

(11) Número de Publicação: **PT 1012274 E**

(51) Classificação Internacional:
C12N 15/12 (2006.01) **C07K 16/28** (2006.01)
C07K 14/705 (2006.01)

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

(22) Data de pedido: 1998.01.27	(73) Titular(es): CRAIG A. ROSEN 22400 ROLLING HILL ROAD LAYTONSVILLE, MD 20882 US THE REGENTS OF THE UNIVERSITY OF MICHIGAN US HUMAN GENOME SCIENCES, INC. US
(30) Prioridade(s): 1997.01.28 US 35722 P 1997.02.05 US 37829 P	
(43) Data de publicação do pedido: 2000.06.28	
(45) Data e BPI da concessão: 2007.05.23 058/2007	(72) Inventor(es): VISHVA M. DIXIT US JAMES G. PAN US JIAN NI US REINER L. GENTZ US
	(74) Mandatário: JOSÉ EDUARDO LOPES VIEIRA DE SAMPAIO R DO SALITRE 195 RC DTO 1250-199 LISBOA PT

(54) Epígrafe: **RECEPTOR 4 CONTENDO UM DOMÍNIO DE MORTE (DM4; RECEPTOR DE MORTE 4), MEMBRO DA SUPER-FAMÍLIA DO FTN E LIGAÇÃO A LIART (AP02)**

(57) Resumo:

DESCRIÇÃO

**RECEPTOR 4 CONTENDO UM DOMÍNIO DE MORTE (DM4;
RECEPTOR DE MORTE 4), MEMBRO DA SUPER-FAMÍLIA DO
FNT E LIGAÇÃO A LIART (AP02)**

Domínio da Invenção

A presente invenção tem por objecto um novo elemento da família dos receptores do factor de necrose do tumor. Mais especificamente, tem por objecto moléculas isoladas de ácidos nucleicos que codificam o receptor 4 humano que contém um domínio de morte, algumas vezes referido aqui como "DM4". Também tem por objecto polipéptidos de DM4, assim como, vectores, células hospedeiras e processos recombinantes para a produção dos mesmos. A presente invenção tem ainda por objecto processos de rastreio para identificar agonistas e antagonistas da actividade de DM4.

Antecedentes da Invenção

Numerosas acções biológicas, por exemplo, as que respondem a certos estímulos e processos biológicos naturais, são controladas por factores, tais como, as citocinas. Muitas citocinas actuam através de receptores envolvendo o receptor e produzindo uma resposta intracelular.

Por exemplo, os factores de necrose do tumor (FNT) alfa e beta são citocinas, que actuam através dos receptores de FNT para regular numerosos processos biológicos, incluindo a protecção contra a infecção e a indução de choque e de doenças inflamatórias. As moléculas do FNT pertencem a super-família dos "ligandos

de FNT" e actuam em conjunto com os seus receptores ou contra-ligandos, a super-família dos "receptores de FNT". Até agora, foram identificados nove membros da super-família dos ligandos de FNT e dez membros da super-família dos receptores de FNT foram caracterizados.

Entre os ligandos, estão incluídos FNT- α , linfotoxina- α (LT- α , também conhecida por FNT- β), LT- β (encontrada no heterotrímico complexo LT- α 2- β), FasL, CD40L, CD27L, CD30L, 4-1BBL, OX40L e o factor de crescimento do nervo (FCN). A super-família dos receptores de FNT inclui o receptor de p55FNT, o receptor de p75FNT, a proteína relacionada com o receptor de FNT, o antígeno de FAS ou APO-1, CD40, CD27, CD30, 4-1BB, OX40, p75 de baixa afinidade e o receptor de FCN (Meager, A., *Biologicals*, 22: 291-295 (1994)).

Muitos membros da super-família dos ligandos do FNT são expressos pelas células T activadas, implicando isso que eles são necessários para as interacções das células T com outros tipos de células subjacentes a ontogenia e as funções das células (Meager, A., supra).

Já se ganhou uma considerável visão sobre as funções essenciais de vários membros da família dos receptores de FNT a partir da identificação e criação de mutantes que vão abolir a expressão dessas proteínas. Por exemplo, as mutações de ocorrência natural no antígeno FAS e no seu ligando causam uma doença linfoproliferativa (Watanabe-Fukunaga, R., et al., *Nature* 356: 314 (1992)), reflectindo talvez uma falha da morte programada das células. As mutações do ligando CD40 causam um estado de imunodeficiência ligado a X, caracterizado por elevados níveis de imunoglobulina M e baixos níveis de imunoglobulina G no plasma,

indicando uma activação defeituosa das células E dependentes das células T (Allen, R. C. et al., *Science* 259: 990 (1993)). As mutações pretendidas do receptor do factor de crescimento do nervo de baixa afinidade causam um distúrbio, caracterizada por uma deficiência da inovação sensória das estruturas periféricas (Lee, K. i. et al., *Cell* 69: 737 (1992)).

O FNT e a LT- α são capazes de se ligarem aos dois receptores de FNT (os receptores de FNT de 55 e 75 kd). Um grande número de efeitos biológicos provocados pelo FNT e pela LT- α , actuando através dos seus receptores, incluem a necrose hemorrágica de tumores transplantados, a citotoxicidade, um papel na choque endotóxico, inflamação, imuno-regulação, proliferação e respostas antivirais, assim como, protecção contra os efeitos prejudiciais da radiação por ionização. O FNT e a LT- α estão envolvidos na patogenese de uma vasta gama de doenças, incluindo o choque endotóxico, a malária cerebral, tumores, doenças auto-imunes, SIDA e rejeição de enxertos pelo hospedeiro (Beutler, B. e Von Huffel, C., *Science* 264: 667-668 (1994)). As mutações no receptor p55 causam uma susceptibilidade acrescida para infecções microbianas.

Além disso, foi referenciado um domínio de cerca de 80 aminoácidos próximo da terminação C do R-1 de FNT (p55) e Fas, como o "domínio de morte", que é responsável pelos sinais de transdução para a morte programada das células (Tartaglia et al., *Cell* 74: 845 (1993)).

A apoptose ou morte programada das células é um processo fisiológico essencial para o normal desenvolvimento e a homeostase dos organismos multicelulares (H. Steller, *Science* 267: 1445-1449

(1995)). Alterações da apoptose contribuem para a patogénese de várias doenças humanas incluindo o cancro, distúrbios neurodegenerativos e síndrome de imunodeficiência adquirida (C.B. Thompson, *Science* 267: 1456-1462 (1995)). Recentemente, tem-se focado muitíssima atenção na transdução ao sinal e na função biológica dos dois receptores do domínio da morte da superfície das células, Fas/APO-1 e R-1 de FNT (J. L. Cleveland et al., *Cell* 81: 479-482 (1995); A. Fraser, et al., *Cell* 85: 781-784 (1996); S. Nagata et al., *Science* 267: 1449-56 (1995)). Ambos são membros da família dos receptores de FNT que também incluem R-2 de FNT, R-FCN de baixa afinidade, CD40 e CD30, entre outros (C. A. Smith et al., *Science* 248: 1019-23 (1990); M. Tewari et al., em *Modular Texts in Molecular and Cell Biology* M. Purton, Heldin, Carl, Ed. (Chapman and Hall, London, 1995). Embora os membros da família sejam definidos pela presença de repetições ricas em cisteína nos seus domínios extracelulares, Fas/APO-1 e R-1 de FNT também partilham uma região de homologia intracelular, apropriadamente designada por "domínio da morte", que está relacionada, de uma forma distante, com o gene suicida *reaper* de *Drosophila*, (P. Golstein, et al., *Cell* 81: 185-186 (1995); K. White et al., *Science* 264: 677-83 (1994)). Este domínio de morte partilhado sugere que ambos os receptores interagem com um conjunto relacionado de moléculas de transdução do sinal que, até recentemente, permaneciam não identificadas. A activação de Fas/APO-1 recruta a molécula do adaptador contendo o domínio de morte FADD/MORT1 (A. M. Chinnaiyan et al., *Cell* 81: 505-12 (1995); M. P. Boldin et al., *J. Biol Chem* 270: 7795-8 (1995); F. C. Kischkel et al., *EMBO* 14: 5579-5588 (1995)), que, por sua vez, se liga e presumivelmente activa FLICE/MACH1, um membro da família de ICE/CED-3 das proteases pró-apoptóticas (M. Muzio et al., *Cell* 85:

817-827 (1996); M. P. Eoidin *et al.*, *Cell* 85: 803-815 (1996)). Embora o papel central de Fas/APO-1 seja alavancar a morte da célula, R-1 de FNT pode assinalar um conjunto de diversas actividades biológicas, muitas delas caracterizadas pela sua capacidade para activar NF-kB (L. N. Tartaglia *et al.*, *Immunol Today* 13: 151-3 (1992)). De acordo com isto, R-1 de FNT recruta a molécula adaptadora multivalente TRADD, que, tal como FADD, também contém um domínio de morte (H. Hsu *et al.*, *Cell* 81: 495-504 (1995); H. Hsu, *et al.*, *Cell* 84: 299-308 (1996)). Através das suas associações com um certo número de moléculas de sinalização, incluindo FADD, TRAF2 e RIP, TRADD pode assinalar tanto a apoptose como a activação de NF-kB (H. Hsu *et al.*, *Cell* 84: 299-308 (1996); H. Hsu, *et al.*, *Immunity* 4:387-396 (1996)).

Recentemente, descobriu-se um novo ligando de FNT que induz a apoptose. S. R. Wiley *et al.* (*Immunity* 3: 673-682 (1995)) designando a molécula por "ligando que induz a apoptose relacionada com o FNT" ou, simplesmente, "LIART", TRAIL na terminologia inglesa. A molécula foi também designada por "ligando de Apo-2" ou "LApo-2". " R. M. Pitt *et al.*, *J. Biol. Chem.* 271: 12687-12690 (1996). Esta molécula está também descrita no pedido de patente provisional co-pendente, norte-americano N°. 60/013.405, correspondente à patente de invenção WO 97/33889. Por conveniência, a molécula será então aqui referida como LIART.

Ao contrário do ligando de FAS, cujo os transcriptos parecem estar largamente restritos às células T estimuladas, detectam-se níveis significativos de LIART em muitos tecidos humanos (por exemplo, baço, pulmões, próstata, timo, ovários, intestino delgado, cólon, linfócitos do sangue periférico, placenta, rim) e é

transcripto sob o ponto de vista constitutivo por algumas linhas de células. Tem-se demonstrado que LIART actua independentemente do ligando de Fas (Wiley et al., supra). Também tem sido demonstrado que LIART activa rapidamente a apóptose, dentro de um intervalo de tempo que é semelhante à sinalização de morte por Fas/LApo-1, mas muito mais rápida do que a apóptose induzida por FNT. S. A. Marsters et al., *Current Biology* 6: 750-752 (1996). Todo o trabalho até à data sugere que o receptor para LIART não é um dos muitos receptores do FNT.

Os efeitos dos ligandos da família de FNT e dos receptores da família de FNT são variados e influenciam numerosas funções, tanto normais como anormais, nos processos biológicos do sistema dos mamíferos. Há uma necessidade clara, por isso, de identificar e caracterizar esses receptores e ligandos que influenciam a actividade biológica, tanto normalmente como em estados de doença. Em particular, há uma necessidade de isolar e caracterizar o receptor para o ligando de LIART recentemente descoberto.

Sumário da Invenção

A presente invenção tem por objecto moléculas de ácido nucleico isoladas, compreendendo sequências de ácidos nucleicos que codificam a sequência de aminoácidos mostrada na fig. 1 (SEQ ID N.º. 2) ou a sequência de aminoácidos codificada pelo clone de ADNc depositado na ATCC com o N.º. de depósito 97853, em 21 de Janeiro de 1997.

A presente invenção também tem por objecto vectores e células hospedeiras para a expressão recombinante das moléculas de ácidos nucleicos aqui descritas, assim como

os processos de produzir esses vectores e essas células hospedeiras e para a sua utilização para a produção de polipéptidos de DM4 ou péptidos de DM4, por meio de técnicas recombinantes.

A presente invenção tem ainda por objecto um polipéptido de DM4 isolado com uma sequência de aminoácidos codificada por um polinucleótido aqui descrito.

A presente invenção também tem por objecto ensaios de diagnóstico, tal como, ensaios quantitativos e ensaios de diagnóstico para a detecção dos níveis da proteína do DM4. Assim, por exemplo, um ensaio de diagnóstico de acordo com a presente invenção para detectar a sobre-expressão de DM4 ou de uma sua forma solúvel, comparando com as amostras de controlo de tecido normal, pode ser utilizado para detectar a presença de tumores.

Os ligandos da família do factor de necrose de tumor (FNT) são conhecidos por estarem entre as citocinas mais pleiotrópicas, induzindo um grande número de respostas celulares, incluindo citotoxicidade, actividade antiviral, actividades irnuno-reguladoras e a regulação transcricional de vários genes. A resposta celular aos ligandos da família dos FNT inclui não apenas as respostas fisiológicas normais, mas também doenças associadas com o aumento da apoptose ou a inibição da apoptose. A apoptose - morte programada das células - é um mecanismo fisiológico envolvido na eliminação dos linfócitos T periféricos do sistema imunitário e a sua desregulação pode levar a um certo número de diferentes processos patogénicos. As doenças associadas com o aumento da sobrevivência das células ou a inibição da apoptose incluem cancro, distúrbios auto-imunes, infecções virais, inflamações, enxertos versus doença do

hospedeiro, rejeição aguda de enxertos e rejeição crónica de enxertos. As doenças associadas com um aumento da apoptose incluem a SIDA, doenças neurodegenerativas, síndromas mielodisplásicos, traumas isquémicos, choque séptico, doenças do fígado induzidas por toxinas, caquexia e anorexia.

Assim, a presente invenção tem ainda por objecto um processo para aumentar a apoptose induzida por um ligando da família dos FNT, que envolve a administração a uma célula que expressa o polipéptido de DR4 uma quantidade efectiva de um agonista capaz de aumentar a sinalização mediada por DM4. Preferencialmente, aumenta-se a sinalização mediada por DM4 para tratar uma doença em é evidente a diminuição da apoptose.

Num outro aspecto, a presente invenção tem ainda por objecto um processo para inibir a apoptose induzida por um ligando da família do FNT, o que envolve a administração a uma célula que expressa o polipeptido de DM4, de uma quantidade de um agonista capaz de diminuir a sinalização mediada por DM4. Preferencialmente, diminui-se a sinalização mediada por DM4 para tratar uma doença em que se verificou que a apoptose estava aumentada.

Quando um qualquer candidato a "agonista" ou a "antagonista" da presente invenção pode aumentar ou inibir a apoptose, pode-se fazer a respectiva determinação utilizando ensaios de resposta celular do receptor/ligando da família do FNT, conhecidos na técnica, incluindo os descreites com mais detalhe a seguir. Assim, num outro aspecto, descreve-se um processo de avaliação para determinar se um agonista candidato ou um antagonista candidato é capaz de aumentar ou de inibir uma resposta celular a um ligando da família do FNT. O

processo envolve o contacto das células que expressam o polipéptido DM4 com um composto candidato e um Ligando da família dos FNT, ensaiando uma resposta celular e comparando a resposta celular com uma resposta celular padrão, sendo o padrão ensaiado quando o contacto é feito com o ligando na ausência do composto candidato, em que um aumento da resposta celular em relação ao padrão, indica que o composto candidato é um agonista da via de sinalização do ligando/receptor e uma diminuição da resposta celular comparada com o padrão, indica que o composto candidato é um antagonista da via de sinalização do ligando/receptor. Pela presente invenção, pode fazer-se contactar uma célula que expressa o polipéptido de DM4 com um qualquer dos ligandos da família dos FNL, administrado quer endogenamente quer exogenamente.

Breve Descrição das Figuras

- A figura **1** mostra a sequência de nucleótidos e a sequência de aminoácido deduzida de DM4. É previsível que os aminoácidos 1-23 constituam o péptido de sinal, os aminoácidos 24-238 constituem o domínio extracelular, os aminoácidos 239-264 constituem o domínio da transmembrana e os aminoácidos 265-468 constituem o domínio intracelular do qual os aminoácidos 379-422 constituem o domínio de morte.

A **figura 2** mostra as regiões de semelhança entre as sequências de aminoácidos de DM4, do receptor 1 do factor humano de necrose do tumor (SEQ ID NO: 3), a proteína Fas humana (SEQ ID NO: 4) e o receptor 3 contendo o domínio de morte (DM5) (SEQ ID NO: 5).

A **figura 3** mostra uma análise da sequência de aminoácidos de DM4. Regiões alfa, beta e de anéis;

hidrofilicidade e hidrofobicidade; regiões anfipáticas; regiões flexíveis; índice antigénico e probabilidade da superfície são todos parâmetros mostrados. No gráfico "Índice antigénico - Jameson-Wolf" os resíduos de aminoácidos 35-92, 114-160, 169-240, 267-298, 330-364, 391-404 e 418-465 na figura 1, correspondem às regiões altamente antigénicas mostiadas, da proteína de DM4.

A **figura 4** mostra as sequências de nucleótidos de fragmentos de ácidos nucleicos relacionados HTOIY07R (SEQ ID NO:6) e HTXEY8OR (SEQ ID NO:?).

As **figuras 5A e 5B** mostram a capacidade de DM4 para induzir a apoptose nas linhas de células MCF7 e 293. A **figura 5C** mostra a capacidade dos inibidores da protease da morte, CrmA e z-VAD-fmk, para inibir a acção apoptótica de DM4.

A **figura 6A** mostra a capacidade de uma fusão de DM4-Fc extracelular, solúvel para bloquear a capacidade de indução apoptótica de LIART.

A **figura 6B** mostra a incapacidade de uma fusão de DM4-Fc extracelular, solúvel para bloquear a capacidade de indução apoptótica de FNT α .

Descrição Detalhada dos Enquadramentos Preferidos

A presente invenção tem por objecto moléculas isoladas de ácidos nucleicos, compreendendo uma sequência de ácidos nucleicos que codifica o polipéptido de DM4 cuja sequência de aminoácidos se mostra na figura I (SEQ ID NO: 2) ou um fragmento do polipéptido. O polipéptido

de DM4 da presente invenção partilha uma homologia da sequência com R-1 de FNT humano, DM3 e o ligando de Las (figura 2). A sequência de nucleótidos mostrada na figura 1 (SEQ ID NO: 1) foi obtida sequenciando clones de ADN, tal como, HCUDS60, que foi depositado em 21 de Janeiro de 1997, na American Type Culture Collection, 12301 Park Lawn Drive, Rockville, Maryland 20852 e a que foi dado o número de acesso 97853. O clone depositado estava contido no plasmido pBK (Stratagene, LaJolla, CA).

Moléculas de Ácidos Nucleicos

A menos que seja indicado de outra forma, todas as sequências de nucleótidos determinadas pela sequenciação de uma molécula de ADN, foram aqui determinadas utilizando um sequenciador automático de ADN (tal como o modelo 373 da Applied Biosystems, Inc.) e todas as sequências de amino-ácidos dos polipeptidos codificados pelas moléculas de ADN aqui determinadas, foram previstos por tradução de uma sequência de ADN determinada tal como antes. Por isso, tal como é sabido na técnica para qualquer sequência de ADN determinada por esta abordagem automatizada, qualquer sequência de nucleótidos aqui determinada pode conter alguns erros. As sequências de nucleótidos determinadas por um processo automático são normalmente pelo menos cerca de 90 % idênticas, mais normalmente, pelo menos cerca de 95 % até pelo menos cerca de 99,9 % idênticas a sequência de nucleótidos actual da molécula de ADN sequenciada. A sequência actual pode ser determinada mais precisamente por meio de outras abordagens incluindo processos manuais de sequenciação de ADN bem conhecidos na técnica. Também como é sabido na técnica, uma inserção ou eliminação simples numa determinada sequência de nucleótidos comparada com a sequência actual, vai causar uma deslocação dum fragmento

na translação da sequência de nucleótidos, de tal maneira que a previsível sequência de aminoácidos codificada por determinada sequência de nucleótidos, será completamente diferente da sequência de aminoácidos actualmente codificada pela molécula de ADN sequenciada, começando no ponto dessa inserção ou eliminação.

Por polipéptido ou proteína "isolados" entende-se um polipéptido ou uma proteína retirada do seu ambiente natural. Por exemplo, os polipéptidos e as proteínas produzidos por recombinação, expressos nas células hospedeiras, são considerados isolados para os fins da presente invenção quando são polipeptidos naturais ou recombinantes ~~que~~ tenham sido substancialmente purificados por qualquer técnica apropriada tal como, por exemplo, o processo de purificação numa única etapa descrito em Smith e Johnson, Gene 67: 31-4C (1988).

Utilizando a informação dada aqui, tal como a sequência de ácidos nucleicos estabelecida na fig. 1, uma molécula de ácidos nucleicos da presente invenção que codifica um polipéptido de DM4, pode obter-se utilizando processos de clonagem e avaliação normalizados, tal como os da clonagem de ADNcs utilizando ARNm como material inicial. Ilustrativo da presente invenção, o gene da presente invenção tem sido identificado em bibliotecas de ADNc dos seguintes tecidos: células amnióticas, coração, cancro do fígado, rim, leucócitos, células T activadas, células K562 ~~mais~~ PMA, células W138, ~~células~~ Th2, amígdalas humanas e pele cor de couro com falta de CD34 (sangue do cordão).

O gene de DM4 ~~contém~~ um fragmento de leitura aberta que codifica ~~uma~~ proteína de cerca de 445 resíduos de aminoácidos cujo ~~codão~~ códon de iniciação está na posição 19-21

da sequência de nucleótidos mostrada na figura 1 (SEQ ID NO: 1), com uma sequência líder de cerca de 23 resíduos de aminoácidos (isto é, um comprimento total da proteína de 468 aminoácidos) e um peso molecular deduzido de 50 kDa. Dos membros conhecidos da família de receptores de FNT, o polipéptido de DMÇ da presente invenção partilha o mais alto grau de homologia com o R de FNT1 humano e os polipéptidos de DM3 mostrados na figura 2, incluindo uma significativa homologia da sequência em relação aos múltiplos domínios ricos em cisteína.

Para além da homologia da sequência exibida entre DM4 e outros receptores contendo o domínio de morte, DM4 tem mostrado que se ligava a LIART e que induz a apoptose quando expresso de forma transiente. As células MCF7 do carcinoma humano da mama e as células 293 foram transfectadas de forma transiente com uma estrutura que expressa DM4, tal como descrito no exemplo 5. Como se vê nas figuras 5A e 5B, uma porção substancial das células transfectadas sofre as alterações morfológicas características de apoptose. Como já se antecipou, a eliminação do domínio de morte aboliu a capacidade de DM4 para aproveitar a via de morte. Como se pode ver na figura 5C, a apoptose induzida por DM4 foi bloqueada de forma eficiente por inibidores de proteases do domínio de morte, incluindo z-VAD-fmk, um inibidor de caspase de largo espectro, irreversível e CrmA, um vírus de vacínia codificado por serpina que inibe preferencialmente as caspases apicais tais como FLICE/DMACH-1 (caspase-8). Dado que a apoptose induzida por R de FNT-1, CD-95 e DM3 está também atenuada por estes mesmos inibidores, é provável que as moléculas do effector do domínio de morte a jusante sejam similares na sua natureza.

Para determinar se DM4 era capaz de se ligar a LIART, o domínio de ligação do ligando extracelular de DM4 foi expresso como a região Fc de IgG (DM4-Fc) humana. LIART liga-se seletivamente a DM4-Fc mas não aos domínios extracelulares correspondentes do R de FNT-1 ou C3-95, também expresso como fusões de Fc, dados não mostrados. Adicionalmente, DM4-Fc não se liga nem a FNT alfa nem ao ligando Faz, nas condições em que ambos estes ligandos se ligam aos seus receptores de cognatos.

A capacidade de LIART para induzir a apoptose nas células MCF7 foi especificamente bloqueada por DM4-Fc mas não foi influenciada por R de FNT-1-Fc, CD95-Fc ou apenas Fc (Figura 6A). Além disso, conforme se esperava, a apoptose induzida por FNT alfa foi inibida por FNT-1-Fc mas não por DM4-Fc, CD95-Fc e Fc isoladamente (Figura 6B).

Considerados em conjunto, os dados descritos antes indicam que DM4 é um domínio de morte contendo um receptor com a capacidade para Induzir a apoptose e é um receptor para LIART, um ligando conhecido que induz a apoptose.

Tal como indicado, a presente invenção também tem por objecto as formas maduras da proteína DM4 da presente invenção. De acordo com a hipótese de sinal, as proteínas segregadas por células de mamíferos têm uma sequência de sinal ou uma sequência líder secretora, que é clivada a partir da proteína madura logo que se exporta da cadeia da proteína de crescimento quando o retículo endoplasmático rugoso é iniciado. A maior parte das células de mamíferos e mesmo as células de insectos clivam as proteínas segregadas com a mesma especificidade. Contudo, nalguns casos, a clivagem de uma proteína segregada não é

completamente uniforme, o que resulta em duas ou mais espécies maduras na proteína. Além disso, sabe-se desde há muito tempo que a especificidade da clivagem de uma proteína segregada é em última análise determinada pela estrutura primária da proteína completa, isto é, inerente à sequência de aminoácidos do polipéptido. Por isso, a presente invenção tem por objecto uma sequência de nucleótidos que codifica o polipéptido de DM4 maduro, contendo a sequência de aminoácidos codificada pelo clone de ADNc contido no hospedeiro identificado com o número de depósito na ATCC 97853 e, tal como se mostra na figura 1 (SEQ ID NO: 2). Por proteína de DM4 madura com a sequência de aminoácidos codificada pelos clones de ADNc contidos no hospedeiro identificado com o número de depósito na ATCC 97853, entende-se as formas maduras da proteína de DM4 produzida por expressão numa célula de mamíferos (por exemplo, células de COS, tal como se descreve a seguir) do fragmento completo de leitura aberta codificado pela sequência de ADN humano do clone contido no vector no hospedeiro depositado. Tal como se indica a seguir, o DM4 maduro comportando a sequência de aminoácidos codificada pelo clone de ADNc com o número de depósito na ALCC 97853, pode ou não diferir da prevista proteína de DM4 "madura" mostrada na figura 1 (aminoácidos desde cerca de 24 até próximo de 468) consoante a precisão do sitio de clivagem previsto com base na análise do computador.

