

(11) Número de Publicação: **PT 1141297 E**

(51) Classificação Internacional:

C12N 15/24 (2007.10) **C07K 14/54** (2007.10)

A61K 38/20 (2007.10) **C07K 16/24** (2007.10)

G01N 33/68 (2007.10)

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

(22) Data de pedido: **2000.01.10**

(30) Prioridade(s): **1999.01.11 US 228822**

(43) Data de publicação do pedido: **2001.10.10**

(45) Data e BPI da concessão: **2009.09.30**
003/2010

(73) Titular(es):

SCHERING CORPORATION

2000 GALLOPING HILL ROAD, KENILWORTH

NEW JERSEY 07033-0530 US

(72) Inventor(es):

J. FERNANDO BAZAN

US

ROBERT A. KASTELEIN

US

DANIEL M. GORMAN

US

GERARD ZURAWSKI

US

(74) Mandatário:

ANTÓNIO JOÃO COIMBRA DA CUNHA FERREIRA

RUA DAS FLORES, Nº 74, 4º AND 1249-235 LISBOA

PT

(54) Epígrafe: **CITOQUINAS DE MAMÍFEROS RELACIONADAS COM A INTERLEUCINA-17.
POLINUCLEÓTIDOS QUE AS CODIFICAM. UTILIZAÇÕES**

(57) Resumo:

DESCRIÇÃO

"Citoquinas de mamíferos relacionadas com a interleucina-17. Polinucleótidos que as codificam. Utilizações."

CAMPO DO INVENTO

O presente invento refere-se a composições relacionadas com proteínas que actuam no controlo da fisiologia, do desenvolvimento e da diferenciação das células de mamíferos, por exemplo, as células do sistema imunitário de um mamífero. Em particular, proporciona ácidos nucleicos e proteínas que regulam a fisiologia, o desenvolvimento e a diferenciação celulares ou a função de vários tipos de células, incluindo as células hematopoiéticas.

ANTECEDENTES DO INVENTO

O sistema imunitário dos vertebrados consiste num determinado número de órgãos e vários tipos diferentes de células. Dois tipos principais de células incluem as linhagens mielóide e linfóide. Entre a linhagem de células linfóides encontram-se as células B, que foram originalmente caracterizadas como sofrendo diferenciação no fígado fetal ou na medula óssea do adulto, e os linfócitos T, que foram originalmente caracterizados como sofrendo diferenciação no timo. Consultar, por exemplo, Paul (ed. 1998) Fundamental Immunology (4th ed.) Raven Press, New York.

Em muitos aspectos do desenvolvimento de uma resposta imunitária ou da diferenciação celular, proteínas solúveis conhecidas por citocinas desempenham um papel crítico na regulação das interações celulares. Aparentemente, estas citocinas medeiam as actividades celulares de muitas formas. Em muitos casos, demonstrou-se que modulam a proliferação, o crescimento e a diferenciação das células estaminais hematopoiéticas no grande número de progenitores que constitui as linhagens responsáveis por uma resposta imunitária.

Contudo, as moléculas celulares que são expressas pelas células em diferentes fases do desenvolvimento, nestas vias de maturação, ainda estão identificadas de forma incompleta. Além disso, os papéis e os mecanismos de acção das moléculas de

sinalização que induzem, sustentam ou modulam os vários estados fisiológicos, do desenvolvimento ou proliferativos destas células são mal compreendidos. Manifestamente, o sistema imunitário e a sua resposta a várias pressões foram relevantes para a medicina, por exemplo, doenças infecciosas, respostas e tratamentos relacionados com o cancro, respostas alérgicas e de rejeição de transplantes. Ver, por exemplo, Thorn et al., Harrison's Principles of Internal Medicine McGraw/Hill, New York.

A ciência médica confia, em larga medida, no recrutamento ou supressão apropriados do sistema imunitário para a obtenção de curas para as respostas fisiológicas insuficientes ou inadequadas a factores ambientais. Contudo, a falta de conhecimento sobre a forma como o sistema imunitário é regulado ou se diferencia bloqueou a possibilidade de modular de forma vantajosa os mecanismos normais de defesa a provocações biológicas. As condições médicas caracterizadas por uma regulação anormal ou inapropriada do desenvolvimento ou da fisiologia das células relevantes permanecem, assim, pouco manejáveis. O encontro e a caracterização de citocinas específicas contribuirão para o desenvolvimento de terapias destinadas a uma ampla gama de condições degenerativas ou de outros tipos que afectam o sistema imunitário e as células hematopoiéticas, bem como outros tipos de células. O presente invento proporciona soluções para alguns destes e muitos outros problemas.

SUMÁRIO DO INVENTO

O presente invento baseia-se, em parte, na identificação de clones de ADNc que codificam várias proteínas semelhantes a citocinas que apresentam uma similaridade de sequência significativa com a citocina designada por CTLA-8.

O invento contempla genes isolados que codificam a proteína IL-174 do invento, variantes da proteína codificada, por exemplo, mutações (muteínas) das sequências naturais, variantes alélicas e variantes de espécies, proteínas de fusão, compostos químicos miméticos e outros análogos estruturais ou funcionais. Várias utilizações destas diferentes composições de ácidos nucleicos ou proteínas são também descritas.

