sequência SEQ ID N°: 4 ou com um seu fragmento (ou derivado) e a selecção de moléculas capazes de se ligar à sequência SEQ ID N°: 2 (nomeadamente, a região compreendida entre os resíduos 1 e 344) ou à sequência SEQ ID N°: 4. As moléculas testadas podem ser de natureza variada (péptido, nucleico, lípido, açúcar, etc., ou as misturas dessas moléculas, por exemplo das bibliotecas combinatórias, etc.). Como aqui anteriormente indicado, as moléculas identificadas deste modo podem ser utilizadas para modular a actividade da proteína Parkina e representam potenciais agentes terapêuticos para o tratamento de patologias neurodegenerativas.

Outras vantagens da presente invenção irão surgir com a leitura dos exemplos e figura que se seguem, que deve ser consideradas como ilustrativas e não limitantes.

#### **LEGENDAS DA FIGURA:**

Figura 1: Representação do vector pLex9-Parkina (135-290)

Figura 2: Resultados da primeira experiência 5'-RACE. Foram obtidos 8 clones. A sequência electrónica inicial é indicada no final da figura.

Figura 3: Resultados da segunda experiência 5'-RACE. Foram apenas validados dois dos 8 clones obtidos na primeira experiência (clones A12 e D5). A sequência electrónica inicial é indicada no final da figura. A sequência completa de ADN e proteínas é proporcionada nas Sequências 12-15.

Figura 4: Detalhe da organização dos clones C5 e D4 da segunda experiência 5'-RACE. A sequência de consenso resultante é indicada no cimo da figura.

Figura 5: Estrutura dos transcritos isolados a partir de cérebro humano.

Figura 6: Sequência nucleica e proteica de LY111 (completa) de cérebro humano. Duplo sublinhado: cisteínas conservadas do domínio em dedo de zinco. Negrito: Domínio C<sub>2</sub>I, Itálico: domínio C<sub>2</sub>2.

Figura 7: sequência nucleica e proteica de LY111 (versão curta) do cérebro humano. Duplo sublinhado: cisteínas conservadas do domínio em dedo de zinco. Negrito: Domínio C<sub>2</sub>I, Itálico: domínio C<sub>2</sub>2.

Figura 8: Localização da proteína LY111 curta (8b) ou longa (8a) após expressão em células Cos-7.

Figura 9: Sequência nucleica e proteica de LY111 (versão longa) de pulmão humano.

Figura 10: Sequência nucleica e proteica de LY 111 (versão curta) de cérebro humano.

## **MATERIAIS E TÉCNICAS REALIZADAS**

### **1) Estirpes de levedura:**

Foi utilizada a estirpe L40 do género *S. cerevisiae* (*Mata, his3D200, trp1-901, leu2-3,112, ade2, LYS2:: (lexAop)4-HIS3, URA3::(lexAop)8LacZ, GAL4, GAL80*) para verificar as interações proteína-proteína quando um dos parceiros proteicos está fundido com a proteína LexA. Este último é capaz de reconhecer o elemento de resposta LexA controlando a expressão dos genes repórteres *LacZ* e *His3*.

Esta foi cultivada no seguinte meio de cultura:

#### Meio YPD completo:

- Extracto de levedura (10 g/L) (Difco)
- Bactopeptona (20 g/L) (Difco)
- Glucose (20 g/L) (Merck)

Este meio foi tornado sólido por adição de 20 g/L de agar (Difco).

Meio mínimo YNB:

- Base de Azoto de Levedura (sem aminoácidos) (6,7 g/L) (Difco)
- Glucose (20 g/L) (Merck)

Este meio pode ser tornado sólido por adição de agar 20 g/L (Difco). Pode ser também suplementado com aminoácidos e/ou 3-amino-1,2,4-triazole por adição de meios CSM [CSM-Leu, -Trp, -His (620 mg/L), CSM-Trp (740 mg/L) ou CSM-Leu, -Trp (640 mg/L) (Bio101)] e/ou de 3-amino-1,2,4-triazole 2,5 mM.

**2) Estirpes de bactérias:**

Foi utilizada a estirpe TG1 de *Escherichia coli*, do genótipo supE, hsdΔ5, thi, Δ(lac-proAB), F'[tra D36 pro A+B+ lacIq lacZΔM15], na construção de plasmídeos, como meio de amplificação e isolamento de plasmídeos recombinantes utilizados. Esta foi cultivada no seguinte meio:

Meio LB:

- NaCl (5 g/L) (Prolabo)
- Bactotriptonna (10 g/L) (Difco)
- Extracto de levedura (5 g/L) (Difco)

Este meio é tornado sólido por adição de agar (Difco) 15 g/L.

Foi utilizada ampicilina a 100 µg/mL, este antibiótico é utilizado para seleccionar as bactérias que receberam os plasmídeos contendo como marcador o gene da resistência a este antibiótico.

Foi utilizada a estirpe HB101 de *Escherichia coli* do genótipo supE44, ara14, galK2, lacY1, Δ(gpt-proA)62, rpsL20(Strr), xyl-5, mtl-1, recA13, Δ(mcrC-mrr), HsdS-(r-m-) como meios de amplificação e isolamento de plasmídeos provenientes da biblioteca de ADNc do linfócito humano. Esta foi cultivada em

- Meio M9:
- Na<sub>2</sub>HPO<sub>4</sub> (7 g/L) (Prolabo)
  - KH<sub>2</sub>PO<sub>4</sub> (3 g/L) (Prolabo)
  - NH<sub>4</sub>Cl (1 g/L) (Prolabo)
  - NaCl (0,5 g/L) (Prolabo)
  - Glucose (20 g/L) (Sigma)
  - MgSO<sub>4</sub> (1 mM) (Prolabo)
  - Tiamina (0,001%) (Sigma)

Este meio é tornado sólido por adição de agar (Difco) 15 g/L.

Devem ser adicionadas leucina (50 mg/L) (Sigma) e prolina (50 mg/L) (Sigma) ao meio M9 para permitir o crescimento da estirpe HB101.

Durante a selecção de plasmídeos provenientes da biblioteca de dois híbridos de ADNc de linfócito, não foi adicionada leucina ao meio dado que os plasmídeos contêm um marcador de selecção Leu2.

### **3) plasmídeos:**

O vector pLex9 (pBTM116) (Bartel *et al.*, 1993) de 5 kb homólogo ao pGBT10 que contém local de clonagem múltipla localizado a jusante da sequência codificante do repressor bacteriano LexA e a montante de um terminador para formar uma proteína de fusão.

pLex-HaRasVal12, plasmídeo pLex9, tal como descrito no pedido W098/21327, que contém a sequência codificante da proteína HaRas mutada na posição Val12 conhecida por interagir com a proteína Raf de mamífero (Vojtek *et al.*, 1993). Este plasmídeo foi utilizado para testar a especificidade de interacção da proteína PAP1 na estirpe L40.

pLex9-cAPP, plasmídeo pLex9 que contém a sequência codificante para o domínio citoplasmático da proteína APP conhecido por interagir com o domínio PTB2 de FE65. Este plasmídeo foi utilizado para testar a especificidade de interacção da proteína PAP1 na estirpe L40.

#### 4) Oligonucleótidos de síntese:

TTAAGAATTC GGAAGTCCAG CAGGTAG (SEQ ID N° 5)

ATTAGGATCC CTACACACAA GGCAGGGAG (SEQ ID N° 6)

Oligonucleótidos que permitiram a obtenção do fragmento de PCR correspondente à região central da Parkina flanqueado pelos locais *EcoRI* e *BamHI*.

GCGTTTGGAA TCACTACAG (SEQ ID N° 7)

GGTCTCGGTG TGGCATC (SEQ ID N° 8)

CCGCTTGCTT GGAGGAAC (SEQ ID N° 9)

CGTATTTCTC CGCCTTGG (SEQ ID N° 10)

AATAGCTCGA GTCAGTGCAG GACAAGAG (SEQ ID N° 11)

Oligonucleótidos que foram utilizados para sequenciar a inserção correspondente ao gene PAP1.

Os oligonucleótidos são sintetizados no equipamento Applied System ABI 394-08. Estes são libertados da matriz de síntese com amónia e são precipitados duas vezes com 10 volumes de n-butanol depois são retomados em água. A quantificação é realizada por medição da densidade óptica (1 DO<sub>260</sub> corresponde a 30 µg/mL).

## **5) Preparação de ADN plasmídicos.**

Foram realizadas preparações de ADN plasmídico em pequena quantidade e em grande quantidade de acordo com os protocolos recomendados pelo fabricante Quiagen dos kits de purificação de ADN:

- kit Quiaprep Spin Miniprep, ref : 27106
- kit Quiaprep Plasmid Maxiprep, ref : 12163.

## **6) Amplificação enzimática de ADN por PCR (Reacção em Cadeia da Polimerase):**

As reacções de PCR são realizadas num volume final de 100 µL na presença do molde de ADN, dNTP (0,2 mM), tampão PCR (Tris-HCl pH 8,5 10 mM, MgCl<sub>2</sub> 1 mM, KCl 5 mM, gelatina a 0,01%), 10-20 pmoles de cada oligonucleótido e 2,5 UI de polimerase de ADN Ampli Taq (Perkin Elmer). A mistura é coberta novamente com 2 gotas de óleo de parafina, para limitar a evaporação da amostra. O aparelho utilizado é o "Crocodile II" de Appligene.

A requerente utilizou uma temperatura de desnaturação do molde de 94 °C, uma temperatura de hibridação de 52 °C e uma temperatura de alongamento pela enzima de 72 °C.

## **7) Ligações:**

Todas as reacções de ligação são realizadas a 37 °C durante uma hora num volume final de 20 µL na presença de 100 a 200 ng

de vector, 0,1 a 0,5 µg de inserção, 40 UI de enzima ligase de ADN de T4 (Biolabs) e um tampão de ligação (Tris-HCl pH 7,8 50 mM; MgCl<sub>2</sub> 10 mM; DTT 10 mM; ATP 1 mM). O controlo negativo é constituído por ligação do vector na ausência de inserção.

#### **8) Transformação das bactérias:**

A transformação das bactérias por um plasmídeo é realizada de acordo com o protocolo seguinte: são utilizados 10 µL de volume de ligação para transformar bactérias TG1 de acordo com o método de Chung (Chung *et al.*, 1989). Após transformação as bactérias são espalhadas sobre meio LB + ampicilina e são incubadas, durante 16 h, a 37 °C.

#### **9) Separação e extracção do ADN:**

A separação do ADN é realizada de acordo com o seu tamanho por electroforese em gel de agarose de acordo com Maniatis (Maniatis *et al.*, 1989): gel de agarose a 1% (Gibco BRL) num tampão TBE (Tris bases 90 mM; Borato 90 mM; EDTA 2 mM)

#### **10) Sequenciação fluorescente de ADN plasmídicos:**

A técnica de sequenciação utilizada é derivada do método de Sanger (Sanger *et al.*, 1977) e é adaptada para a sequenciação por fluorescência desenvolvida por Applied Biosystems. O protocolo utilizado é o descrito pelos criadores do sistema (Perkin Elmer, 1997).

### **11) Transformação da levedura:**

Os plasmídeos são introduzidos na levedura por uma técnica clássica da transformação de leveduras, desenvolvido por Gietz (Gietz *et al.*, 1992) e modificado do modo seguinte:

No caso particular da transformação da levedura pela biblioteca de ADNc do linfócito, a levedura utilizada contém o plasmídeo pLex9-Parkina (135-290) codificando para a parte central da Parkina fundida com a proteína LexA. Esta é cultivada em 200 mL de meio mínimo YNB suplementado com aminoácidos CSM-Trp a 30 °C sob agitação até uma densidade de  $10^7$  células/mL. Para realizar a transformação de leveduras de acordo com o protocolo anterior, a suspensão celular foi separada em 10 tubos de 50 µL aos quais foram adicionadas 5 µg da biblioteca. O choque térmico foi realizado durante 20 minutos, depois as células foram recolhidas por centrifugação e foram ressuspensas em 100 mL de meio YPD, durante 1 h, com 30 °C e em 100 mL de meio YNB suplementado com CSM- Leu, -Trp durante 3 h 30, a 30 °C. A eficácia da transformação é determinada espalhando várias diluições de células transformadas sobre meio YNB sólido suplementado com CSM-Trp, -Leu. As colónias obtidas foram contadas após uma cultura, a 30 °C, durante 3 dias e foi determinada a taxa de transformação por µg de ADN da biblioteca de linfócitos.

### **12) Isolamento de plasmídeos extraídos de levedura:**

São centrifugados 5 mL de uma cultura de levedura incubada durante 16 h, a 30 °C e foram retomados em 200 µL de um tampão de lise (Sorbitol 1 M,  $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$  0,1 M pH 7,4, zymolyase

12,5 mg/mL) e são incubados, durante 1 h, a 37 °C. O lisado é, então, tratado de acordo com o protocolo recomendado pelo fabricante Quiagen do kit de purificação de ADN, kit Quiaprep Spin Miniprep, ref 27106.

### **13) Teste de actividade da $\beta$ -galactosidase:**

É previamente depositada uma folha de nitrocelulose numa placa de Petri contendo os clones de levedura individualizados. Esta folha, então, é submersa em azoto líquido durante 30 segundos para fazer rebentar as leveduras e libertar deste modo a actividade  $\beta$ -galactosidásica. Após descongelação, a folha de nitrocelulose é depositada, com as colónias para cima, numa outra placa de Petri contendo um papel Whatman previamente embebido com 1,5 mL de solução de PBS ( $\text{Na}_2\text{HPO}_4$  60 mM,  $\text{NaH}_2\text{PO}_4$  40 mM, KCl 10 mM,  $\text{MgSO}_4$  1 mM, pH 7) contendo 15  $\mu\text{L}$  de X-Gal (5-bromo-4-cloro-3-indoil- $\beta$ -D-galactoside) com 40 mg/mL de N,N-dimetilformamida. A placa é de seguida colocada numa estufa a 37 °C. O teste é considerado positivo quando as colónias ficam azuis na membrana no final de 12 horas.

**EXEMPLO 1:** CONSTRUÇÃO DE UM VECTOR PERMITINDO A EXPRESSÃO DE UMA PROTEÍNA DE FUSÃO ENTRE A PARTE CENTRAL DA PARKINA E O REPRESSOR BACTERIANO LEXA.

O rastreio de uma biblioteca utilizando o sistema duplo híbrido requer que a região central da Parkina esteja fundida com uma proteína de ligação ao ADN como o repressor bacteriano LexA. A expressão desta proteína de fusão é realizada utilizando o vector pLex9 (ver materiais e métodos), em que foi

introduzida, dentro da mesma grelha de leitura que a sequência correspondente à proteína LexA, a sequência codificante da região central da Parkina que figura na sequência apresentada na SEQ ID N°3 ou 4.

Foi obtido o fragmento de ADN com 468 pb correspondente aos 156 aminoácidos da região central da Parkina que se inicia no aminoácido 135 por PCR a partir de oligonucleótidos (SEQ ID N° 5 e N° 6) que permitiram igualmente introduzir o local *EcoRI* na extremidade 5' e um codão de terminação e um local *BamHI* na extremidade 3'. O fragmento de PCR foi introduzido entre os locais *EcoRI* e *BamHI* do local de clonagem múltipla do plasmídeo pLex9 a jusante da sequência codificante da proteína LexA para originar o vector pLex9-Parkina (135-290) (Fig. 1).

A construção foi verificada por sequenciação do ADN. Esta verificação permitiu mostrar que este fragmento não apresentava mutações produzidas durante a reacção de PCR e que foi fundido na mesma grelha de leitura aberta que a do fragmento correspondente a LexA.

**EXEMPLO 2:** RASTREIO DE UMA BIBLIOTECA DE FUSÃO DE LINFÓCITOS.

A requerente utilizou o método do Duplo-Híbrido (Fields e Song, 1989).

O rastreio de uma biblioteca de fusão permite identificar clones produzindo proteínas fundidas com o domínio transactivador de GAL4, podendo interagir com a proteína do interesse descrita no exemplo 1 (região central da parkina).

Esta interacção permite reconstituir um transactivador que será então capaz de induzir a forma dos genes repórteres His3 e LacZ na estirpe L40.

Para realizar este rastreio a requerente seleccionou uma biblioteca de fusão preparada a partir de ADNc proveniente de linfócitos humanos periféricos proporcionada por Richard Benarous (Peytavi *et al.*, 1999). As leveduras foram transformadas pela biblioteca de linfócitos e foram seleccionados clones positivos como aqui descrito abaixo.