Estão disponíveis processos para prever se uma proteína tem um líder secretor, assim como, o ponto de clivagem para essa sequência líder. Por exemplo, os processos de McGeech (*Virus Res.* 3: 271-286 (1988)) e de von Heinje (*Nucleic Acids Res.* 14: 4683-4690 (1986)) podem ser utilizados. A precisão para prever os pontos de clivagem de proteínas secretoras conhecidas de mamíferos

para cada um destes processos, está no intervalo de cerca de 75-80 %. von Heinje, *supra*. Contudo, os dois processos não produzem sempre os mesmos pontos de clivagem para uma dada proteína.

No presente caso, a sequência de aminoácidos prevista do polipéptido de DM4 completo da presente invenção foi analisada por um programa de computador ("PSORT"). (Ver, K. Nakai e M. Kanehisa, *Genomics* 14: 897-911 (1992)) que é um sistema pericial para prever a localização celular de uma proteína com base na sequência de aminoácidos. Como parte desta previsão computacional da localização, incorporam-se aqui os processos de McGeoch e de von Heinje. A análise por meio do programa de previsão PSORT prevê os sítios de clivagem entre os aminoácidos 23 e 24 na figura 1 (SEQ ID NO: 2). Depois, as sequências de aminoácidos completas foram ainda analisadas por meio de inspeção visual, aplicando uma forma simples da regra (-1, -3) de von Heinje. von Heinje, *supra*. Assim, a sequência líder para a proteína DM4 previsivelmente consiste nos resíduos de aminoácidos 1-21, sublinhados na figura I (SEQ ID NO: 2), embora a proteína madura DM4 prevista consista nos resíduos 24-468.

Tal como se indicou, as moléculas de ácido nucleico da presente invenção, podem estar sob a forma de ARN, tal como, ARNm ou sob a forma de ADN, incluindo, por exemplo, ADNc e ADN genômico obtido por clonagem ou produzido por síntese. O ADN pode ser de estrutura helicoidal dupla ou de estrutura helicoidal simples. O ADN de estrutura helicoidal simples pode ser a estrutura de codificação, também conhecido como a estrutura de um sentido (sense) ou paralela ou pode ser a estrutura helicoidal de não

codificação, também referida como a estrutura de sentido contrario (anti-sense) ou anti-paralela.

Por moléculas de ácido nucleico "isoladas" entende-se uma molécula de ácido nucleico, ADN ou ARN, que tenha sido retirada do seu ambiente original. Por exemplo, as moléculas de ADN recombinante contidas num vector são consideradas isoladas para os fins da presente invenção. Outros exemplos de moléculas de ADN isoladas incluem moléculas de ADN recombinante mantidas nas células hospedeiras heterólogas ou moléculas de ADN purificadas (parcialmente ou totalmente) em solução. As moléculas de ARN isoladas incluídas nos transcriptos de ARN *in vivo* ou *in vitro* das moléculas de ADN da presente invenção. As moléculas de ácido nucleico isoladas, de acordo com a presente invenção, incluem ainda essas moléculas produzidas por síntese.

As moléculas de ácido nucleico isoladas da presente invenção incluem moléculas de ADN de DM4, que compreendem um fragmento de leitura aberta (FLA) mostrado na fig. 1 (SEQ ID NO: 1) e incluem ainda as moléculas de ADN, que compreendem uma sequência substancialmente diferente de toda ou parte da FLA cujo codão de iniciação está na posição 19-21 da sequência de nucleótidos mostrada na fig. 1 (SEQ ID NO: 1) mas que, devido à degenerescência do código genético, ainda codifica a proteína do polipéptido de DM4 ou a um seu fragmento. Obviamente, o código genético é bem conhecido na técnica. Assim, será uma rotina para um especialista na matéria gerar essas variantes degeneradas.

Num outro aspecto, a presente invenção tem por objecto moléculas isoladas de ácido nucleico, que codificam o polipéptido de DM4 comportando uma sequência

de aminoácidos codificada pelo clone de ADNc contido no plasmido depositado na ATCC com o número de depósito 97853, em 21 de Janeiro de 1997. Preferencialmente, estas moléculas de ácido nucleico irão codificar o polipéptido maduro ocodificado pelo clone de ADNc depositado, descrito antes. A presente invenção tem ainda por objecto uma molécula de ácido nucleico isolada comportando a sequência de nucleótidos mostrada na figura 1 (SEQ ID NO: 1) ou a sequência de nucleótidos do ADNc de DM4 contida no clone depositado descrito antes ou uma molécula de ácido nucleico comportando uma sequência complementar das sequências anteriores. Essas moléculas isoladas e os seus fragmentos são úteis como sondas para o mapeamento dos genes por meio da hibridação *in situ* com o gene de DM4 em tecido humano, por meio de análise de Northern, de transferência para uma matriz imobilizada (ADN/ARNm).

A presente invenção tem ainda por objecto fragmentos das moléculas de ácido nucleico isoladas aqui descritas. Por um fragmento de uma molécula de ADN isolada com a sequência de nucleótidos mostrada na figura 1 (SEQ ID NO: 1) entende-se os fragmentos de ADN com pelo menos 20 pb e, mais preferencialmente, pelo menos 30 pb de comprimento, que são úteis como sondas de ADN conforme discutido antes, obviamente que fragmentos de ADN maiores, com 50-1500 pb de comprimento, são também úteis como sondas de ADN de acordo com a presente invenção, dado que são fragmentos de ADN que correspondem à maior parte, se não a toda, a sequência de nucleótidos que se mostra na fig. 1 (SEQ ID NO: 1). Por um fragmento com pelo menos 29 pb de comprimento, por exemplo, entende-se fragmentos que incluem 20 ou mais bases da sequência de nucleótidos da fig. 1.

Os fragmentos de ácido nucleico preferidos da presente invenção incluem moléculas de ácidos nucleicos que codificam: um polipéptido que compreende o domínio extracelular de DM4 (resíduos de aminoácidos desde cerca de 24 até cerca de 238 na figura 1 (SEQ ID NO:2)); um polipéptido que compreende um domínio da membrana de DM4 (resíduos de aminoácidos desde cerca de 239 até cerca de 264 na figura 1 (SEQ ID NO: 2)); um polipéptido que compreende o domínio intracelular de DM4 (resíduos de aminoácidos desde cerca de 379 até cerca de 422 na figura 1 (SEQ ID NO: 2)). Dado que a localização destes domínios foi prevista por computação gráfica, um especialista na matéria compreenderá que os resíduos de aminoácidos que constituem estes domínios podem variar ligeiramente (por exemplo, cerca de 1 a 15 resíduos) consoante os critérios utilizados para definir cada domínio.

Os fragmentos de ácidos nucleicos preferidos da presente invenção codificam um polipéptido de DM4 de comprimento completo a que faltam os nucleótidos que codificam a metionina de terminal amino (nucleótidos 19-21 na SEQ ID NO: 1); dado que se sabe que a metionina é clivada naturalmente e que essas sequências podem ser úteis em engenharia genética nos vectores de expressão de DM4. Os polipéptidos codificados por esses polinucleótidos estão todos contemplados na presente invenção.

Os fragmentos de ácidos nucleicos preferidos da presente invenção incluem ainda moléculas de ácidos nucleicos que codificam as porções que se ligam ao epítipo da proteína de DM4. Em particular, esses fragmentos de ácidos nucleicos da presente invenção incluem moléculas de ácidos nucleicos que codificam: um polipéptido que compreende os resíduos de aminoácidos

desde cerca de 35 até cerca de 920 na figura 1 (SEQ ID NO: 2); um polipéptido que compreende os resíduos de aminoácidos desde cerca de 114 até cerca de 160 na figura 1 (SEQ ID NO: 2); um polipéptido que compreende os resíduos de aminoácidos desde cerca de 169 até cerca de 240 na figura 1 (SEQ ID NO: 2); um polipéptido que compreende os resíduos de aminoácidos desde cerca de 267 até cerca de 298 na figura 1 (SEQ ID NO: 2); um polipéptido que compreende os resíduos de aminoácidos desde cerca de 330 até cerca de 364 na figura 1 (SEQ ID NO: 2); um polipéptido que compreende os resíduos de aminoácidos desde cerca de 391 até cerca de 404 na figura 1 (SEQ ID NO: 2); e um polipéptido que compreende os resíduos de aminoácidos desde cerca de 418 até cerca de 485 na figura 1 (SEQ ID NO: 2). Os requerentes determinaram que os fragmentos de polipéptidos anteriores são regiões antigénicas da proteína de DM4. Os processos para a determinação, para além das porções que se ligam ao epítipo da proteína de DM4, estão descritos com detalhe a seguir.

Além disso, a presente invenção tem por objecto moléculas de ácidos nucleicos que têm sequências de nucleótidos relacionadas com porções extensas da SEQ ID NO: 1, tal como se segue: HTQIY07R (SEQ ID NO:6) e HTXEY80R (SEQ ID NO:7) que estão ambas indicadas na figura 4.

Além disso, a presente invenção inclui um polinucleótido que compreende qualquer porção de pelo menos cerca de 30 nucleótidos, preferencialmente, pelo menos cerca de 50 nucleótidos da SEQ ID NO: 1, do resíduo 365 a 1,424.

Num outero aspecto, a presente invenção tem por objecto uma molécula de ácido nucleico isolada, que compreende um polinucleótido, que hibrida, em condições de hibridação severas, com uma porção do polinucleótido numa molécula de ácido nucleico da presente invenção descrita antes, por exemplo, os clones de ADNc depositados na ATCC com a número de depósito 97853. Por "condições de hibridação severas" entende-se a incubação durante a noite, a 42 °C, numa solução que compreende: formamida a 50 , 5 x SSC (NaCl 150 mM, citrato trisódico 15 mM), fosfato de sódio 50 mM (a pH 7,6), 5 x de solução de Denhardt, sulfato de dextrano a 10 % e 20 g/mL de ADN de esperma de salmão cortado e desnaturado, seguido da lavagem dos filtros em 0,1 x SSC, a cerca de 65 °C.

Por um polinucleótido que hibrida com uma "porção" de um polinucleótido, entende-se um polinucleótido (quer ADN, quer ARN), que hibrida com pelo menos cerca de 15 nucleótidos (nt) e, mais preferencialmente, com pelo menos cerca de 20 nt, ainda mais preferencialmente, com pelo menos cerca de 30 nt e, ainda mais preferencialmente, pelo menos cerca de 30-70 de polinucleótidos de referência. Eles são úteis como sondas de diagnóstico e iniciadores, conforme se discutiu antes e em mais detalhe a seguir.

Por uma porção de um polinucleótido de "pelo menos 20 nt de comprimento", por exemplo, entende-se 20 ou mais nucleótidos contíguos da sequência de nucleótidos do polinucleótido de referência (por exemplo, o ADNc depositado ou a sequência de nucleótidos conforme se mostra na fig. 1 (SEQ ID NO: 1) ou na Figura 2 (SEQ ID NO: 3).

Obviamente, um polinucleótido que híbrida apenas com uma sequência de poli A (tal como, por exemplo, o tracto de poli A do terminal 3' do ADNc de ~~DM4~~ mostrado não figura 1 (SEQ ID NO: 1)) ou com uma extensão complementar de resíduos de T (ou de U), não estaria incluído num polinucleótido da presente invenção utilizado para hibridar com uma porção de um ácido nucleico da presente invenção, dado que esse polinucleótido iria hibridar com qualquer molécula de ácido nucleico contendo a extensão de poli A do seu complemento (por exemplo, praticamente qualquer clone de ADNc de estrutura helicoidal dupla).

Tal como se indicou, as moléculas de ácido nucleico da presente invenção que codificam um polipéptido de DM4, podem incluir, mas não se limitam, a sequências de codificação do polipéptido maduro, por ele próprio; a sequência de codificação para o polipéptido maduro e as sequências adicionais, tais como as que codificam uma sequência líder ou segregadora, tais como, sequências de pré- ou pró- ou pré-pró- proteína; a sequência de codificação do polipéptido maduro com ou sem as sequências de codificação adicionais mencionadas antes, em conjunto com sequências adicionais, não codificantes, incluindo, por exemplo, mas não se limitando, a intrões e sequências de 3' e 5' não codificantes, tal como, as sequências não traduzidas, as transcriptas que desempenham um papel na transcrição, fragmentação incluindo o tratamento de ARNm e de sinais de polidensilação, por exemplo, - ligação de ribossoma e estabilidade de ARNm -; sequências de codificação adicionais que codificam para aminoácidos adicionais, tal como as que providenciam funcionalidades adicionais. Assim, por exemplo, o polipéptido pode fundir-se com uma sequência marcadora tal como um péptido, que facilita a purificação do polipéptido fundido. Em certos enquadra-

mentos preferidos deste aspecto da presente invenção, a sequência marcadora é um péptido de hexa-histidina, tal como um marcador providenciado por um vector pQE (Qiagen, Inc.), entre outros, muitos dos quais estão comercialmente disponíveis. Tal como descrito em Gentz *et al.*, *Proc. Natl. Acad. Sci. EUA* 86: 821-824 (1989), por exemplo, a hexa-histidina providencia uma purificação conveniente da proteína de fusão. O marcador "HA" corresponde a um epítipo derivado da proteína da hemaglutinina da gripe, que já foi descrita, por exemplo, por Wilson *et al.*, *Cell* 37: 767-778 (1984).

A presente invenção tem ainda por objecto variantes das moléculas de ácidos nucleicos da presente invenção que codificam para fragmentos, análogos ou derivados do polipéptido de DM4. As variantes podem ocorrer naturalmente, tal como uma variante alélica. Por uma "variante alélica" entende-se uma das várias formas alternativas de um gene, que ocupa um dado local num cromossoma de um organismo. *Genes II*, Lewin, B., ed., John Wiley & Sons, New York (1985). As variantes de ocorrência não natural podem ser produzidas utilizando técnicas de mutagénesis conhecidas na técnica.

Essas variantes incluem as produzidas por substituições, eliminações ou adições de nucleótidos, que podem envolver um ou mais nucleótidos. As variantes podem ser alteradas nas regiões de codificação ou de não codificação ou em ambas. As alterações nas regiões de codificação podem produzir substituições, eliminações ou adições de aminoácidos conservadoras ou não conservadoras,

Outros enquadramentos da presente invenção incluem moléculas isoladas de ácidos nucleicos, que são pelo

menos 90 % idênticas e mais preferencialmente pelo menos 95 %, 96 %, 97 %, 98 % ou 99 % idênticas a !a) uma sequência de nucleótidos, que codifica o polipéptido completo de DM4, que comporta a sequência de aminoácidos na fig. 1 (SEQ ID NO: 2), incluindo a a sequência líder prevista; (b) uma sequência de nucleótidos, que codifica todo o comprimento do polipéptido de DM4, que comporta a sequência de aminoácidos completa na figura 1 (SEQ ID NO: 2), incluindo a a sequência líder prevista mas a que falta a metionina do terminal amino; (c) uma sequência de nucleótidos, que codifica o polipéptido maduro de DM4 (polipéptido de comprimento completo com a sequência líder eliminada) que comporta a sequência de aminoácidos nas posições próximas de 24 até as próximas de 468 na figura 1 (SEQ ID NO: 2); (d) uma sequência de nucleótidos, que codifica o polipéptido de comprimento completo de DM4, que comporta a sequência completa de aminoácidos incluindo a sequência líder codificada pelo clone de ADNc depositado na ATCC com o número de depósito 97853; (e) uma sequência de nucleótidos, que codifica o polipéptido de comprimento completo de DM4, que comporta a sequência de aminoácidos incluindo a sequência líder mas a que falta a metionina do terminal amino codificada pelo clone de ADNc depositado na ATCC com o número de depósito 97853; (f) uma sequência de nucleótidos, que codifica o polipéptido maduro de DM4, que comporta as sequências de aminoácidos codificadas pelo clone de ADNc depositado na ATCC com o número de depósito 97853; (g) uma sequência de nucleótidos, que codifica o domínio da transmembrana de DM4; (h) uma sequência de nucleótidos que codifica o domínio da transmembrana de DM4, (i) uma sequência de nucleótidos que codifica o domínio intracelular de DM4; (j) uma sequência de nucleótidos que codifica o domínio de morte DM4; ou (k) uma sequência de nucleótidos complementar de qualquer uma das sequências

de nucleótidos nas alíneas (a), (b), (c), (d), (e), (f), (g), (h), (i) ou (j) anteriores.

For um polinucleótido comportando uma sequência de nucleótido pelo menos, por exemplo, 95 % "idêntica" a uma sequência de nucleótidos de referência que codifica um polipéptido de DM4, entende-se que a sequência de nucleótidos do polinucleótido é idêntica à sequência de referência, excepto no facto da sequência de polinucleótidos poder incluir até cinco pontos de mutações por cada cem nucleótidos da sequência de nucleótidos de referência, que codifica o polipéptido de DM4. For outras palavras, para se obter um polinucleótido comportando uma sequência de nucleótidos pelo menos 95 % idêntica a uma sequência de nucleótidos de referência, até 5 % dos nucleótidos na sequência de referência podem ser eliminados ou substituídos por outros nucleótidos ou um certo número de nucleótidos até 5 % dos nucleótidos totais na sequência de referência, podem ser inseridos na sequência de referência. Estas mutações da sequência de referência podem ocorrer nas posições terminais 5 ou 3 da sequência de nucleótidos de referência ou em qualquer parte entre estas duas posições terminais, dispersas quer individualmente entre os nucleótidos na sequência de referência ou em um ou mais grupos contíguos dentro da sequência de referência.

Como uma matéria prática, se qualquer molécula particular de ácido nucleico é pelo menos 90 %, 95 %, 96 %, 97 %, 98 % ou 99 % idêntica a, por exemplo, a sequência de nucleótidos mostrada na figura 1 ou à sequência de nucleótidos do clone de ADNc depositado, pode ser determinada de uma forma convencional utilizando programas de computador conhecidos, tal como o programa Bestfit (Wisconsin Sequence Analysis Package, Versão 8

para Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). O Bestfit utiliza o algoritmo de homologia local de Smith and Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981), para encontrar o melhor segmento de homologia entre duas sequências. Quando se utiliza o Bestfit ou qualquer outro programa de alinhamento de sequências para determinar se uma sequência particular é, por exemplo, 95 % idêntica a uma sequência de referência, de acordo com a presente invenção, os parâmetros são estabelecidos, obviamente, de tal modo que a percentagem de identidade é calculada em relação ao comprimento completo da sequência de nucleótidos de referência e essas falhas na homologia até 5 % no número total de nucleótidos na sequência de referência são permitidas.

O presente pedido de patente de invenção tem por objecto moléculas de ácido nucleico pelo menos 90 %, 95 %, 96 %, 97 %, 98 % ou 99 % idênticas à sequência de ácidos nucleicos mostrada na figura i (SEQ ID NO: 1) ou a sequência de ácidos nucleicos dos ADNcs depositados, independentemente do facto de eles codificarem um polipéptido que tem actividade de DM4. Isto deve-se ao facto de quando uma molécula particular de ácido nucleico não codifica um polipéptido que tem actividade de DM4, um especialista na matéria saberá ainda como utilizar a molécula de ácido nucleico, por exemplo, como uma sonda de hibridação ou como um iniciador de uma reacção em cadeia de polimerase (RCP). As utilizações das moléculas de ácidos nucleicos da presente invenção, que não codificam um polipéptido que tem uma actividade de DM4, incluem, *inter alia*: (1) o isolamento do gene de DM4 ou das suas variantes alélicas numa biblioteca de ADNc; (2) hibridação *in situ* (por exemplo, "FISH") com extensões cromossómicas de metafase para providenciar uma

localização cromossômica precisa do gene de DM4, conforme descrito em Verma *et al.*, *Human Chromosomes: A Manual of Basic Techniques*, Pergamon Press, New York (1988); e (3) análise de transferência psra uma matriz imobilizada de Northern (ADN/ARNm) para detectar a expressão de ARNm de DM4 em tecidos específicos,

Preferidas, contudo, são as moléculas de ácidos nucleicos que têm sequências pelo menos 90 %, 95 %, 96 %, 97 %, 98 % ou 99 % idênticas a sequência de ácidos nucleicos mostrada na figura 1 (SEQ ID NO: h) ou a sequência de ácidos nucleicos dos ADNcs depositados que, de facto, codificam um polipéptido que tem actividade da proteína DM4. Por "um polipéptido que tem a actividade de DM4" entende-se polipéptidos que exibem uma actividade semelhante mas não necessariamente idêntica a uma actividade da proteína de DM4 da presente invenção (quer a proteína de comprimento completo ou, preferencialmente, a proteína madura) conforme se mede num ensaio biológico particular. Por exemplo, a actividade da proteína DM4 pode ser medida utilizando os ensaios de morte de células realizado essencialmente conforme descrito previamente (A. M. Chinnaiyan, *et al.*, *Cell* 81: 505-12 (1995); M. P. Boldin, *et al.*, *J Biol Chem* 270: 7795-8 (1995); F. C. Kischkel, *et al.*, *EMBO* 14: 5579-5588 (1995); A. M. Chinnaiyan, *et al.*, *J Biol Chem* 271: 4961-4965 (1996)) ou conforme estabelecido no exemplo 5 a seguir. Nas células MCF7, os plasmidos que codificam DM4 de comprimento completo ou um candidato a domínio de norte contendo receptores, são co-transfectados com a estrutura relatora de pLantern codificando uma proteína fluorescente verde. Os núcleos das células transfectadas com DM4 vão exibir uma morfologia apoptótica conforme pode ser analisado por meio da coloração por DAPI. Semelhante à apoptose induzida por B de FNT-1 e Fas/APO-1 (M. Muzio, *et al.*,

Cell 85: 817-827 (1996); M. P. Boldin, et al., *Cell* 85: 803-815 (1996); M. Tewari, et al., *J Biol Chem* 270: 3255-60 (1995)), a apoptose induzida por DM4 é bloqueada pelos inibidores das proteases semelhantes a ICE, CrmA e z-VAD-fmk.

Obviamente, devido a degenerescência do código genético, um especialista na matéria reconheceria imediatamente que um grande número de moléculas de ácido nucleico tem uma sequência pelo menos 90 %, 95 %, 96 %, 97 %, 98 % ou 99 % idêntica a sequência de ácidos nucleicos do ADNc depositado ou a sequência de ácidos nucleicos mostradas na figura 1 (SEQ ID NO: 1) que vão codificar um polipéptido "com actividade da proteína de DM4". De facto, dado que as variantes degeneradas destas sequências de nucleótidos codificam todas o mesmo polipéptido, estará bem claro para um técnico da matéria, mesmo sem a realização do ensaio de comparação descrito antes. É ainda reconhecido na técnica que, para essas moléculas de ácidos nucleicos que não são variantes degeneradas, um número razoável entre elas vai também codificar um polipéptido que tem actividade da proteína DM4. Isto verifica-se porque um técnico da matéria está completamente dentro das substituições de aminoácidos que são menos prováveis ou que é provável que afectem significativamente a função da proteína (por exemplo, substituindo um aminoácido alifático por um segundo aminoácido alifático).

Por exemplo, as indicações respeitantes à forma de produzir fenotipicamente substituições de aminoácidos silenciosas sob o ponto de vista do fenotipo, são dadas em Boxie, J. U. et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions", *Science* 247: 1306-1310 (1990), em que os

autores indicam que as proteínas são surpreendentemente tolerantes às substituições dos aminoácidos.

Ensaio de Polinucleótidos

A presente invenção também tem por objecto a utilização dos polinucleótidos de DM4 para detectar polinucleótidos complementares, tais como, por exemplo, como um reagente de diagnóstico. A detecção de uma forma mutada de DM4 associada com uma disfunção fornecerá uma ferramenta de diagnóstico que pode adicionar ou definir um diagnóstico de uma doença ou de uma susceptibilidade para uma doença, que resulta da sub-expressão ou da sobre-expressão da expressão alterada de DM4 ou de uma sua forma solúvel, tal como, por exemplo, tumores ou doenças auto-imunes.

Os indivíduos que comportam mutações no gene de DM4 podem ser detectados ao nível do ADN por uma variedade de técnicas. Os ácidos nucleicos para diagnóstico podem ser obtidos de células de pacientes, tais como, células de sangue, urina, saliva, biópsia de tecido e material de autópsia. O ADN genómico pode ser utilizado directamente para a detecção ou pode ser amplificado enzimaticamente utilizando RCP antes da análise. (Saiki *et al.*, *Nature* 324: 163-166 (1986)). Pode-se também utilizar ARN ou ADNc da mesma forma. Como exemplo, os iniciadores de RCP complementares do ácido nucleico que codificam DM4 podem ser utilizados para identificar e analisar as expressões e as mutações de DM4. Por exemplo, as eliminações e as inserções podem ser detectadas por meio de uma alteração da dimensão do produto amplificado em comparação com o genótipo normal. Podem-se identificar as mutações pontuais por hibridação do ADN amplificado com o AKN de DM4 marcado com rádio ou, alternativamente, sequências de

ACN de sentido oposto de DM4 marcadas com rádio. As sequências perfeitamente emparelhadas podem distinguir-se das sequências duplas desemparelhadas por digestão de RNase A ou pelas diferenças nas temperaturas de fusão.