Em determinadas concretizações de ácidos nucleicos, o invento proporciona um polinucleótido isolado ou recombinante compreendendo a sequência de uma IL-174 de mamífero que: codifica pelo menos 16 aminoácidos contíguos da SEQ ID NO:14; ou compreende um ou mais segmentos com pelo menos 33 nucleótidos contíguos da SEQ ID NO:13. Outras concretizações incluem um polinucleótido deste tipo num vector de expressão, compreendendo uma sequência que: codifica pelo menos 16 aminoácidos contíguos da SEQ ID NO:14 ou compreende pelo menos 33 nucleótidos contíguos da SEQ ID NO:13. Determinadas concretizações incluirão aqueles polinucleótidos que codificam pelo menos 16 resíduos de aminoácidos contíguos da SEQ ID NO:14; compreendem pelo menos 33 nucleótidos contíguos da SEQ ID NO:13 ou compreendem a totalidade da porção codificante madura da SEQ ID NO:13.

Vários métodos são proporcionados, por exemplo, a produção de: a) um polipéptido, que compreende a expressão do vector de expressão descrito, produzindo deste modo o polipéptido; b) um ácido nucleico de cadeia dupla, que compreende o contacto de um polinucleótido descrito com um ácido nucleico complementar, resultando desta forma na produção do ácido nucleico de cadeia dupla; ou c) um polinucleótido descrito, que compreende a amplificação utilizando um método de PCR.

Encontra-se aqui também descrito um polinucleótido isolado ou recombinante que hibrida, em condições rigorosas de lavagem consistindo em pelo menos 55°C e uma concentração de sal inferior a 400 mM, com o polinucleótido descrito (IL-174) que consiste na porção codificante da SEQ ID NO:13. Estes polinucleótidos descritos incluem polinucleótidos em que as condições de lavagem são pelo menos 65°C e uma concentração de sal inferior a 300 mM. Em outras concretizações, são proporcionados polinucleótidos que compreendem pelo menos 50 nucleótidos contíguos da SEQ ID NO:13 (IL-174).

Determinados *kits* são descritos, por exemplo, *kits* que compreendem um polinucleótido descrito e: a) instruções para a utilização do polinucleótido para detecção; b) instruções para

a eliminação do polinucleótido ou de outros reagentes do *kit*; ou c) ambas a e b.

Várias células são também proporcionadas, por exemplo, uma célula contendo o vector de expressão descrito, em que a célula é: uma célula procariótica; uma célula eucariótica; uma célula bacteriana; uma célula de levedura; uma célula de insecto; uma célula de mamífero; uma célula de ratinho; uma célula de primata ou uma célula humana.

As concretizações do polipéptido incluem, por exemplo, um polipéptido IL-174 antigénico isolado ou recombinante compreendendo pelo menos um segmento de 16 aminoácidos contíguos e idênticos da SEQ ID NO:14. Outras concretizações incluem um polipéptido descrito, em que o polipéptido IL-174: a) compreende a SEQ ID NO:14 madura; b) se liga com selectividade a um anticorpo policlonal produzido contra um imunogénio da SEQ ID NO:14 madura; c) compreende uma pluralidade de segmentos polipeptídicos distintos de 16 aminoácidos contíguos da SEQ ID NO:14; d) possui um comprimento de pelo menos 30 aminoácidos; ou e) apresenta pelo menos dois epítomos não sobreponíveis que são selectivos para a SEQ ID NO:14 madura. Várias outras concretizações incluem um polipéptido descrito deste tipo que: a) está numa composição estéril; b) não é glicosilado; c) está desnaturado; d) é um polipéptido sintético; e) está ligado a um substrato sólido; ou f) é uma proteína de fusão com uma marca de detecção ou de purificação.

Métodos de utilização dos polipéptidos descritos são também proporcionados, por exemplo, a) para marcar o polipéptido, que compreende a marcação do polipéptido com um marcador radioactivo; b) para separar o polipéptido de outro polipéptido numa mistura, que compreende correr a mistura numa matriz de cromatografia, separando desta forma os polipéptidos; c) para identificar um composto que se liga selectivamente ao polipéptido, que compreende a incubação do composto com o polipéptido sob condições apropriadas, fazendo, assim, com que o composto se ligue ao polipéptido; ou d) para conjugar o polipéptido com uma matriz, que compreende derivatizar o polipéptido com um reagente activo e conjugar o polipéptido com a matriz.

São igualmente descritos anticorpos, incluindo um composto de ligação que compreende uma porção de ligação ao antigénio de um anticorpo que se liga com selectividade a um destes polipéptidos descritos, em que o polipéptido IL-174 compreende a SEQ ID NO:14. Um composto de ligação pode ser um anticorpo policlonal que é produzido contra o polipéptido IL-174 da SEQ ID NO:14. Outros aspectos incluem um composto de ligação descrito deste tipo, em que: a) o anticorpo; i) é imuno-seleccionado; ii) se liga a uma proteína desnaturada; ou iii) apresenta uma K_d para o polipéptido de pelo menos 30 mM; ou b) o composto de ligação: i) está ligado a um substrato sólido, incluindo uma esfera ou uma membrana de plástico; ii) encontra-se numa composição estéril; ou iii) está marcado de forma detectável, incluindo um marcador radioactivo ou fluorescente.