Durante o rastreio é necessário preservar a probabilidade de cada plasmídeo independente da biblioteca de fusão estar presente em, pelo menos, uma levedura, simultaneamente com o plasmídeo pLex9-Parkina (135-290). Para conservar esta probabilidade é importante ter uma boa eficácia de transformação da levedura. Como resultado a requerente seleccionou um protocolo de transformação da levedura proporcionando uma eficácia de  $2,6 \cdot 10^5$  células transformadas por  $\mu\text{g}$  de ADN. Além disso como a cotransformação da levedura por dois plasmídeos diferente reduz esta eficácia, a requerente preferiram utilizar uma levedura transformada previamente pelo plasmídeo pLex9-Parkina (135-290). Esta estirpe L40 pLex9-Parkina (135-290) de fenótipo His-, Lys-, Leu-, Ade- foi transformada com 50  $\mu\text{g}$  de ADN plasmídico da biblioteca de fusão. Esta quantidade de ADN permitiu aos requerentes obter após estimativa  $1,3 \cdot 10^7$  células transformadas, o que corresponde a um número ligeiramente superior ao número de plasmídeos independentes que constituem a biblioteca. De acordo com este resultado a requerente acreditaram que foi utilizada a quase totalidade dos plasmídeos da biblioteca para transformar as leveduras. A selecção das células transformadas, capazes de reconstituir um

transactivador funcional, foi realizada num meio YNB suplementado com 3-amino-1,2,4-triazole 2,5 mM e CSM (Bio101) 620 mg/L não contendo histidina, leucina e triptofano.

Foram obtidos numerosos clones com fenótipo His<sup>+</sup> como resultado desta selecção. Foi realizado um teste da actividade  $\beta$ -galactosidase destes transformantes para validar este número de clones obtidos pela expressão do outro gene repórter, *LacZ*. 115 clones apresentando o fenótipo duplo His<sup>+</sup>,  $\beta$ -Gal<sup>+</sup> podem corresponder a uma interacção proteína-proteína.

**EXEMPLO 3:** ISOLAMENTO DE PLASMÍDEOS DA BIBLIOTECA NOS CLONES SELECIONADOS.

Foram extraídos os plasmídeos da biblioteca de fusão contidos nas leveduras seleccionadas durante o rastreio de duplos híbridos, para identificar as proteínas podendo interagir com a região central da Parkina. Para os poder obter em grande quantidade, este isolamento requer previamente uma transformação de *E. coli* por um extracto de ADN das estirpes de levedura positivas. Como o plasmídeo da biblioteca contido neste extracto é um plasmídeo transportador de levedura/*E. coli* este pode ser facilmente replicado na bactéria. O plasmídeo da biblioteca foi seleccionado por complementação da bactéria HB101 auxotrófica para a leucina em meio desprovido de leucina.

Foram analisados os ADN plasmídico das colónias bacterianas obtidas após transformação com extractos de ADN de leveduras por digestão com enzimas de restrição e separação dos fragmentos de ADN em gel de agarose. De entre os 115 clones analisados, foi obtido um clone contendo um plasmídeo da biblioteca apresentando

um perfil diferente dos outros. Este plasmídeo, designado pGAD-Ly1111b, foi estudado mais detalhadamente.

**EXEMPLO 4:** DETERMINAÇÃO DA SEQUÊNCIA DA INSERÇÃO CONTIDA NO PLASMÍDEO IDENTIFICADO.

A sequenciação da inserção contida no plasmídeo identificado foi realizada inicialmente a partir do oligonucleótido SEQ ID N° 7 complementar à sequência de GAL4TA na proximidade do local *EcoRI* da inserção da biblioteca de ADN de linfócitos. Depois, numa segunda fase, segunda a partir dos oligonucleótidos SEQ ID N°8 a SEQ ID N° 11, correspondentes à sequência da inserção obtida durante a progressão da sequenciação. A sequência obtida é apresentada na sequência SEQ ID N°: 1. A proteína identificada deste modo foi designada PAP1 (Parkin-Associated Protein 1).

A comparação entre a sequência desta inserção e as sequências contidas nas bases de dados EMBL e GENBank (European Molecular Biology Lab) mostrou uma homologia de 25%, ao nível proteico, com vários membros da família das Sinaptotagminas. As Sinaptotagminas pertencem a uma família de proteínas membranares codificadas por, pelo menos, onze genes diferentes expressos no cérebro e em outros tecidos. Estas contêm um domínio transmembranar único e dois domínios regulados pelo cálcio designados C2. É neste domínio que se encontra a homologia entre Sinaptotagminas e a proteína PAP1. Não foi observada qualquer outra homologia significativa.

**EXEMPLO 5:** ANÁLISE DA ESPECIFICIDADE DA INTERACÇÃO ENTRE A REGIÃO CENTRAL DA PARKINA E A PROTEÍNA PAP1.

Foi realizado um teste de interacção específica de dois híbridos com outras proteínas não-relevantes para determinar a especificidade da interacção entre o fragmento correspondente à proteína PAP1 e a região central da Parkina. Para realizar este teste os requerentes transformaram a estirpe L40 com os plasmídeos de controlo plex9-APP ou pLex9-HaRasVall12 em vez do plasmídeo pLex9-Parkin (135-290) codificando respectivamente para o domínio citoplasmático da APP ou a proteína HaRasVall12 fundidos com o domínio de ligação ao ADN de LexA e para o plasmídeo isolado durante o rastreio da biblioteca de dois híbridos. Foi realizado um teste de actividade  $\beta$ -Gal nas células transformadas pelos diferentes plasmídeos para determinar uma interacção proteína-proteína. De acordo com o resultado do teste, apenas as leveduras transformadas pelo plasmídeo isolado durante o rastreio da biblioteca de dois híbridos e pelo plasmídeo pLex9-Parkina (135-290) apresentaram uma actividade  $\beta$ -Gal+, demonstrando, deste modo, uma interacção entre a região central da Parkina e a proteína PAP1. Esta interacção revela, consequentemente, ser específica uma vez que este fragmento de PAP1 não parece interagir com as proteínas APP ou HaRasVall12.

Estes resultados mostram, consequentemente, a existência de uma nova proteína, designada PAP1, capaz de interagir de um modo específico com a Parkina. Esta proteína, relacionada com as sinaptotagminas, não tem qualquer homologia significativa com proteínas conhecidas e pode ser utilizada em aplicações terapêuticas ou de diagnóstico, para a produção de anticorpos, sondas ou péptidos ou para o rastreio de moléculas activas.

**EXEMPLO 6:** CLONAGEM DO GENE PAP1 A PARTIR DE UMA BIBLIOTECA DE ADN DE PULMÃO HUMANO

Com o objectivo de identificar a sequência completa do gene PAP1 humano e de caracterizar a existência de formas variantes, foram realizadas duas abordagens de alongamento electrónico a partir da sequência SEQ ID N°: 1. Deste modo, foram obtidas duas sequências electrónicas, com respectivamente 1644 pb e 1646 pb, compreendendo um alongamento de 330 pb em comparação com a sequência SEQ ID N°: 1. No entanto, a análise destas sequências revelou diferenças na região de consenso, aparentes após tradução. Deste modo, é obtida uma ORF com 420 aa num caso e uma ORF com 230 aa para a outra sequência. A sequência proteica obtida foi comparada com as sequências conhecidas e revelou uma homologia de 24% nos 293 aminoácidos sobrepostos com a sinaptogamina humana 1 (p65) (p21579). A função da sinaptogamina I pode consistir num papel regulador nas interacções membranares durante o tráfego de vesículas sinápticas na zona da sinapse. Esta liga-se aos fosfolipídios ácidos com alguma especificidade. Além disso, foi descrita uma interacção dependente do cálcio entre a sinaptogamina e os receptores da cinase da proteína C activada. A sinaptogamina pode-se igualmente ligar a três outras proteínas que são neurexina, syntaxina e ap2. Tendo em consideração o desaparecimento brutal e precoce de qualquer homologia entre as sequências identificadas e a família das sinaptogaminas, é possível que a sequência identificada apresente uma eliminação em relação à sequência natural. Foi realizada uma experiência de RT-PCR e sequenciação utilizando a sequência de 1644 pb para verificar esta hipótese e validar as sequências. A sequência a se obtida compreende uma ORF com 420 aa apresentando uma homologia da mesma ordem com as sinaptogaminas.

Para tentar obter uma sequência maior e verificar se a sequência obtida pode corresponder a uma forma de processamento, foi iniciada uma experiência de alongamento por 5'-RACE a partir da região 3' da sequência validada, utilizando os oligonucleótidos L1 e L2 numa preparação de ADNc de pulmão humano.

Os resultados obtidos são apresentados na figura 2 e mostram a identificação de 8 clones correspondentes em 6 extremidades 5' terminais diferentes. Três destas contêm um codão de terminação que interrompe a ORF (clones A12, F2, F12) e o clone A3 não contém qualquer ORF. Foi confirmada a presença de diferentes transcritos por RT-PCR e RT-PCR com iniciadores internos (Tabela 1).

Tabela 1

RT-PCR	Primária	PCR Secundária U3-L3	PCR Secundária U1-L4	PCR Secundária C-B
U3-L3	170			
A-L4	153		+	
A-L3	Smear	+		
U1-L4	130			
U1-L3	Smear	+		
U1-B	415			+
U2-B	515			+
Tamanho esperado		170	130	120

Os pares de iniciadores U3-L3 e C-B são específicos para o fragmento comum da sequência, os oligonucleótidos A e U1 são específicos para a sequência inicial e para o clone C11, o oligonucleótido L4 é específico para a sequência inicial e o iniciador U2 é específico para o clone A3. Foi realizada uma

segunda 5'-RACE com os oligonucleótidos L3 e L7 localizados na região comum dos diferentes clones (Figura 2). Os resultados obtidos são apresentados nas Figuras 3 e 4. A presença dos diferentes transcritos foi confirmada por RT-PCR e RT-PCR com iniciadores internos (Tabela 2).

Tabela 2

RT-PCR	Resultado	PCR Secundária C-B	PCR Secundária U3-B	PCR Secundária U5-L7
U4-F	Smear	+	+	+
U5-F	Smear	+	+	+
U3-F	1550 pb	+	+	
Tamanho esperado		120	385	530

A sequência dos iniciadores e oligonucleótidos é proporcionada nas Tabelas 3 e 4 (SEQ ID N°: 16-37).

Tabela 3

				SEQ ID
LY111_U4	CCAGTTCTGCCTGTTTCATC	23	a 41	16
LY111_U5	TTCAAAACACAGAGGAGGAG	319	a 338	17
LY111_U3	GAATTTGGTCAGTTTAGAGG	759	a 778	18
LY111_L7	TTCTGGGATTTGGAGAGCTTTTTCAC	851	a 825	19
LY111_L6	TCTGTCTGTCCCACACACTGCC	914	a 892	20
LY111_L3	GACTGGCTCCGTCTCTCTG	928	a 910	21
LY111_C	AAGCAACAGAATCTCCCATCC	1029	a 1049	22
LY111_B	GCATTGTCAAAATTGCCCATC	1147	a 1127	23
LY111_E	AGGCGGAGAAATACGAAGAC	1543	a 1562	24
LY111_D	GCAGAGTGAGACAGCCCTTAAC	1767	a 1746	25
Ly111_L2	CTTCCTCAGGACTGGCGACTTCAG	1811	a 1782	26

(continuação)

				SEQ ID
Ly111_L1	CAAGCGGTCGTTTCATTCCAAAGAG	1934	a 1913	27
LY111_F	AAGAGGAGATAACCCACCAGAG	2288	a 2269	28

Tabela 4

LY111_A	TCGTAGAGCAGCAGGTCCAAG	14	a 34	46
LY111_U1	AGGGCTGCTGGCTATTTTC	36	a 55	29
LY111_L4	TAAGAAATGGGTTGTGAAC	148	a 166	30
LY111_C	AAGCAACAGAATCTCCCATCC	1029	a 1049	31
LY111_B	GCATTGTCAAATGCCCCATC	1147	a 1127	32
LY111_E	AGGCGGAGAAATACGAAGAC	1543	a 1562	33
LY111_D	GCAGAGTGAGACAGCCCTAAC	1767	a 1746	34
Ly111_L2	CTTCCTCAGGACTGGCGACTTCAG	1811	a 1782	35
Ly111_L1	CAAGCGGTCGTTTCATTCCAAAGAG	1934	a 1913	36
LY111_F	AAGAGGAGATAACCCACCAGAG	2288	a 2269	37

O conjunto destes resultados permite validar a sequência de consenso correspondente à isoforma longa (Figura 9, SEQ ID N°: 12 e 13) e à isoforma curta (Figura 10, SEQ ID N°: 14 e 15) da proteína PAP1 identificada a partir de pulmão humano. Esta proteína é igualmente designada ao longo dos exemplos pelo termo LY111. A isoforma longa é codificada por uma ORF com 1833 pb localizada nos resíduos 237-2069 da SEQ ID N°: 12 e compreende 610 aminoácidos. O sinal de poliadenilação está localizado a partir do nucleótido 2315. A isoforma curta é codificada por uma ORF com 942 pb localizada nos resíduos 429-1370 da SEQ ID N°: 14 e compreende 313 aminoácidos. O sinal de poliadenilação está localizado a partir do nucleótido 1616.

Foram, em seguida, realizadas experiências de transferência de Northern, em diferentes tecidos humanos, com sondas (amplimer CD e E-F) e que permitiram revelar um transcrito de 6 kb no músculo, transcrito no coração (3 kb), assim como um transcrito de 6 kb no fígado fetal. O exemplo 7 descreveu, adicionalmente, a clonagem de um transcrito no cérebro fetal humano.

Foram realizados diferentes estudos de homologia nas diferentes bases de dados proteicas, cujos resultados são apresentados na tabela 5 abaixo.

Tabela 5

Biblioteca	Homologia
Genpept116	G5926736 (AB025258) granufilina-a Identidade: 31% (215/679), Homologia (POS): 46% (322/679)
	G5926738 (AB025259) granufilina-b Identidade: 31% (150/479), Homologia (POS): 47% (230/479)
	G1235722 (D70830) Doc2 beta (homo sapiens) Identidade: 25% (74/292), Homologia (POS): 43% (127/292)
	G289718 (L 15302) Sinaptogamina-I Identidade: 26% (77/293), Homologia (POS): 45% (133/293)

(continuação)

Biblioteca	Homologia
Swissprot	SP:SYT1_CAEEL Sinaptogamina I Identidade: 26% (77/293), Homologia (POS): 45% (133/293)
	SP:SYT2_MOUSE Sinaptogamina II Identidade: 24% (72/293), Homologia (POS): 44% (131/293)

EXEMPLO 7: CLONAGEM DE DOIS TRANSCRITOS COMPLETOS DE PAP1 (LY111B) A PARTIR DE ADN COMPLEMENTAR DE CÉREBRO FETAL HUMANO

Para confirmar a presença de um transcrito Ly111b *full-length* (“completo”) no cérebro humano, foi realizada uma PCR partir de ADN complementar proveniente de cérebro fetal humano (Marathon Ready cDNA, Clontech), utilizando como iniciadores os oligonucleótidos LyF1 (AAT GGA AGG GCG TGA CGC, figura 5, SEQ ID N°: 38) e HA71 (CCT CAC GCC TGC TGC AAC CTG, SEQ ID N°: 39). Foi amplificado um fragmento de ADN pouco representado com aproximadamente duas kilobases. Foi utilizado o produto desta primeira PCR como molde para uma PCR com iniciadores internos, realizada com os oligonucleótidos LyEcoF (GCACGAATTC TGA GCC CAA GAA ATA GAT CTG, SEQ ID N°: 40) e HA72 (CTG TCT TCG TAT TTC TCC GCC TTG, SEQ ID N°: 41). Os produtos amplificados foram digeridos com as enzimas da restrição EcoRI (integrado no oligonucleótido LyEcoF) e BstEII (figura 5) e foram inseridos no vector de expressão pcDNA3, depois foi determinada a sua sequência. A análise da sequência dos clones obtidos revelou a presença de dois transcritos potenciais Ly111b completos no cérebro fetal humano (figura 5). O primeiro destes transcritos (Ly111b<sub>full1A</sub>) corresponde a ARNm identificado no

pulmão humano (Exemplo 6) e codifica uma proteína de 609 aminoácidos (pLy111b<sub>fullA</sub>; figuras 5, 6, SEQ ID N°: 42-43). O segundo (Ly111b<sub>fullB</sub>) representa provavelmente um produto de processamento alternativo de um ARNm primário comum. Neste transcrito, idêntico a Ly111b<sub>fullA</sub>, a sequência entre os nucleótidos 752 e 956 da sequência validada no pulmão humano está ausente (SEQ ID N°: 42). Ly111b<sub>fullB</sub> codifica, deste modo, para uma proteína com 541 aminoácidos (pLy111b<sub>fullB</sub>) idêntica a pLy111b<sub>fullA</sub>, na qual no entanto o domínio incluído entre os aminoácidos 172 e 240 (figura 5, 7, SEQ ID N°: 44-45) está ausente. As duas proteínas pLy111b<sub>fullA/fullB</sub> integram o domínio da interação com o fragmento da Parkina compreendendo os aminoácidos 135 a 290, identificado na levedura (sequência inicial Ly111b, a figura 5) e podem, deste modo, manter teoricamente esta interação.