As diferenças das sequências entre um gene de referência e os genes com mutações podem também ser reveladas por sequenciação directa do ADN. Além disso, podem utilizar-se segmentos de ADN clonados como sondas para detectar segmentos específicos de ADN. A sensibilidade desses processos pode ser muitíssimo melhorada pela utilização apropriada de RCP ou de outros processos de amplificação. Por exemplo, utiliza-se um iniciador de sequenciação com um produto de RCP de estrutura helicoidal dupla ou uma molécula matriz com uma estrutura helicoidal simples, gerada por uma RCP modificada. A determinação da sequência realiza-se por processos convencionais com nucleótidos marcados com rádio ou por processos automáticos de sequenciação com marcadores fluorescentes.

O ensaio genético com base nas diferenças das sequências de ADN pode ser feito por detecção da alteração na mobilidade electroforética de fragmentos de ADN em géis, com ou sem agentes de desnaturação. Podem visualizar-se eliminações e as inserções em sequências pequenas por electroforese em gel de elevada resolução. Os fragmentos de ADN das diferentes sequências podem distinguir-se nos géis de gradientes de formamida de desnaturação, em que as mobilidades dos diferentes fragmentos de ADN estão retardadas no gel em diferentes posições de acordo com as suas temperaturas de fusão específicas ou com as temperaturas de fusão parciais (ver, por exemplo, Myers et al., *Science* 230: 1242 (1985)).

As alterações da sequência em localizações específicas podem também ser reveladas por ensaios de protecção da nuclease, tais como protecção de RNase e de SI ou o processo de clivagem química (por exemplo, Cotton *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 4397-4401 (1985)).

Assim, a detecção de uma sequência de ADN específica, pode ser conseguida por processos tais como hibridação, protecção de RNase, clivagem química, sequenciação directa do ADN ou a utilização de enzimas de restrição (por exemplo, polimorfismos compridos de fragmentos de restrição ("PCR") e transferência para uma matriz imobilizada (ADN/ADN) de Southern do ADN genómico.

Para além da electroforese em gel mais convencional e da sequenciação de ADN, também se podem detectar as mutações por análise *in situ*.

Ensaio dos cromossomas

As sequências da presente invenção são também válidas para a identificação de cromossomas. A sequência é especificamente visada e pode hibridar com uma localização particular num cromossoma individual humano. O mapeamento dos ADNs no que respeita aos cromossomas, de acordo com a presente invenção, é um primeiro passo importante na correlação dessas sequências com os genes associados com a doença.

Em certos enquadramentos preferidos, sob este ponto de vista, o ADN aqui descrito é utilizado para clonar ADN genómico de um gene de DNA. Isto pode ser conseguido utilizando uma variedade de técnicas e bibliotecas bem conhecidas, que normalmente estão disponíveis comercialmente. O ADN genómico é utilizado para o

mapeamento de cromossomas *in situ*, utilizando técnicas bem conhecidas para este fim.

Além disso, as sequências podem ser mapeadas no que respeita aos cromossomas preparando Iniciadores de RCP (preferencialmente de 15-25 pb) a partir do ADN. A análise por computador da região 3' não traduzida do gene é utilizada para seleccionar rapidamente iniciadores que não medem mais do que um exão no ADN genómico, complicando assim o processo de amplificação. Estes iniciadores são então utilizados para a avaliação por RCP de híbridos de células somáticas contendo cromossomas individuais humanos.

A hibridação por fluorescência *in situ* ("FISH") de um clone de ADN com uma dispersão cromossómica em metáfase pode ser utilizada para providenciar uma localização precisa numa etapa. Esta técnica pode ser utilizada com ADN tão curto quanto 50 ou 60. Para uma revisão desta técnica, ver Verma e tal., *Human Chromosomes: a Manual of Basic Techniques*, Pergamon Press, New York (1988).

Logo que uma sequência tenha sido mapeada até a uma localização cromossómica precisa, a posição física da sequência no cromossoma pode ser correlacionada com os dados do mapa genético. Esses dados encontram-se, por exemplo, em V. McKusick. *Mendelian Inheritance in Man*, disponível em linha através da Johns Hopkins University, Welch Medical Library. A relação entre os genes e as doenças tem vindo a ser mapeada com a mesma região cromossómica são então identificadas através da análise de ligação (co-herança de genes fisicamente adjacentes),

Em seguida, é necessário determinar as diferenças no ADNc ou na sequência genómica entre indivíduos afectados e não afectados. Se se observa uma mutação nalgum ou em todos os indivíduos afectados, mas não se observa em nenhum dos indivíduos normais, então a mutação é provavelmente o agente causador da doença.

Vectores e Células Hospedeiras

A presente invenção também tem por objecto vectores que incluem moléculas de ADN da presente invenção, células hospedeiras que são tratadas por engenharia genética com os vectores da presente Invenção e a produção de polipéptidos da presente invenção por meio de técnicas recombinantes.

As células hospedeiras podem ser alteradas por engenharia genética para incorporar moléculas de ácidos nucleicos e expressar os polipéptidos da presente invenção. Os polinucleótidos podem ser introduzidos isoladamente ou com outros polinucleótidos. Esses outros polinucleótidos podem ser introduzidos independentemente, co-introduzidos ou introduzidos em conjunto com os polinucleótidos da presente invenção.

De acordo com este aspecto da presente invenção, o vector pode ser, por exemplo, um vector de plasmido, um vector de fago de estrutura helicoidal simples ou dupla, um ARN de estrutura helicoidal simples ou dupla ou um vector viral de ADN. Esses vectores podem ser introduzidos nas células como polinucleótidos, preferencialmente ADN, por meio de técnicas bem conhecidas para a introdução de ADN e de ARN nas células. Os vectores virais podem ser eficazes na replicação ou defeituosos na replicação. Neste último caso, a

propagação viral geralmente ocorre apenas nas células hospedeiras complementares.

Entre os vários vectores, preferem-se, em certos aspectos, aqueles que servem para a expressão de polinucleótidos e polipéptidos da presente invenção. Geralmente, esses vectores compreendem regiões de controlo activas em cis, efectivas para a expressão num hospedeiro ligado operacionalmente ao polinucleótido a ser expresso. Os factores de actuação "trans" apropriados, são fornecidos quer pelo hospedeiro ou são fornecidos por um vector complementar ou fornecidos pelo próprio vector após introdução no hospedeiro.

Pode utilizar-se uma grande variedade de vectores de expressão para expressar um polipéptido da presente invenção. Esses vectores incluem vectores cromossómicos, episómicos e derivados de vírus, por exemplo, vectores derivados de plasmídeos bacterianos, de bacteriófagos, de episomas de fungos, de elementos cromossómicos de fungos, de vírus, tais como, baculovírus, vírus papova, tais como SV40, vírus de vaccínia, adenovírus, poxvírus de aves, vírus pseudorrábico e retrovírus e vectores derivados da combinação dos anteriores, tais como aqueles que derivam de plasmídeos e elementos genéticos bacteriófagos, tais como, cosmídeos e fagemídeos, podendo todos eles ser utilizados para a expressão de acordo com este aspecto da presente invenção. Geralmente, qualquer vector apropriado para manter, propagar ou expressar polinucleótidos para expressar um polipéptido num hospedeiro, pode ser utilizado para a expressão aqui contemplada.

A sequência de ADN no vector de expressão está ligada operacionalmente à sequência ou às sequências de

controle de expressão apropriadas, incluindo, por exemplo, um promotor para dirigir a transcrição de ARNm. Representantes desses promotores incluem o promotor de fago lambda PL, os promotores *lac*, *trp* e *tac* de *E. coli*, os promotores SV40 precoces e tardios e os promotores de LTRs retrovirais, para enumerar ~~apenas~~ alguns dos ~~promotores~~ bem conhecidos. Em geral, as estruturas de expressão deverão conter sítios para a transcrição, a iniciação e a terminação ~~o~~, na região trans-crita, um sítio de ligação do ribossoma para a tradução. A porção de codificação dos transcriptos maduros expressa pelas estruturas incluirá uma iniciação da tradução AUG no início e um ~~codão~~ de terminação (UAA, UGA ou UAG) apropriadamente posicionado na extremidade do polipéptido a ser traduzido.

Além disso, as estruturas podem conter regiões de controle, que regulam, assim como controlam a expressão. Geralmente, essas regiões ~~vão~~ operar por transcrição do controle, tal como sítios de ligação do repressor e potenciadores, entre outros.

Os vectores para a propagação e a expressão geralmente incluirão marcadores seleccionáveis. Esses marcadores também podem ser apropriados para a amplificação ou os vectores podem conter marcadores adicionais ~~para~~ este fim. A este respeito, ~~os~~ vectores de expressão ~~contêm~~, preferencialmente, um ~~ou~~ mais genes de marcadores seleccionáveis para providenciar um traço fenotípico para a selecção de células hospedeiras transformadas. ~~Os marcadores preferidos~~ incluem, a di-hidrofolato redutase ou a resistência à ~~medicina~~ para culturas de células eucarióticas e genes de resistência a tetraciclina ou à ampicilina para a cultura de *E. coli* e outras bactérias.

O vector que contém a sequência de ADN apropriada, tal como se descrever; noutro sítio aqui, assim como um promotor apropriado e outras sequências de controlo apropriadas, podem ser introduzidos num hospedeiro apropriado utilizando uma variedade de técnicas bem conhecidas apropriadas para a expressão de um polipéptido desejado. Exemplos representativos de hospedeiros apropriados incluem células bacterianas, tais como células de *E. coli*, *Streptomyces* e *Salmonella typhimurium*; células de fungos, tal como, células de levedura; células de insectos, tais como células S2 de *Drosophila* e células Sf9 de *Spodoptera*; células de animais, tais como, células de OHC, COS e células do melanoma de Bowes; e células de plantas. Os hospedeiros para uma grande variedade de estruturas de expressão são bem conhecidos e os especialistas na matéria estarão habilitados com a presente memória descritiva, para seleccionarem facilmente um hospedeiro para expressar um polipéptido de acordo com este aspecto da presente invenção.

Entre os vectores preferidos para serem utilizados em bactérias estão pQE70, pQE60 e pQE-9, disponíveis na Qiagen; vectores de pBS, vectores de Phagescript, vectores de Bluescript, pNH55A, pNH16a, pNH18A, pNH46A, disponíveis na Stratagene; e ptrc99a, pKK223-3, pKK233-3, pDR4540, pRITS, disponíveis na Pharmacia. Entre os vectores eucarióticos preferidos estão pWLNEO, pSV2CAT, pOG44, pXT1 e pSG, disponíveis na Stratagene; e pSVK3, pBPV, pMSG e pSVL, disponíveis na Pharmacia. Estes vectores são listados apenas a título ilustrativo de muitos dos disponíveis comercialmente e que são vectores bem conhecidos dos especialistas na matéria.

A selecção aos vectores e promotores apropriados para a expressão numa célula hospedeira é um processo bem conhecido e os requisitos técnicos para a construção do vector de expressão, a introdução do vector num hospedeiro e a expressão no hospedeiro, são rotinas dos técnicos da matéria.

A presente invenção também tem por objecto células hospedeiras contendo as estruturas descritas antes e discutidas antes. A célula hospedeira pode ser uma célula eucariótica superior, tal como, uma célula de mamífero ou uma célula eucariótica inferior, tal como, uma célula de levedura ou a célula hospedeira pode ser uma célula procariótica, tal como, uma célula bacteriana.

A introdução da estrutura na célula hospedeira pode ser efectuada por transfecção com fosfato de cálcio, transfecção mediada por DEAE-dextrano, transfecção mediada por lípidos catiónicos, electroporação, transdução, infecção ou outros processos. Esses processos estão descritos em muitos manuais de normalização laboratorial, tal como, Davis *et al.*, *Basic Methods in Molecular Biology* (1986).

C polipéptido pode ser expresso de uma forma modificada, tal como, uma proteína de fusão e pode incluir não apenas os sinais de secreção mas também regiões funcionais heterólogas adicionais. Assim, por exemplo, pode-se adicionar uma região de aminoácidos adicionais, particularmente, aminoácidos carregados, à terminação N do polipéptido para melhorar a estabilidade e a persistência na célula hospedeira, durante a purificação ou durante o manuseamento e armazenagem subseqüentes. Adicionalmente, pode também adicionar-se uma região ao polipéptido para facilitar a purificação.

Essas regiões podem ser eliminadas antes da preparação final do polipéptido. A adição de partes de péptido aos polipéptidos para provocar a secreção ou a excreção, para melhorar a estabilidade e para facilitar a purificação, entre outras, são técnicas familiares e de rotina na técnica. Uma proteína de fusão preferida compreende uma região heteróloga da imunoglobulina que é útil para solubilizar proteínas. Por exemplo, a patente de invenção europeia EP-A-0 464 533 (contraparte canadiana 2045869) descreve proteínas de fusão que compreendem várias porções da região constante das moléculas de imunoglobulina em conjunto com outras proteínas humanas ou partes delas. Em muitos casos, a parte Fc numa proteína de fusão, é fortemente vantajosa para ser utilizada na terapia e no diagnóstico e isto resulta, por exemplo, em melhores propriedades farmacocinéticas (patente de invenção europeia EP-A 0232 262). Por outro lado, para algumas utilizações, será desejável ser capaz de eliminar a parte Fc da proteína de fusão à medida que tenha sido expressa, detectada e purificada, de uma forma vantajosa já descrita. Este é o caso quando a porção de Fc prova ser um problema para a utilização em terapia e diagnóstico, por exemplo, quando a proteína de fusão vai ser usada como um antigénio para imunizações. Na descoberta do fármaco, por exemplo, as proteínas humanas, tal como, o receptor de IL-5 humana têm sido fundidas com porções de Fc para fins de ensaios de avaliação de alto rendimento para identificar antagonistas de IL-5 humana. Ver, D. Bennett *et al.*, *Journal of Molecular Recognition*, 8: 52-58 (1995) e K. Johanson *et al.*, *The Journal of Biological Chemistry*, 270: 9459-9471 (1995).

Os polipeptídeos de D₄ podem ser recuperados e purificados a partir de culturas de células recombinantes, por processos bem conhecidos que incluem a

precipitação no seio de sulfato de amónio ou de etanol, extracção por ácidos, a cromatografia de permuta aniónica ou catiónica, a cromatografia em fosfocelulose, a cromatografia de interacção hidrofóbica, a cromatografia de afinidade, a cromatografia em hidroxilapatite e a cromatografia em leptina. Mais preferencialmente utiliza-se para a purificação a cromatografia líquida de elevada resolução ("CLER"). Podem utilizar-se técnicas bem conhecidas para a redobragem das proteínas para regenerar a conformação activa quando o polipéptido é desnaturado durante o isolamento e/ou a purificação.

Os polipéptidos da presente invenção incluem produtos purificados de forma natural, produtos resultantes de processos de síntese química e produtos produzidos por técnicas recombinantes a partir de um hospedeiro procariótico ou eucariótico, incluindo, por exemplo, células de bactérias, de fungos, de plantas superiores, de insectos e de mamíferos. Consoante o hospedeiro utilizado num processo de produção recombinante, os polipéptidos da presente invenção podem ser glicosilados ou podem ser não glicosilados. Além disso, os polipéptidos da presente invenção podem também incluir um resíduo inicial de metionina modificada, nalguns casos, como resultado de processos mediados pelo hospedeiro.

Os polinucleótidos e os polipéptidos de DM4 podem ser utilizados de acordo com a presente invenção, para uma variedade de aplicações, particularmente aquelas que fazem uso das propriedades químicas e biológicas de DM4. Entre estas, estão as aplicações em tratamento de tumores, resistência a parasitas, bactérias e vírus, para induzir a proliferação de células T e células endoteliais e certas células hematopoiéticas, para tratar restenose,

enxertos versus doenças do hospedeiro, para regular respostas antivirais e para prevenir certas doenças auto-imunes após a estimulação de DM4 por um agonista. As aplicações adicionais estão relacionadas com o diagnóstico e com o tratamento de distúrbios de células, tecidos e organismos. Estes aspectos da presente invenção serão discutidos melhor a seguir.

Polipéptidos e Fragmentos de DM4

A presente invenção tem ainda por objecto um polipéptido de DM4 isolado, comportando a sequência de aminoácidos mostrada na fig. 1 [SEQ ID NO: 2] ou um polipéptido ou um péptido que compreende uma porção das polipéptidos anteriores.

Para melhorar ou alerar as características dos polipéptidos de DM4 pode-se utilizar engenharia de proteínas. A tecnologia do ADN recombinante, conhecida pelos especialistas na matéria, pode ser utilizada para criar novas proteínas mutantes ou "muteínas que incluem substituições, eliminações, adições simples ou múltiplas de aminoácidos ou proteínas de fusão, Esses polipéptidos modificados podem mostrar, por exemplo, uma actividade aumentada ou uma estabilidade aumentada. Além disso, podem ser purificados com rendimentos mais elevados e mostram uma melhor estabilidade do que os polipéptidos naturais correspondentes, pelo menos sob certas condições de purificação e de armazenagem.

Por exemplo, para muitas proteínas, incluindo o domínio extracelular de uma proteína associada a uma membrana ou as formas maduras de uma proteína segregada, sabe-se na técnica que se pode eliminar um ou mais

aminoácidos da terminação N ou da terminação C sem uma perda substancial da função biológica. Por exemplo, Ron e tal., J. Biol. Chem., 268: 2984-2988 (1993) relatou proteínas KGF modificadas que têm actividade de ligação a heparina mesmo se estiverem a faltar 3, 8 ou 27 resíduos de aminoácidos de terminação amino. No caso presente, dada que a proteína da presente invenção é um elemento da família dos polipéptidos dos receptores contendo o domínio de morte (RCDM), ~~eliminações~~ dos aminoácidos de terminal N até ao resíduo de cisteína na posição 132 na SEQ IC NO: 2, ela pode reter alguma actividade biológica tal como a capacidade para induzir a apoptose. Os polipéptidos que, além disso, ainda contêm eliminações do terminal N incluindo o resíduo de cisteína na posição 132 (C-132) na SEQ ID NO: 2, não é provável que retenham essas actividades biológicas porque este resíduo que está conservado entre os elementos da família, ver figura 2, pode ser necessário para a formação de uma ponte de dissulfureto para providenciar a estabilidade estrutural que é necessária para a ligação do receptor.

Contudo, mesmo se a eliminação de um ou mais aminoácidos da terminação N de uma proteína resultar na modificação da perda de uma ou mais funções biológicas da proteína, podem ainda ser relidas outras actividades biológicas, Assim, a capacidade da proteína encurtada para induzir ~~e/ou~~ ligar-se a anticorpos que reconhecem o domínio completo ~~ou~~ extracelular da proteína geralmente serão retidos quando menos do que a maioria dos resíduos da proteína do domínio completo ou extracelular são eliminados da terminação N. Se um polipéptido particular a que faltam resíduos da terminação N ~~de~~ uma proteína completa retém essas actividades biológicas, pode ser facilmente determinado por processos ~~de~~ rotina aqui descritos e de alguma forma conhecidos na técnica.

De acordo com isto, a presente invenção ainda tem por objecto polipéptidos em que foram eliminados um ou mais resíduos das terminações amino da sequência de aminoácidos do DM4 mostrado na SEQ ID NC: 2, até ao resíduo C-132 e polinucleótidos que codificam esses polipéptidos. Em particular, a presente invenção tem por objecto polipéptidos que compreendem a sequência de aminoácidos dos resíduos n-468 da SEQ ID NO: 2 em que o símbolo n representa um número inteiro no intervalo de 1-132, em que C-132 é o primeiro resíduo da terminação N do domínio extracelular do polipéptido de DM4 (mostrado na SEQ ID NO: 2) que se crê que seja necessário para a actividade de ligação do ligando do receptor (por exemplo, a ligação de LIART) da proteína de DM4. A presente invenção ainda tem por objecto os polinucleótidos que codificam estes polipéptidos.

Do mesmo modo, conhecem-se muitos exemplos de muteínas com eliminações no terminal C funcionais sob o ponto de vista biológico. Por exemplo, o interferão gama exibe actividades até dez vezes mais elevadas eliminando 8-10 resíduos de aminoácidos da terminação carboxi da proteína (Döbeli et al., J. Biotechnology 7: 199-216 (1988)). No presente caso, dado que a proteína da presente invenção é um elemento da família dos polipéptidos dos RCDM, as eliminações de aminoácidos do terminal C até à cisteína na posição 221 (C-221) da SEQ ID NO: 2, podem reter alguma actividade biológica tal como a ligação do receptor. Não se espera que os polipéptidos que têm ainda eliminações de aminoácidos do terminal V, incluindo C-221 da SEQ ID NO: 2, retenham essas actividades biológicas porque este resíduo está conservado entre os elementos da família dos RCCM e é necessário para a formação de uma ponte de di-sulfureto para providenciar a estabilidade

estrutural que é necessária para a ligação receptor-ligando.

Contudo, mesmo se a eliminação de um ou mais aminoácidos da terminação C de uma proteína resultar na modificação da perda de uma ou mais funções biológicas da proteína, podem ainda ser retidas outras actividades biológicas. Assim, a capacidade da proteína encurtada para induzir e/ou para se ligar a anticorpos que reconhecem o domínio completo ou extracelular geralmente estará retida quando menos do que a maioria dos resíduos do domínio extracelular ou completo são eliminados da terminação C. Se um polipéptido particular a que faltam resíduos do terminal C de uma proteína completa retém essas actividades imunológicas pode ser facilmente determinado por processos de rotina aqui descritos e outros conhecidos na técnica.

De acordo com isto, a presente invenção tem ainda por objecto polipéptidos que têm um ou mais resíduos da terminação carboxi da sequência de aminoácidos da DM4 mostrada na SEQ ID NO: 2, até C-221 da SEQ ID NO: 2 e polinucleótidos que codificam esses polipéptidos. Em particular, a presente invenção tem ainda por objecto polipéptidos que têm a sequência de aminoácidos dos resíduos i-m da sequência de aminoácidos da SEQ ID NO: 2, em que o símbolo m representa qualquer número inteiro no intervalo de 221-468 e o resíduo C-221 é a posição do primeiro resíduo na terminação C do polipéptido completo de DM4 (mostrado na SEQ ID NO: 2) crê-se ser necessário para a actividade de ligação ao receptor da proteína de DM4. Também tem por objecto polinucleótidos que codificam estes polipéptidos.

A presente invenção também tem por objecto polipéptidos a que foram eliminados um ou mais aminoácidos de ambos os terminais amino e carboxi, que podem ser descritos, de uma forma geral, como tendo os resíduos n-m da SEQ ID NO: 2, em que os símbolos n e m são números inteiros como se descreveu antes.

Também está incluída uma sequência de nucleótido que codifica um polipéptido que consiste numa porção de uma sequência de aminoácidos completa de DM4 codificada pelo clone de ADN contido no depósito nº 97951 da ATCC, em que esta porção exclui de 1 a cerca de 108 aminoácidos da terminação amino da sequência completa de aminoácidos codificada pelo clone de ADN da ATCC com o nº de depósito 97853 ou de 1 a cerca de 247 aminoácidos da terminação carboxi ou qualquer combinação das eliminações dos terminais amino e carboxi anteriores da sequência de aminoácidos completa codificada pelo clone de ADN contido na ALCC com o nº de depósito 97853. Também tem por objecto os polinucleótidos que codificam todas as formas de polipéptidos dos mutantes anteriores com eliminações.

Entre os mutantes de eliminação preferidos dos terminais N e C estão aqueles que compreendem apenas uma porção do domínio extracelular, isto é, dentro dos resíduos 24-238, dado que é expectável que qualquer uma das suas porções seja solúvel.

É sabido da técnica que algumas sequências de aminoácidos de DM4 podem variar sem um efeito significativo na estrutura ou na função da proteína. Se essas diferenças na sequência estão contempladas, deve ser recordado que haverá áreas críticas nas proteínas que determinam a actividade. Essas áreas compreenderão,

normalmente, resíduos que fazem o sítio de ligação do ligando ou o domínio de morte cu que formam estruturas terciárias que afectam estes domínios.

Assim, a presente invenção inclui ainda variações da proteína de DM4 que mostram uma actividade substancial da proteína de DM4 ou que incluem regiões de DM4, tal como os fragmentos de proteínas discutidas a seguir. Alguns mutantes incluem eliminações, inserções, inversões, repetições e substituições de vários tipos. Tal como se indicou antes, as normas respeitantes às alterações dos aminoácidos, são provavelmente silenciosas sob o ponto de vista fenotípico e podem ser encontradas em Bowie, J. U. *et al.*, *Science* 247: 1306-1310 (1990).

De particular Interesse, são as substituições dos aminoácidos carregados por outros aminoácidos carregados e com aminoácidos carregados negativamente ou neutros. Estes últimos resultam em proteínas com cargas positivas reduzidas, para melhorar as características da proteína de DML. A prevenção da agregação é altamente desejável. A agregação das proteínas não resulta apenas numa perda de actividade mas pode também ser problemática quando se preparam formulações farmacêuticas, porque podem ser imunogénicas. (Pinckard *et al.*, *Clin. Exp. Immunol.* 2: 331-340 (1967); Robbins *et al.*, *Diabetes* 36: 838-845 (1987); Cleland *et al.*, *Crit. Rev. Therapeutic Drug Carrier Systems* 10: 307-377 (1993)).

A substituição dos aminoácidos pode também alterar a selectividade da ligação aos receptores da superfície das células. Ostade *et al.*, *Nature* 361: 266-268 (1993) descreve certas mutações que resultam na ligação selectiva de FNT-alfa apenas a um dos dois tipos de receptores de FNT conhecidos. Assim, o receptor de DM4 da

presente invenção, pode incluir una ou mais substituições, eliminações ou adições de aminoácidos, quer de mutações naturais ou de manipulação humana,

Tal como se indicou, as alterações são preferencialmente de natureza menor, tal como, as substituições conservadoras de aminoácidos, que não afectam significativamente a dobragem ou a actividade da proteína (ver quadro 1).

QUADRO 1. Substituições Conservadoras de Aminoácidos

Aromático	Fenilalanina Tryptofano Tirosina
Hidrofóbico	Leucina Isoleucina Valina
Polar	Glutamina Asparagina
Básico	Arginina Lisina Histidina
Ácido	Ácido Aspartático Ácido Glutâmico
Pequeno	Alanina Serina Trionina Metionina Glicina

Os aminoácidos da proteína de DM4 da presente invenção que são essenciais para a função, podem ser identificados por processos conhecidos nesta técnica, tal como, a mutagénese dirigida ao sítio ou a mutagénese de avaliação da alanina (Cunningham e Wells, *Science* 244: 1081-1085 (1989)). Este último processo introduz mutações apenas de alanina em cada resíduo da molécula. As moléculas mutantes resultantes são então ensaiadas para pesquisar a sua actividade biológica, tal como, a ligação do receptor ou a actividade proliferativa *in vitro* ou *in vivo*. Os sítios que são críticos para a ligação de ligando-receptor, podem também ser determinados por análise estrutural, tal como cristalização, ressonância magnética nuclear ou marcação por foto-afinidade (Smith et al., *S. Mol. Biol.* 224: 899-904 (1992) e de Vos et al. *Science* 255: 306-312 (1992)).

Os polipéptidos da presente invenção são preferencialmente produzidos sob uma forma isolada e preferencialmente são praticamente puros. Uma versão do polipéptido DM4 produzido de forma recombinante pode ser praticamente purificada por um processo numa única etapa, descrito em Smith e Johnson, *Gene* 67: 31-40 (1988).

Os polipéptidos da presente invenção também incluem o polipéptido codificado pelo ADNc depositado, incluindo o polipéptido líder, maduro, codificada pelo ADNc depositado menos o líder (isto é, a proteína madura); o polipéptido da figura I (SEQ ID NO: 2), incluindo o líder, o polipéptido da figura 1 (SEQ ID NO: 2) menos a metionina do terminal amino, o polipéptido da figura 1 (SEQ ID NO: 2) menos o líder, o domínio extracelular, o domínio da membrana, o domínio intracelular, o domínio de morte, os polipéptidos solúveis que compreendem todo ou parte dos domínios extracelular e

intracelular mas a que falta o domínio da transmembrana, assim como os polipéptidos que são pelo menos 80 % idênticos, mais preferencialmente que são pelo menos 90 % ou 95 % idênticos, ainda mais preferencialmente, pelo menos 96 , 97 , 38 % ou 99 % idênticos aos polipéptidos codificados pelos clones de ADMc depositados, com os da figura 1 (SEQ ID NO: 2) e também incluem porções desses polipéptidos com, pelo menos 30 aminoácidos e, mais preferencialmente, pelo menos 50 aminoácidos.

Por um polipéptido com uma sequência de aminoácidos pelo menos, por exemplo, 95 % "idêntica" a uma sequência de aminoácidos de referência de um polipéptido de DM4, entende-se que a sequência de aminoácidos do polipéptido é idêntica à sequência de referência, excepto no facto da sequência de polipéptido poder incluir até cinco alterações de aminoácidos por cada 100 aminoácidos da sequência de aminoácidos de referência do polipéptido DM4. Por outras palavras, para se obter um polipéptido com uma sequência de aminoácidos pelo menos 95 % idêntica a uma sequência de aminoácidos de referência, pode-se eliminar ou substituir até 5 % dos resíduos de aminoácidos na sequência de referência, com outros aminoácidos ou um certo número de aminoácidos até 5 % do total dos resíduos de aminoácidos na sequência de referência, podem ser inseridos na sequência de referência. Estas alterações da sequência de referência podem ocorrer nas posições terminais amino ou carboxi da sequência de aminoácidos de referência ou em qualquer parte entre estas posições terminais, espalhados entre resíduos individualmente na sequência de referência ou em um ou mais grupos contíguos dentro da sequência de referência.

Como matéria prática, se qualquer polipéptido em particular, é pelo menos 90 %, 95 %, 96 %, 97 %, 98 % ou 99 % idêntico a, por exemplo, à sequência de aminoácidos mostrada na figura 1 (SEQ ID NC: 2), ou a sequência de aminoácidos codificada pelos clones do ADNC depositado, pode ser determinada de uma forma convencional utilizando programas de computador conhecidos, tal como o programa Bestfit (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). Quando se utiliza o Bestfit ou qualquer outro programa de alinhamento de sequências para determinarse uma sequência particular é, por exemplo, 95 % idêntica a uma sequência de referência, de acordo com a presente invenção, os parâmetros são estabelecidos de tal modo que obviamente a percentagem de identidade é calculada em relação ao comprimento total da sequência de referência de aminoácidos e que as falhas ou faltas na homologia até 5 % do número total de resíduos de aminoácidos na sequência de referência são permitidas.

Os requerentes descobriram que o polipéptido DM4 é uma proteína com 468 resíduos, que exhibe três domínios estruturais principais. O primeiro, o domínio de ligação ao ligando, foi identificado dentro dos resíduos de cerca de 42 até cerca de 238 na figura I (SEQ ID MC: 2). O segundo, o domínio da transmembrana, foi identificado dentro dos resíduos de cerca de 239 até cerca de 468 na figura 1 (SEQ ID NO: 2). O terceiro, o domínio intracelular, foi identificado dentro dos resíduos de cerca de 265 até cerca de 468 na figura 1 (SEQ ID NO: 2). É importante notar que os domínios intracelulares incluem um domínio de morte nos resíduos de cerca de 379 até cerca de 422. Outros fragmentos preferidos do polipéptido mostrado na figura 1 (SEQ ID NO: 2), incluem a proteína

madura dos resíduos de cerca de 24 até cerca de 468 e os polipéptidos solúveis que compreendem todo ou parte dos domínios extracelular e intracelular, mas a que falta o domínio da transmembrana.

A presente invenção tem ainda por objecto polipéptidos de DM4 codificados pelo clone de ADN depositado, incluindo o líder e fragmentos de polipéptidos de DM4 seleccionados da proteína madura, do domínio extracelular, do domínio da transmembrana, do domínio intracelular e do domínio de morte.

Num outro aspecto, a presente invenção tem por objecto um péptido ou um polipéptido que compreende uma porção ligada ao epitopo de um polipéptido aqui descrito. O epitopo desta porção de polipéptido é um epitopo imunogénico ou antigénico de um polipéptido da presente invenção. Um "epitopo imunogénico" define-se como uma parte de uma proteína que provoca uma resposta dos anticorpos quando toda a proteína é o imunogénio. Por outro lado, uma região de uma molécula da proteína a qual um anticorpo se pode ligar, define-se como um "epitopo antigénico". O número de epítopos imunogénicos de uma proteína é geralmente inferior ao número de epítopos antigénicos. Ver, por exemplo, Geysen et al., *Proc. Natl. Acad. Sci. EUA* 81: 3998-4002 (1983).

Como para a selecção de péptidos ou de polipéptidos que se ligam a um epitopo antigénico (isto é, que contém uma região de uma molécula de proteína à qual um anticorpo se pode ligar), é bem conhecido na técnica, que relativamente poucos péptidos sintéticos que simulam parte de uma sequência de proteína, são capazes, por rotina, de provocar um anti-soro que reage com a proteína parcialmente simulada. Ver, por exemplo, Sutcliffe,

J. G., Shinnick, T. M., Green, N. e Learner, R.A., (1983) Antibodies That React With Predetermined Sites on Proteins, *Science* 219: 660-666. Os péptidos capazes de provocar soros reactivos à proteína, estão frequentemente representados na sequência primária de uma proteína e podem ser caracterizados por um conjunto de regras químicas simples e não estão confinados às regiões imuno-dominantes das proteínas intactas (isto é, epítomos imunogénicos) nem aos terminais amino ou carboxi.

Os péptidos e os polipéptidos da presente Invenção, que se ligam ao epítomo antigénico, são por isso úteis para produzir anticorpos, incluindo anticorpos monoclonais, que se ligam especificamente a um polipéptido da presente invenção. Ver, por exemplo, Wilson et al., *Cell* 37: 767-778 (1984) em 777.

Os peptidos e polipéptidos da presente invenção, que se ligam ao epítomo antigénico contêm, preferencialmente, uma sequência de pelo menos sete, mais preferencialmente, pelo menos nove e, ainda mais preferencialmente, entre pelo menos cerca de 15 até cerca de 30 aminoácidos, contidos dentro da sequência de aminoácidos de um polipéptido da presente invenção.

Exemplos não limitativos de polipéptidos antigénicos ou de péptidos antigénicos que podem ser utilizados para gerar anticorpos específicos de DM4 incluem: um polipéptido que compreende resíduos de aminoácidos de cerca de 35 até cerca de 92C na figura I (SEQ ID NO: 2); um polipéptido que compreende resíduos de aminoácidos de cerca de 114 até cerca de 180 na figura I (SEQ ID NO: 2); um polipéptido que compreende resíduos de aminoácidos de cerca de 169 até cerca de 240 na figura I (na SEQ ID NC: 2); um polipéptido que compreende resíduos de aminoácidos

de cerca de 267 até cerca de 298 na figura 1 (SEQ ID NO: 2), um polipéptido que compreende resíduos de aminoácidos de cerca de 330 até cerca de 364 na figura 1 (SEQ ID NO: 2); um polipéptido que compreende resíduos de aminoácidos de cerca de 391 até cerca de 404 na figura 1 (SEQ ID NO: 2); e um polipéptido que compreende resíduos de aminoácidos de cerca de 418 até cerca de 465 na figura 1 (SEQ 13 NO: 2). Tal como se indicou antes, os requerentes determinaram que os fragmentos dos polipéptidos anteriores, são regiões antigénicas da proteína de DM4.

Os péptidos e polipéptidos da presente invenção que se ligam ao epítipo podem ser produzidos por quaisquer meios convencionais. Houghthen, R. A., "General Method for the Rapid Solid-Phase Synthesis of Large Numbers of Peptides: Specificity of Antigen-Antibody Interaction at the Level of Individual Amino Acids", *Proc. Natl. Acad. Sci. USA* 82: 5131-5135 (1985). Este processo de "síntese simultânea de múltiplos péptidos (SSMP)", está ainda descrito na patente de invenção norte-americana U.S. NO: 4.631.211 para Houghthen et al. (1986).

Como será evidente para um especialista na matéria, os polipéptidos de DM4 da presente invenção e os seus fragmentos ligados a epítopos descritos antes, podem combinar-se com partes do domínio constante das imunoglobulinas (IgG), resultando em polipéptidos quiméricos. Estas proteínas de fusão facilitam a purificação e mostram um aumento do ~~semi-período~~ de vida *in vivo*. Isto foi já mostrado, por exemplo, para proteínas quiméricas que consistem nos primeiros dois domínios do polipéptido humano CD4 e nos vários domínios ~~das~~ regiões constantes das cadeias pesadas ou leves das imunoglobulinas de mamíferos (patente de invenção europeia EPA 394.827; Trauneker et al., *Nature* 331:

84-86 (1988)). As proteínas de fusão que têm uma estrutura dimérica ligada a um di-sulfureto devido a parte de IgG, podem também ser mais eficientes na ligação e na neutralização de outras moléculas do que a proteína monomérica DM4 ou um fragmento de proteína isolado (Fountoulakis et al., *J Biochem* 270: 3958-3964 (1995)).

Ensaio dos Polipéptidos

A presente invenção também tem por objecto ensaios de diagnóstico, tal como ensaios quantitativos e de diagnóstico para a detecção dos níveis da proteína de DM4 ou das suas formas solúveis, em células e tecidos, incluindo a determinação dos níveis normais e anormais. Assim, por exemplo, um ensaio de diagnóstico de acordo com a presente invenção, para detectar a sobre-expressão de DM4 ou da sua forma solúvel, comparado com amostras de tecido normal de controlo, pode ser utilizado para detectar, por exemplo, a presença de tumores. As técnicas de ensaio que podem ser utilizadas para determinar os níveis de uma proteína, tal como, uma proteína de DN4 da presente invenção ou de uma sua forma solúvel, numa amostra derivada de um hospedeiro, são bem conhecidos dos especialistas na matéria. Esses processos de ensaio incluem ensaios rádio-imunológicos, ensaios comparativos da ligação, análise de Western de transferência de anticorpo/proteína para uma matriz imobilizada e ensaios por ELISA.

A análise dos níveis da proteína de DM4 numa amostra biológica pode ocorrer utilizando qualquer processo conhecido na técnica. As técnicas preferidas para o ensaio dos níveis de proteína de DM4 numa amostra biológica, são as técnicas à base de anticorpos. Por exemplo, a expressão da proteína de DM4 em tecidos pode

ser estudada com processos imuno-histológicos clássicos. (Jalkanen, M. et al., *J. Cell. Biol.* 101: 976-985 (1985); Jalkanen, M. et al., *J. Cell. Biol.* 105: 3087-3096 (1987)).

Outros processos, à base de anticorpos, úteis para a detecção da expressão dos genes da proteína de DM4 incluem imuno-ensaios, tais como, o ensaio imuno-absorvente ligado a enzimas (ELISA) e o rádio-imuno-ensaio (RIA).

Conhecem-se na técnica marcadores apropriados que incluem marcadores de enzimas, tal como, glicose-oxidase, rádio-isótopos, tal como os marcadores de iodo (I^{125} , I^{131}), de carbono (C^{14}), de enxofre (S^{35}), de trítio (H^3), de índio (In^{111}) e de tecnécio (Tc^{99m}) e marcadores fluorescentes, tal como, fluoresceína e rodamina e biotina.

Terapêutica

Os ligantes da família do factor de necrose do tumor (FNT) são conhecidos por estarem entre as citocinas mais pleiotrópicas, induzindo um grande número de respostas celulares, incluindo citotoxicidade, actividade antiviral, actividades imuno-reguladoras e a regulação transcricional de vários genes (Goeddel, D.V. et al., "Tumor Necrosis Factors: Gene Structure and Biological Activities", *Quant. Biol.* 51: 597-609 (1986), Cold Spring Harbor; Beutler, B., e Cerami, A., *Annu. Rev. Biochem.* 57: 505-518 (1988); Old, L. J., *Sci. Am.* 258: 59-75 (1988); Fiers, W., *FEBS Lett.* 285: 199-224 (1991)). Os ligantes da família dos FNT induzem essas várias respostas celulares por meio da ligação aos receptores da família do FNT, incluindo o DM4 da presente invenção. As

células que expressam o polipéptido DM4 e que se crê que tenham uma potente resposta celular aos ligandos de DM4 incluem células amnióticas, leucócitos do sangue periférico, do coração, do cancro do fígado, do rim, células T activadas, tecido correspondendo a células Th2, amígdalas humanas e camadas de células brancas com falta de CD34 (sangue do cordão umbilical). Por "uma resposta celular a um ligando da família do FNT" entende-se qualquer alteração genotípica, fenotípica e/ou morfológica de uma célula, de uma linha de células, de tecido, de cultura de tecido ou de um paciente, que é induzida por um ligando da família do FNT. Tal como indicado, essas respostas celulares incluem não apenas respostas fisiológicas normais aos ligandos da família do FNT, mas também doenças associadas com um aumento da apoptose ou a inibição da apoptose. A apoptose (morte programada das células) é um mecanismo fisiológico envolvido na eliminação de linfócitos T periféricos do sistema imunitário e a sua desregulação pode levar a um certo número de diferentes processos patogénicos (Ameisen, J. C., *AIDS* 8: 1197-1213 (1994); Krammer, P. H. et al., *Curr. Opin. Immunol.* 6: 279-289 (1994)).

As doenças associadas ao aumento da sobrevivência das células ou à inibição da apoptose incluem cancros (tais como, linfomas foliculares, carcinomas com mutações de p53 e tumores dependentes de hormonas, tais como, cancro da mama, cancro da próstata, sarcoma de Kaposi e cancro do ovário); distúrbios auto-imunitários (tais como, lúpus eritematoso agudo e artrite reumatóide da glomerulonefrite relacionada com o sistema imunitário) e infecções virais (tais como, vírus da herpes poxvirus e adenovírus), inflamação, enxerto versus doença de hospedeiro, rejeição aguda de enxerto e rejeição crónica de enxerto, Doenças associadas com o aumento da apoptose

i SIDA; distúrbios neurodegenerativos (tais como, doença de Alzheimer, doença de Parkinson, esclerose lateral amiotrófica, retinite pigmentosa, degeneração cerebelar); síndromas mielodisplásicos (tais como, anemia aplásica), traumas isquémicos (tais como os causados pelo enfarte do miocárdio, acidente vascular e traumas de reperfusão), doenças do fígado induzidas por toxinas (tais como as causadas pelo álcool:, choque séptico, caquexia e anorexia.

Assim, num aspecto, a presente invenção tem por objecto um processo para aumentar a apoptose induzida por um ligando da família dos FNT, que envolve a administração a uma célula, que expressa o polipéptido de DM4, de uma quantidade de ligando, de um seu análogo ou de um agonista de DM4 capaz de aumentar a sinalização mediada por DM4. Preferencialmente, a sinalização mediada por DM4 é aumentada para tratar uma doença em que existe uma diminuição da apoptose ou uma diminuição da citocina e da expressão da molécula de adesão, Um agonista pode incluir formas solúveis de DM4 e anticorpos monoclonais dirigidos contra o polipéptido de DM4.

Num outro aspecto, a presente invenção tem por objecto um processo para inibir a apoptose, induzido por um ligando da família dos FNT, que envolve a administração a uma célula, que expressa o polipéptido de DM4, de uma quantidade efectiva de um antagonista capaz de diminuir a sinalização mediada por DM4. Preferencialmente para tratar uma doença em que existe um aumento da apoptose ou em que há expressão de NFkB. Um agonista pode incluir formas solúveis de DM4 e anticorpos monoclonais dirigidos contra o polipéptido de DM4.

Por "agonista" entende-se compostos sintéticos e de ocorrência natural, capazes de aumentarem o; de potenciarem a apoptose. Por "antagonista" entende-se compostos sintéticos e de ocorrência natural, capazes de inibir a apoptose. Pode-se determinar se qualquer candidato a "agonista" ou a "antagonista", da presente invenção pode aumentar ou inibir a apoptose utilizando ensaios de resposta celular ao receptor/ligando da família do FNT conhecidos na técnica, incluindo aqueles que são descritos com mais detalhe a seguir.

Um desses processos de rastreio envolve a utilização de melanóforos que são transfectados para expressar o receptor da presente invenção. Essas técnicas de rastreio estão descritas na patente de invenção YCT WO 92/01810, publicada em 06 de Fevereiro de 1992. Esse pode ser utilizado, por exemplo, para rastrear um composto que inibe (ou potencia) a activação do polipéptido do receptor da presente invenção, por meio do contacto das células do melanóforo que codificam o receptor, tanto com um ligando da família do FNT como com um antagonista (ou agonista) candidato. A inibição ou o aumento do sinal gerado pelo ligando indica que o composto é um antagonista ou agonista da via de sinalização do ligando/receptor.

Outras técnicas de rastreio incluem a utilização de células que expressam o receptor (por exemplo, células de OHC (ovário de hamster chinês) transfectadas) num sistema que mede as alterações extracelulares ao pH causadas pela activação do receptor, por exemplo, conforme descrito em *Science* 246: 181-296 (Outubro de 1989). Por exemplo, pode-se fazer contactar os compostos com uma célula que expressa o polipéptido do receptor da presente invenção e uma segunda resposta do mensageiro, por exemplo,

transdução do sinal ou alterações do pH, que podem ser medidas para determinar se o potencial composto activa ou inibe o receptor.

Outras técnicas de rastreio envolvem a introdução do ADN que codifica o receptor em oócitos de *Xenopus*, para expressar transientemente o receptor. Os oócitos do receptor podem então por-se em contacto com o ligando do receptor e um composto a ser rastreado seguido da detecção da inibição ou da activação de um sinal de cálcio no caso do rastreio de compostos que se pensa que inibem a activação do receptor.

Outra técnica de rastreio envolve a expressão, em células, de uma estrutura em que o receptor está ligado a uma fosfolipase C ou D. Essas células incluem células endoteliais, células do músculo liso, células embrionárias do rim, etc. O rastreio pode ser realizado conforme se descreveu aqui antes, por meio da detecção da activação do receptor ou da inibição da activação do receptor a partir do sinal de fosfolipase.

Outro processo envolve o rastreio de compostos que inibem a activação do polipéptido do receptor dos antagonistas da presente invenção, por meio da determinação da inibição da ligação do ligando marcado às células que têm o receptor na sua superfície. Esse processo envolve a transfecção de uma célula eucariótica com ADN, que codifica o receptor, de tal modo que a célula expressa o receptor na sua superfície e contacta a célula com um composto na presença de uma forma marcada de um ligando conhecido. O ligando pode ser marcado, por exemplo, por radioactividade. A quantidade de ligando marcado ligado aos receptores é medida, por exemplo, medindo a radioactividade dos receptores. Se o composto

se liga ao receptor conforme determinado por uma redução do ligando marcado que se liga aos receptores, a ligação do ligando marcado aos receptores está inibida.

Outros ensaios de rastreio para os agonistas e os antagonistas da presente invenção, estão descritos em Tartaglia, L. A. e Goeddel, D. V., *J. Biol. Chem.* 267: 4304-4307 (1992).

Assim, num outro aspecto, descreve-se um processo de rastreio para determinar se um agonista ou um antagonista candidato é capaz de aumentar ou de inibir uma resposta celular o um ligando da família do FNT. O processo envolve o contacto das células que expressam o polipéptido de DM4 com um composto candidato como um ligando da família do FNT, ensaiando uma resposta celular e comparando a resposta celular a uma resposta celular padrão, sendo o padrão analisado quando se faz o contacto com o ligando na ausência do composto candidato, em que um aumento da resposta celular em relação ao padrão, indica que o composto candidato é um agonista da via de sinalização do ligando/receptor e uma diminuição da resposta celular comparada com a do padrão, indica que o composto candidato é um antagonista da via de sinalização do ligando/receptor. Por "análise de uma resposta celular" entende-se a medição qualitativa ou quantitativa de uma resposta celular a um composto candidato e/ou um ligando da família do FNT (por exemplo, determinação ou estimativa de um aumento ou de uma diminuição na proliferação das células T ou na marcação de timidina tritiada). De acordo com a presente invenção, uma célula que expressa o polipéptido de DM4 pode ser contactada com um ligando da família do FNT administrado quer endogenamente, quer exogenamente.

Agonista, de acordo com a presente invenção, inclui compostos sintéticos e de ocorrência natural, tal como, por exemplo, fragmentos de péptidos de ligando da família do FNT, factor de crescimento da transformação, neurotransmissores (tais como, glutamato, dopamina, N-metil-D-aspartato), supressores de tumor (p53), células T citolíticas e antimetabolitos. Os agonistas preferidos incluem fármacos quimioterapêuticos, tais como, por exemplo, cisplatina, doxorubicina, bleomicina, arabinósido de citosina, mostarda de azoto, metotrexato e vincristina. Outros incluem péptidos de etanol e de amilóide. (*Science* 267: 1457-1458 (1995)). Outros agonistas preferidos incluem anticorpos policlonais e monoclonais dirigidos contra o polipéptido de DM4 ou um seu fragmento. Esses anticorpos de agonistas dirigidos contra um receptor da família ao FNT estão descritos em Tartaglia, L. A., et al., Proc. Natl. Acad. Sci. EUA 88: 9292-9296 (1991); e Tartaglia, L. A. e Goeddel, D. V., J. Biol. Chem. 267 (7): 4304-4307 (1992). Ver também, o pedido de patente de invenção PCT WO 94/09137.

Antagonista, de acordo com a presente invenção, inclui compostos sintéticos e de ocorrência natural, tal como, por exemplo, o ligando CD40, aminoácidos neutros, zinco, estrogénio, androgénios, genes virais (tais como adenovírus E1B, baculovírus p35 e IAP, poxvírus de vacas crmA, vírus de Epstein-Barr BHRF1, LMP-1, vírus da febre suína africana LMW5-HL, e vírus do herpes yl 34.5), inibidores de calpaína, inibidores de cisteína protease e promotores de tumor (tais como, PMA, fenobarbital e alfa-hexaclorociclo-hexano).

Outros antagonistas potenciais incluem moléculas de sentido inverso (anti-paralelas). A tecnologia de sentido inverso (anti-sense ou anti-paralela) pode ser utilizada

para controlar a expressão dos genes através do ADN ou do ARN anti-paralelo ou através da formação de hélices triplas. As técnicas de sentido inverso estão discutidas, por exemplo, em Okano, *J. Neurochem.* 56: 560 (1991); *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988). A formação de hélices triplas discute-se, por exemplo, em Lee et al., *Nucleic Acids Research* 6: 3073 (1979); Cooney et al., *Science* 241: 456 (1988); e Dervan et al., *Science* 251: 1360 (1991). Os processos baseiam-se na ligação de um polinucleótido a um ADN ou a um ARN complementar.

Por exemplo, a porção de codificação de 5' de um polinucleótido que codifica o polipéptido maduro da presente invenção, pode ser utilizada para desenhar um oligonucleótido de ARN de sentido inverso, com desde cerca de 10 até cerca de 40 pares de bases de comprimento. Gesenha-se um oligonucleótido de ADN para ser complementar de uma região do gene envolvido na transcrição, prevenindo assim a transcrição e a produção do receptor. O oligonucleótido de ARN de sentido inverso hibrida com o ARNm *in vivo* e bloqueia a tradução da molécula de ARNm no polipéptido do receptor. Os oligonucleótidos descritos antes podem também ser libertados para as células de tal modo que o ARN ou o ADN anti-paralelo pode ser expresso *in vivo* para inibir a produção do receptor.