As proteínas pLy111b<sub>fullA/fullB</sub> pertencem à família RIM/Rabfilina

A pLy111b<sub>fullA/fullB</sub> tem uma homologia com as proteínas da família RIM/rabfilina (Wang Y, Sugita S e Südhof TG. The RIM/NIM family of neuronal C2 domain proteins. *J Biol Chem* (2000) **275**, 20033-20044) e, nomeadamente com as granulofilinas (Wang Jie, Takeuchi T, Yokota H e Izumi T Novel Rabphilin-3-like protein associates with insulin-containing granules in pancreatic beta cells. *J Biol Chem* (1999) **274**, 28542-28548). Estas são caracterizados pela presença de um domínio dedo de zinco na parte do terminal N e de dois domínios C<sub>2</sub> na parte do terminal C (Figuras 6 e 7). O domínio dedo de zinco das proteínas da família RIM/rabfilina foi envolvido na interação com as proteínas Rab. Estas últimas, proteínas que fixam o GTP, são componentes essenciais da maquinaria do tráfego membranar nas

células eucariotas. Além disso, foi descrito que os domínios C<sub>2</sub> de proteínas da família RIM/rabfilina se podem ligar às membranas por interacção com fosfolípidos.

Expressão das proteínas  $ply111b_{fullA/fullB}$  nas células da linha cos-7: co-localização com a Parkina

A sequência codificante para os transcritos  $Ly111b_{fullA/B}$  foi inserida no vector de expressão eucariota pcDNA3 em fase com a sequência codificando para um epitopo myc do terminal N ( $pcDNA3-mycLy111b_{fullA/B}$ ). As células da linha cos-7 transfectadas utilizando estes vectores produzem proteínas com um peso molecular aparente de aproximadamente 67 kDa ( $pcDNA3-mycLy111b_{fullA}$ ) e 60 kDa ( $pcDNA3-mycLy111b_{fullB}$ ), correspondentes ao peso molecular esperado. Estas proteínas, detectadas por imunomarcacão utilizando um anticorpo dirigido contra o epitopo myc do terminal N, distribuem-se de uma forma não-homogénea, dispersa, no citoplasma, nos prolongamentos e por vezes no núcleo das células da linha cos-7 (figura 8a, b, coluna A). Quando estas são sobre-expressas com a Parkina, revelada utilizando o anticorpo anti-Parkina Asp5 nas células da linha cos-7 (figura 8a, b, coluna B), pode ser observada uma distribuição semelhante e uma co-localização destas proteínas (figura 8a, b, coluna C).

## REFERÊNCIAS BIBLIOGRÁFICAS

**Abbas**, N. *et al.* (1999). A wide variety of mutations in the parkin gene are responsible for autosomal recessive parkinsonism in Europe. *Hum Mol Genet.* **8**, 567-574.

**Bartel**, P. L. *et al.* (1993). D. A Hartley Ed, Oxford University press, 153

**Chung**, CT. *et al.* (1989). One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. USA.* **86**, 2172-2175.

**Fields**, S. and Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature.* **340**, 245-246.

**Gietz**, RD. *et al.* (1992). Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.*, **20**, 1425.

**Hattori**, N. *et al.* (1998). Molecular genetic analysis of a novel Parkin gene in Japanese families with autosomal recessive juvenile parkinsonism: evidence for variable homozygous deletions in the Parkin gene in affected individuals. *Ann Neurol.* **6**, 935-941.

**Kahn**, A. *et al.*, (1991) Thérapie génique: espoirs et limites. *Médecine et Sciences.* **7**, 705-714.

**Kitada**, T. *et al.* (1998). Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature.* **392**, 605-608.

**Leroy**, E. *et al.* (1998). The ubiquitin pathway in Parkinson's disease. *Nature.* **395**, 451-452.

Lewy, F. H. (1912). in *Handbuch der Neurologie* (Lewandowski, M., ed) pp 920 - 933, Springer, Berlin

Lucking, C. *et al.* (1998). Homozygous deletions in parkin gene in European and North African families with autosomal recessive juvenile parkinsonism. *Lancet*. **352**, 1355-1356.

Maniatis, T. *et al.* (1989). *Molecular cloning*, second edition. Cold Spring Harbor Laboratory. Cold Spring Harbor, N. I.

**Morett**, E. (1999). A novel transactivation domain in parkin. *Trends Biochem Sci.* 24, 229-231.

**Peytavi**, R. *et al.* (1999). HEED, the product of the human homolog of the murine eed gene, binds to the matrix protein of HIV-1. *J. Biol. Chem.* **274**, 1635-1645.

**Polymeropoulos**, M.H. *et al.* (1997). Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science*. **276**, 2045-2047.

**Sanger**, F. *et al.* (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.

**Shimura**, H. *et al.* (1999). Immunohistochemical and subcellular localization of Parkin protein: absence of protein in autosomal recessive juvenile parkinsonism patients. *Ann Neurol.* **45**, 668-672.

**Sunada, Y. et al.** (1998). Differential expression of the parkin gene in the human brain and peripheral leukocytes. *Neurosci Lett.* 254, 180-182.

**Vojtek, AB. et al.,** (1993). Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell.* **74**, 205-214.

#### LISTAGEM DE SEQUÊNCIAS

<110> AVENTIS PHARMA

INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE M

<120> COMPOSTOS CAPAZES DE MODULAR A ACTIVIDADE DA PARKINA,  
SEQUÊNCIAS NUCLEOTÍDICAS E UTILIZAÇÕES

<130> PRJ00004

<140>

<141>

<160> 46

<170> PatentIn Ver. 2.1

<210> 1

<211> 1313

<212> ADN

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(1032)

<400> 1

```
cag aat ctc cca tcc agt ccg gca ccc agt acc ata ttc tct gga ggt 48
Gln Asn Leu Pro Ser Ser Pro Ala Pro Ser Thr Ile Phe Ser Gly Gly
  1             5             10             15

ttt aga cac gga agt tta att agc att gac agc acc tgt aca gag atg 96
Phe Arg His Gly Ser Leu Ile Ser Ile Asp Ser Thr Cys Thr Glu Met
          20             25             30

ggc aat ttt gac aat gct aat gtc act gga gaa ata gaa ttt gcc att 144
Gly Asn Phe Asp Asn Ala Asn Val Thr Gly Glu Ile Glu Phe Ala Ile
          35             40             45

cat tat tgc ttc aaa acc cat tct tta gaa ata tgc atc aag gcc tgt 192
His Tyr Cys Phe Lys Thr His Ser Leu Glu Ile Cys Ile Lys Ala Cys
          50             55             60

aag aac ctt gcc tat gga gaa gaa aag aag aaa aag tgc aat ccg tat 240
Lys Asn Leu Ala Tyr Gly Glu Glu Lys Lys Lys Lys Cys Asn Pro Tyr
          65             70             75             80

gtg aag acc tac ctg ttg ccc gac aga tcc tcc cag gga aag cgc aag 288
Val Lys Thr Tyr Leu Leu Pro Asp Arg Ser Ser Gln Gly Lys Arg Lys
          85             90             95

act gga gtc caa agg aac acc gtg gac ccg acc ttt cag gag acc ttg 336
Thr Gly Val Gln Arg Asn Thr Val Asp Pro Thr Phe Gln Glu Thr Leu
          100             105             110

aag tat cag gtg gcc cct gcc cag ctg gtg acc cgg cag ctg cag gtc 384
Lys Tyr Gln Val Ala Pro Ala Gln Leu Val Thr Arg Gln Leu Gln Val
          115             120             125

tcg gtg tgg cat ctg ggc acg ctg gcc cgg aga gtg ttt ctt gga gaa 432
Ser Val Trp His Leu Gly Thr Leu Ala Arg Arg Val Phe Leu Gly Glu
          130             135             140

gtg atc att tct ctg gcc acg tgg gac ttt gaa gac agc aca aca cag 480
Val Ile Ile Ser Leu Ala Thr Trp Asp Phe Glu Asp Ser Thr Thr Gln
          145             150             155             160
```

tcc ttc cgc tgg cat ccg ctc cgg gcc aag gcg gag aaa tac gaa gac	528
Ser Phe Arg Trp His Pro Leu Arg Ala Lys Ala Glu Lys Tyr Glu Asp	
165 170 175	
agc gtt cct cag agt aat gga gag ctc aca gtc cgg gct aag ctg gtt	576
Ser Val Pro Gln Ser Asn Gly Glu Leu Thr Val Arg Ala Lys Leu Val	
180 185 190	
ctc cct tca cgg ccc aga aaa ctc caa gag gct caa gaa ggg aca gat	624
Leu Pro Ser Arg Pro Arg Lys Leu Gln Glu Ala Gln Glu Gly Thr Asp	
195 200 205	
cag cca tca ctt cat ggt caa ctt tgt ttg gta gtg cta gga gcc aag	672
Gln Pro Ser Leu His Gly Gln Leu Cys Leu Val Val Leu Gly Ala Lys	
210 215 220	
aat tta cct gtg cgg cca gat ggc acc ttg aac tca ttt gtt aag ggc	720
Asn Leu Pro Val Arg Pro Asp Gly Thr Leu Asn Ser Phe Val Lys Gly	
225 230 235 240	
tgt ctc act ctg cca gac caa caa aaa ctg aga ctg aag tcg cca gtc	768
Cys Leu Thr Leu Pro Asp Gln Gln Lys Leu Arg Leu Lys Ser Pro Val	
245 250 255	
ctg agg aag cag gct tgc ccc cag tgg aaa cac tca ttt gtc ttc agt	816
Leu Arg Lys Gln Ala Cys Pro Gln Trp Lys His Ser Phe Val Phe Ser	
260 265 270	
ggc gta acc cca gct cag ctg agg cag tcg agc ttg gag tta act gtc	864
Gly Val Thr Pro Ala Gln Leu Arg Gln Ser Ser Leu Glu Leu Thr Val	
275 280 285	
tgg gat cag gcc ctc ttt gga atg aat gac cgc ttg ctt gga gga acc	912
Trp Asp Gln Ala Leu Phe Gly Met Asn Asp Arg Leu Leu Gly Gly Thr	
290 295 300	
aga ctt ggt tca aag gga gac aca gct gtt ggc ggg gat gca tgc tca	960
Arg Leu Gly Ser Lys Gly Asp Thr Ala Val Gly Gly Asp Ala Cys Ser	
305 310 315 320	
cta tcg aag ctc cag tgg cag aaa gtc ctt tcc agc ccc aat cta tgg	1008
Leu Ser Lys Leu Gln Trp Gln Lys Val Leu Ser Ser Pro Asn Leu Trp	
325 330 335	

```

aca gac atg act ctt gtc ctg cac tgacatgaag gcctcaaggt tccaggttgc 1062
Thr Asp Met Thr Leu Val Leu His
340

agcaggcgtg aggcactgtg cgtctgcaga ggggctacga accaggtgca ggggccagc 1122

tggagacccc tttgacctg agcagctctcc atctgcgggc ctgtcccatg gcttaaccgc 1182

ctattggtat ctgtgtatat ttacgttaaa cacaattatg ttacctaagc ctctggtggg 1242

ttatctcctc tttgagatgt agaaaatggc cagattttaa taaacgttgt tacccatgaa 1302

aaaaaaaaa a 1313

```

<210> 2

<211> 344

<212> PRT

<213> Homo sapiens

<400> 2

```

Gln Asn Leu Pro Ser Ser Pro Ala Pro Ser Thr Ile Phe Ser Gly Gly
 1           5           10           15

Phe Arg His Gly Ser Leu Ile Ser Ile Asp Ser Thr Cys Thr Glu Met
           20           25           30

Gly Asn Phe Asp Asn Ala Asn Val Thr Gly Glu Ile Glu Phe Ala Ile
 35           40           45

His Tyr Cys Phe Lys Thr His Ser Leu Glu Ile Cys Ile Lys Ala Cys
 50           55           60

Lys Asn Leu Ala Tyr Gly Glu Glu Lys Lys Lys Lys Cys Asn Pro Tyr
 65           70           75           80

Val Lys Thr Tyr Leu Leu Pro Asp Arg Ser Ser Gln Gly Lys Arg Lys
           85           90           95

Thr Gly Val Gln Arg Asn Thr Val Asp Pro Thr Phe Gln Glu Thr Leu
 100           105           110

```

Lys Tyr Gln Val Ala Pro Ala Gln Leu Val Thr Arg Gln Leu Gln Val  
 115 120 125

Ser Val Trp His Leu Gly Thr Leu Ala Arg Arg Val Phe Leu Gly Glu  
 130 135 140

Val Ile Ile Ser Leu Ala Thr Trp Asp Phe Glu Asp Ser Thr Thr Gln  
 145 150 155 160

Ser Phe Arg Trp His Pro Leu Arg Ala Lys Ala Glu Lys Tyr Glu Asp  
 165 170 175

Ser Val Pro Gln Ser Asn Gly Glu Leu Thr Val Arg Ala Lys Leu Val  
 180 185 190

Leu Pro Ser Arg Pro Arg Lys Leu Gln Glu Ala Gln Glu Gly Thr Asp  
 195 200 205

Gln Pro Ser Leu His Gly Gln Leu Cys Leu Val Val Leu Gly Ala Lys  
 210 215 220

Asn Leu Pro Val Arg Pro Asp Gly Thr Leu Asn Ser Phe Val Lys Gly  
 225 230 235 240

Cys Leu Thr Leu Pro Asp Gln Gln Lys Leu Arg Leu Lys Ser Pro Val  
 245 250 255

Leu Arg Lys Gln Ala Cys Pro Gln Trp Lys His Ser Phe Val Phe Ser  
 260 265 270

Gly Val Thr Pro Ala Gln Leu Arg Gln Ser Ser Leu Glu Leu Thr Val  
 275 280 285

Trp Asp Gln Ala Leu Phe Gly Met Asn Asp Arg Leu Leu Gly Gly Thr  
 290 295 300

Arg Leu Gly Ser Lys Gly Asp Thr Ala Val Gly Gly Asp Ala Cys Ser  
 305 310 315 320

Leu Ser Lys Leu Gln Trp Gln Lys Val Leu Ser Ser Pro Asn Leu Trp  
 325 330 335

Thr Asp Met Thr Leu Val Leu His  
 340

<210> 3

<211> 471

<212> ADN

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(471)

<400> 3

```
gga agt cca gca ggt aga tca atc tac aac agc ttt tat gtg tat tgc 48
Gly Ser Pro Ala Gly Arg Ser Ile Tyr Asn Ser Phe Tyr Val Tyr Cys
 1             5             10             15

aaa ggc ccc tgt caa aga gtg cag ccg gga aaa ctc agg gta cag tgc 96
Lys Gly Pro Cys Gln Arg Val Gln Pro Gly Lys Leu Arg Val Gln Cys
          20             25             30

agc acc tgc agg cag gca acg ctc acc ttg acc cag ggt cca tct tgc 144
Ser Thr Cys Arg Gln Ala Thr Leu Thr Leu Thr Gln Gly Pro Ser Cys
          35             40             45

tgg gat gat gtt tta att cca aac cgg atg agt ggt gaa tgc caa tcc 192
Trp Asp Asp Val Leu Ile Pro Asn Arg Met Ser Gly Glu Cys Gln Ser
          50             55             60

cca cac tgc cct ggg act agt gca gaa ttt ttc ttt aaa tgt gga gca 240
Pro His Cys Pro Gly Thr Ser Ala Glu Phe Phe Phe Lys Cys Gly Ala
          65             70             75             80

cac ccc acc tct gac aag gaa aca tca gta gct ttg cac ctg atc gca 288
His Pro Thr Ser Asp Lys Glu Thr Ser Val Ala Leu His Leu Ile Ala
          85             90             95

aca aat agt cgg aac atc act tgc att acg tgc aca gac gtc agg agc 336
Thr Asn Ser Arg Asn Ile Thr Cys Ile Thr Cys Thr Asp Val Arg Ser
          100             105             110

ccc gtc ctg gtt ttc cag tgc aac tcc cgc cac gtg att tgc tta gac 384
Pro Val Leu Val Phe Gln Cys Asn Ser Arg His Val Ile Cys Leu Asp
          115             120             125
```

tgt ttc cac tta tac tgt gtg aca aga ctc aat gat cgg cag ttt gtt 432  
 Cys Phe His Leu Tyr Cys Val Thr Arg Leu Asn Asp Arg Gln Phe Val  
 130 135 140

cac gac cct caa ctt ggc tac tcc ctg cct tgt gtg tag 471  
 His Asp Pro Gln Leu Gly Tyr Ser Leu Pro Cys Val  
 145 150 155