Outros antagonistas, de acordo com a presente invenção, incluem formas solúveis de DM4, isto é, fragmentos de DM4 que incluem o domínio de ligação do ligando da região extracelular do receptor de comprimento completo. Essas formas solúveis do receptor, que podem ser sintéticas ou de ocorrência natural, antagonizam a sinalização mediada por DM4 competindo com DM4 da

superfície das células para a ligação aos ligandos da família do FNT. Assim, as formas solúveis do receptor que incluem o domínio de ligação do ligando, são citocinas novas capazes de inibir a apoptose induzida pelos ligandos da família do FNT. Estes são preferencialmente expressos como dímeros ou trímeros, dado que estes se têm mostrado como sendo superiores às formas monoméricas do receptor solúvel, como antagonistas, por exemplo, fusões da família do receptor de FNT com IgG-Fc. Outras dessas citocinas são conhecidas na técnica e incluem Fas B (uma forma solúvel do receptor Fas de rato), que actua fisiologicamente para limitar a apoptose induzida pelo ligando de Fas (Hughes, D. P. e Crispe, I. N., *J. Exp. Med.* 182: 1395-1401 (1995)).

As experiências estabelecidas no exemplo 5 demonstram que DM4 é uma molécula contendo o domínio de morte capaz de provocar a apoptose que é importante na regulação do sistema imunitário. Além disso, as experiências estabelecidas a seguir demonstram que a apoptose induzida por DM4 foi bloqueada pelos inibidores das proteases tal como ICE, Cma e z-VAD-fmk. Assim, os inibidores de proteases tal como ICE, FADD-DN e FLICE-DN/MACH103608 podem ser utilizados como antagonistas da actividade de DM4.

O termo "anticorpo" (Ac) ou "anticorpo monoclonal" (Acm), tal como se utiliza aqui, é suposto incluir moléculas intactas, assim como os seus fragmentos (tais como, por exemplo, fragmentos de Fab e F(ab')₂), que são capazes de se ligar a um antigénio. Os fragmentos de Fab e F(ab')₂ a que falta o fragmento Fc do anticorpo Intacto, desaparecem mais rapidamente da circulação e podem ter menos ligação ao tecido não específico de um

anticorpo intacto (Wahl et al., *J. Nucl. Med.* 24: 316-325 (1983)).

Os anticorpos de acordo com a presente invenção podem ser preparados por qualquer de uma da variedade de processos utilizando os imunogénios de DM4 da presente invenção. Tal como se indicou, esses imunogénios de DM4 incluem o polipéptido de DM4 de comprimento completo (que pode ou não incluir a sequência líder) e fragmentos do polipéptido de DM4, tal como o domínio de ligação do ligando, o domínio da transmembrana, o domínio intracelular e o domínio de morte.

As proteínas e outros compostos que se ligam aos domínios de DM4 são também agonistas e antagonistas candidatos, de acordo com a presente invenção. Esses compostos de ligação podem ser "capturados" utilizando o sistema híbrido de dois fungos (Fields e Song, *Nature* 340: 245-246 (1989)). Uma versão modificada do sistema híbrido de dois fungos foi descrita por Roger Brent e os seus colegas (Gyuris, J. et al., *Cell* 75: 791-803 (1993); Zervos, A. S. et al., *Cell* 72: 223-232 (1993)). Preferencialmente, utiliza-se o sistema híbrido de dois fungos de acordo com a presente invenção para capturar compostos que se ligam quer ao domínio de ligação do ligando de DM4, quer ao domínio intracelular de DM4. Esses compostos são candidatos a agonistas e a antagonistas do presente invenção.

Por um "ligando da família do FNT" entende-se ligandos de ocorrência natural, recombinantes e sintéticos, que são capazes de se ligarem a um membro da família dos receptores de FNT e induzir a via de sinalização do ligando/receptor. Os membros da família do FNT incluem, mas não se limitam, a ligandos de DM4,

incluindo LIART, FNT- α , linfotóxina- α (LT- α , também conhecida como FNT- β), LT- β (encontrado no heterotrímico complexo LT- α 2- β), FasL, CD40, CD27, CD30, 4-1BB, OX40 e factor de crescimento do nervo (FCN).

Aplicações terapêuticas representativas da presente invenção são descritas com mais detalhe a seguir. O estado de imunodeficiência que define a SIDA é secundário em relação a uma diminuição do número e da função dos linfócitos T de CD4⁺. Recentes relatórios estimam que a perda diária das células T de CD4⁺ está entre $3,5 \times 10^7$ e 2×10^9 células (Wei X. et al., *Nature* 373: 117-122 (1995)). Crê-se que uma causa da diminuição das células T de CD4⁺ no estabelecimento da infecção por VIH, seja a apoptose induzida por VIH. Na verdade, a morte apoptótica das células induzida por VIH, foi demonstrada não só *in vitro* mas também, e o que é mais importante, em indivíduos infectados (Ameisen, J. C., *AIDS* 8: 1197-1213 (1994); Finkel, T. H., e Banda, N. K., *Curr. Opin. Immunol.* 6: 605-615 (1995); Muro-Cacho, C. A. et al., *J. Immunol.* 154: 5555-5566 (1995)). Além disso, a apoptose e a diminuição dos linfócitos T de CD4⁺ está fortemente correlacionada em diferentes modelos de animais, com a SIDA (Brunner, T., et al., *Nature* 373: 441-444 (1995); Gougeon, M. L., et al., *AIDS Res. Hum. Retroviruses* 9: 553-563 (1993)) e não se observou apoptose nesses modelos de animais em que a replicação viral não resultou em SIDA (Gougeon, M. L. et al., *AIDS Res. Hum. Retroviruses* 9: 553-563 (1993)). Outros dados indicam que linfócitos T não infectados mas iniciados ou activados a partir de indivíduos infectados com VIH sofrem apoptose depois de encontrarem o ligando da família do FNT FasL. Utilizando linhas de células monocíticas que resultam na morte, no seguimento da infecção por VIH, tem-se demonstrado que a infecção das células U937 com VIH, resulta na expressão

de novo de FasL e que FasL medeia a apoptose induzida por VIH (Badley, A. D. et al., *J. Virol.* 70: 199-206 (1996)). Além disso, o ligando da família do FNT foi detectável em macrófagos não infectados e a sua expressão foi sobrerregulada no seguimento da infecção por VIH, resultando na morte selectiva de linfócitos T ~~de~~ CD4 não infectados (Badley, A. D. et al., *J. Virol.* 70: 199-206 (1996)). Assim, por meio da presente invenção, providencia-se para tratar indivíduos positivos ao VIH, que envolve a administração de um antagonista da presente invenção para reduzir a morte selectiva de linfócitos T de CD4. Os modos de administração e as dosagens estão discutidos em detalhe a seguir.

Na rejeição de um alo-enxerto, o sistema imunitário do animal receptor não foi previamente iniciado para responder porque o sistema imunitário, na sua maior parte, é apenas desencadeado por antigénios ambientais. Os tecidos de outros membros ou das mesmas espécies não ~~têm~~ sido apresentados da mesma maneira que, por exemplo, os vírus e as bactérias se têm apresentado. No caso de rejeição dos alo-enxertos, os regimes imunossupressores são desenhados de forma a prevenir que o sistema imunitário atinja o estado de efector. Contudo, o perfil imunitário da rejeição ao xeno-enxerto pode simular a recorrência da doença mais do que a rejeição do alo-enxerto. No caso de recorrência da doença, o sistema imunitário já tinha sido activado, como ~~se~~ pode ~~ver~~ em evidência pela destruição das células de ilhetos originais. Por isso, na ~~recorrência~~ da doença, o sistema imunitário está ~~já~~ no estado de efector. Os agonistas da presente invenção são capazes de suprimir a resposta ~~imunitária~~ tanto aos alo-enxertos como aos xeno-enxertos, porque os linfócitos activados e diferenciados nas células efectoras ~~vão~~ expressar o polipéptido de DM4 e

assim são susceptíveis em relação aos compostos que aumentam a apoptose. Assim, a presente invenção tem ainda por objecto um processo para a criação de tecidos imunitários privilegiados. O antagonista da presente invenção pode ainda ser utilizado no tratamento da doença inflamatória do intestino.

Os antagonistas de DM4 podem ser úteis para o tratamento de doenças inflamatórias, tais como, artrite reumatóide, osteoartrite, psoríase, septicémia e doença inflamatória de intestino.

Modos de Administração

O agonista ou os antagonistas descritos aqui podem ser administrados *in vitro*, *ex vivo* ou *in vivo* as células que expressam o receptor da presente invenção. Por administração de uma "quantidade efectiva" de um agonista ou de um antagonista da presente invenção, entende-se uma quantidade do composto que é suficiente para aumentar ou inibir uma resposta celular a um ligando da família dos FNT e incluindo os polipéptidos. Em particular, por administração de uma "quantidade efectiva" de um agonista ou de antagonistas entende-se uma quantidade efectiva para aumentar ou inibir a apoptose mediada por DM4. Obviamente, quando é desejável que a apoptose seja aumentada, pode-se co-administrar um agonista de acordo com a presente invenção com um ligando da família do FNT. Um especialista na matéria entenderá que as quantidades efectivas de um agonista ou de um antagonista podem ser determinadas empiricamente e podem ser utilizadas na forma pura ou sob a forma de um sal, um éster ou um pró-fármaco aceitáveis sob o ponto de vista farmacêutico. O agonista ou o antagonista podem ser administrados em

composições, em combinação com um ou mais excipientes aceitáveis sob o ponto de vista farmacêutico.

Deve entender-se que, quando administrados a um paciente humano, a utilização diária total dos compostos e das composições da presente invenção, será seguida pelo médico assistente dentro do âmbito da decisão médica. O nível da dose efectiva específica, sob o ponto de vista terapêutico, para qualquer paciente em particular, dependerá de factores bem conhecidos na medicina.

Gomo uma proposta genérica, a quantidade efectiva total do ponto de vista farmacêutico, do polipéptido de DM4, administrado parentericamente, por dose, estará no intervalo de cerca de 1 µg/kg/dia até cerca de 10 mg/kg/dia do peso do corpo do paciente, embora, tal como se fez nctar antes, isto esteja sujeito a uma decisão terapêutica. Mais preferencialmente, esta dose é de peio menos 0,01 mg/kg/dia e, mais preferencialmente, para seres humanos, deve estar entre cerca de 0,01 e 1 mg/kg/dia para a hormona. Se forem dados continuamente, os agonistas ou os ant-agonistas de DM4 são normalmente administrados numa dose de cerca de 1 µg/kg/hora até cerca de 50 µg/kg/hora, quer por meio de 1-4 injeções por dia ou por infusões subcutâneas contínuas, por exemplo, utilizando uma mini-bomba. Também se pode utilizar a solução do saco intravenoso.

A dosagem pode também ser arranjada de uma forma específica para o paciente para providenciar uma concentração pré-determinada de um agonista ou de um antagonista no sangue, ~~conforme~~ determinado pela técnica de RIA. Assim, a dosagem para um paciente pode sei ajustada de modo a regular a entrada nos níveis de

sangue, conforme medido por RIA, por ordem de 50 até 1.000 ng/mL, preferencialmente, 150 até 500 ng/mL.

As composições farmacêuticas são dadas incluindo um agonista ou um antagonista e um veículo ou ~~um~~ excipiente, aceitáveis sob o ponto de vista farmacêutico, que podem ser administrados oralmente, rectalmente, parentericamente, intracisternalmente, intravaginalmente, intraperitonealmente, topicamente (sob a forma de pós, pomadas, gotas ou adesivos transdérmicos), bucalmente ou sob a forma de um aerossol oral ou nasal. É importante notar que, por co-administração de um agonista e de um ligando da família dos FNT os efeitos clínicos colaterais podem ser reduzidos, utilizando doses mais baixas tanto do ligando como do agonista. Deve entender-se que o agonista pode ser "co-administrado" quer antes, quer depois ou simultaneamente com o ligando da família do FNT, consoante as exigências de uma aplicação terapêutica em particular. Por "veículo aceitável sob o ponto de vista farmacêutico", entende-se uma carga não tóxica, sólida, semi-sólida ou líquida, um diluente, um material de encapsulação ou um auxiliar de formulação de qualquer tipo. O termo "parentérico", tal como se utiliza aqui, refere-se aos modos de administração que incluem a administração intravenosa, intramuscular, intraperitoneal, intra-esternal, subcutânea e intra-articular * por infusão.

As composições farmacêuticas da presente invenção para injeção parentérica podem compreender soluções, dispersões, suspensões ou emulsões aquosas esterilizadas ou não aquosas, aceitáveis sob o ponto de vista farmacêutico, assim como, pós esterilizados para reconstituir sob a forma de soluções ou dispersões

esterilizadas, injectáveis, para reconstituição imediatamente antes da sua utilização.

Para além dos polipéptidos solúveis de DM4, o polipéptido de DM4 contendo a região da transmembrana pode também ser utilizado, quando apropriado, sob a forma solubilizada, incluindo detergentes, tais como, CHAPS ou NP-40, com um tampão.

Exemplo 1

Expressão e Purificação em *E. coli*

A sequência de ADN que codifica a proteína de DM4 madura no clone de ADNc depositado (ATCC NO: 97853) é amplificada utilizando iniciadores de oligonucleótidos na RCP, específicos para as sequências de terminal amino da proteína de DM4 e específicas das sequências 3' do vector com o gene. Adicionam-se mais nucleótidos contendo sítios de restrição para facilitar as clonagens com as sequências 5' e 3', respectivamente.

Os iniciadores que se seguem são utilizados para a expressão do domínio extracelular de DM4 em *E. coli*: o iniciador 5' tem a sequência 5'-GCGGCAT-GCATGATCAATCAATfGGCAC-3' (SEQ ID NO: 8) contém o sítio de SphI sublinhado. O iniciador de 3' 5'-GCGAGCCTTTCAT-TATGTCCATTGCCTG-3' (SEQ ID NO: 9) e contém o sítio de HindIII sublinhado. O vector é pQE60.

Os sítios de restrição são convenientes para a restrição dos sítios da enzima no vector de expressão bacteriana pQE60, e são utilizados para a expressão bacteriana nestes exemplos (Qiagen, Inc. 9259 Eton Avenue, Chatsworth, CA, 91311). pQE60 codifica a

resistência ao antibiótico ampicilina ("Amp") e contém uma origem bacteriana de replicação ("ori"), um promotor indutível de IPTG e um sítio de ligação do ribossoma ("SLR").

O ADN de DM4 amplificado e o vector pQE60, são ambos digeridos com SphI e HindIII e os ADNs digeridos são então ligados em conjunto. A inserção do ADN da proteína de DDCR nos locais do vector pQE60, coloca a região de codificação da proteína de DM4 a jusante e ligada operacionamente ao promotor indutível do IPTG do vector e na mesma zona com uma AUG apropriadamente posicionada para a tradução da proteína de DM4.

A mistura de ligação é transformada nas células de *E. coli* concorrentes utilizando processos padrão. Esses processos estão descritos em Sambrook et al., *Molecular Cloning: a Laboratory Manual*, 2ª Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). A estirpe m15/rep4 de *E. coli*, contendo múltiplas cópias do plasmido pREP4, que expressa o repressor lac e confere resistência a canamicina ("rCan"), é utilizada na realização do exemplo ilustrativo aqui descrito. Esta estirpe, que é apenas uma das muitas que são apropriadas para a expressão da proteína de DM4, está disponível comercialmente na Qiagen, supra.

Os transformantes são identificados pela sua capacidade de crescer em placas de LB na presença de ampicilina e canamicina. Isola-se o ADN do plasmido a partir de colónias resistentes e a identidade do ADN clonado é confirmada por análise de restrição.

Os clones que contêm as estruturas desejadas crescem durante a noite ("D/N") em cultura líquida em meio LB

complementado tanto com ampicilina (100 µg/mL) como com canamicina (25 µg/mL).

A cultura D/N é utilizada para inocular uma grande cultura, numa diluição de aproximadamente 1:100 até 1:250. As células crescem até a uma densidade óptica a 600 nm ("OD₆₀₀") entre 0,4 e 0,6. Adiciona-se então o isopropil-B-D-tiogalactopiranosídeo ("IPTG") até a uma concentração final de 1 mM para induzir a transcrição dos promotores sensíveis ao repressor *lac*, por inativação do repressor *lacI*. Em seguida faz-se a incubação das células durante mais 3 a 4 horas. Colhem-se então as células por centrifugação e abrem-se por processos normalizados. Os corpos de inclusão são purificados a partir das células abertas utilizando técnicas de recolha de rotina e solubiliza-se a proteína a partir dos corpos de inclusão em ureia 8 M. A solução de ureia 8 M contendo a proteína solubilizada, é passada numa coluna PD-10 em solução salina tampada com fosfato ("STF"), duas vezes, eliminando-se assim a ureia, permutando o tampão e redobrando a proteína. Purifica-se a proteína por mais uma etapa de cromatografia para eliminar a endotoxina. Depois, filtra-se em meio esterilizado. A preparação da proteína filtrada em meio esterilizado é armazenada no seio de 2 X STF a uma concentração de 95 µ/mL.

Exemplo 2: Expressão em Células de Mamíferos

A maior parte dos vectores utilizados para a expressão transiente de uma dada sequência de genes em células de mamíferos comporta a origem de replicação SV40. Isto permite a replicação do vector até a um número elevado de cópias em células (por exemplo células de COS) que expressam o antigénio T requerido para a iniciação da

síntese do ADN viral. Pode-se utilizar para este fim qualquer outra linha de células de mamíferos.

Um vector de expressão típico, em mamíferos, contém o elemento promotor, que medeia a iniciação da transcrição de ARNm, a sequência de codificação da proteína e os sinais necessários para a terminação da transcrição e a poliadenilação do transcripto. Os elementos adicionais incluem melhoradores, sequências de Kozak e sequências de intervenção flanqueadas pelos sítios dadores e receptores para a segmentação de ARNm. Pode-se conseguir uma transcrição altamente eficiente com os promotores precoces e tardios de SV40, as repetições de terminais longos (RTLs), de retrovírus, por exemplo, VSR, VI de HTL, HIVI e o promotor precoce do citomegalovírus (CMV). Contudo, podem utilizar-se sinais celulares (por exemplo, o promotor de actina humana). Os vectores de expressão apropriados para serem utilizados na prática da presente invenção, incluem, por exemplo, vectores, tais como, pSVL e pMSG (Pharmacia, Uppsala, Suécia), pRSVcat (ATCC NO: 37152), pSV2dhfr (ATCC NO: 37146) e pBC12MI (ATCC NO: 67109). As células dos hospedeiros em mamíferos que podem ser utilizadas incluem Hela 293 humana, células H9 e de Jurkat, células NIH3T3 e C127 de rato, Cos 1, Cos 7 e células CV1 de macaco verde africano, células QC1-3 de coelho, células L de rato e células de ovário de hamster chinês (OHC).

Alternativamente, o gene de Interesse pode ser expresso em linhas de células estáveis, que contêm o gene integrado num cromossoma. A co-transfecção com um marcador seleccionável, tal como, dhfr, gpt, ~~neomicina~~, higromicina, permite a identificação e o isolamento das células transfectadas.

O gene transfectado pode também ser amplificado para expressar grandes quantidades da proteína codificada. A di-hidrofolato redutase (DHFR) é um marcador útil para desenvolver as linhas de células que comportam várias centenas ou mesmo vários milhares de cópias do gene de interesse. Utilizando este marcador, as células de mamífero crescem em quantidades crescentes de metotrexato para selecção e seleccionam-se as células com a resistência mais elevada. Outro marcador de selecção útil é a enzima glutamina sintase (GS) (Murphy *et al.*, *Biochem. J.* 227: 277-279 (1991); Bebbington *et al.*, *Bio/Technology* 10: 169-175 (1992)). Utilizando estes marcadores, as células de mamífero crescem meio selectivo e seleccionam-se as células com a resistência mais elevada. Estas linhas de células contêm os genes amplificados integrados num cromossoma. As células de ovário de hamster chinês (OHC) são muitas vezes utilizadas para a produção de proteínas.

Os vectores de expressão pC1 e pC4 contêm o promotor forte (RTL) do vírus do sarcoma de Rous (Cullen *et al.*, *Molecular and Cellular Biology* 5: 438-447 (Março de 1985)), mais um fragmenta do melhorador de CMV (Boshart *et al.*, *Cell* 42: 521-530 (1985)). Múltiplos sítios de clonagem, por exemplo, como os sítios de clivagem das enzimas de restrição BamHI, XbaI e Asp718, facilitam a clonagem do gene de interesse. Os vectores contêm, além disso, o intrão 3', a poliadenilação e o sinal de terminação do gene de pré-pró-insulina do rato.

Clonagem e Expressão em células de OHC

O vector pC4 é utilizado para a expressão do polipéptido de DM4. O plasmido pC4 é um derivado do plasmido pSV2-dhfr (número de acesso na ATCC 37146). O

plasmido contém o gene DHFR de rato sob controlo do promotor precoce de SV40. As células de ovário de hamster chinês ou outras células que não têm actividade de di-hidrofolato, que são transfectadas com estes plásmidos, podem ser seleccionadas fazendo crescer as células num meio selectivo (MEM alfa menos, Life Technologies), complementado com o agente quimioterapêutico metotrexato (MTX). A amplificação dos genes de DHFR em células resistentes a metotrexato (MTX) tem sido bem documentada (ver, por exemplo, Alt, F. W., Kellems, R. M., Bertino, J. R., e Schímke, R. T., *J. Biol. Chem.* 253: 1357-1370 (1978); Hamlin, J. L. e Ma, C., *Biochem. et Biophys. Acta* 1097: 107-143 (1990); Page, M. J. e Sydenham, M. A. 1991, *Biotechnology* 9: 64-68 (1991)). As células que crescem em concentrações crescentes de MTX desenvolvem resistência ao fármaco pela sobre-produção da enzima alvo, DHFR, como resultado da amplificação do gene de DHFR. Se um segundo gene está ligado ao gene de DHFR, ele é normalmente co-amplificado e sobre-expresso. Sabe-se na técnica que esta abordagem pode ser utilizada para desenvolver linhas de células que comportam mais do que 1.000 cópias dos genes amplificados. Em seguida, quando o metotrexato desaparece, obtêm-se linhas de células que contêm o gene amplificado integrado em um ou mais dos cromossomas da célula hospedeira.

O plasmido pC4 contém, para a expressão do gene de interesse, o promotor forte de repetição terminal longa (RTL) do vírus do sarcoma de Rous (Cullen et al., *Molecular and Cellular Biology* 5: 438-447 (Março de 1985)), mais um fragmento isolado do melhorador do gene precoce imediato do citomegalovírus (CMV) humano (Boshart et al., *Cell* 41: 521-530 (1985)). A jusante do promotor estão os seguintes únicos sítios de clivagem da enzima de restrição que permitem a integração dos genes: BamHI,

XbaI e Asp718. Para além destes sítios de clonagem, o plasmido contém o intrão 3' e o sítio de poliadenilação do gene da pré-pró-insulina de rato. Podem utilizar-se, para a expressão, outros promotores altamente eficientes, por exemplo, o promotor de β -actina humana, os promotores precoces e tardios de SV40 ou as repetições de terminal longo de outros retrovírus, por exemplo, VII e VI de HTLV. Os sistemas de expressão de genes Tet-Off e Tet-On da Clontech e sistemas similares podem ser utilizados para expressar o polipéptido de DM4, de uma forma regulada, em células de mamíferos (Gossen, M., & Bujard, H., *Proc. Natl. Acad. Sci. USA* 89: 5547-5551 (1992)). Para a poliadenilação de ARNm, podem utilizar-se também outros sinais, por exemplo, dos genes da hormona do crescimento humano ou dos genes da globina. As linhas de células estáveis que comportam um gene de interesse integrado nos cromossomas, podem também ser seleccionadas após co-transfecção com um marcador seleccionável, tal como, gpt, G418 ou higromicina. É vantajoso utilizar mais do que um marcador seleccionável no princípio, por exemplo, G418 mais metotrexato.

O plasmido pC4 é digerido com a enzima de restrição BamHI e depois é desfosforilado utilizando fosfatos de intestino de bezerro por processos conhecidos na técnica. O vector é então isolado a partir de um gel de agarose a 1 %.

A sequência de ADN que codifica o polipéptido completo é amplificada utilizando iniciadores de oligonucleótidos para RCP, correspondendo às sequências 5' e 3' da porção desejada do gene. O iniciador 5' contendo o sítio de BamHI sublinhado, uma sequência de Kozak e um codão inicial de AUG, tem a seguinte sequência:

5' GCGGGATCCGCCATCATGGCGCCACCACCAGCTAGA 3' (SEQ ID NO: 10). O iniciador 5', contendo o sítio BamHI site, tem a seguinte sequência: 5' GCGGGATCCTCACTCCAAGGACACOGCAGAGCC 3' (SEQ ID NO: 11).

O fragmento amplificado é digerido com a endonuclease BamHI e depois purificado num gel de agarose a 1 %. O fragmento isolado e o vector desfosforilado, são então ligados com ligase de ADN de T4. As células HB 101 ou XL-1 Blue de *E. coli* são então transformadas e identificam-se as bactérias que contêm o fragmento inserido no plasmido pC4 utilizando, por exemplo, as análises de restrição de enzimas.