<210> 4

<211> 156

<212> PRT

<213> Homo sapiens

<400> 4

Gly Ser Pro Ala Gly Arg Ser Ile Tyr Asn Ser Phe Tyr Val Tyr Cys  
 1 5 10 15  
 Lys Gly Pro Cys Gln Arg Val Gln Pro Gly Lys Leu Arg Val Gln Cys  
 20 25 30  
 Ser Thr Cys Arg Gln Ala Thr Leu Thr Leu Thr Gln Gly Pro Ser Cys  
 35 40 45  
 Trp Asp Asp Val Leu Ile Pro Asn Arg Met Ser Gly Glu Cys Gln Ser  
 50 55 60  
 Pro His Cys Pro Gly Thr Ser Ala Glu Phe Phe Phe Lys Cys Gly Ala  
 65 70 75 80  
 His Pro Thr Ser Asp Lys Glu Thr Ser Val Ala Leu His Leu Ile Ala  
 85 90 95  
 Thr Asn Ser Arg Asn Ile Thr Cys Ile Thr Cys Thr Asp Val Arg Ser  
 100 105 110  
 Pro Val Leu Val Phe Gln Cys Asn Ser Arg His Val Ile Cys Leu Asp  
 115 120 125  
 Cys Phe His Leu Tyr Cys Val Thr Arg Leu Asn Asp Arg Gln Phe Val  
 130 135 140  
 His Asp Pro Gln Leu Gly Tyr Ser Leu Pro Cys Val  
 145 150 155

<210> 5

<211> 27

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: Oligonucleótido

<400> 5

ttaagaattc ggaagtccag caggtag 27

<210> 6

<211> 29

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: Oligonucleótido

<400> 6

ataggatcc ctacacacaa ggcagggag 29

<210> 7

<211> 19

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: Oligonucleótido

<400> 7

gcgtttggaa tcactacag 19

<210> 8

<211> 17

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: Oligonucleótido

<400> 8

ggctctcgggtg tggcatc 17

<210> 9

<211> 18

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: Oligonucleótido

<400> 9

ccgcttgctt ggaggaac 18

<210> 10

<211> 18

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: Oligonucleótido

<400> 10

cgtattfctc cgccttgg 18

<210> 11

<211> 28

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: Oligonucleótido

<400> 11

aatagctcga gtcagtgcag gacaagag 28

<210> 12

<211> 2347

<212> ADN

<213> Homo sapiens

<400> 12

---

ggccttgggg cactgagga tgccagttct gcctgttcat ctggaacctg gatctaagga 60  
gggaagaggc gttgccctg ctggcatagt caggtaccag cccagccagg tattgaacgg 120  
gctgagcttt tcatgatggt tcctgctgac ctggaaacat cttaaattga agggcgtgag 180  
cgcttggtcc atgcagtgaa gctcttccaa cctgggtcaa cgaaaacgga gaagaaatgg 240  
cccaagaaat agatctgagt gctctcaagg agttagaacg cgaggccatt ctccaggtcc 300  
tgtaccgaga ccaggcggtt caaaacacag aggaggagag gacacggaaa ctgaaaacac 360  
acctgcagca tctccggtgg aaaggagcga agaacacgga ctgggagcac aaagagaagt 420

```

gctgtgcgcg ctgccagcag gtgctggggg tectgtgca cgggggcgcc gtgtgccggg 480
gctgcagcca ccgcggtgtgt gccagtgcc gagtgttctt gagggggacc catgcctgga 540
agtgcacggg gtgcttcgag gacaggaatg tcaaaataaa aactggagaa tggttctatg 600
aggaacgagc caagaaatth ccaactggag gcaaacatga gacagttgga gggcagctct 660
tgcaatctta tcagaagctg agcaaaatth ctgtggttcc tcctactcca cctcctgtca 720
gcgagagcca gtgcagccgc agtcctggca ggttacagga atttggtcag tttagaggat 780
ttaataagtc cgtggaaaat ttgtttctgt ctcttgctac ccacgtgaaa aagctctcca 840
aatcccagaa tgatatgact tctgagaagc atcttctcgc cacgggcccc aggcagtgtg 900
tgggacagac agagagacgg agccagtctg aactgctggg caacgtcacc accaggaagg 960
tcagtgcacc agatattctg aaacctctca atcaagagga tcccaaatgc tctactaacc 1020
ctatthtgaa gcaacagaat ctccatcca gtccggcacc cagtaccata ttctctggag 1080
gtthtagaca cggagthta attagcattg acagcacctg tacagagatg ggcaatthg 1140
acaatgctaa tgtcactgga gaaatagaat ttgccattca ttattgcttc aaaaccatt 1200
ctthgaaat atgcatcaag gcctgtaaga acctgctca tggagaagaa aagaagaaa 1260
agtgcaatcc gtatgtgaag acctacctgt tgcccgacag atctcccag gaaagcgca 1320
agactggagt ccaaaggaac accgtggacc cgaccttca ggagacctg aagtatcagg 1380
tggcccctgc ccagctggtg acccggcagc tgcaagtctc ggtgtggcat ctgggcacgc 1440
tggcccggag agtgtthctt ggagaagtga tcattctctt ggccacgtgg gactthgaa 1500
acagcacaac acagtcctc cgctggcatc cgctccgggc caaggcggag aaatacgaag 1560
acagcgttcc tcagagtaat ggagagctca cagtccgggc taagctggtt ctcccttcac 1620
ggcccagaaa actccaagag gctcaagaag ggacagatca gccatcactt catggtcaac 1680
thtghthgt agtgctagga gccagaath tacctgtgcg gccagatggc acctgaaact 1740
cattghthaa gggctgtctc actctgccag accaacaata actgagactg aagtgcgccag 1800
tcttgaggaa gcaggcttgc cccagtgga aactctatt tgtcttcagt ggcgtaacct 1860
cagctcagct gaggcagctg agcttggagt taactgtctg ggatcaggcc ctctthggaa 1920
tgaacgaccg cttgcttggg ggaaccagac ttggttcaaa gggagacaca gctgthggcg 1980
gggatgcatg ctcaaatcg aagctccagt ggcagaaagt cctthccagc ccaatctat 2040
ggacagacat gactcttgtc ctgcaactgac atgaaggcct caaggthcca ggttgcagca 2100
ggcgtgaggc actgtgcgtc tgcaagggg ctacgaacca ggtgcagggt cccagctgga 2160
gaccctthg acctgagca gtctccatct gcggccctgt cccatggctt aaccgcctat 2220
tggatctgt gtatthtac gthaaacaca attatgthac ctaagcctct ggtgggtht 2280
ctcctctthg agatgtagaa aatggccaga ththtaataa cgttghthacc catgaaaaaa 2340
aaaaaaa 2347

```

<210> 13

<211> 610

<212> PRT

<213> Homo sapiens

<400> 13

Met Ala Gln Glu Ile Asp Leu Ser Ala Leu Lys Glu Leu Glu Arg Glu  
1                    5                    10                    15

Ala Ile Leu Gln Val Leu Tyr Arg Asp Gln Ala Val Gln Asn Thr Glu  
                  20                    25                    30

Glu Glu Arg Thr Arg Lys Leu Lys Thr His Leu Gln His Leu Arg Trp  
                  35                    40                    45

Lys Gly Ala Lys Asn Thr Asp Trp Glu His Lys Glu Lys Cys Cys Ala  
                  50                    55                    60

Arg Cys Gln Gln Val Leu Gly Phe Leu Leu His Arg Gly Ala Val Cys  
65                    70                    75                    80

Arg Gly Cys Ser His Arg Val Cys Ala Gln Cys Arg Val Phe Leu Arg  
                  85                    90                    95

Gly Thr His Ala Trp Lys Cys Thr Val Cys Phe Glu Asp Arg Asn Val  
                  100                    105                    110

Lys Ile Lys Thr Gly Glu Trp Phe Tyr Glu Glu Arg Ala Lys Lys Phe  
                  115                    120                    125

Pro Thr Gly Gly Lys His Glu Thr Val Gly Gly Gln Leu Leu Gln Ser  
130                    135                    140

Tyr Gln Lys Leu Ser Lys Ile Ser Val Val Pro Pro Thr Pro Pro Pro  
145                    150                    155                    160

Val Ser Glu Ser Gln Cys Ser Arg Ser Pro Gly Arg Leu Gln Glu Phe  
                  165                    170                    175

Gly Gln Phe Arg Gly Phe Asn Lys Ser Val Glu Asn Leu Phe Leu Ser  
                  180                    185                    190

Leu Ala Thr His Val Lys Lys Leu Ser Lys Ser Gln Asn Asp Met Thr  
                  195                    200                    205

Ser Glu Lys His Leu Leu Ala Thr Gly Pro Arg Gln Cys Val Gly Gln  
210                    215                    220

Thr Glu Arg Arg Ser Gln Ser Asp Thr Ala Val Asn Val Thr Thr Arg  
 225 230 235 240

Lys Val Ser Ala Pro Asp Ile Leu Lys Pro Leu Asn Gln Glu Asp Pro  
 245 250 255

Lys Cys Ser Thr Asn Pro Ile Leu Lys Gln Gln Asn Leu Pro Ser Ser  
 260 265 270

Pro Ala Pro Ser Thr Ile Phe Ser Gly Gly Phe Arg His Gly Ser Leu  
 275 280 285

Ile Ser Ile Asp Ser Thr Cys Thr Glu Met Gly Asn Phe Asp Asn Ala  
 290 295 300

Asn Val Thr Gly Glu Ile Glu Phe Ala Ile His Tyr Cys Phe Lys Thr  
 305 310 315 320

His Ser Leu Glu Ile Cys Ile Lys Ala Cys Lys Asn Leu Ala Tyr Gly  
 325 330 335

Glu Glu Lys Lys Lys Lys Cys Asn Pro Tyr Val Lys Thr Tyr Leu Leu  
 340 345 350

Pro Asp Arg Ser Ser Gln Gly Lys Arg Lys Thr Gly Val Gln Arg Asn  
 355 360 365

Thr Val Asp Pro Thr Phe Gln Glu Thr Leu Lys Tyr Gln Val Ala Pro  
 370 375 380

Ala Gln Leu Val Thr Arg Gln Leu Gln Val Ser Val Trp His Leu Gly  
 385 390 395 400

Thr Leu Ala Arg Arg Val Phe Leu Gly Glu Val Ile Ile Pro Leu Ala  
 405 410 415

Thr Trp Asp Phe Glu Asp Ser Thr Thr Gln Ser Phe Arg Trp His Pro  
 420 425 430

Leu Arg Ala Lys Ala Glu Lys Tyr Glu Asp Ser Val Pro Gln Ser Asn  
 435 440 445

Gly Glu Leu Thr Val Arg Ala Lys Leu Val Leu Pro Ser Arg Pro Arg  
 450 455 460

Lys Leu Gln Glu Ala Gln Glu Gly Thr Asp Gln Pro Ser Leu His Gly  
 465 470 475 480

Gln Leu Cys Leu Val Val Leu Gly Ala Lys Asn Leu Pro Val Arg Pro  
 485 490 495

Asp Gly Thr Leu Asn Ser Phe Val Lys Gly Cys Leu Thr Leu Pro Asp  
 500 505 510

Gln Gln Lys Leu Arg Leu Lys Ser Pro Val Leu Arg Lys Gln Ala Cys  
 515 520 525

Pro Gln Trp Lys His Ser Phe Val Phe Ser Gly Val Thr Pro Ala Gln  
 530 535 540

Leu Arg Gln Ser Ser Leu Glu Leu Thr Val Trp Asp Gln Ala Leu Phe  
 545 550 555 560

Gly Met Asn Asp Arg Leu Leu Gly Gly Thr Arg Leu Gly Ser Lys Gly  
 565 570 575

Asp Thr Ala Val Gly Gly Asp Ala Cys Ser Gln Ser Lys Leu Gln Trp  
 580 585 590

Gln Lys Val Leu Ser Ser Pro Asn Leu Trp Thr Asp Met Thr Leu Val  
 595 600 605

Leu His  
 610

<210> 14  
 <211> 1648  
 <212> ADN  
 <213> Homo sapiens

<400> 14

```
gaaatcatgc ccctcgtaga gcagcaggtc caagcagggc tgctggctat ttttccaaaa 60
agtgaggcag tttaaaaaaaa aggcggagaa ctagaattat agaataatgg cacattttgt 120
gtattttgtaa aactaacggc ttgcatggtt cacaacccat ttcttatgcc tgggttttcc 180
ttggcagcaa aatttctgtg gttcctccta ctccacctcc tgtcagcgag agccagtgca 240
gccgcagtcc tggcaggaag gtcagtgcac cagatattct gaaacctctc aatcaagagg 300
atcccaaatg ctctactaac cctatthtga agcaacagaa tctcccatcc agtccggcac 360
ccagtaccat attctctgga ggttttagac acggaagttt aattagcatt gacagcacct 420
gtacagagat gggcaattht gacaatgcta atgtcactgg agaaatagaa tttgccattc 480
attattgctt caaaacccat tcttttagaaa tatgcatcaa ggctgtgtaag aaccttgect 540
atggagaaga aaagaagaaa aagtgcaatc cgtatgtgaa gacctacctg ttgcccagaca 600
gatcctccca gggaaagcgc aagactggag tccaaaggaa caccgtggac ccgacctttc 660
aggagacctt gaagtatcag gtggcccttg cccagctggt gaccggcag ctgcaggtct 720
cgggtgtggca tctgggcacg ctggcccgga gagtgtttct tggagaagtg atcattcctc 780
tggccacgtg ggactttgaa gacagcacia cacagtcctt ccgctggcat ccgctccggg 840
ccaaggcgga gaaatacгаа gacagcgttc ctcaagtaa tggagagctc acagtccggg 900
ctaagctggt tctcccttca cggcccagaa aactccaaga ggctcaagaa gggacagatc 960
agccatcact tcatggtcaa ctttgthtgg tagtgctagg agccaagaat ttacctgtgc 1020
ggccagatgg caccttgaac tcatttgta agggctgtct cactctgcca gaccaaaaa 1080
aactgagact gaagtgcgca gtcttgagga agcaggcttg cccccagtgg aaactcat 1140
ttgtcttcag tggcgtaacc ccagctcagc tgaggcagtc gagcttggag ttaactgtct 1200
gggatcaggc cctctthtgg atgaacgacc gcttgcttg aggaaccaga cttggttcaa 1260
agggagacac agctgttggc ggggatgcat gctcacaatc gaagctccag tggcagaaag 1320
tcctttccag cccaatcta tggacagaca tgactcttgt cctgcaactga catgaaggcc 1380
tcaaggttcc aggttgacgc aggcgtgagg cactgtgcgt ctgcagaggg gctacgaacc 1440
aggtgcaggg tcccagctgg agacccttt gaccttgagc agtctccatc tggggcctg 1500
tcccatggct taaccgccta ttggtatctg tgtatattta cgttaaacac aattatgta 1560
cctaagctc tggtgggtta tctcctctt gagatgtaga aaatggccag attttaataa 1620
acgttggtac ccatgaaaaa aaaaaaaaa 1648
```

<210> 15

<211> 313

<212> PRT

<213> Homo sapiens

<400> 15

```
Met Gly Asn Phe Asp Asn Ala Asn Val Thr Gly Glu Ile Glu Phe Ala
  1             5             10             15
Ile His Tyr Cys Phe Lys Thr His Ser Leu Glu Ile Cys Ile Lys Ala
          20             25             30
```

Cys Lys Asn Leu Ala Tyr Gly Glu Glu Lys Lys Lys Lys Cys Asn Pro  
 35 40 45

Tyr Val Lys Thr Tyr Leu Leu Pro Asp Arg Ser Ser Gln Gly Lys Arg  
 50 55 60

Lys Thr Gly Val Gln Arg Asn Thr Val Asp Pro Thr Phe Gln Glu Thr  
 65 70 75 80

Leu Lys Tyr Gln Val Ala Pro Ala Gln Leu Val Thr Arg Gln Leu Gln  
 85 90 95

Val Ser Val Trp His Leu Gly Thr Leu Ala Arg Arg Val Phe Leu Gly  
 100 105 110

Glu Val Ile Ile Pro Leu Ala Thr Trp Asp Phe Glu Asp Ser Thr Thr  
 115 120 125

Gln Ser Phe Arg Trp His Pro Leu Arg Ala Lys Ala Glu Lys Tyr Glu  
 130 135 140

Asp Ser Val Pro Gln Ser Asn Gly Glu Leu Thr Val Arg Ala Lys Leu  
 145 150 155 160

Val Leu Pro Ser Arg Pro Arg Lys Leu Gln Glu Ala Gln Glu Gly Thr  
 165 170 175

Asp Gln Pro Ser Leu His Gly Gln Leu Cys Leu Val Val Leu Gly Ala  
 180 185 190

Lys Asn Leu Pro Val Arg Pro Asp Gly Thr Leu Asn Ser Phe Val Lys  
 195 200 205

---

Gly Cys Leu Thr Leu Pro Asp Gln Gln Lys Leu Arg Leu Lys Ser Pro  
 210 215 220

Val Leu Arg Lys Gln Ala Cys Pro Gln Trp Lys His Ser Phe Val Phe  
 225 230 235 240

Ser Gly Val Thr Pro Ala Gln Leu Arg Gln Ser Ser Leu Glu Leu Thr  
 245 250 255

Val Trp Asp Gln Ala Leu Phe Gly Met Asn Asp Arg Leu Leu Gly Gly  
 260 265 270

Thr Arg Leu Gly Ser Lys Gly Asp Thr Ala Val Gly Gly Asp Ala Cys  
275 280 285

Ser Gln Ser Lys Leu Gln Trp Gln Lys Val Leu Ser Ser Pro Asn Leu  
290 295 300

Trp Thr Asp Met Thr Leu Val Leu His  
305 310

<210> 16

<211> 19

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 16

ccagtctgc ctgtcatc 19

<210> 17

<211> 20

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 17

ttcaaaacac agaggaggag 20

<210> 18

<211> 20

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 18

**gaattggtc agtttagagg** 20

<210> 19

<211> 26

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 19

**ttctgggatt tggagagctt tttcac** 26

<210> 20

<211> 22

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 20

tctgtctgtc ccacacactg cc 22

<210> 21

<211> 19

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 21

gactggctcc gtctctctg 19

<210> 22

<211> 21

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 22

aagcaacaga atctccatc c 21

<210> 23

<211> 21

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 23

gcattgtcaa aattgccat c 21

<210> 24

<211> 20

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 24

aggcggagaa atacgaagac 20

<210> 25

<211> 22

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 25

gcagagtgag acagcccta ac 22

<210> 26  
<211> 24  
<212> ADN  
<213> Sequência artificial

<220>  
<223> Descrição da sequência artificial: oligonucleótido

<400> 26  
cttcctcagg actggcgact tcag 24

<210> 27  
<211> 24  
<212> ADN  
<213> Sequência artificial

<220>  
<223> Descrição da sequência artificial: oligonucleótido

<400> 27  
caagcggtcg ttcattcaa agag 24

<210> 28  
<211> 22  
<212> ADN  
<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 28

aagaggagat aaccaccag ag 22

<210> 29

<211> 20

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 29

agggctgctg gctattttc 20

<210> 30

<211> 19

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 30

taagaaatgg gttgtgaac 19

<210> 31

<211> 21

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 31

**aagcaacaga atctccatc c            21**

<210> 32

<211> 21

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 32

**gcattgtcaa aattgccat c            21**

<210> 33

<211> 20

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 33

**aggcggagaa atacgaagac** 20

<210> 34

<211> 22

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 34

**gcagagtgag acagcccta ac** 22

<210> 35

<211> 24

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 35

**cttcctcagg actggcgact tcag** 24

<210> 36

<211> 24

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 36

caagcggtcg ttcattcaa agag 24

<210> 37

<211> 22

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 37

aagaggagat aaccaccag ag 22

<210> 38

<211> 18

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 38

aatggaaggg cgtgacgc 18

<210> 39

<211> 21

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 39

cctcagcct gctgcaacct g 21

<210> 40

<211> 31

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 40

gcacgaattc atggccaag aaatagatct g 31

<210> 41

<211> 24

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial: oligonucleótido

<400> 41

ctgtcttcgt atttctccgc ctg 24

<210> 42

<211> 2347

<212> ADN

<213> Homo sapiens

<400> 42

ggccttgggg cactgagggg tgccagttct gcctgttcat ctggaacctg gatctaagga 60  
gggaagagggc gttgcccctg ctggcatagt caggtaccag cccagccagg tattgaacgg 120  
gctgagcttt tcatgatggt tctgctgac ctggaacat cttaaatgga agggcgtgag 180  
cgcttggtec atgcagtgaa gctcttcaa cctgggtcaa cgaaaacgga gaagaaatgg 240  
cccaagaaat agatctgagt gctctcaagg agttagaacg cgaggccatt ctccaggtcc 300  
tgtaccgaga ccaggcgggt caaaacacag aggaggagag gacacggaaa ctgaaaacac 360  
acctgcagca tctccgggtg aaaggagcga agaacacgga ctgggagcac aaagagaagt 420  
gctgtgcgcg ctgccagcag gtgctggggg tctgctgca cgggggcgcc gtgtgccggg 480  
gctgcagcca ccgcgtgtgt gccagtgcc gagtgttccct gagggggacc catgcctgga 540  
agtgcacggt gtgcttcgag gacaggaatg tcaaaataaa aactggagaa tggttctatg 600  
aggaacgagc caagaaattt ccaactggag gcaaacatga gacagttgga gggcagctct 660  
tgcaatctta tcagaagctg agcaaaattt ctgtggttcc tctactcca cctcctgtca 720  
gcgagagcca gtgcagccgc agtcctggca ggttacagga atttggtcag tttagaggat 780  
ttaataagtc cgtggaaaat ttgtttctgt ctcttctac ccacgtgaaa aagctctcca 840  
aatcccagaa tgatatgact tctgagaagc atcttctcgc cacgggcccc aggcagtgtg 900  
tgggacagac agagagacgg agccagtctg aactgagggt caacgtcacc accaggaagg 960  
tcagtgcacc agatattctg aaacctctca atcaagagga tcccaaatgc tctactaacc 1020  
ctattttgaa gcaacagaat ctcccatcca gtccggcacc cagtaccata ttctctggag 1080

```

gttttagaca cggaagttta attagcattg acagcacctg tacagagatg ggcaatthttg 1140
acaatgctaa tgtcactgga gaaatagaat ttgccattca ttattgcttc aaaaccatt 1200
ctttagaaat atgcatcaag gcctgtaaga accttgccca tggagaagaa aagaagaaaa 1260
agtgcaatcc gtatgtgaag acctacctgt tgcccgcacag atcctcccag ggaaagcgca 1320
agactggagt ccaaaggaac accgtggacc cgacctttca ggagaccttg aagtatcagg 1380
tggcccctgc ccagctggtg acccggcagc tgcaggtctc ggtgtggcat ctgggcacgc 1440
tggcccggag agtgtttctt ggagaagtga tcattcctct ggccacgtgg gactttgaag 1500
acagcacaac acagtecttc cgctggcatc cgctccgggc caaggcggag aaatacgaag 1560
acagcgttcc tcagagtaat ggagagctca cagtccgggc taagctgggt ctcccttcac 1620
ggcccagaaa actccaagag gctcaagaag ggacagatca gccatcactt catggtcaac 1680
tttgtttggt agtgctagga gccaagaatt tacctgtgcg gccagatggc accttgaact 1740
catttgtaa gggctgtctc actctgccag accaacaanaa actgagactg aagtgcagg 1800
tcctgaggaa gcaggcttgc cccagtgga aacactcatt tgtcttcagt ggcgtaacc 1860
cagctcagct gaggcagtcg agcttgaggt taactgtctg ggatcaggcc ctctttgaa 1920
tgaacgaccg cttgcttgga ggaaccagac ttggttcaaa gggagacaca gctgttggcg 1980
gggatgcatg ctcaaatcg aagctccagt ggcagaaagt cctttccagc cccaatctat 2040
ggacagacat gactcttgtc ctgcaactgac atgaaggcct caaggttcca ggttgagca 2100
ggcgtgaggc actgtgcgtc tgcagagggg ctacgaacca ggtgcagggt cccagctgga 2160
gaccccttg accttgagca gtctccatct gcggccctgt cccatggctt aaccgcctat 2220
tggtatctgt gtatatttac gttaaacaca attatgttac ctaagcctct ggtgggttat 2280
ctcctctttg agatgtagaa aatggccaga ttttaataaa cgttgttacc catgaaaaa 2340
aaaaaaa 2347

```

<210> 43

<211> 610

<212> PRT

<213> Homo sapiens

<400> 43

```

Met Ala Gln Glu Ile Asp Leu Ser Ala Leu Lys Glu Leu Glu Arg Glu
 1             5             10            15

Ala Ile Leu Gln Val Leu Tyr Arg Asp Gln Ala Val Gln Asn Thr Glu
          20            25            30

Glu Glu Arg Thr Arg Lys Leu Lys Thr His Leu Gln His Leu Arg Trp
          35            40            45

Lys Gly Ala Lys Asn Thr Asp Trp Glu His Lys Glu Lys Cys Cys Ala
          50            55            60

```

Arg Cys Gln Gln Val Leu Gly Phe Leu Leu His Arg Gly Ala Val Cys  
 65 70 75 80

Arg Gly Cys Ser His Arg Val Cys Ala Gln Cys Arg Val Phe Leu Arg  
 85 90 95

Gly Thr His Ala Trp Lys Cys Thr Val Cys Phe Glu Asp Arg Asn Val  
 100 105 110

Lys Ile Lys Thr Gly Glu Trp Phe Tyr Glu Glu Arg Ala Lys Lys Phe  
 115 120 125

Pro Thr Gly Gly Lys His Glu Thr Val Gly Gly Gln Leu Leu Gln Ser  
 130 135 140

Tyr Gln Lys Leu Ser Lys Ile Ser Val Val Pro Pro Thr Pro Pro Pro  
 145 150 155 160

Val Ser Glu Ser Gln Cys Ser Arg Ser Pro Gly Arg Leu Gln Glu Phe  
 165 170 175

Gly Gln Phe Arg Gly Phe Asn Lys Ser Val Glu Asn Leu Phe Leu Ser  
 180 185 190

Leu Ala Thr His Val Lys Lys Leu Ser Lys Ser Gln Asn Asp Met Thr  
 195 200 205

Ser Glu Lys His Leu Leu Ala Thr Gly Pro Arg Gln Cys Val Gly Gln  
 210 215 220

Thr Glu Arg Arg Ser Gln Ser Asp Thr Ala Val Asn Val Thr Thr Arg  
 225 230 235 240

Lys Val Ser Ala Pro Asp Ile Leu Lys Pro Leu Asn Gln Glu Asp Pro  
 245 250 255

Lys Cys Ser Thr Asn Pro Ile Leu Lys Gln Gln Asn Leu Pro Ser Ser  
 260 265 270

Pro Ala Pro Ser Thr Ile Phe Ser Gly Gly Phe Arg His Gly Ser Leu  
 275 280 285

Ile Ser Ile Asp Ser Thr Cys Thr Glu Met Gly Asn Phe Asp Asn Ala  
 290 295 300

Asn Val Thr Gly Glu Ile Glu Phe Ala Ile His Tyr Cys Phe Lys Thr  
 305 310 315 320

His Ser Leu Glu Ile Cys Ile Lys Ala Cys Lys Asn Leu Ala Tyr Gly  
 325 330 335

Glu Glu Lys Lys Lys Lys Cys Asn Pro Tyr Val Lys Thr Tyr Leu Leu  
 340 345 350

Pro Asp Arg Ser Ser Gln Gly Lys Arg Lys Thr Gly Val Gln Arg Asn  
 355 360 365

Thr Val Asp Pro Thr Phe Gln Glu Thr Leu Lys Tyr Gln Val Ala Pro  
 370 375 380

Ala Gln Leu Val Thr Arg Gln Leu Gln Val Ser Val Trp His Leu Gly  
 385 390 395 400

Thr Leu Ala Arg Arg Val Phe Leu Gly Glu Val Ile Ile Pro Leu Ala  
 405 410 415

Thr Trp Asp Phe Glu Asp Ser Thr Thr Gln Ser Phe Arg Trp His Pro  
 420 425 430

Leu Arg Ala Lys Ala Glu Lys Tyr Glu Asp Ser Val Pro Gln Ser Asn  
 435 440 445

Gly Glu Leu Thr Val Arg Ala Lys Leu Val Leu Pro Ser Arg Pro Arg  
 450 455 460

Lys Leu Gln Glu Ala Gln Glu Gly Thr Asp Gln Pro Ser Leu His Gly  
 465 470 475 480

Gln Leu Cys Leu Val Val Leu Gly Ala Lys Asn Leu Pro Val Arg Pro  
 485 490 495

Asp Gly Thr Leu Asn Ser Phe Val Lys Gly Cys Leu Thr Leu Pro Asp  
 500 505 510

Gln Gln Lys Leu Arg Leu Lys Ser Pro Val Leu Arg Lys Gln Ala Cys  
 515 520 525

Pro Gln Trp Lys His Ser Phe Val Phe Ser Gly Val Thr Pro Ala Gln  
530 535 540

Leu Arg Gln Ser Ser Leu Glu Leu Thr Val Trp Asp Gln Ala Leu Phe  
545 550 555 560

Gly Met Asn Asp Arg Leu Leu Gly Gly Thr Arg Leu Gly Ser Lys Gly  
565 570 575

Asp Thr Ala Val Gly Gly Asp Ala Cys Ser Gln Ser Lys Leu Gln Trp  
580 585 590

Gln Lys Val Leu Ser Ser Pro Asn Leu Trp Thr Asp Met Thr Leu Val  
595 600 605

Leu His  
610

<210> 44

<211> 1648

<212> ADN

<213> Homo sapiens

<400> 44

gaaatcatgc cctcgtaga gcagcaggtc caagcagggc tgctggctat ttttcaaaa 60  
agtgaggcag tttaaaaaaa aggcgagaa ctagaattat agaataatgg cacattttgt 120  
gtatttgtaa aactaacggc ttgcatgggt cacaacccat ttcttatgcc tgtgttttcc 180  
ttggcagcaa aatttctgtg gttcctccta ctccacctcc tgtcagcgag agccagtgca 240  
gccgcagtcc tggcaggaag gtcagtgcac cagatattct gaaacctctc aatcaagagg 300  
atcccaaatg ctctactaac cctattttga agcaacagaa tctcccatcc agtccggcac 360  
ccagtacatc attctctgga ggttttagac acggaagtct aattagcatt gacagcacct 420  
gtacagagat gggcaatctt gacaatgcta atgtcactgg agaaatagaa tttgccattc 480  
attattgctt caaaacccat tctttagaaa tatgcatcaa ggctgtgtaag aaccttgctc 540  
atggagaaga aaagaagaaa aagtgcaatc cgtatgtgaa gacctacctg ttgcccgaca 600  
gatcctccca gggaaagcgc aagactggag tccaaaggaa caccgtggac ccgaccttcc 660  
aggagacctt gaagtatcag gtggcccctg cccagctggg gaccggcag ctgcaggctc 720  
cgggtgtggca tctgggcacg ctggcccggg gagtgcttct tggagaagtg atcattcctc 780  
tggccacgtg ggactttgaa gacagcacia cacagtcctt ccgctggcat ccgctccggg 840  
ccaaggcggg gaaatacgaa gacagcgttc ctcaagagtaa tggagagctc acagtcgggg 900  
ctaagctggt tctcccttca cggcccagaa aactccaaga ggctcaagaa gggacagatc 960

```

agccatcact tcatggtaa ctttgtttg tagtgctagg agccaagaat ttacctgtgc 1020
ggccagatgg caccttgaac tcatttgta agggctgtct cactctgcca gaccaacaaa 1080
aactgagact gaagtcgcca gtctgagga agcaggcttg cccccagtgg aaacactcat 1140
ttgtcttcag tggcgtaacc ccagctcagc tgaggcagtc gagcttgag ttaactgtct 1200
gggatcaggc cctctttgga atgaacgacc gcttgcttgg aggaaccaga cttggttcaa 1260
agggagacac agctgttggc ggggatgcat gtcacaatc gaagctccag tggcagaaaag 1320
tcctttccag cccaatcta tggacagaca tgactcttgt cctgcaactga catgaaggcc 1380
tcaaggttcc aggttgacgc aggcgtgagg cactgtgcgt ctgcagaggg gctacgaacc 1440
aggtgcaggg tcccagctgg agacccttt gaccttgagc agtctccatc tgcggccctg 1500
tcccatggct taaccgccta ttggtatctg tgtatattta cgttaaacac aattatgtta 1560
cctaagcctc tgggtgggta tctctcttt gagatgtaga aaatggccag attttaataa 1620
acgttgttac ccatgaaaaa aaaaaaaaa . 1648

```

<210> 45

<211> 313

<212> PRT

<213> Homo sapiens

<400> 45

```

Met Gly Asn Phe Asp Asn Ala Asn Val Thr Gly Glu Ile Glu Phe Ala
  1             5             10             15

Ile His Tyr Cys Phe Lys Thr His Ser Leu Glu Ile Cys Ile Lys Ala
             20             25             30

Cys Lys Asn Leu Ala Tyr Gly Glu Glu Lys Lys Lys Lys Cys Asn Pro
             35             40             45

Tyr Val Lys Thr Tyr Leu Leu Pro Asp Arg Ser Ser Gln Gly Lys Arg
             50             55             60

Lys Thr Gly Val Gln Arg Asn Thr Val Asp Pro Thr Phe Gln Glu Thr
             65             70             75             80

Leu Lys Tyr Gln Val Ala Pro Ala Gln Leu Val Thr Arg Gln Leu Gln
             85             90             95

Val Ser Val Trp His Leu Gly Thr Leu Ala Arg Arg Val Phe Leu Gly
             100             105             110

Glu Val Ile Ile Pro Leu Ala Thr Trp Asp Phe Glu Asp Ser Thr Thr
             115             120             125