As células de ovário de hamster chinês a que falta um gene de DHFR activo, são utilizadas para a transfecção. Faz-se a co-transfecção de 5 µg do plasmido de expressão pC4 com 0,5 µg do plasmido pSV-neo utilizando o processo da lipofectina (Felgner et al., supra). O plasmido pSV2-neo contém um marcador seleccionável dominante, o gene neo de Tn5, que codifica uma enzima que confere resistência a um grupo de antibióticos incluindo G418. As células são então semeadas em meio MEM alfa menos, complementado com 3 mg/mL de G418. Passados 2 dias, as células são tripsinizadas e semeadas em placas de clonagem de hibridoma (Greiner, Alemanha) em meio MEM alfa menos complementado com 10, 25 ou 50 ng/mL de metotrexato mais 1 mg/mL de G418. Passados cerca de 10-14 dias, os clones isolados são tripsinizados e depois semeados em placas de Petri de 6 microtubos ou frascos de 10 mL, utilizando diferentes concentrações de metotrexato (50 nM, 100 , 200 nM, 400 nM, 800 nM). Os clones que crescem nas concentrações mais elevadas de metotrexato, são então transferidos para placas de 6 microtubos contendo concentrações ainda mais elevadas de metotrexato (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). Repete-se o mesmo

processo até se obterem clones que crescem a uma concentração de 100-200 µM. Analisa-se a expressão do produto de gene desejado, por exemplo, por EGPA-SDS, análise de transferência do anticorpo/proteína para uma matriz imobilizada (Western) e por CLER de fase inversa.

Exemplo 3

Clonagem e Expressão do domínio Extracelular Solúvel de DM4 num Sistema de Expressão de Baculovírus

A sequência de ADNc que codifica o domínio extracelular solúvel da proteína de DM4 no clone depositado (nº de ATCC 97853) é amplificada utilizando iniciadores de oligonucleótidos de RCP que correspondem às sequências 5' e 3' do gene.

O iniciador 5' para DM4 tem a sequência 5' GCGGGATCCGCCATCATGGCGCCACGACCAGCTAGA 3' (SEQ ID NO: 10), contendo o sítio de restrição de enzima BamHI sublinhado. Inserido num vector de expressão, tal como se descreve a seguir, a extremidade 5' do fragmento amplificado que codifica DM4, origina um péptido de sinal de clivagem eficiente. Um sinal eficiente para a iniciação da tradução nas células eucarióticas, conforme descrito por Kozak, M., J. o *Biol.* 196: 947-950 (1987), está apropriadamente localizado na porção de vector da estrutura.

O iniciador 3' para DM4 tem a sequência 5'- CCGGAATCCCAATATATGTCCTTCCTG -3' (SEQ ID NO: 12), contendo a restrição de BamHI sublinhada seguida de nucleótidos complementares à sequência de nucleótidos de DM4 na figura i, seguido de um codão de paragem.

O fragmento amplificado é isolado a partir de um gel de agarose a 1 %, utilizando um kit disponível comercialmente ("GeneClean", BIO 101 Inc., La Jolla, Ca.). O fragmento é então digerido com BamHI e Asp718 e é novamente purificado em gel de agarose a 1 %.

O vector pA2 é utilizado para expressar a proteína de DM4 no sistema de expressão de baculovírus, utilizando processos normalizados, tais como os descritos em Summers *et al.*, *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*. Texas Agriculture Experimental Station Bulletin No. 1555 (1987). Este vector de expressão contém o promotor forte de poli-hedron do vírus de poli-hedrose nuclear californica Autograph (ACMNPV) seguido dos sítios de restrição convenientes. Para uma selecção fácil do vírus recombinante, insere-se o gene de beta-galactosidase de *E. coli* na mesma orientação que o o promotor de poli-hedron e & seguido pelo sinal de poliadenilação do gene de poli-hedron. As sequências de poli-hedron são flanqueadas de ambas os lados por sequências virais para a recombinação de homólogos mediada por células com o ADN viral de tipo selvagem para gerar vírus viáveis que expressam o polinucleótido clonado.

Podem utilizar-se muitos outros vectores de baculovirus em lugar de pA2, tais como pAc373, pVL941 e pAcIM1 desde que, como será evidente para os especialistas na matéria, essa construção providencie sinais para transcrição, tradução, tráfego e similares, localizados apropriadamente, tais como numa secção da AOC e num péptido de sinal, conforme necessário. Esses vectores estão descritos em Luckow *et al.*, *Virology* 170: 31-39, entre outros.

O plasmido é digerido com as enzimas de restrição Bam HI e depois é desfosforilado utilizando fosfatase intestinal de bovino, utilizando os processos de rotina conhecidos na técnica. O ADN é então isolado do gel de agarose a 1 % utilizando um kit disponível comercialmente ("GeneClean", BIO 101 Inc., La Jolla, Ca.).

O fragmento e o plasmido desfosforilado, são ligados em conjunto com ligase de ADN de T4. As células HB101 de *E. coli* são transformadas com a mistura de ligação e espalhadas nas placas de cultura. Identificam-se as bactérias que contêm o plasmido com o gene de DDCR humano por digestão do ADN de colónias individuais utilizando BamHI e depois analisando o produto da digestão por meio de electroforese em gel. Esta sequência dos fragmentos clonados é confirmada pela sequenciação do ADN. Este plasmido é aqui designado por pBac DM4.

Faz-se a co-transfecção de 5 µg do plasmido pBac DM4 com 1,0 µg de ADN de um baculovírus linearizado, disponível comercialmente, ("ADN de baculovírus BaculoGold™", Pharmingen, San Diego, CA.), utilizando o processo da lipofectina descrito por Felgner *et al.*, *Proc. Natl. Acad. Sci. EUA* 84: 7413-7417 (1987). Mistura-se 1 µg do ADN do vírus BaculoGold™ e 5 µg do plasmido pBac DM4 num microtubo esterilizado de uma placa de microlitros contendo 50 µL de meio de Grace isento de soro (Life Technologies Inc., Gaithersburg, MD). Depois adiciona-se 10 µL de Lipofectina plus, 90 microlitros de meio de Grace, mistura-se e incuba-se durante 15 minutos à temperatura ambiente. Depois a mistura de transfecção é adicionada, gota a gota, às células de insectos Sf9 (ATCC CRL 1711), semeadas numa placa de cultura de tecidos de 35 mm com 1 mL de meio de Grace sem soro. A placa é agitada e mistura-se novamente com a solução recentemente

adicionada. A placa é então incubada durante 5 horas, a 27 °C. Passadas 5 horas, a solução de transfecção é retirada da placa e adiciona-se 1 mL de meio de insectos de Grace, complementado com soro bovino fetal a 10 %. Volta a por-se a placa num incubador e a cultura continua a 27 °C, durante 4 dias.

Passados 4 dias, recolhe-se o sobrenadante e realiza-se um ensaio da placa, conforme descrita por Summers e Smith, citados antes. Utiliza-se um gel de agarose com "Blue Gal" (Life Technologies Inc., Gaithersburg, MD) para permitir uma fácil identificação e o isolamento dos clones que expressam gal, que produzem placas coradas de azul. (Pode-se encontrar uma descrição detalhada de um "ensaio de placa" deste tipo, no guia do utilizador para as culturas de células de insectos e baculovirologia distribuído por Life Technologies Inc., Gaithersburg, MD, paginas 9-10).

Quatro dias depois da diluição em série, adiciona-se o vírus às células. Depois de uma incubação apropriada, as placas coradas de azul são picadas com a ponta de uma micropipeta de Eppendorf. O agar contendo os vírus recombinantes é então novamente suspenso num tubo de Eppendorf contendo 200 µL de meio de Grace. Retira-se o agar por meio de uma centrifugação breve e utiliza-se o sobrenadante contendo o baculovírus recombinante para infectar as células Sf9 semeadas em placas de 35 mm. Quatro dias mais tarde, os sobrenadantes dos pratos destas culturas são colhidos e depois são armazenados a 4 °C. Um clone contendo o DM4 inserido apropriadamente é identificado por análise do ADN incluindo o mapeamento de restrição e a sequenciação. É designado aqui por V-DM4.

Faz-se crescer células Sf9 em meio de Grace complementado com SBF a 10 % inactivado pelo calor. As células são infectadas com o baculovírus recombinante V-DM4 com uma multiplicidade de infecção ("MOI") de cerca de 2 (cerca de 1 até cerca de 3). Seis horas mais tarde, o meio é eliminado e é substituído por meio SF900 II menos metionina e cisteína (disponível na Life Technologies Inc., Gaithersburg, MD). 42 horas mais tarde, adiciona-se 5 µCi de ³⁵S-metionina e 5 µCi de ³⁵S-cisteína (disponíveis na Amersham). As células são depois Incubadas durante 16 horas e depois são colhidas por centrifugaçã faz-se a respectiva lise e visualizam-se as proteínas marcaçã por EGPA-SDS seguida de autoradiografia.

Exemplo 4

Distribuição ao Tecido da Expressão do Gene de DM4

Realizou-se a análise de transferência de ADN/ARNm para uma matriz imobilizada (Northern) para examinar a expressão do gene DM4 (ATCC No. 97853) em tecidos humanos, utilizando os processos descritos, entre outros, por Sambrook et al., citado antes. Marcou-se uma sonda de ADNc contendo toda a sequência de nucleótidos da proteína de DM4 (SEQ ID NO: 3) com ³²P, utilizando o sistema de marcação de ADN com rediprimeTM (Amersham Life Science), de acordo com as instruções do fabricante. Depois da marcação, purificou-se a sonda utilizando uma coluna CHROMA SPIN-100TM (Clontech Laboratories, Inc.), de acordo com o número PT1200-1 do protocolo do fabricante. A sonda marcada, purificada, foi então utilizada para examinar vários tecidos humanos quanto ao ARNm de DM4.

As manchas de Northern em múltiplos tecidos (MNT), contendo vários tecidos humanos (H) ou tecidos do sistema imunitário humano (IM), foram obtidas da Clontech e foram examinadas com sondas marcadas, utilizando a solução de hibridação ExpressHyb™ (Clontech), de acordo com o número PT1190-1 do protocolo do fabricante. No seguimento da hibridação e da lavagem, montaram-se as manchas e expôs-se o filme a -70 °C, durante a noite e revelaram-se os filmes de acordo com processos padrão. Detectou-se a expressão de DM4 em tecidos enriquecidos com linfócitos incluindo células amnióticas, no coração, cancro do fígado, rim, leucócitos do sangue periférico, células T activadas, células K562 plus PMA, W 138 cenç, Th2, amígdalas humanas e cordão umbilical com falta de CD34. Pode-se verificar que DM4 desempenha um papel na homeostase de linfócitos.

Exemplo 5

Apóptose Induzida por DM4

A sobre-expressão de Fas/APO-1 e R de FNT-1 em células de mamíferos simula a activação do receptor (M. Muzio *et al.*, *Cell* 85: 817-827 (1996); M. P. Boldin *et al.*, *Cell* 85: 803-815 (1996)). Assim, utilizou-se este sistema para estudar o papel funcional de DM4. A expressão transiente de DM4 em células humanas MCF7 do carcinoma da mama e células 293 do rim embriónico humano induziram uma rápida apoptose.

Os ensaios de morte das células foram realizados praticamente como foi descrito antes (A. M. Chinnaiyan, *et al.*, *Cell* 81: 505-12 (1995); M. . Boldin, *et al.*, *J Biol Chem* 270: 7795-8 (1995); E. C. Kischkel, *et al.*, *EMBO* 14: 5579-5588 (1995); A. M. Chinnaiyan, *et al.*, *J*

Biol Chem 271: 4961-4965 (1996)). Em resumo, fez-se a transfecção estável de linhas de células clonais do carcinoma da mama humana MCF-7 quer apenas com vector ou com uma estrutura de expressão de CrmA (M. Tewari, *et al.*, *J. Biol Chem* 270, 3255-60 (1995)), que foram transfectadas de forma transiente com pCMV-DM4-galactosidase (ou pCMV-DM4-galactosidase (a que falta o domínio de morte)) na presença de um excesso de 10 vezes das estruturas de expressão de pcADN3 que codificam as proteínas indicadas, utilizando a lipofectamina (GIBCO-BRL). Transfectaram-se do mesmo modo células 293 utilizando o processo do ~~CaPO₄~~. Adiciona-se as células o inibidor da família dos ICE z-VAD-fmk (Enzyme Systems Products, Dublin, CA) e fixa-se e cora-se com X-Gal como foi previamente descrito (A. M. Chinnaiyan, *et al.*, *Cell* 81, 505-12 (1995); M. P. Boldin, *et al.*, *J Biol Chem* 270, 7795-8 (1995); F. C. Kischkel, *et al.*, *EMBO* 14, 5579-5588 (1995)).

As células exibiram alterações morfológicas típicas das células que sofrem apoptose, tornando-se arredondadas, condensadas e destacando-se do prato. Do mesmo modo que com o R de FNT-1 e com Faz/APO-1 (M. Muzio, *et al.*, *Cell* 85, 817-827 (1996); M. P. Boldin, *et al.*, *Cell* 85, 803-815 (1996); M. Tewari, *et al.*, *J Biol Chem* 270, 3255-60 (1995)), a apoptose induzida por DM4 foi bloqueada pelos inibidores de proteases semelhantes a ICE, CrmA e z-VAD-fmk.

Será claro que a presente invenção pode ser praticada de ~~forma~~ diferentes da que esta descrita em forma particular na descrição e nos exemplos anteriores.

LISTAGEM DE SEQUÊNCIAS

(1) INFORMAÇÃO GERAL:

(ii) REQUERENTE: Ni, Jian
Rosen, Craig A.
Pan, James G.
Gentz, Reiner L.
Dixit, Vishva M.

(iii) TÍTULO DA INVENÇÃO: Receptor 4 Contendo o
Domínio de Morte

(iii) NÚMERO DE SEQUÊNCIAS: 12

(iv) ENDEREÇO PARA CORRESPONDÊNCIA:

(A) ENDEREÇO: Human Genome Sciences, Inc.
(B) RUA: 9410 Key West Avenue
(C) CIDADE: Rockville
(D) ESTADO: MD
(E) PAÍS: EUA
(F) CÓDIGO POSTAL: 20850

(v) FORMA LIGÍVEL EM COMPUTADOR:

(A) TIPO DE MEIO: Floppy disk
(B) COMPUTADOR: Compatível com IBM PC
(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.C,
Versão #1.30

(vi) DADOS DO PRESENTE PEDIDO DE PATENTE DE
INVENÇÃO:

(A) NÚMERO DO PEDIDO DE PATENTE DE
INVENÇÃO: US

(B) DATA DE REGISTO: 28 de Janeiro de
1997

(C) CLASSIFICAÇÃO:

(viii) ADVOGADO/AGENTE DE INFORMAÇÃO:

(A) NOME: BROOKES, ANDERS A.

(B) NÚMERO DE REGISTO: 36.373

(C) REFERÊNCIA/NÚMERO DE ETIQUETA: PF355

(ix) INFORMAÇÃO SOBRE TELECOMUNICAÇÕES:

(A) TELEFONE: (301) 309-8504

(B) TELEFAX: (301) 309-8512

(2) INFORMAÇÃO PARA A SEQ ID N°. 1:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 2152 pares de bases

(B) TIPO: Ácido Nucleico

(C) ESTRUTURA HELICOIDAL: simples

(D) TOPOLOGIA: Linear

(ii) TIPO DE MOLÉCULA: ADN (genômico)

(ix) CARACTERÍSTICA:

(A) NOME/CHAVE: CDS

(E) LOCALIZAÇÃO: 19..1422

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°. 1:

Met Ala Pro Pro Pro Ala Arg Val His Leu Gln
1 18 19

GCG TTC CTG GCA GTG ACT CCG AAT CCC GGG AGC GCA GCG AGI GGG KCA 99
Ala Fen Leu Ala Val Tre Pro Asn Pro Gln Ser Ala Ale Ser Gln Tre
15 20 25

AAA GCA GGC CCC GGT ACA GCT AGC AAA GTC TGG GGC TCT TAT GAG GGG 107
Glu Ala Ala Ala Ala Ire Pro Ser Lis Val Trp Gln Ser Ser Ala Gln
30 35 40

AGG ATT GAA CCA CGA GSC GGG GGC CGA GGA GCG CTC CCT ACC TCC ATG 105
Arg Ile Glu Pro Arg Gln Gln Gln Arg Gln Ala Leu Pro Tre Ser Met
45 50 55

GGA CAG CAC GGA GCG AGT GCC CGG GCC CGG GCA GAG GAT GCT GCA GAA 113
Gln Gln His Gln Pro Ser Ala Arg Ala Arg Ala Gln Arg Ala Pro Gln
E3 65 70 75

CCC AGG SCG GCG CGG GAA GCC GGC CCT CGG CTC CGG GTC CAC AAG ACC 111
Pro Arg Pro Ala Arg Glu Ala Ser Pro Arg Leu Arg Val His Lis Tre
80 85 90

TTC AAG TST GTC GTC GTC GGG GTC CTG CTG CAG GTC GTA CCT AGC TCA 119
Fen Lis Fen Val Val Val Gln Val Leu Leu Gln Val Val Pro Ber Ser
95 100 135

GCT GCA ACC ATC AAA CTT CAT GAT CAA TCA ATT GGC ACA CAG CAA TGG 107
Ala Ala Tre Ile Lis Leu His Asp Gln Ser Ile Gln Tre Gln Gln Trp
110 115 120

GAA CAT AGC CGT TTG GGA GAG TIG TGT CCA CCA GGA TCT CAT AGA TCA 105
Glu His Ser Ero Leu Gln Gln Leu Cys Pro Pro Gln Ser His Arg Ser
125 130 135

GAA CGT CCT GGA GCC TGT AAC CGG TGC ACA GAG GGT GTG GGT TAC ACC 103
Glu Arg Pro Gln Ala Cys Asn Arg Cys Tre Glu Gln Val Gln Tir Tre
140 143 150 155

AAA GCT TCA AAC AAT TCG TTT GCT TTT GAA CCA TGT ACA GCS TGT AAA 111
Asn Ala Ser Asn Asn Leu Fen Ala Cys Leu Pro Cys Tre Ala Cys Lis
160 165 170

TCA GAT GAA GAA GAG AGA AGT CCC TGC ACC ACG ACC AGG AAC ACA GCA 119
Ser Arg Glu Glu Glu Arg Ser Pro Val Tre Tre Tre Arg Asn Tre Ala
175 180 185

TCT CAG TGC AAA CCA GGA ACT TTC CGG AAT GAC AAT TCT GCT GAG ATG
 Cis Gin Cis Lis Pro Gli ire Fen Arg Asn Asp Asn Ser Ala Gln Met
 190 195 200

TCC CCG AAC ACC ACA GGG TCC CCA AGA GAG ATG CTC AAG CTC AAG 878
 Cis Arg Lis Cis Phe Trp Gln Cys Phe Arg Gln Met Val Lis Val Lis
 205 210 215

GAT TGT ACC ACC TGG AGT GAC ACC GAG TGT GTC CAC AAA GAA TCA ACC 723
 Asp Cys Trp Phe Trp Ser Asp Ile Glu Cys Val His Lis Glu Ser Gln
 220 225 230 235

AAT GGA GAT AAT ATA TGG GTG AAT TGG GAT GTC ACC TGG GTT GTT CCG
 Asn Gln His Asn Ile Trp Val Ile Leu Val Val Tre Leu Val Val Pro
 240 245 250

TTC CTC TTC CTC GAT GGG GAT ACC GTC TGT TGT TGC ATC GGC TCA GGT 819
 Leu Leu Leu Val Ala Val Leu Ile Val Cys Cys Cys Ile Gln Ser Gln
 255 263 265

TGT GGA GGG GAC CCC AAG TGC ATG GAC AGG GTG TGT TTC TGG CGC TTG 867
 Cis Gln Gln Asp Pro Lis Cys Met Asp Arg Val Cys Fen Trp Arg Leu
 270 275 280

GAT CTC GTA CGA ACC CCT GGG GCT GAG GAC AAT GCT CAC AAC GAG AAT 818
 Gln Leu Leu Arg Gln Pro Gln Ala Glu Asp Asn Ala His Asn Glu Lie
 285 290 295

CTG ACC AAC GGA GAC TCG CTC TCC ACC ACC TTC CTC TCT GAG CAG GAA ATG 868
 Leu Ser Asn Ala Asp Ser Leu Ser Tre Fen Val Ser Glu Gln Gln Met
 300 305 310 315

GAA ACC CAG GAG TCG ACC GAT TTG ACA GAT GTC ACT GAA CAG TCC CCA 1011
 Gln Ser Cys Cys Phe Ala Asp Leu Tre Gln Val Tre Val Gln Ser Pro
 320 325 330

GAT CAG GCA GAG TGT CTC CTC GAA CCA GCA GAA GAT GAA GGG TCT CAG 1068
 Gln Cys Ala Cys Cys Leu Leu Gln Pro Ala Cys Ala Cys Gln Ser Cys
 335 340 345

AGC ACC AAG CTC CTC GAT GCA ACC AAT GAT GAT GAC ACC ACC ACC ACC 1107
 Arg Arg Arg Leu Leu Val Phe Ala Asn Gln Ala Asp Pro Trp Cys Trp
 350 355 360

CTG ATG CCG TTC TTT GAC AAG TTT GCA AAC ATC GTG CCC TTT GAC TCC 1188
 Leu Met Leu Fen Fen Asp Lis Fen Ala Asn Ile Val Pro Fen Asp Ser
 365 370 375

TGG GAG CAG CTC ATG AGG CAG CTG GAG CTC ACG AAA AAT GAG ATC GAT 1193
 Trp Arg Glu Leu Met Arg Gln Leu Asp Leu Tre Lis Met Glu Ile Asp
 383 385 390 395

GCG CTC AAA GGT GGT ACA GCA GGC CCA GGG GAT GCC TTG TAT GCA ATG 1201
 Val Val Arg Ala Glu Tre Ala Gli Pro Gli Asp Ala Leu Tir Ala Met

400 405 410

CYS ATG AAA TGG GTC AAC AAA ACT GGA CGG AAC GCC TCG ATC CAC ACC 1299
 Cys Met Lis Trp Val Asn Lis Tre Glu Arg Asn Ala Ser Ile His Tre
 415 420 425

CTG CTG GAT GCC TTG GAG AGG AIG CAA GAG AGA CAT GCA AAA GAG AAC 1347
 Leu Leu Asp Ala Leu Glu Arg Met Glu Glu Arg His Ala Lis Glu Lis
 430 435 440

ATT CAG GAC CTC TTG GTG GAC TCT GGA AAG TTC ATC TAC TTA GAA GAT 1335
 Ile Gln Asp Leu Leu Val Asp Ser Gli Lis Fen Ile Tir Leu Glu Asp
 445 450 455

GGC ASA GGC TCF GCC GTG TCC TTG GXG TTAGGCTC ITTTTACCAG 1348
 Gli Ire Gli Ser Ala Val Ser Leu Glu
 160 465

AGGTTTCCTC TTAGGTGTTA GGAGTTAATA CATATTAGGT TTTTTTTTTT TTTAACATGT 1500

ATACAAAGTA AATTTTAAAC CACGTGTATT GCCTCCTGCC TGTAATCCCA TCACCTTGGG 1502

AGGCTGACGC CGGTGGATCC ACTTGAGGTC CGAAGTTCCA AGACGAGCCC TGAACCAACA 1522

TCGTGAAAAT GCCCGTCTTT TACAAATAA TACCAAAAT TCACTGAAA TGTCTATGAT 1502

GTGTGCCATC APTTCTGAG CTAACACGG GAGGTCTGAG GCCAGGAGAA TCCACTTGAA 1742

CCCTACAAA GCACTGAAA ACTGCAGATT GCACTGCTG AATTCAGGAC TGGACACAA 1502

GGGACAGCT CTGTTCTGAG AAAAAAATA ATAACTTAA AAAAAATTT GCGGACGCA 1560

GGTAACTG CCAAGGAAA AACTGCT TCT GCGTAAAT GCGGCTGCT TGTCTCTT 1512

TGTCTCTGAT CATGCGGAG AATTAATTT AATGCTTAA GAAACAAATA GAAACATGAT 1502

AAGCAATAT TTTTGGGCA TTAATTTG GATGCTGTC GAGGCTTTT CACAGGACA 2042

TTCATATAT AAGGAGGCA CTGATGAAA AAAAAATAA ATTAATTTT GCGGCTGCT 2102

TCCGACAGAC TGGATATCA CTATAGAAA AAAAAAATA AAAAAAATA 2152

Das ist die erste Seite des Buches.

300 295 290

Das ist die zweite Seite des Buches.

305 300 295

Das ist die dritte Seite des Buches.

310 305 300
Das ist die vierte Seite des Buches.

300 295 290

Das ist die fünfte Seite des Buches.

315 310 305

Das ist die sechste Seite des Buches.

320 315 310

Das ist die siebte Seite des Buches.

255 250 245

Das ist die achte Seite des Buches.

330 325 320

Das ist die neunte Seite des Buches.

335 330 325

Das ist die zehnte Seite des Buches.

340 335 330

Das ist die elfte Seite des Buches.

345 340 335

Das ist die zwölfte Seite des Buches.

350 345 340

Das ist die dreizehnte Seite des Buches.

355 350 345

Das ist die vierzehnte Seite des Buches.