```

Gln Ser Phe Arg Trp His Pro Leu Arg Ala Lys Ala Glu Lys Tyr Glu  
130 135 140

Asp Ser Val Pro Gln Ser Asn Gly Glu Leu Thr Val Arg Ala Lys Leu  
145 150 155 160

Val Leu Pro Ser Arg Pro Arg Lys Leu Gln Glu Ala Gln Glu Gly Thr  
165 170 175

Asp Gln Pro Ser Leu His Gly Gln Leu Cys Leu Val Val Leu Gly Ala  
180 185 190

Lys Asn Leu Pro Val Arg Pro Asp Gly Thr Leu Asn Ser Phe Val Lys  
195 200 205

Gly Cys Leu Thr Leu Pro Asp Gln Gln Lys Leu Arg Leu Lys Ser Pro  
210 215 220

Val Leu Arg Lys Gln Ala Cys Pro Gln Trp Lys His Ser Phe Val Phe  
225 230 235 240

Ser Gly Val Thr Pro Ala Gln Leu Arg Gln Ser Ser Leu Glu Leu Thr  
245 250 255

Val Trp Asp Gln Ala Leu Phe Gly Met Asn Asp Arg Leu Leu Gly Gly  
260 265 270

---

Thr Arg Leu Gly Ser Lys Gly Asp Thr Ala Val Gly Gly Asp Ala Cys  
275 280 285

Ser Gln Ser Lys Leu Gln Trp Gln Lys Val Leu Ser Ser Pro Asn Leu  
290 295 300

Trp Thr Asp Met Thr Leu Val Leu His  
305 310

<210> 46

<211> 21

<212> ADN

<213> Sequência artificial

<220>

<223> Descrição da sequência artificial:

oligonucleótido

<400> 46

tcgtagagca gcaggtccaa g        21

Lisboa, 26 de Julho de 2010

## REIVINDICAÇÕES

1. Polipéptido interagindo com a Parkina, apresentando, em todo o seu comprimento uma identidade de, pelo menos, 95% com a sequência SEQ ID N°: 2.
2. Polipéptido apresentando uma das sequências SEQ ID N°: 13 e SEQ ID N°: 43.
3. Polipéptido apresentando uma das sequências SEQ ID N°: 15 e SEQ ID N°: 45.
4. Ácido nucleico codificando para um polipéptido de acordo com uma das reivindicações 1 a 3.
5. Ácido nucleico de acordo com a reivindicação 4, caracterizado por esta apresentar uma identidade de, pelo menos 95%, com a sequência SEQ ID N°: 1.
6. Ácido nucleico de acordo com a reivindicação 4, caracterizado por este apresentar uma das sequências SEQ ID N°: 12 e SEQ ID N°: 42.
7. Ácido nucleico de acordo com a reivindicação 4, apresentando uma das sequências SEQ ID N°: 14 e SEQ ID N°: 44.
8. Vector compreendendo um ácido nucleico de acordo com uma das reivindicações 4 a 7.

9. Vírus recombinante deficiente compreendendo um ácido nucleico de acordo com uma das reivindicações 4 a 7.
10. Composição farmacêutica compreendendo, pelo menos, um polipéptido de acordo com uma das reivindicações 1 a 3.
11. Composição farmacêutica compreendendo, pelo menos, um ácido nucleico de acordo com uma das reivindicações 4 a 7 ou um vector de acordo com uma das reivindicações 8 ou 9.
12. Composição de acordo com uma das reivindicações 10 ou 11 destinados ao tratamento de patologias neurodegenerativas.
13. Processo para o rastreio ou a caracterização de moléculas destinadas ao tratamento de patologias neurodegenerativas, compreendendo uma etapa de selecção de moléculas capazes de se ligarem à sequência SEQ ID N°: 2 ou a um seu fragmento.
14. Processo para o rastreio ou a caracterização de moléculas destinadas ao tratamento de patologia neurodegenerativas, compreendendo uma etapa da selecção de moléculas capazes de se ligarem a uma sequência seleccionada das sequências SEQ ID N°: 13, SEQ ID N°: 15, SEQ ID N°: 43 e SEQ ID N°: 45 ou um seu fragmento.
15. Processo de produção de um polipéptido de acordo com uma das reivindicações 1 a 3, compreendendo a cultura de uma célula contendo um ácido nucleico de acordo com uma das reivindicações 4 a 7 ou um vector de acordo com a reivindicação 8 ou 9, nas condições de expressão do referido ácido nucleico e a recuperação do composto peptídico produzido.

16. Célula contendo um ácido nucleico de acordo com uma das reivindicações 4 a 7 ou um vector de acordo com a reivindicação 8 ou 9.
17. Mamífero não humano compreendendo, nas suas células, um ácido nucleico de acordo com uma das reivindicações 4 a 7.
18. Anticorpos, fragmentos ou derivados dos anticorpos dirigidos contra um polipéptido de acordo com uma das reivindicações 1 a 3.

Lisboa, 26 de Julho de 2010

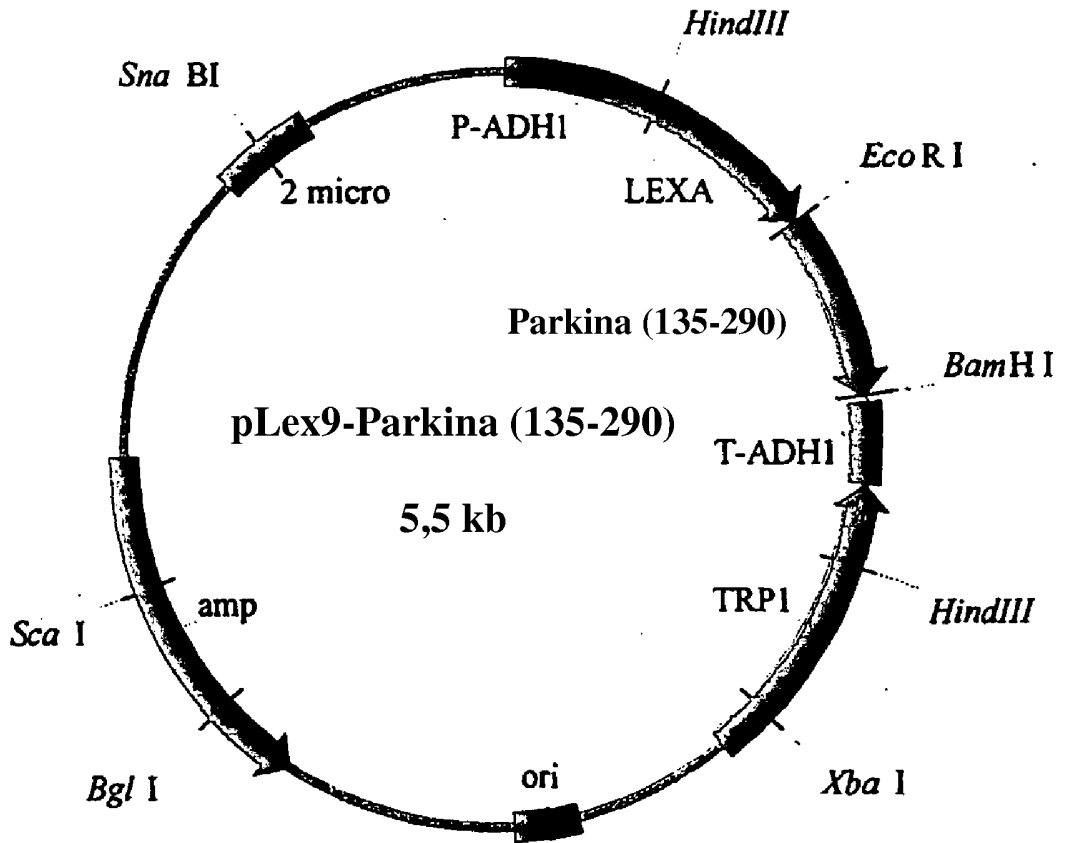


FIGURA 1

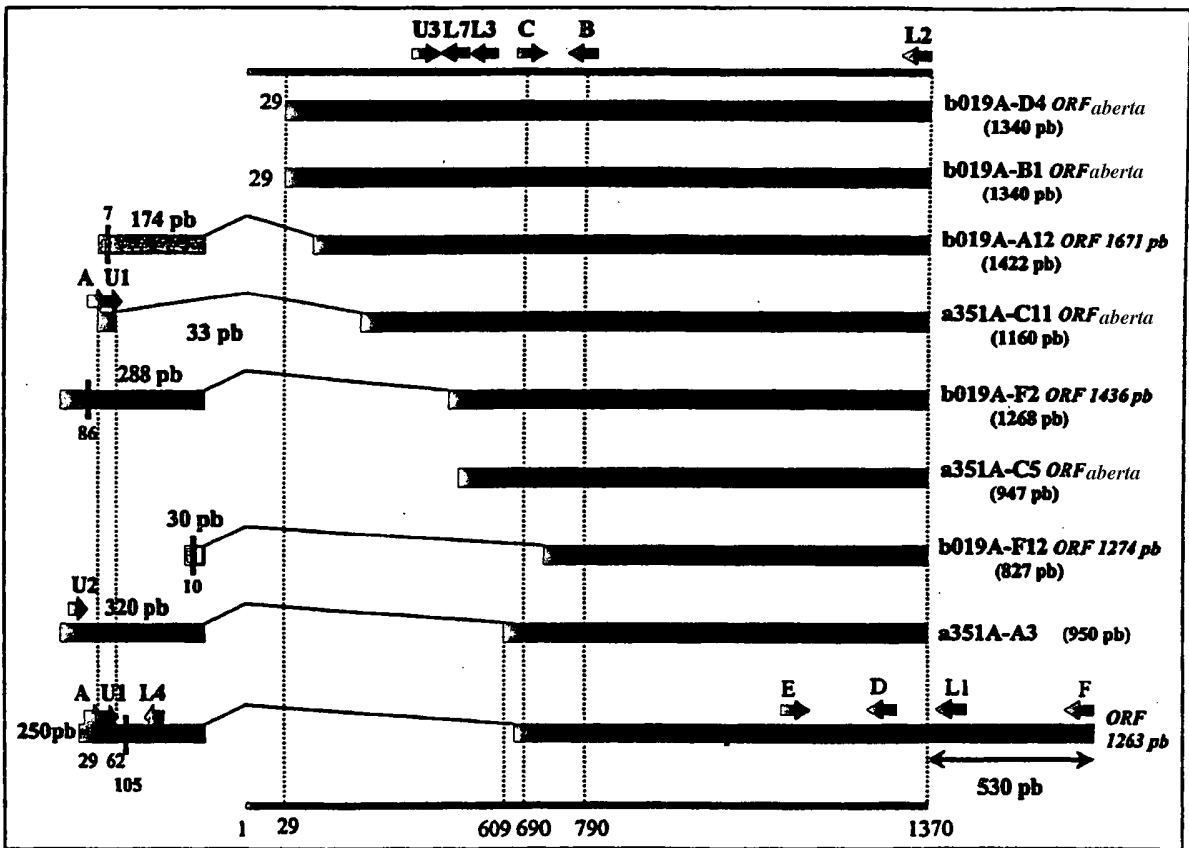


FIGURA 2

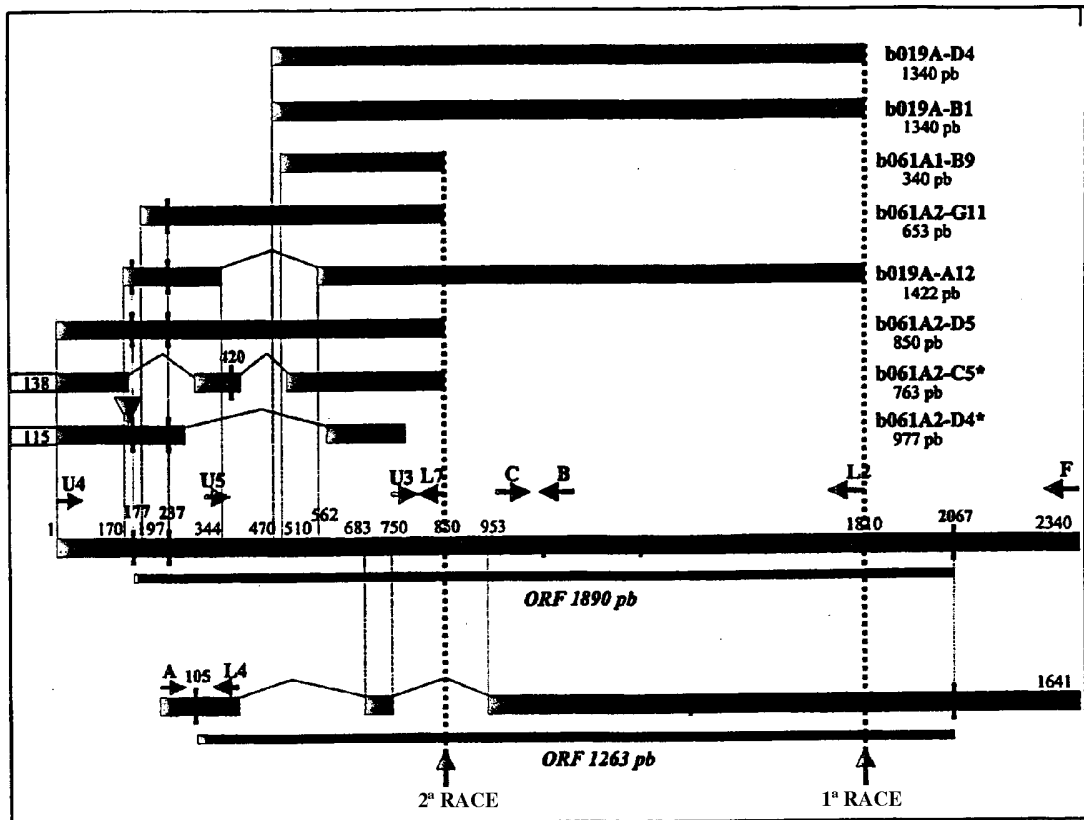


FIGURA 3

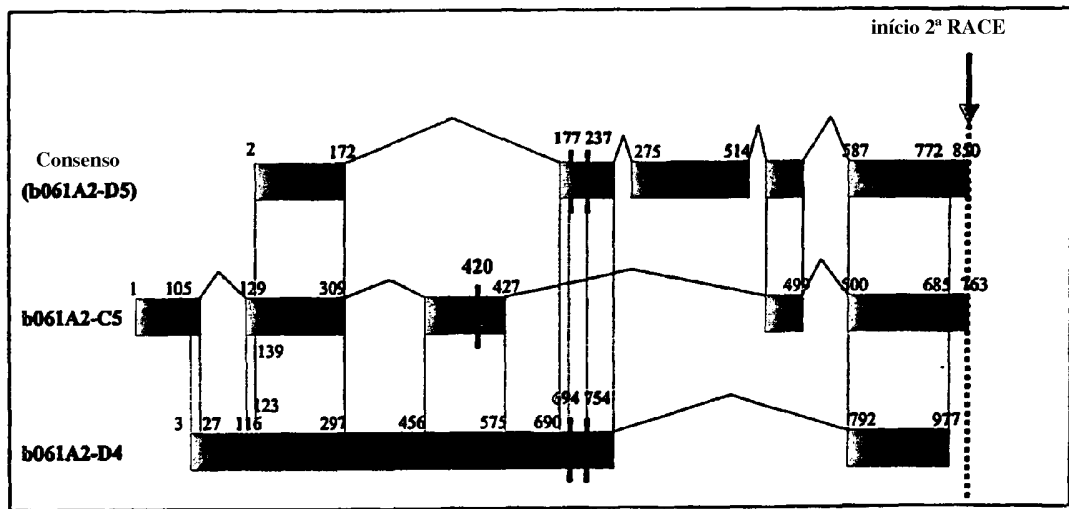


FIGURA 4

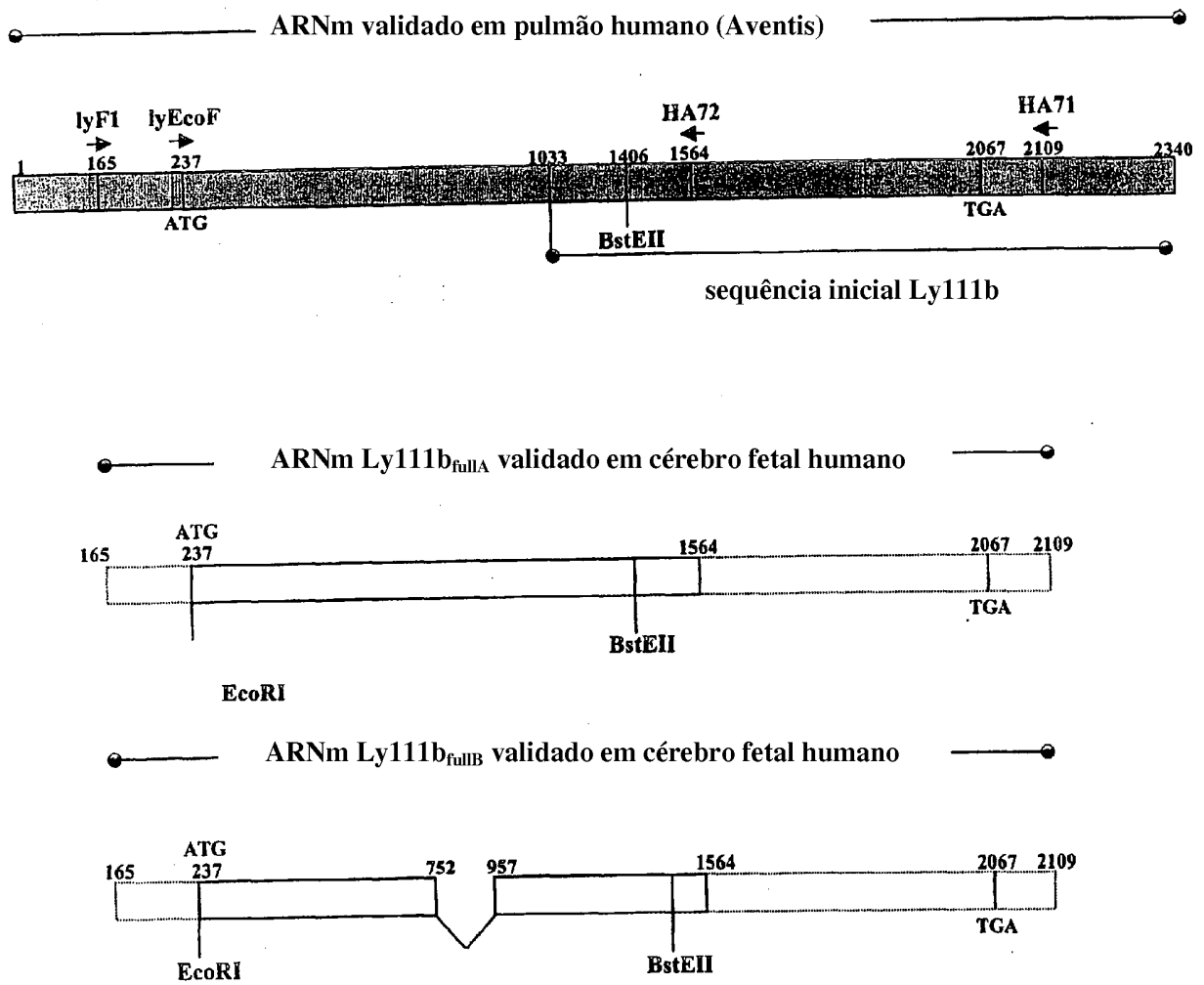


FIGURA 5

## Ly111b-fullA: o transcrito

AATGGAAGGGCGTGAGCGCTTGGTCCATGCAGTGAAGCTCTTCCAACCTGGGTCAACGAAAACG  
 GAGAAGAAATGGCCCAAGAAATAGATCTGAGTGTCTCAAGGAGTTAGAACGCGAGGCCATTCT  
 CCAGGTCCCTGTACCGGAGACCAGGCGGTTCAAAACACAGAGGAGGAGAGGACACGGAACTGAAA  
 ACACACCTGCAGCATCTCCGGTGGAAAGGAGCGAAGAACACGGACTGGGAGCACAAAGAGAAGT  
 GCTGTGCGCGCTGCCAGCAGGTGCTGGGGTTCCTGCTGCACCGGGGCGCCGTGTGCCGGGGCTG  
 CAGCCACCGCGTGTGTGCCAGTGCCGAGTGTTCCTGAGGGGGACCCATGCCTGGAAGTGCAGG  
 GTGTGCTTCGAGGACAGGAATGTCAAAATAAAAACTGGAGAATGGTTCTATGAGGAACGAGCCA  
 AGAAATTTCCAAC TGGAGGCAAAACATGAGACAGTTGGAGGGCAGCTCTTGCAATCTTATCAGAA  
 GCTGAGCAAAATTTCTGTGGTTCCTCCTACTCCACCTCCTGTGTCAGCGAGAGCCAGTGCAGCCGC  
 AGTCTTGGCAGGTTACAGGAATTTGGTCAAGTTTAGAGGATTTAATAAGTCCGTGGAAAATTTGT  
TTCTGTCTCTTGTACCCACGTGAAAAGCTCTCCAATCCAGAATGATATGACTTCTGAGAA  
GCATCTTCTCGCCACGGGCCCCAGGCAGTGTGTGGGACAGACAGAGAGACGGAGCCAGTCTGAC  
ACTGCGGTCAACGTCACCACCAGGAAGGTCAGTGCACCAGATATTCTGAAACCTCTCAATCAAG  
 AGGATCCCAAATGCTCTACTAACCCTATTTGAAGCAACAGAATCTCCCATCCAGTCCGGCACC  
 CAGTACCATATTCTCTGGAGGTTTAGACACGGAAGTTAATTAGCATTGACAGCACCTGTACA  
 GAGATGGCAATTTGACAATGCTAATGTCACCTGGAGAAATAGAATTTGCCATTCAATATTGCT  
 TCAAAACCCATTCTTAGAAATATGCATCAAGGCCTGTAAGAACCCTGCCTATGGAGAAGAAAA  
 GAAGAAAAAGTGAATCCGTATGTGAAGACCTACCTGTTGCCGACAGATCCTCCAGGGAAAG  
 CGCAAGACTGGAGTCCAAAGGAACACCGTGGACCCGACC'TTTCAGGAGACCTTGAAGTATCAGG  
 TGGCCCTGCCAGCTGGTGACCCGGCAGCTGCAGGTCTCGGTGTGGCATCTGGGCACGCTGGC  
 CCGGAGAGTGTCTTGGAGAAGTGATCATTCCTCTGGCCACGTGGGACTTTGAAGACAGCACA  
 ACACAGTCTTCCGCTGGCATCCGCTCCGGGCCAAGGCGGAGAAATACGAAGACAGCGTTCCTC  
 AGAGTAATGGAGAGCTCACAGTCCGGGCTAAGCTGGTTCCTCCCTCACGGCCAGAAAACCTCA  
 AGAGGCTCAAGAAGGGACAGATCAGCCATCACTTCATGGTCAACTTTGTTTGGTAGTGCTAGGA  
 GCCAAGAATTTACCTGTGCGGCCAGATGGCACCTTGAACCTATTGTTAAGGGCTGTCTCACTC  
 TGCCAGACCAACAAAACTGAGACTGAAGTCCGAGTCTGAGGAAGCAGGCTTGCCCCCAGTG  
 GAAACTCATTTGTCTTTCAGTGGCGTAACCCAGCTCAGCTGAGGCAGTCCAGCTTGGAGTTA  
 ACTGTCTGGGATCAGGCCCTCTTTGGAATGAACGACCGCTTGCTTGGAGGAACCAGACTTGGTT  
 CAAAGGGAGACACAGCTGTTGGCGGGGATGCATGCTCACAAATCGAAGCTCCAGTGGCAGAAAGT  
 CCTTCCAGCCCAATCTATGGACAGACATGACTCTTGTCTGCACTGACATGAAGGCCTCAAG  
 GTTCCAGGTTGCAGCAGGCGTGAGG

## pLy111b-fullA: a proteína

MAQEIDLSALKELEREAILQVLYRDQAVQNTTEEERTRKCLKHLQHLRWKGAKNTDWEHKEKCCARQVVLGFLLRG  
 AVCRGCSHRVCAQCRVFLRGTHAWKCTVCFEDRNVKIKTGEWFYERAKKFPPTGGKHETVGGQLLQSYQKLSKISVV  
 PPTPPPVSQCSRSFGRLEFGQFRGFNKSVENLFLSLATHVKKLSKSNQDMTSEKHLLATGPROCVGQTERRSQS  
 DTAVNVVTRKVSAPDILKPLNQEDPKCSTNPILKQQLPSSPAPSTIFSGGFRHGSLISIDSTCTEMGNFDNANVTG  
 EIEFAIHYCFKTHSLEICIKACKNLAYGEEKKKCNPVYVKTLLPDRSSQGRKRTGVQRNTVDPTFQETLKYQVAPA  
 QLVTRQLQVSVVHLGTLARRVFLGEVIIPLATWDFEDSTTQSFRWHPLRAKAKEYEDSVPQSNGELTVRAKLVLP  
 PRKLQEAQEGTDQPSLHGQLCLVVLGAKNLPVVRPDGTLNSFVKGCLTLPDQQLRLKSPVLRKQACQWKHSFVFSG  
 VTPAQLRQSSLELTVWDQALFGMNDRLGGTRLGSKGDTAVGGDACSQSKLQWQVLSNPLWDTMLVLH

FIGURA 6

## Ly11b-fullB: o transcrito

AATGGAAGGGCGTGAGCGCTTGGTCCATGCAGTGAAGCTCTTCCAACCTGGGTCAACGAAAA  
 CGGAGAAGAAATGGCCCAAGAAATAGATCTGAGTGCTCTCAAGGAGTTAGAACGCGAGGCCA  
 TTCTCCAGGTCTGTACCGAGACCAGGCGGTTCAAAACACAGAGGAGGAGAGGACACGGAAA  
 CTGAAAACACACCTGCAGCATCTCCGGTGGAAAGGAGCGAAGAACACGGACTGGGAGCACAA  
 AGAGAAGTGCTGTGCGCGCTGCCAGCAGGTGCTGGGGTTCTTGCTGCACCGGGGCGCCGTGT  
 GCCGGGGCTGCAGCCACCGCGTGTGTGCCAGTGCCGAGTGTTCTTGAGGGGGACCCATGCC  
 TGGAAGTGCACGGTGTGCTTCGAGGACAGGAATGTCAAAATAAAAACTGGAGAATGGTTCTA  
 TGAGGAACGAGCCAAGAAATTTCCAACCTGGAGGCAAACATGAGACAGTTGGAGGGCAGCTCT  
 TGCAATCTTATCAGAAGCTGAGCAAATTTCTGTGGTTCTCCTACTCCACCTCCTGTGAGC  
 GAGAGCCAGTGCAGCCGAGTCTGGCAGGAAGGTGAGTGCACCAGATATTCTGAAACCTCT  
 CAATCAAGAGGATCCCAAATGCTCTACTAACCTATTTTGAAGCAACAGAATCTCCCATCCA  
 GTCCGGCACCCAGTACCATATCTCTGGAGTTTTAGACACGGAAGTTAATTAGCATTGAC  
 AGCACCTGTACAGAGATGGGCAATTTGACAATGCTAATGTCACTGGAGAAATAGAATTTGC  
 CATTATTATTGCTTCAAACCCATCTTTAGAAATATGCATCAAGGCCTGTAAGAACCCTTG  
 CCTATGGAGAAGAAAAGAAGAAAAGTGCAATCCGTATGTGAAGACCTACCTGTTGCCCGAC  
 AGATCCTCCAGGGAAAGCGCAAGACTGGAGTCAAAGGAACACCGTGGACCCGACCTTTCA  
 GGAGACCTTGAAGTATCAGGTGGCCCTGCCAGCTGGTGACCCGGCAGCTGCAGGTCTCGG  
 TGTGGCATCTGGGCACGCTGGCCCGGAGAGTGTCTTGGAGAAGTGATCATTCTCTGGCC  
 ACGTGGGACTTTGAAGACAGCACAAACACAGTCTTCCGCTGGCATCCGCTCCGGGCCAAGGC  
 GGAGAAATACGAAGACAGCGTTCTCAGAGTAATGGAGAGCTCACAGTCCGGGCTAAGCTGG  
 TTCTCCCTTACGGCCCAAGAAACTCCAAGAGGCTCAAGAAGGGACAGATCAGCCATCATT  
 CATGGTCAACTTTGTTTGGTAGTGCTAGGAGCCAAGAATTTACCTGTGCGGCCAGATGGCAC  
 CTTGAACTCATTGTTAAGGGCTGTCTCACTCTGCCAGACCAAAAACTGAGACTGAAGT  
 CGCCAGTCTGAGGAAGCAGGCTTGCCCCAGTGGAAACACTCATTGTCTTCAAGTGGCGTA  
 ACCCCAGCTCAGCTGAGGCAGTCCAGCTTGGAGTTAACTGTCTGGGATCAGGCCCTCTTTGG  
 AATGAACGACCCTTGTCTGGAGGAACCAGACTTGGTTCAAAGGGAGACACAGCTGTTGGCG  
 GGGATGCATGCTCACAATCGAAGCTCCAGTGGCAGAAAGTCCTTTCCAGCCCCAATCTATGG  
 ACAGACATGACTCTTGTCTGCACTGACATGAAGGCCTCAAGGTTCCAGGTTGCAGCAGGCG  
 TGAGG

## pLy11b-fullB: a proteína

MAQEIDLSALKELEREAIIQVLYRDQAVQNTTEERTRKLLKTHLQHLRWKGAKNTDWEHKEKCCARCOQVVLGFLHRG  
 AVCRGCSHRVCAQCRVFLRGTHAWKCTVCFEDRNVKIKTGEWFYEERAKKFPTGGKHETVGGQLLQSYQKLSKISV  
 PPTPPVSESQCSRSPGRKVSAPDILKPLNQEDPKCSTNPIKQNLPSPPAPSTIFSGGFRHGLISIDSTCTEMG  
 NFDNANVTGEIEFAIHCFKTHSLEICIKACKNLAYGEEKKKCNPYKTYLLPDRSSQGRKRTGVQRNTVDPTTFQE  
 TLKYQVAPAQLVTRQLQVSVVHLGLTARRVFLGEVLIPLATWDFEDSTQSFWRHPLRAKAEKYEDSVQSNGLT  
 RAKLVLPSPRPRKLQEAQEGTDQPSLHGQLCLVVLGAKNLPVRPDGTLNSFVKGCLTLPDQQLRLKSPVLRKQACPQ  
 WKHSFVFSGVTPAQLRQSSLELTVWDQALFGMNDRLGGTRLSKSGDTAVGGDACSQSKLQWQKVLSSPNLWTDMTL  
 VLH