Arg Leu Ser Ser Lis Ser Val Asn Aia Gln Val Tre Asp Ile Asn Ser

20 25 30

Lis Gln Leu Gln Leu Arg Lis Tre Val Tre Tre Val Glu Tre Gln Asn

35 40 45

Leu Gln Gln Leu His His Asp Gln Gln Fen Cis His His Pro Cis Pro

50 55 60

Pro Gln Gln Arg Lis Ala Arg Asp Cis Tre Val Asn Gln Asp Glu Pro

65 70 75 80

Asp Gln Val Pro Cis Gln Gln Gln Lis Gln Val Tre Asp Lis Ala His

85 90 95

His Ser Ser Lis Cis Arg Arg Cis Arg Leu His Asp Gln Gln His Gln

100 105 110

Leu Glu Val Glu Ile Asn Cis Tre Arg Tre Gln Asn Tre Lis Cis Arg

115 120 125

Cis Lis Pro Asn Fen Fen Cis Asn Ser Tre Val Cis Glu His Cis Asp

130 135 140

Pro Cis Tre His Cis Gln His Gln Ile Ile Lis Glu Cis Tre Leu Tre

145 150 155 160

Ser Asn Tre Lis Cis Lis Glu Glu Gln Ser Arg Sei Asn Leu Gln Trp

165 170 175

Leu Gln Leu Leu Leu Leu Pro Ile Pro Leu Ile Val Trp Val His Arg

180 185 190

Lis Gln Val Gln Lis Tre Cis Arg Lis His Arg Lis Glu Asn Gln Gln

195 200 205

Ser His Gln Ser Pro Tre Leu Asn Pro Gln Tre Val Ala Ile Asn Leu

210 215 220

Ser Asp Val Asp Leu Ser Lis Val Ile Ile Tre Ile Aia Gln Val Met

225 230 235 240

Tre Leu Ser Gln Val Lis Glu Fen Val Arg Lis Asn Glu Val Asn Glu
 245 250 255

Ala Lis Ile Asp Glu Ile Lis Asn Asp Asn Val Gln Asp Tre Ala Glu
 260 265 270

Glu Lis Val Glu Leu Leu Arg Asn Trp His Gln Leu His Glu Lis Ile
 275 283 285

Glu Ala Tir Asp Tre Leu Ile Lis Asp Leu Lis Ile Ala Asn Leu Glu
 290 295 300

Tre Leu Ala Glu Lis Ile Gln Tre Ile Ile Leu Lis Asp Ile Tre Ser
 305 310 315 320

Asp Ser Glu Asn Ser Asn Fen Arg Asn Glu Ile Gln Ser Leu Val Met
 325 330 335

Leu Glu Ile Trp Tre Leu Teu Pro Leu Val Lei: Tre Ser Val Ala Arg
 340 345 353

Leu Ser Ser Lis Ser Val Asn Ala Gln Val Tre Asp Ile Asn Ser Lis
 355 360 365

Glu Leu Glu Leu Arg Lis Tre Val Tre Tre Val Glu Tre Gln Asn Leu
 370 375 380

Glu Glu Leu His Eis Asp Glu Gln Fen Cis His Lis Pro Cis Pro Pro
 385 390 395 400

Glu Glu Arg Lis Ala Arg Asp Cis Tre Val Asn Glu Asp Glu Pro Asp
 405 410 415

Glu Val Pro Cis Glu Glu Glu Lis Glu Tir Tre Asp Lis Ala His Fen
 420 425 430

Ser Ser Ile Cis Arg Arg Cis Arg Leu Cis Asp Glu Glu Ile Glu Leu
 435 440 445

Glu Val Glu Ile Asp Cis Tre Arg Tre Glu Asn Tre Lis Cis Arg Cis
 450 455 460

Lis Pro Asn Fen Fen Cys Asn Ser Tre Val Cys Glu His Cys Asp Pro
 465 470 475 480

Cys Tre Ile Cys Glu His Gln Ile Ile Ile Val Glu Cys Tre Leu Tre Ser
 485 490 495

Asn Tre Lis Cys His Glu Glu Gln Ser Arg Ser Asn Leu Gln Trp Leu
 500 505 510

Cys Leu Leu Leu Leu Pro Ile Pro Leu Ile Val Val Lis Arg Lis Glu
 515 520 525

Val Gln Lis Arg Cys Arg Lis His Arg Lis Glu Asn Gln Gln Ser His
 530 535 340

Glu Ser Pro Tre Leu Asn Pro Glu Tre Val Ala Ile Asn Leu Ser Asp
 545 550 555 560

Val Asp Leu Ser Lis Trp Ile Tre Tre Ile Ala Gln Val Met Tre Leu
 565 570 575

Ser Gln Val Lis Gln Fen Val Arg Val Asn Gln Val Asn Glu Ala Lis
 580 585 590

Ile Asp Glu Ile Lis Asn Asp Asn Val Gln Asp Tre Ala Glu Gln Lis
 595 600 605

Val Gln Leu Leu Arg Asn Trp His Gln Leu His Gln Lis Lis Glu Ala
 610 615 620

Trp Asp Tre Leu Ile Lis Asp Leu Lis Lis Ala Asn Leu Cys Arg Leu
 625 630 635 640

Ala Cys His Ile Gln Tre Ile Ile Leu His Asp Ile Tre Ser Asp Ser
 645 650 655

Glu Asn Ser Asn Ser Arg Asn Glu Ile Gln Ser Leu Val
 660 665

(2) INFORMAÇÃO PARA A SEQ ID Nº. 4:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 909 aminoácidos
- (B) TIPO: aminoácido
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°. 4:

Met Gln Leu Ser Trp Val Pro Asp Leu Leu Pro Leu Val Leu Leu Glu
5 10 15
Leu Leu Val Gln Ile Tyr Pro Ser Gln Val Ile Gln Leu Val Pro His
20 25 30
Leu Gln Asp Arg Glu Lis Arg Asp Ser Val Cys Pro Gln Gln Lis Tyr
35 40 45
Ile His Pro Gln Asn Asn Ser Ile Cys Cys Trp Lis Cys His Lis Gln
50 55 60
Ile Tyr Leu Tyr Asn Asp Cys Pro Gln Pro Gln Asp Trp Asp Cys Arg
65 70 75 80
Gln Cys Glu Ser Gln Ser Phe Trp Ala Ser Glu Asn His Leu Arg His
85 90 95
Gln Leu Ser Cys Ser His Cys Arg His Glu Met Gln Gln Val Gln Tyr
100 105 110
Ser Ser Cys Trp Val Asp Arg Asp Trp Val Cys Gln Cys Arg Lis Asn
115 120 125
Gln Tyr Arg His Tyr Trp Ser Glu Asn Leu Phe Gln Cys Phe Asn Cys
130 135 140
Ser Leu Cys Leu Asn Gln Trp Val His Leu Ser Cys Gln Glu Lis Gln
145 150 155 160

Asn Tre Val Cis Ire Cis His Ala Gln Fen Fen Leu ~~Asn~~ Glu Asn Gln
6 5 170 175

Cis Val Ser Cis Ser Asn Cis Lis Lis Ser ~~Leu~~ Gln Cis Tre Lis Leu
180 735 190

Cis Leu Pro ~~Gln~~ Ile Gln Asn Val Lis Gln Tre Gln Asp Ser Gln Tre
195 200 205

Tre Val Leu ~~Leu~~ Pro Leu Val Ile Fen Fen Gln Leu Cis Leu Leu Ser
210 215 220

Leu Leu Fen Ile Gln Leu ~~Met~~ Tir Arg Tir Gln Arg Trp Lis Ser Lis
225 230 235 240

Leu Tir Ser Ile Val Cis Gln Lis Ser ~~Tyr~~ Pro Glu Lis Gln Gln Gln
245 250 255

Leu Glu Gln Tre Tre Tre Lis Pro Leu Ala Pro Asn Pro Ser Fen Ser
260 265 270

Pro Tre Pro Gln Fen Tre Pro Tre Leu Gln Fen Ser Pro ~~Val~~ Pro Ser
275 280 285

Ser Tre Fen Tre Ser Ser Ser Tre Tir Tre Pro Gln Asp Cis Pro ~~Asn~~
290 295 300

Fen Ala Ala Ero ~~Arg~~ Arg Glu Val Ala Ero Pro Tir Gln Gln Ala Asp
305 310 315 320

Pro Ile Leu Ala Tre Ala Leu Ala Ser Asp Pro Ile Fro Asn Pro ~~Leu~~
325 330 335

Gln ~~Met~~ Tyr Glu Asp Ser Ala His Lis Pro Gln Ser ~~Leu~~ Asp Tre Asp
340 345 350

Arg Pro Ala Tre Leu Tir Ala Val Val Gln ~~Asn~~ Val Pro Pro Leu Arg
355 360 365

Tyr Lis Gln Fen Val Arg Arg Leu Gln Leu Ser Asp Gln Gln Ile Asp
370 375 380

Arg Leu Glu Leu Glu Asn Gli Arg Cys Leu Arg Glu Ala Gln Tir Ser
383 390 395 400

Met Leu Ala Tre Trp Arg Arg Arg Tre Pro Arg Arg Glu Ala Tre Leu
405 410 415

Glu Leu Leu Glu Arg Val Leu Arg Asp Met Asp Leu Leu Glu Cys Leu
420 425 430

Glu Asp Ile Glu Glu Ala Leu Cys Glu Pro Ala Ala Leu Pro Pro Ala
435 440 445

Pro Ser Leu Leu Arg Met Glu Leu Ser Tre Val Pro Asp Leu Leu Leu
450 455 460

Pro Leu Val Leu Leu Glu Leu Leu Val Glu Ile Tir Pro Ser Glu Val
465 470 475 480

Ile Glu Leu Val Pro His Leu Glu Asp Arg Glu Lis Arg Asp Ser Val
465 490 495

Cys Pro Gln Glu Lis Tir Ile His Pro Gln Asn Asn Ser Ile Cys Cys
500 505 510

Tre Lis Cys His Lis Glu Tre Tir Len Tir Asn Asp Cys Pro Glu Pro
515 520 525

Ala Glu Asp Tre Asp Cys Arg Glu Glu Glu Ser Glu Ser Fen Tre Ala
530 535 540

Ser Glu Asn His Leu Arg Eis Cys Leu Ser Cys Ser Lis Cys Arg Glu
545 550 555 560

Ile Glu Met Glu Glu Val Glu Ile Ser Ser Cys Tre Val Asp Arg Asp
565 570 575

Tre Val Glu Glu Glu Arg Lis Asn Glu Tir Arg Glu Tir Trp Ser Glu
580 585 590

Asn Leu Fen Glu Cys Fen Asp Cys Ser Leu Cys Leu Asp Glu Tre Val
600 605 610

His Leu Ser Cys Gln Glu Lis Gln Asn Tre Val Cys Tre Cys His Ala His
615 623

Gli Fen Fen Leu Arg Glu Asn Glu Cys Val Ser Cys Ser Asn Cys His
625 630 635 640

His Ser Leu Glu Cys Trp Lis Leu Cys Leu Pro Gln Ile Glu Asn Val
645 650 655

His Gli Tre Glu Asp Ser Gli Tre Tre Val Leu Leu Pro Leu Val His
660 665 670

Fen Fen Gli Leu Cys Leu Leu Ser Leu Leu Fen Ile Gli Leu Met Tir
675 683 683

Arg Tir Glu Arg Trp Cys Ser Asp Leu Tir Ser Ile Val Cys His His
690 695 700

Ser Tre Pro Glu Lis Glu Gli Glu Leu Glu Gli Tre Tre Lys Lis Pro
705 710 715 720

Leu Ala Pro Asn Pro Ser Fen Ser Pro Tre Pro Gli Fen Tre Pro Tre
725 730 733

Leu Gli Fen Ser Pro Val Pro Ser Ser Tre Fen Tre Ser Ser Ser Tre
740 745 753

Tir Tre Pro Gli Asp Cys Pro Asn Fen Ala Ala Pro Arg Arg Glu Val
755 760 765

His Pro Pro Trp Gln Gli Ala Asp Pro Ile Leu Ala Tre Ala Leu His
770 775 780

Ser Asp Pro Ile Pro Asn Pro Leu Gln Trp Trp Glu Asp Ser His His
785 790 795 800

His Pro Glu Ser Leu Asp Tre Asp Asp Pro Ala Tre Leu Tir His Val
805 810 815

Val Glu Asn Val Pro Pro Leu Arg Trp His Glu Fen Val Arg Arg Leu
820 825 830

Gli Leu Ser Pro His Glu Ile Asp Arg Leu Glu Leu Glu Asn His Arg
835 840 845

GLL Ala Leu Ser Arg His Tre Arg Leu Leu Cis Ser Arg Arg Asp Tre
 145 159 157 160
 Asp Cis Gli Tre Cis Met Pro Gln Fen Tyr Glu His Glu Asp Gli Cis
 155 163 168
 Val Ser- Cis Pro Tyr Ser Tre Leu Gln Ser Cis Pro Gln Arg His Ala
 180 195 190
 Val Cis Glu Trp Arg Gln Met Fen Trp Val Gln Val Leu Leu Ala
 195 200 205
 Leu Val Val Pro Leu Leu Leu Gln Ala Tre Leu Tre Tyr Tyr Tyr
 210 215 220
 Arg His Cis Trp Pro His Lis Pro Leu Val Tyr Ala Asp Glu Ala Gli
 225 230 235 240
 Met Glu Ala Leu Tre Pro Pro Pro Ala Tre His Leu Ser Pro Met Asp
 245 250 255
 Ser Ala His Tyr Met Leu Ala Pro Pro Asp Ser Ser Glu Lis Ile Cis
 260 265 270
 Tre Val Gln Leu Val Glu Asn Ser Trp Tre Pro Gln Tyr Pro Glu Tre
 275 280 285
 Gln Glu Ala Leu Cis Met Gln Val Tre Tyr Ser Trp Asp Gln Leu Pro
 290 295 300 305
 Ser Ser Ala Met Gln Pro Ala Ala Ala Pro Tre Leu Ser Pro Glu Ser
 310 315 320
 Pro Ala Gln Ser Pro Ala Met Met Leu Gln Pro Gln Pro Gln Met Tyr
 325 330 335
 Asp Val Met Asp Ala Val Pro Ala Ser Arg Tyr Lis Glu Fen Val Arg
 343 345 350
 Tre Leu Gln Leu Arg Glu Ala Glu Ile Glu Ala Val Glu Val Glu Ile
 355 360 365
 Gln Arg Fen Arg Asp Gln Gln Tyr Glu Met Leu Lis Arg Trp Arg Gln
 373 375 380
 Glu Gln Pro Ala Glu Leu Gln Ala Val Tyr Ala Ala Leu Glu Arg Met
 385 390 395 400
 Gln Leu Asp Gln Glu Val Gln Asp Leu Arg Ser Arg Leu Gln Arg Gli
 405 410 415

 Pro Met Glu Gln Arg Pro Arg Gln Cis Ala Ala Val Ala Ala Ala Leu
 420 425 430

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

 Arg Cis Asp Cis Ala Gln Asp Fen His Lis Lis Ile Gln Leu Fen Cis
 450 455 460

 Cis Arg Gln Cis Pro Ala Gln Tyr Tyr Leu His Ala Arg C:~Tre Glu

490 495 500
 Gln Gln Ala Leu Cis Pro Gln Val Tre Trp Ser Trp Asp Gln Leu Pro
 705 710 715 720
 Ser Arg Ala Leu Gln Pro Ala Met Ala Pro Trp Leu Ser Pro Glu Ser
 725 730 735
 Pro Ala Gln Ser Pro Ala Met Met Leu Gln Pro Gln Pro Gln Leu Val
 740 745 750
 Asp Val Met Asp Ala Val Pro Ala Arg Arg Trp Lis Glu Fen Val Arg
 755 760 765
 Tre Leu Gln Leu Arg Glu Ala Glu Ile Gln Ala Val Glu Val Glu Ile
 770 775 780
 Gln Arg Fen Arg Asp Gln Gln Tir Glu Met Leu Lis Arg Trp Arg Gln
 785 790 795 800
 Glu Glu Pro Ala Gln Leu Gln Ala Val Tir Ala Ala Leu Glu Arg Met
 805 810 815
 Gln Leu Asp Gln Cis Val Glu Asp Leu Arg Ser Arg Leu Gln Arg Gln
 820 825 830
 Pro

(2) INFORMAÇÃO PARA A SEQ ID N°. 6:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 426 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°: 6:

GGGAGGATG GATACCTGAC TACCTTCCA GATGAGATG TATGATCCA TCAATTCGA	60
CGACAGATG GAAAGGAGG GAGTTTCCA GAGTTTATG GATGAGATG TCAATTCGA	120
TAAGGATG GAGGATGTT AACCTGTCG GAAAGGATG GATGAGATG TCAATTCGA	180
AGGATGCA AATGATGAG GAGTTTATG AACAGGATG GATGAGATG TCAATTCGA	240
TATGATGCA TTGATGATG GAGTTTATG GATGAGATG GATGAGATG TCAATTCGA	300
TTTATGATG TTGATGATG TTGATGATG GATGAGATG GATGAGATG TCAATTCGA	360
TATGATGCA GATGAGATG GATGAGATG GATGAGATG GATGAGATG TCAATTCGA	420
GGAGATG	438

(2) INFORMAÇÃO PARA A SEQ ID N°. 7:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 339 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA- HELICOIDAL: simples
- (Dj TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°. 7:

TGGGGCTGAG GAGAGGAGG ACNACGAGAS TCTGACCAAC GAGGAGATG GATGAGATG	60
TGCTCTNTGN GAGGAGAGG GAAAGGAGG AGGAGGAGG TTGAGAGATG GATGAGATG	120
AGTCCCCAGG GGAGGCACAG TGTCTCCTGG TGAGTTGGGG ACAGGCCCTT GATGAGATG	180
GTGAGGAGG GATGAGATG GATGAGATG TTGAGAGATG TAAAGGAGG AGGAGATG	240
GAGGAGATG GATGAGATG GATGAGATG TATGAGATG GATGAGATG GATGAGATG	300
AAGGAGATG GATGAGATG GATGAGATG GATGAGATG	360

(2) INFORMAÇÃO PARA A SEQ ID N°. 8:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 28 pares de bases
- (B) TIPC: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°. 8:

GCGGATGCA TGATCATGCA ATTGGGAC

28

(2) INFORMAÇÃO PARA A SEQ ID N°. 9:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 36 pares de bases

(B) TIPO: ácido nucleico

(C) ESTRUTURA HELICOIDAL: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°. 9:

GCGGGATCCG CCATCATGGC GCCACCACCA GCTAGA

34

(2) INFORMAÇÃO PARA A SEQ ID N°. 10:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 33 pares de bases

(B) TIPO: ácido nucleico

(C) ESTRUTURA HELICOIDAL: simples

(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°. 10:

GCGGATGCTT CACTCCACAG ACACGGCAGA GCC

33

(2) INFORMAÇÃO PARA A SEQ ID N°. 11:

CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 29 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID N°. 11:

GCGGGATCCT CAATTATGTC CATTGCCTG 29

Lisboa, 2 de Agosto de 2007

REIVINDICAÇÕES

Molécula de ácido nucleico caracterizada pelo facto de compreender uma sequência de nucleótidos seleccionada no grupo que consiste em:

- (a) uma sequência de nucleótidos que codifica o polipéptido do DM4, de comprimento completo, que compreende a sequência de aminoácidos completa da SEQ ID N°. 2, incluindo a sequência líder prevista;
- (b) uma sequência de nucleótidos que codifica o polipéptido do DM4, de comprimento completo, que compreende a sequência de aminoácidos completa da SEQ ID N°. 2, incluindo a sequência líder prevista, mas a que falta a metionina do terminal amino;
- (c) uma sequência de nucleótidos que codifica o polipéptido de DM4 maduro (polipéptido de comprimento completo com a sequência líder eliminada) que compreende a sequência de aminoácidos nas posições próximas de 24 até às próximas de 468 da SEQ ID N°. 2;
- (d) uma sequência de nucleótidos que codifica o polipéptido do DM4, de comprimento completo, que compreende a sequência de aminoácidos, incluindo a sequência líder codificada pelo clone do ADNC contido no depósito n° 97853 da ATCC;
- (e) uma sequência de nucleótidos que codifica o polipéptido do DM4, de comprimento completo, que compreende a sequência completa de aminoácidos, incluindo a sequência líder mas a que falta a

metionina do terminal amino codificada pelo eione do ADNc contido no depósito n° 97853 da ATCG;

- (f) uma sequência de nucleótidos que codifica o polipeptido maduro do DM4, que compreende a sequência de aminoácidos codificada pelo clone do ADNc contido no depósito n° 97833 da ATCC;
- (g) uma sequência de nucleótidos que codifica o domínio extracelular do DM4, que compreende a sequência de aminoácidos nas posições próximas de 24 até às próximas de 238 da SEQ ID N°. 2 ou conforme codificadas pelo clone do depósito n° 97853 da ATCC;
- (h) uma sequência de nucleótidos que codifica o domínio da trançrnembrana do DM4, que compreende a sequência de aminoácidos nas posições próximas de 239 até as próximas de 264 da SEQ ID N°. 2 ou conforme codificadas pelo clone do depósito n° 97853 da ATCC;
- (i) uma sequência de nucleótidos que codifica o domínio intracelular do DM4, que compreende a sequência de aminoácidos nas posições próximas de 265 até às próximas de 468 da SEQ ID N°. 2 ou conforme codificadas pelo clone do depósito n° 97853 da ATCC;
- (j) uma sequência de ácidos nucleicos que compreende a sequência completa de nucleótidos da SEQ ID N°. 3 ou o clone do ADNc do depósito n° 97853 da ATCC;

- (k) uma sequência de ácidos nucleicos que compreende a sequência de nucleótidos da SEQ ID N°. 1 que codifica o polipéptido do DM4 que comporta a sequência de aminoácidos nas posições 2 a 468 da SEQ ID N°. 2;
- (l) uma sequência de ácidos nucleicos que compreende a sequência de nucleótidos da SEQ ID M°. 1 que codifica o domínio extracelular do polipéptido do DM4 que comporta a sequência de aminoácidos desde cerca de 24 até cerca de 238 da SEQ ID N°. 2;
- (m) uma sequência de nucleótidos que codifica um polipéptido que compreende a sequência de aminoácidos dos resíduos n a 468 da SEQ ID N°. 2, em que c símbolo n representa um número inteiro no intervalo de 1 a 132;
- (n) uma sequência de nucleótidos que codifica um polipéptido que compreende a sequência de aminoácidos dos resíduos 1 a m da SEQ ID N°. 2, em que o símbolo m representa um número inteiro no intervalo de 221 a 468;
- (o) uma sequência de nucleótidos que codifica um polipéptido que compreende a sequência de aminoácidos que consiste nos resíduos n a m da SEQ ID N°. 2, em que os símbolos n e m representam números inteiros, tal como definido antes para (m) e para (n), respectivamente;
- (p) uma sequência de nucleótidos que codifica um polipéptido que consiste numa porção da sequência completa de aminoácidos de DM4 codificada pelo

clone de ADNc do depósito n° 97853 da ATCC, em que a referida porção exclui de 1 a cerca de 131 aminoácidos da terminação amino da referida sequência completa de aminoácidos codificada pelo clone de ADNc contido no depósito n° 97853 da ATCC;

(q) uma sequência de nucleótidos que codifica um polipéptido que consiste numa porção da sequência completa de aminoácidos de DM4 codificada pelo clone de ADNc contido no depósito n° 97853 da ATCC, em que a referida porção exclui de 1 a cerca de 249 aminoácidos da terminação carboxi da referida sequência completa de aminoácidos codificada pelo clone de ADNc contido no depósito n° 97853 da ATCC;

(r) uma sequência de nucleótidos que codifica um polipéptido que consiste numa porção da sequência completa de aminoácidos de DM4 codificada pelo clone de ADNc contido no depósito n° 97853 da ATCC, em que a referida porção inclui uma combinação de qualquer uma das eliminações do terminal amino ou do terminal carboxi nas (p) e (q) anteriores;

(s) uma sequência de nucleótidos que é pelo menos 95 % idêntica a uma sequência de nucleótidos conforme foi definido por (a) a (r) e que codifica um polipéptido que tem actividade da proteína de DM4;

(i) uma sequência de nucleótidos que é pelo menos 95 % idêntica a uma sequência de aminoácidos de um polipéptido codificado por uma sequência de nucleótidos de qualquer uma das (a) a (r) e que

codifica um polipeptido que tem actividade da proteina de DM4;

- (u) uma sequencia de nucleotidos que codifica um polipeptido de DM4 que é, excepto para pelo menos uma substituição conservadoia de aminoácido, um polipeptido codificado pela sequencia de nucleotidos de qualquer uma das (a) a (r); e
- (v) uma sequencia de nucleotidos que hibrida, em condições de hibridação severas, com uma sequencia de nucleotidos conforme definido por (s) e (t), em que a referida sequencia de nucleotidos que hibrida, não hibrida em condições de hibridação severas com um polinucleotido que tem uma sequencia de nucleotidos que consiste apenas em resíduos A ou em apenas resíduos T; e em que a referida sequencia de nucleotidos codifica um polipeptido que e capaz de se ligar a ligandos que induzem a apoptose relacionada com o FNT JLIART; TRAIL na lingua inglesa)

ou a estrutura helicoidal complementar desse polinucleotido.

2. Molécula de ácido nucleico, caracterizada pelo facto de compreender um polinucleotido que codifica a sequencia de aminoácidos de uma porção do polipeptido de DM4 ligado a um epítipo e que comporta uma sequencia de aminoácidos conforme definido em uma qualquer das (a) a (i) e (k) a (r) da reivindicação 1, com a condição de que a referida porção que se liga a um epítipo não esteja contida na seguinte sequencia de aminoácidos:

Met	Gln	Gli	Ala	Leu	Gln	Val	Leu	Ser	Leu	Gli	Ser	Gli	Pro	Asp	Val
1				5					10					15	
Gli	Ser	Ser	Ala	Val	Ser	Glu	Leu	Ser	Pro	Asn	Ser	Ser	Gli	Cis	Leu
			20					25					30		
Cis	Gli	Val	Leu	Ala	Pro	Ala	Leu	Met	Leu	Fen	Fen	Asp	Lis	Fen	Ala
		35					40					45			
ksn	Ile	Val	Pro	Fen	Asp	Ser	Trp	Asp	Gln	Leu	Met	Arg	Glu	Met	Asp
50						55					60				
Leu	Tre	Lis	Asn	Gln	Ile	Asp	Val	Val	Arg	Ala	Gli	Tre	Ala	Gli	Met
65					70					75					80
Gli	Asp	Ala	Leu	Tir	Ala	Bet	Met	Met	Lis	Trp	Val	Met	Ala	Val	Gli
				85					90					95	
Arg	Asn	Ala	Ser	Ile	Ala	Tre	Leu	e	Asp	Ala	Leu	Glu	Arg	e	Glu
			100					105					110		
Glu	Met	His	Ala	Lis	Glu	Lis	Ile	Gln	Asp	Leu	Leu	Val	Asp	Ser	Gli
		115					120					125			
Lis	Fen	Ile	Tir	Leu	Glu	Asp	Gli	Tre	Gli	Met	Ala	Val	Ser	Leu	Glu
	130						135					140			

Molécula de ácido nucleico, de acordo com a reivindicação 2, caracterizada pelo facto de codificar uma porção de um polipéptido de SEQ ID NO. 2 que se liga a um epítopo em que a sequência de aminoácidos da referida porção se selecciona no grupo que consiste em: um polipéptido que compreende os resíduos de aminoácidos desde cerca do 35 até cerca do 92 da SEQ ID NO. 2, um polipéptido que compreende os resíduos de aminoácidos desde cerca do 114 até cerca do 160 da SEQ ID NO. 2, um polipéptido que compreende os resíduos de aminoácidos desde cerca do 160 até cerca de 240 da SEQ ID NO. 2, um polipéptido que compreende os resíduos de aminoácidos desde cerca do 267 até cerca do 298 da SEQ ID NO. 2 e um polipéptido que compreende os resíduos de aminoácidos desde cerca do 330 até cerca do 364 da SEQ ID NO. 2.

- Molécula de ácido nucleico de acordo com uma qualquer das reivindicações 1 a 3, caracterizada pelo facto de ser ADN ou ARN.

5. Processo para produzir um vector, caracterizado pelo facto de compreender a inserção, no vector, de uma molécula de ácido nucleico, de acordo com uma qualquer das reivindicações 7 a 4.
6. Vector, caracterizado pelo facto de compreender a molécula de ácido nucleico, de acordo com uma qualquer das reivindicações 1 a 4 ou produzida pelo processo de acordo com a reivindicação 5.
7. Processo de produção de uma célula hospedeira, caracterizado pelo facto de compreender a introdução do vector, de acordo com a reivindicação 6, numa célula hospedeira.
8. Célula hospedeira, caracterizada pelo facto de compreender a molécula de ácido nucleico de acordo com uma qualquer das reivindicações até 4 ou o vector de acordo com a reivindicação 6 ou produzida de acordo com o processo da reivindicação 7.
9. Processo para a produção de um polipéptido de DM4, caracterizado pelo facto de compreender a cultura da célula hospedeira, de acordo com a reivindicação 8, em condições tais que o referido polipéptido é expresso e se recupera o referido polipéptido.
10. Polipéptido, caracterizado pelo facto de comportar a sequência de aminoácidos codificada pela molécula de ácido nucleico, de acordo com a reivindicação 1 ou que se pode obter por meio ao processo de acordo com a reivindicação 9,

11. Polipéptido caracterizado pelo facto de compreender ~~uma~~ porção da proteína de DM4 produzida por um epítopo, em que a referida porção ~~é~~ codificada pela molécula de ácido nucleico de acordo com a reivindicação 2 ou 3.
12. Anticorpo, caracterizado pelo facto de se ligar especificamente ao polipéptido de acordo com a reivindicação 10 ou 11.
13. Molécula anti-paralela (antisense) caracterizada pelo facto de ser capaz de controlar a expressão do polipéptido de acordo com a reivindicação 10 ou 11, hibridando especificamente a referida molécula anti-paralela com a molécula de ácido nucleico de acordo com uma qualquer das reivindicações 1 a 4.
14. Agonista do polipéptido de acordo com as reivindicações 10 ou 11, caracterizado pelo facto de ser um anticorpo de acordo com a reivindicação 12, capaz de aumentar a actividade apoptótica do referido polipéptido.
15. Composição farmacêutica, caracterizada pelo facto de compreender o polipéptido de acordo com a reivindicação 10 ou 11, o anticorpo de acordo com a reivindicação 12, o agonista de acordo com a reivindicação 14 ou o antagonista de acordo com a reivindicação 13 e, eventualmente, um veículo aceitável sob o ponto de vista farmacêutico.
16. Composição de diagnóstico caracterizada pelo facto de compreender ~~a~~ a molécula de ácido nucleico de acordo com ~~uma~~ qualquer das reivindicações 1 a 4 ou um anticorpo de acordo com a reivindicação 12.

17. Processo para rastrear um antagonista ou um agonista do polipéptido de acordo com as reivindicações 10 ou 11, caracterizado pelo facto de compreender a cultura da célula hospedeira de acordo com a reivindicação 8, em condições tais que o referido polipéptido é expresso, fazendo contactar a referida célula hospedeira com um composto candidato e um ligando com o referido polipéptido, ensaiando uma resposta celular e comparando a resposta ensaiada com uma resposta celular quando a referida célula hospedeira contacta apenas com o referido ligando.

Lisboa, 2 de Agosto de 2007

850 870 890
 ATGGACAGGGTGTGTTTCTGGCGCTTGGSTCTCCTACGAGGGCCTGGGCTGAGSACAAT
 M D R V C F W R L G L L R G P G A E D N
 910 930 950
 GCTCACAACGAGATTCTGAGCAACGCAGACTCGCTGTCCACTTTCGTCTCTGAGCAGCAA
 A H N E I L S N A D S L S T F V S E Q Q
 970 990 1010
 ATGGAAAGCCAGGAGCCCGCAGATTTGACAGTGTCTCACTGTACAETCCCCAGGGGAGCCA
 M E S Q E P A D L T G V T V Q S P G E A
 1030 1050 1070
 CAGTGTCTGCTGGGACCGGCAGAAAGCTGAAGGGTCTCAGAGGAGGAGGCTGCTGGTTCCA
 Q C L L G P A E A E G S Q R R R L L V P
 1090 1110 1130
 GCAAATGGTGTGACCCCACTGAGACTCTGATGCTGTTCTTTGACAAGTTTGCAAACATC
 A N G A D P T E T L M L F F D K F A N I
 1150 1170 1190
 GTGCCCTTTGACTCCTGGGACCAGCTCATGAGGCAGCTGGACCTCACCAGAAAATGAGATC
 V P F D S W D Q L M R Q L D L T K N E I
 1210 1230 1250
 GATGTGGTTCAGAGCTGGTACAGCAGGCCAGGGGATGCCTTGTATGCAATGCTGATGAAA
 D V V R A G T A G P G D A L Y A M L M K
 1270 1290 1310
 TGGGTCAACAAAACCTGCACCGAAGCCTGATCCACACCCTGCTGGATGCCTTGGAGAGG
 W V N K T G R N A S I H T L L D A L E R
 1330 1350 1370
 ATGGAAGAGAGACATGCAAAAGAGAAGATTGAGACCTCTTGGTGGACTCTGAAAAGTTC
 M E E R H A K E K I Q D L L V D S G K F
 1390 1410 1430
 ATCTACTTAGAAGATGGCACAGGCTCTGCCGTGTCTTGGAGTGAAAGACTCTTTTTTACC
 I Y L E D G T F S A V S L E
 1450 1470 1491)
 AGAGGTTTCTCTTAGGTGTTAGGAGTTAATACATATTAGGTTTTTTTTTTTTTTTAAACAT
 1510 1530 1550
 GTATACAAAGTAAATTCTTAGCCACGTGATTGGCTCCTGCCTGTAATCCCATCACTTTG
 1570 1590 1610
 GGAGGCTGACGCCGGTGGATCCACTTGAGGTCCGAAGTTCCAAGACCAGCCCTGAACCAA
 1630 1650 1670
 CATCGTGGAAATGCCCGTCTTTTACAAAAAATACCAAAAATTCACCTGGAATGTGCATG

FIG. 1B

1690	1710	1730
GTGTGTGCCATCATTTCCTGGCTAACTACGGGAGGTCTGAGGCCAGSAGAAATCCACTTG		
1750	1770	1790
AACCCACCGAAGGACAGTGTAGACTGCAGATTGCACCACTGCACTCCCAGCCTGSSAACA		
1810	1830	1850
CAGAGCAAGACTCTGTCTCAAGATAAAATAAAATAAACTTGAAAGAATTATTGCCCGACT		
1870	1890	1910
GAGGCTCACATGCCAAAGGAAAATCTGGTTCCTCCCTGAGCTGGCCTCGCTGTGTTTCT		
1930	1950	1970
TATCATGCTGCTCAATTGGAGGTGTTAATTTGAATGGATTAAGGAACACCTAGAACACTG		
1990	2010	2030
GTAAGGCATTATTTCTGGGACATTATTTCTGGGCATGTCCTCCAGGGTGTTCAGAGGG		
2050	2070	2090
GATTGCCATGCCATCCGGTGGACTGAGTGGAAAAGACCTACCCCTTAATTTGGGGGGCCAC		
2110	2130	2150
CGTCCGACAGACTGGGGAGCAAGATAGAAGAAAACAAAAAAAAAAAAAAAAAAAAA		

FIG. 1C

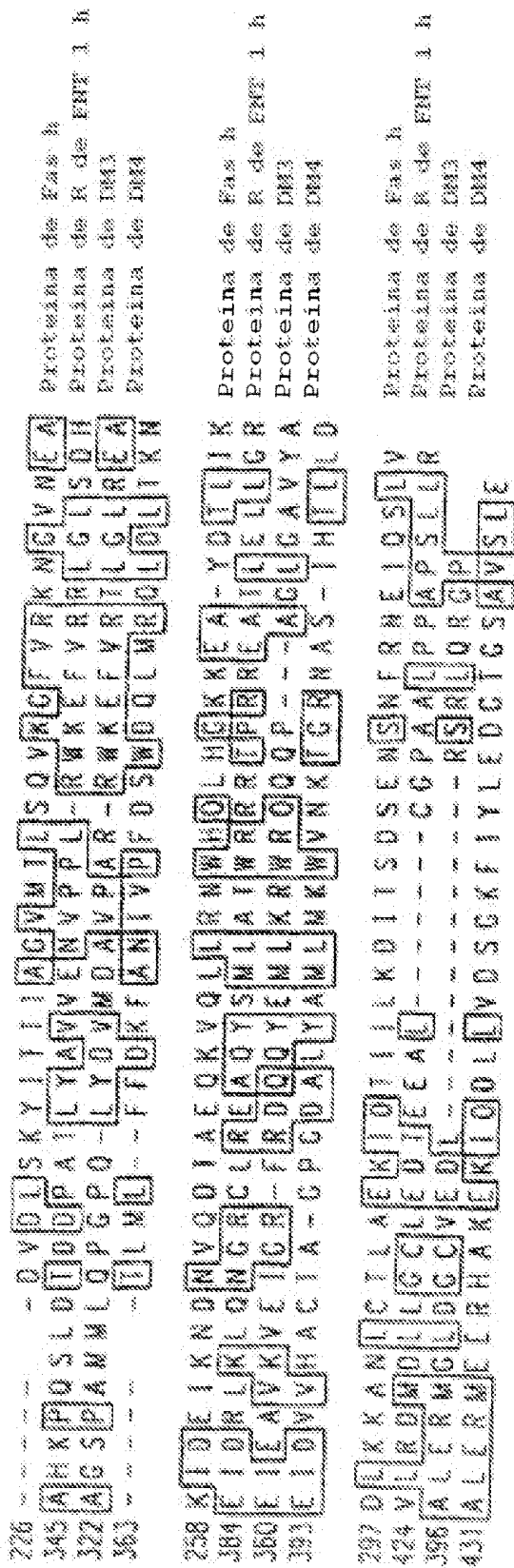


FIG.2F

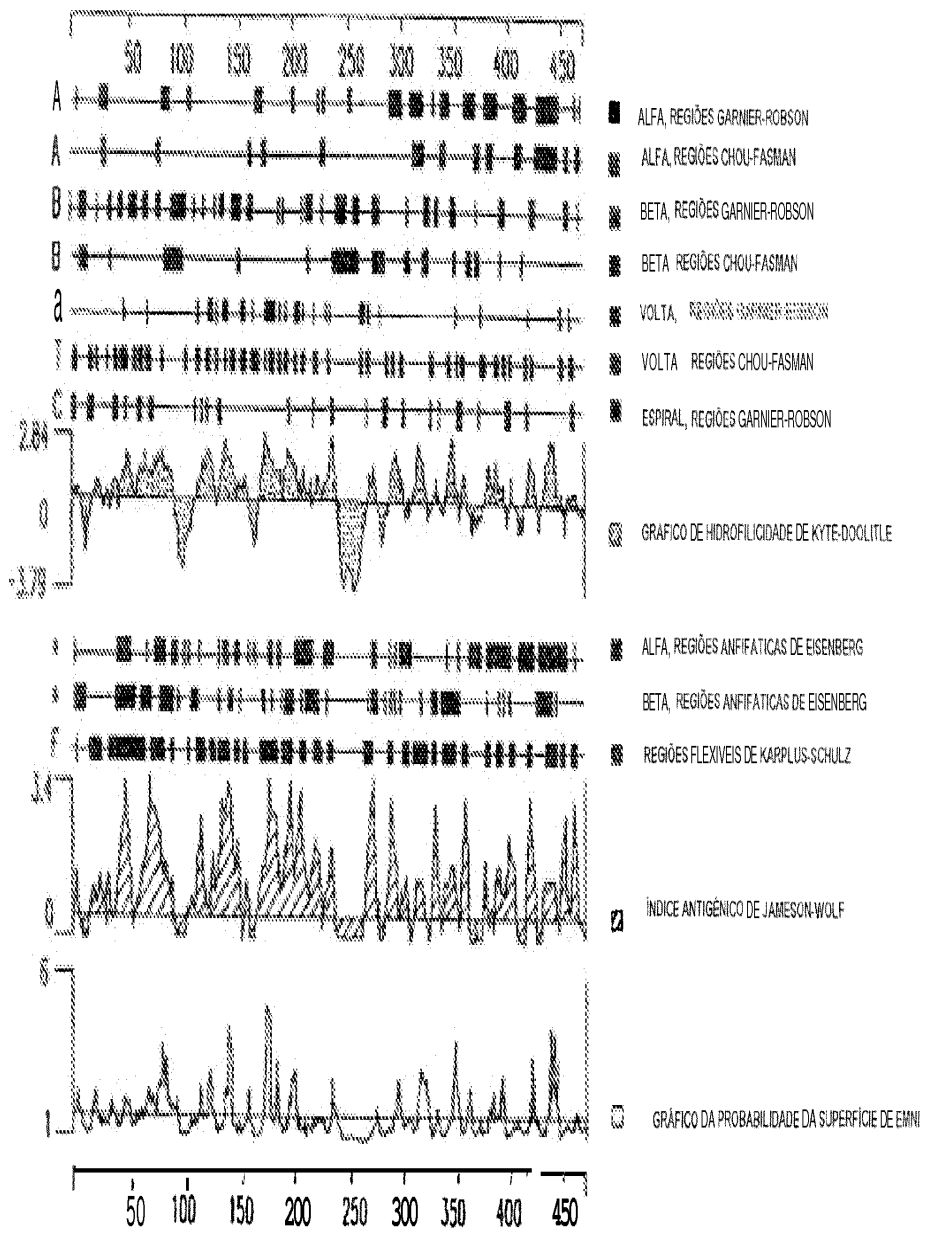


FIG.3

HTDIY07R

```
1  GGCANAGGTN CGTACCTAGC TCACCTGCAA CCATCAAACT TNATGATCAA
51  TCAATTGGCA CACAGCAATG GAAACATAG CCCTTTGGAA GANTTSTNTC
101 CACCAGGATC TCATAGATCA AAACATCCTG GGAGCCTGTT AACCGSTGCC
151 CCAAAGGNTG GTCAAGGTCA AGGATTGTT NCGCCCTGCA AGTBAACATC
201 GAGTGTNTCC ACAAGGATT CAGGCAATGG GACATAAATA TATGGGTGAA
251 TTTTGGTTGT GAACTTTGGT TGNTCCCGTT GNTGTTGNTG GCTGTGCTGA
301 TTGTTTGTG TTGCATCGGC TTCAGGTTNT GGAGGGGGAC CCAAGTGCAT
351 GGACAGGGTG TGTTCCTGGG GTTTGGGTCT CTTAGAGGGC NTGGGTTANG
401 GCANGTTCAC AAGGGTTTTA GCAANG
```

HTXEYBOR

```
1  TGGGGCTGAG GACAATGCTG ACNACGAGAT TCTGAGCAAC GCAGNACTNG
53  CTGTCCACTT TCGTCTNTGN GCAGCAAATG GAAAGCCAGG AGCCGGCAGA
101 TTTGACAGGT GTCACTGTAC AGTCCCCAGG GGAGGCACAG TGTCGTCTGG
151 TGAGTTGGGG ACAGGCCCTT GCAAGACCTT GTGAGGCAGG GGGTGAAGGC
201 CATGNCTCGG CTTCNNNTGG TCAAAGGGGA AGTGGAGCCT GAGGGAGATG
251 GCACTTNAGG GGGACGGNGC TCGTGGGGA AAAAGCACCC ACCNTTTGAC
301 AAGGGGGACA GGCATTTTTN CAAATGTGTG CTTNTTGGT
```

FIG.4

FIG. 5A

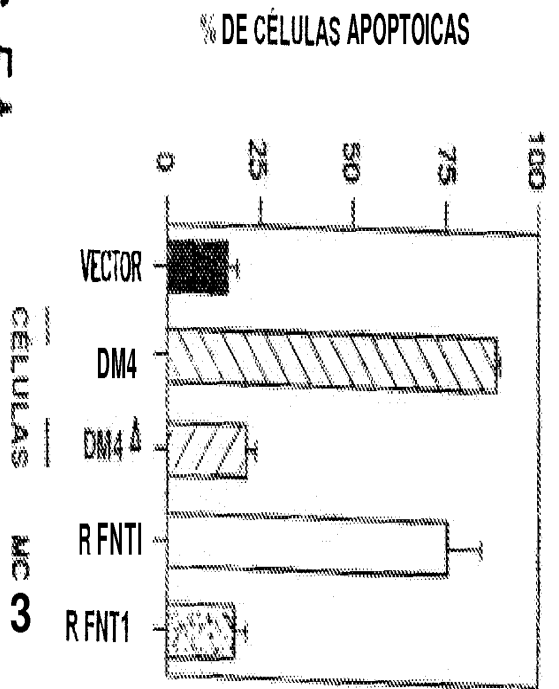
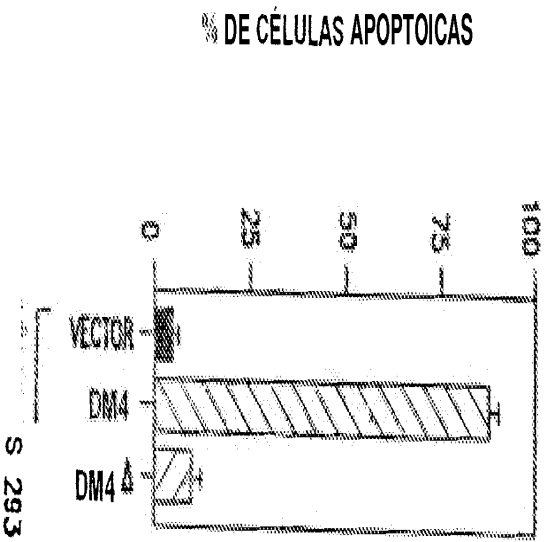


FIG. 5B



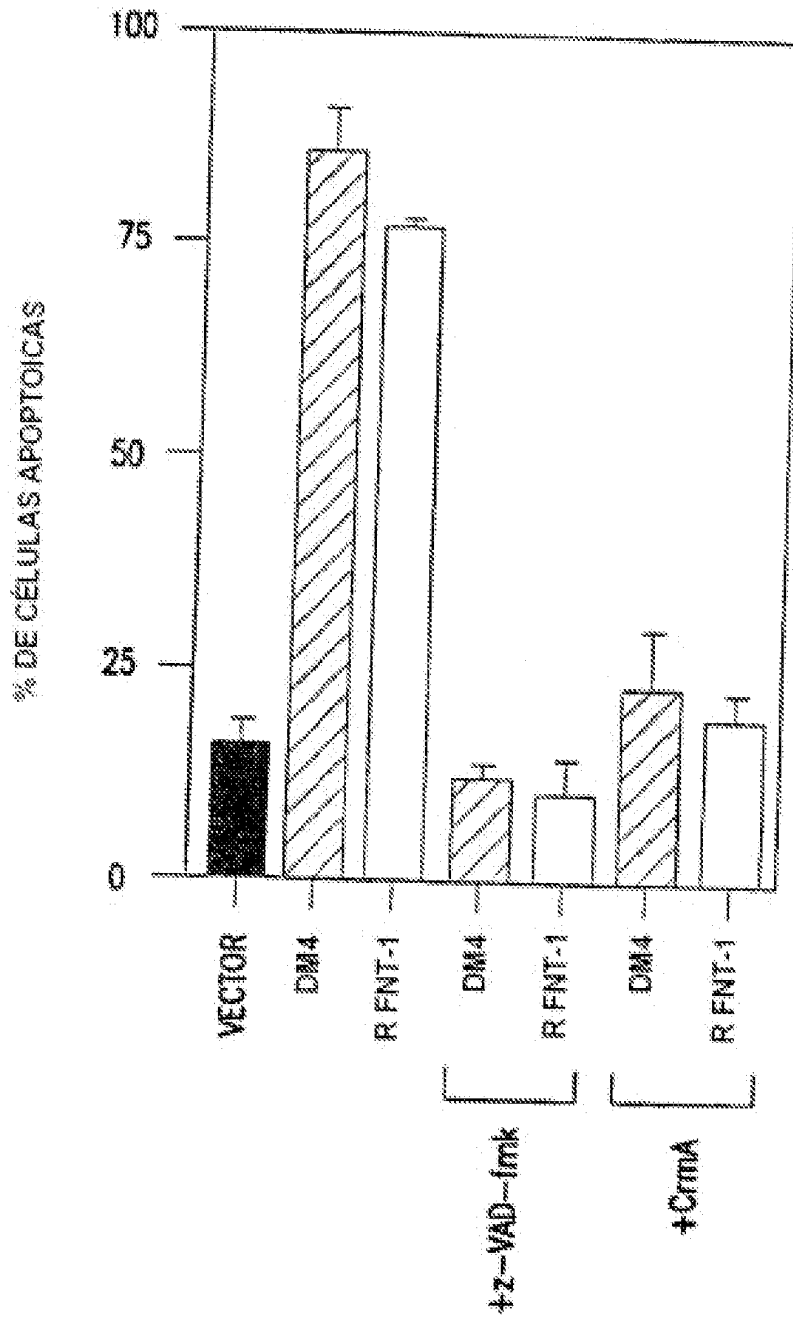


FIG.5C

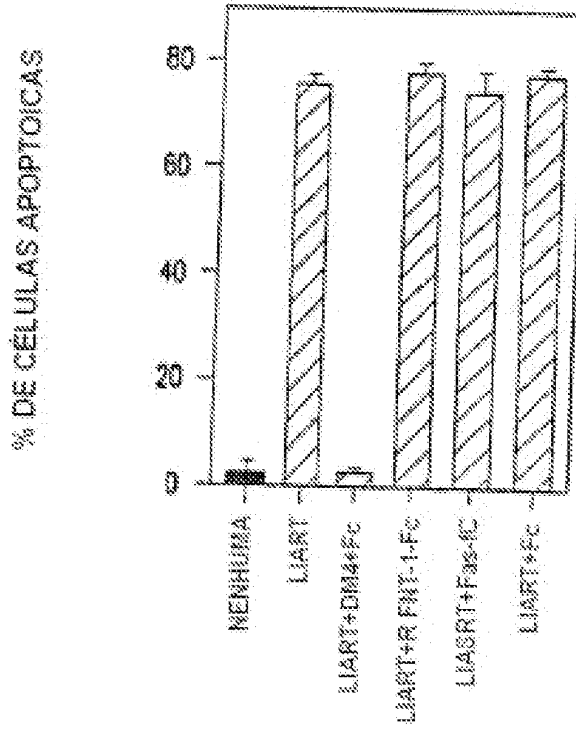


FIG.6A

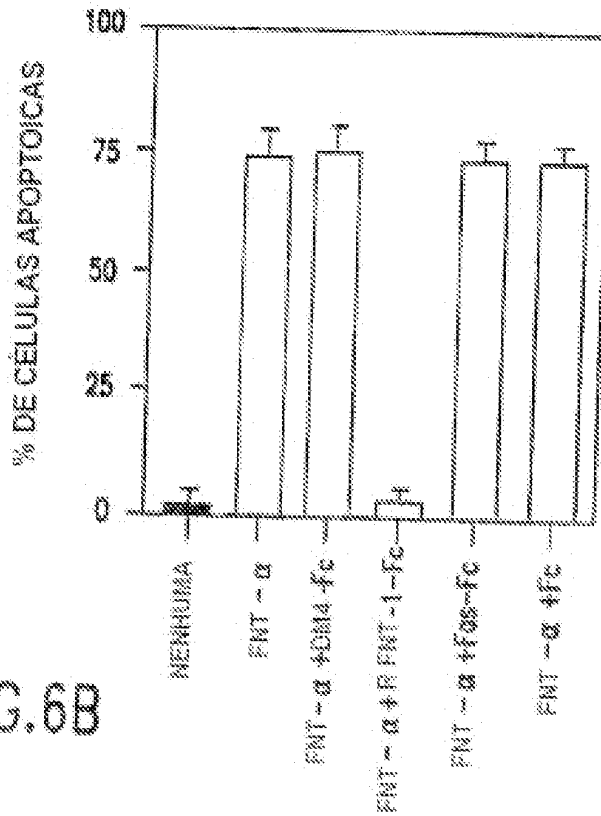


FIG.6B