FIGURA 7

Figura 8a

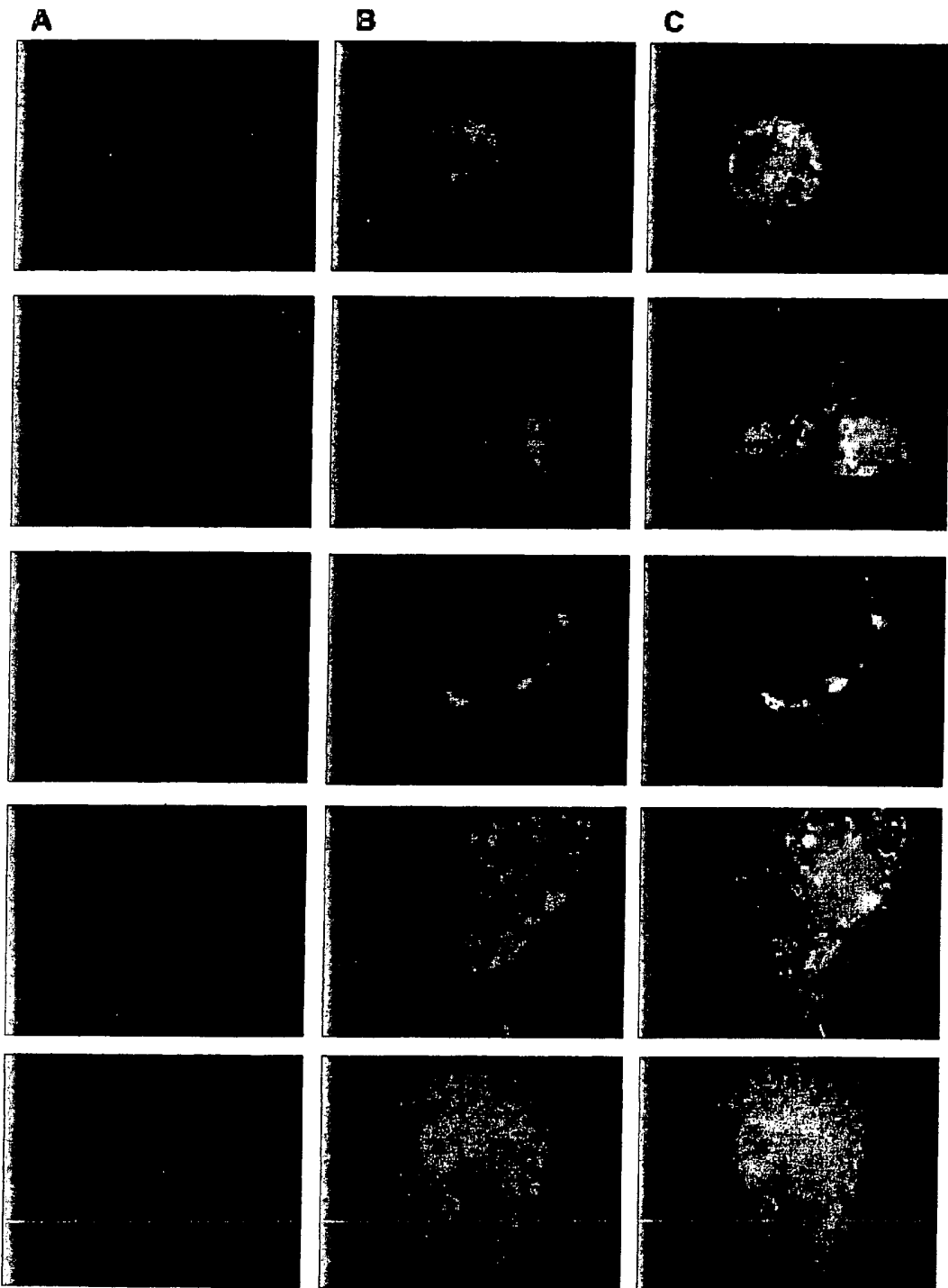
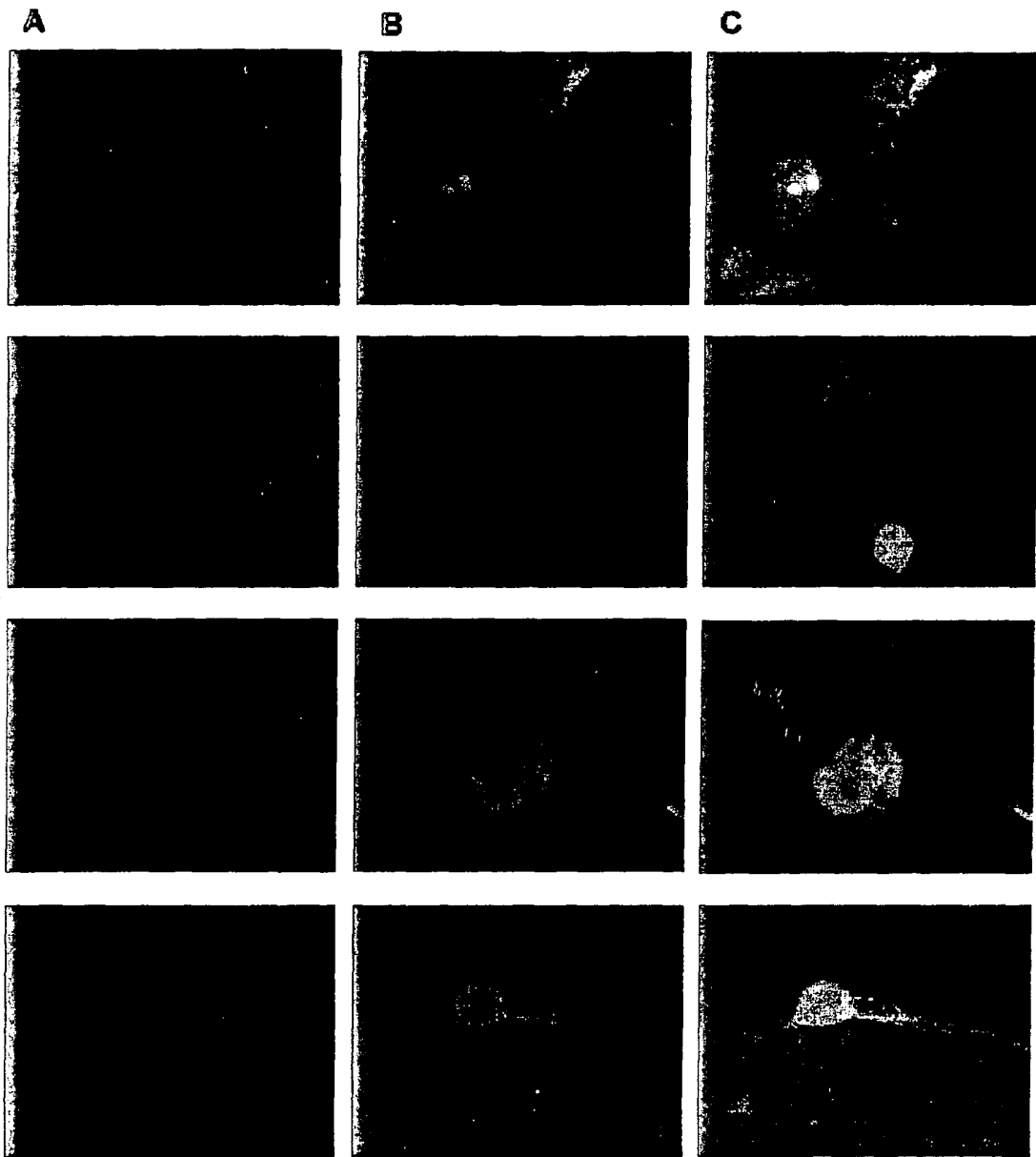


Figura 8b



GGCCTTGGGGCACTGAGGGATGCCAGTTCTGCCTGTTTCATCTGGAACCTGGATCTAAGGAGGGAAGAG  
 GCGTTGCCCCTGCTGGCATAGTCAGGTACCAGCCCAGCCAGGTATTGAACGGGCTGAGCTTTTCATGA  
 TGGTTCCTGCTGACCTGGAAACATCTTAAATGGAAGGGCGTGAGCGCTTGGTCCATGCAGTGAAGCTC  
 TTCCAACCTGGGTCAACGAAAACGGAGAAGAA[REDACTED]GCCAAGAAATAGATCTGAGTGTCTCAAGGAG  
 TTAGAACGCGAGGCCATTCTCCAGGTCTGTACCGAGACCAGGCGGTTCAAAACACAGAGGAGGAGAG  
 GACACGGAAACTGAAAACACACCTGCAGCATCTCCGGTGGAAAGGAGCGAAGAACACGGACTGGGAGC  
 ACAAAGAGAAGTGTGTGCGCGCTGCCAGCAGGTGTGGGGTTCCTGCTGCACCGGGGCGCCGTGTGC  
 CGGGCTGCAGCCACCGGTGTGTGCCAGTGGCCGAGTGTTCCTGAGGGGGACCCATGCCTGGAAGTG  
 CACGGTGTGCTTCGAGGACAGGAATGTCAAAATAAAAACCTGGAGAATGGTTCATGAGGAACGAGCCA  
 AGAAATTTCCAACCTGGAGGCAACATGAGACAGTTGGAGGGCAGCTTTCGCAATCTTATCAGAAGCTG  
 AGCAAAATTTCTGTGGTTCCTCCTACTCCACCTCCTGTCAGCGAGAGCCAGTGCAGCCGAGTCTTG  
 CAGGTTACAGGAATTTGGTCAGTTTAGAGGATTTAATAAGTCCGTGGAAAATTTGTTTCTGTCTCTTG  
 CTACCCACGTGAAAAGCTCTCCAAATCCAGAATGATATGACTTCTGAGAAGCATCTTCTCGCCACG  
 GGCCCCAGGCAGTGTGTGGGACAGACAGAGAGACGGAGCCAGTCTGACACTGCGGTCAACGTCACCAC  
 CAGGAAGTGCAGTGCACCAGATATTCTGAAACCTCTCAATCAAGAGGATCCCAAATGCTCTACTAACC  
 CTATTTTGAAGCAACAGAATCTCCCATCCAGTCCGGCACCCAGTACCATATTCTCTGGAGGTTTTAGA  
 CACGGAAGTTAATTAGCATTGACAGCACCTGTACAGAGATGGGCAATTTTGACAATGCTAATGTCAC  
 TGGAGAAATAGAATTTGCCATTCAATTATTGCTTCAAAAACCCATTCTTTAGAAAATATGCATCAAGGCCT  
 GTAAGAACCTTGCCATATGGAGAAGAAAAGAAGAAAAGTGCAATCCGTATGTGAAGACCTACCTGTTG  
 CCCGACAGATCCTCCAGGGAAAGCGCAAGACTGGAGTCAAAGGAACACCGTGGACCCGACCTTTCA  
 GGAGACCTGAAGTATCAGGTGGCCCTGCCAGTGGTGACCCGGCAGCTGCAGGCTCGGTGTGGC  
 ATCTGGGCACGCTGGCCCGGAGAGTGTTCCTTGGAGAAGTGATCATTCTCTGGCCACTCGGGACTTTT  
 GAAGACAGCACAAACACAGTCTTCCGCTGGCATCCGCTCCGGGCCAAGGCGGAGAAAATACGAAGACAG  
 CGTTCCTCAGAGTAATGGAGAGCTCACAGTCCGGGCTAAGCTGGTTCCTCCCTTACGGCCAGAAAAC  
 TCCAAGAGGCTCAAGAAGGGACAGATCAGCCATCACTTCATGGTCAACTTTGTTTGGTAGTGCTAGGA  
 GCCAAGAAATTTACCTGTGCGGCCAGATGGCACCTTGAACCTATTTGTTAAGGGCTGTCTACTCTGCC  
 AGACCAACAAAAACTGAGACTGAAGTCGCCAGTCTGAGGAAGCAGGCTTGGCCCCAGTGGAAACACT  
 CATTGTCTTCAAGTGGCGTAACCCAGCTCAGCTGAGGCAGTCGAGCTTGGAGTTAACTGTCTGGGAT  
 CAGGCCCTCTTTGGAATGAACGACCGCTTGCTTGGAGGAACCAGACTTGGTTCAAAGGGAGACACAGC  
 TGTTGGCGGGGATGCATGCTCACAATCGAAGCTCCAGTGGCAGAAAGTCTTTCCAGCCCCAATCTAT  
 GGACAGACATGACTCTTGCTCTGACTGACATGAAGGCCTCAAGGTTCCAGGTTGCAGCAGGCGTGAG  
 GCACTGTGCGTCTGCAGAGGGGCTACGAACCAGGTGCAGGGTCCCAGCTGGAGACCCCTTTGACCTTG  
 AGCAGTCTCCATCTGCGGCCCTGTCCCATGGCTTAACCGCCTATTGGTATCTGTGTATATTTACGTTA  
 AACACAATTATGTTACCTAAGCCTCTGGTGGGTTATCTCCTCTTTGAGATGTAGAAAATGGCCAGATT  
 TT[REDACTED]CGTTGTTACCCATGAAAAAAAAAAAAA

#### SEQUÊNCIA PROTEICA

MAQEIDLKELEREAILQVLYRDQAVQNTTEERTRKLKTHLQHLRWKGAKNTDWEHKEKCCARCQQ  
 VLGFLLRHGAVCRGCSHRVCAQCRVFLRGTHAWKCTVCFEDRNVKIKTGEWFYEERAKKFPPTGGKHET  
 VGGQLLQSYQKLSKISVVPPTPPPVSSESQCSRSPGRLOEFGQFRGFNKSVENLFLSLATHVKKLSKSQ  
 NDMTSEKHLLATGPRQCVGQTERRSQSDTAVNVTTTRKVSAPDILKPLNQEDPKCSTNPIILKQQNLPS  
 PAPSTIFSGGFRHGSLISIDSTCTEMGNFDNAVNTGEIEFAIHCFKTHSLEICIKACKNLAYGEEKK  
 KKNPYVKTYLLPDRSSQGRKKTGVQRNTVDPTFQETLKYQVAPQQLVTRQLQVSVWHLGTLARRVFL  
 GEVIIPLATWDFEDSTTQSFRWHPLRAKAEKYEDSVQSNGELTVRAKLVLPSPRKLQEAQEGTDQP  
 SLHGQLCLVVLGAKNLPVRPDGTLNSFVKGCLTLPDQQLRLKSPVLRKQACPOWKHSFVFSGVTPAQ  
 LRQSSLELTVWDQALFGMNDRLGGLRGLSGKGDYAVGGDACSQSKLQWQKVLSSPNLWTDMLVLH

FIGURA 9

GAAATCATGCCCTCGTAGAGCAGCAGGTCCAAGCAGGGCTGCTGGCTATTTTCCAAAAAG  
 TGAGGCAGTTTAAAAAAGGCGGAGAACTAGAATTATAGAATAATGGCACATTTTGTGTAT  
 TTGTAAACTAACGGCTTGCATGGTTACAACCCATTTCTTATGCCTGTGTTTTCTTGGCA  
 GCAAAATTTCTGTGGTTCTCCTACTCCACCTCCTGTCAGCGAGAGCCAGTGCAGCCGCAGT  
 CCTGGCAGGAAGGTGAGTGCACCAGATATTCTGAAACCTCTCAATCAAGAGGATCCCAAATG  
 CTCTACTAACCCATTTTTGAAGCAACAGAATCTCCCATCCAGTCCGGCACCCAGTACCATAT  
 TCTCTGGAGTTTTAGACACGGAAGTTTAATTAGCATTGACAGCACCTGTACAGAG[REDACTED]GGC  
 AATTTGACAATGCTAATGTCACTGGAGAAATAGAATTTGCCATTCAATTATTGCTTCAAAC  
 CCATTTCTTAGAAATATGCATCAAGGCCTGTAAGAACCTTGCCATATGGAGAAGAAAAGAAGA  
 AAAAGTGCATCCGTATGTGAAGACCTACCTGTTGCCCGACAGATCTCCAGGGAAAAGCGC  
 AAGACTGGAGTCAAAGGAACACCGTGGACCCGACCTTTCAGGAGACCTTGAAGTATCAGGT  
 GGCCCTGCCAGCTGGTGACCCGGCAGCTGCAGGTCTCGGTGTGGCATCTGGGCACGCTGG  
 CCCGGAGAGTGTCTTGGAGAAGTGATCATTCTCTGGCCACGTGGGACTTTGAAGACAGC  
 ACAACACAGTCCTTCCGCTGGCATCCGCTCCGGGCCAAGGCGGAGAAATACGAAGACAGCGT  
 TCCTCAGAGTAATGGAGAGCTCACAGTCCGGGCTAAGCTGGTTCTCCCTCACGGCCCAGAA  
 AACTCCAAGAGGCTCAAGAAGGGACAGATCAGCCATCACTTCATGGTCAACTTTGTTTGGTA  
 GTGCTAGGAGCCAAGAATTTACCTGTGCGGCCAGATGGCACCTTGAACCTCATTGTTAAGGG  
 CTGTCTCACTCTGCCAGACCAACAAAACCTGAGACTGAAGTCGCCAGTCCTGAGGAAGCAGG  
 CTTGCCCCAGTGGAAACACTCATTGTCTTCACTGGCGTAACCCAGCTCAGCTGAGGCAG  
 TCGAGCTTGGAGTTAACTGTCTGGGATCAGGCCCTCTTGGGAATGAACGACCCGCTTGTCTGG  
 AGGAACCAGACTTGGTTCAAAGGGAGACACAGCTGTTGGCGGGGATGCATGCTCACAAATCGA  
 AGCTCCAGTGGCAGAAAGTCCTTTCCAGCCCCAATCTATGGACAGACATGACTCTTGTCTCTG  
 CACTGACATGAAGGCCTCAAGGTTCCAGGTTGCAGCAGGCGTGAGGCACTGTGCGTCTGCAG  
 AGGGGCTACGAACCAGGTGCAGGGTCCAGCTGGAGACCCCTTTGACCTTGAGCAGTCTCCA  
 TCTGCGGCCCTGTCCCATGGCTTAACCGCCTATTGGTATCTGTGTATATTTACGTAAACAC  
 AATTATGTTACCTAAGCCTCTGGTGGGTTATCTCCTCTTTGAGATGTAGAAAATGGCCAGAT  
 TTT[REDACTED]CGTTGTTACCCATGAAAAAAAAAAAAA

#### SEQUÊNCIA PROTEICA

MGNFDNANVTGEIEFAIHYCFKTHSLEICIKACKNLAYGEEKKKKCNPYVKTYLLPDRSSQG  
 KRKTGVQRNTVDPTFQETLKYQVAPAQLVTRQLQVSVWHLGTLARRVFLGEVI I PLATWDFE  
 DSTTQSFRRWHLRAKAKEYEDSVQSNGELTVRAKLVLPSPRKLQEAQEGTDQPSLHGQLC  
 LVVLGAKNLPVRPDGTLNSFVKGLTLPDQQKLRLLKSPVLRKQACPQWKHSFVFSGVTPAQL  
 RQSSLELTVWDQALFGMNDRLGGTRLGSKGDTAVGGDACSQSKLQWQKVLSSPNLWTDMTL  
 VLH

FIGURA 10