Métodos são descritos, por exemplo, a produção de um complexo antigénio:anticorpo, que compreende o contacto de uma sequência compreendendo um polipéptido da SEQ ID NO:14 com um composto de ligação descrito, em condições que permitem que o complexo se forme. Preferencialmente, o composto de ligação é um anticorpo, e o polipéptido encontra-se numa amostra biológica.

Kits são descritos, por exemplo, kits compreendendo um composto de ligação descrito e: a) um polipéptido de SEQ ID NO:14; b) instruções para a utilização do composto de ligação para efeitos de detecção; ou c) instruções para a eliminação do composto de ligação ou de outros reagentes do *kit*.

E é descrito um método de avaliação da selectividade da ligação de um anticorpo a uma proteína de SEQ ID NO:14, que compreende o contacto de um anticorpo descrito com a proteína e com outra citoquina e a comparação da ligação do anticorpo à proteína e à citoquina.

DESCRIÇÃO DETALHADA DAS CONCRETIZAÇÕES PREFERIDAS

I. Geral

São aqui proporcionadas sequências de ADN codificando várias proteínas de mamíferos que apresentam particularidades estruturais características das citoquinas, particularmente relacionadas com a citoquina designada por CTLA-8 (também designada por IL-17). As formas de rato, de ratinho e humana e um homólogo viral da citoquina CTLA-8 foram descritos e as suas sequências estão disponíveis no GenBank. Ver Rouvier *et al.* (1993) *J. Immunol.* 150:5445-5456; Yao *et al.* (1995) *Immunity* 3:811-821; Yao *et al.* (1995) *J. Immunol.* 155:5483-5486; e Kennedy *et al.* (1996) *J. Interferon and Cytokine Res.* 16:611-617. A CTLA-8 possui actividades implicadas na artrite, na rejeição de enxertos renais, na tumorigenicidade, nas interações vírus-hospedeiro e na imunidade inata e parece apresentar determinadas funções reguladoras semelhantes a IL-6. Consultar PubMed (pesquisar IL-17); Chabaud *et al.* (1998) *J. Immunol.* 63:139-148; Amin *et al.* (1998) *Curr. Opin. Rheumatol.* 10:263-268; Van Kooten *et al.* (1998) *J. Am. Soc. Nephrol.* 9:1526-1534; Fossiez *et al.* (1998) *Int. Rev. Immunol.* 16:541-551; Knappe *et al.* (1998) *J. Virol.* 72:5797-5801; Seow (1998) *Vet. Immuno. Immunopathol.* 63:139-48; e Teunissen *et al.* (1998) *J. Invest. Dermatol.* 111:645-649. Uma comunicação sobre a sinalização através do factor de transcrição NF κ B implica uma via de sinal que é utilizada na imunidade inata. Shalom-Barak *et al.* (1998) *J. Biol. Chem.* 273:27467-27473.

As sequências de ADN acabadas de apresentar exibem várias particularidades que são características dos ARNm que codificam as citoquinas, os factores de crescimento e os oncogenes. Como a IL-17 é o primeiro membro desta família acabada de reconhecer de citoquinas relacionadas com TGF- β , os requerentes designaram a família por IL-170, com os novos membros IL-172, IL-173, IL-174, IL-176, IL-177 e IL-171 e IL-175. WO 99/61617 descreve polinucleótidos que codificam polipéptidos possuindo homologia com IL-174 e IL-173. WO 99/60127 descreve um polinucleótido que codifica um polipéptido possuindo homologia com IL-172. Prevê-se que o enrolamento para esta família seja o da família TGF- β de citoquinas. A família TGF- β de citoquinas e a família IL-170 partilham a característica comum de um motivo nó de cistina,

caracterizado por um espaçamento particular dos resíduos de cisteína. Ver, por exemplo, Sun & Davies (1995) *Ann. Rev. Biophys. Biomolec. Struct.* 24:269-291; McDonald *et al.* (1993) *Cell* 73:421-424; e Isaacs (1995) *Curr. Op. Struct. Biol.* 5: 391-395. Em particular, as estruturas sugerem um certo número de cisteínas conservadas que correspondem, e estão numeradas, na IL-172 humana (SEQ ID NO:2), a cisteínas nas posições 101, 103, 143, 156 e 158. A primeira cisteína corresponde à posição na Tabela 7 da val19 na IL-172 humana (SEQ ID NO:2). A quarta cisteína corresponde à posição da cys141 na IL-172 de ratinho (SEQ ID NO:4); da cys119 na IL-173 humana (SEQ ID NO:6); da cys104 na IL-174 de ratinho (SEQ ID NO:16); e da cys50 na IL-171 humana (SEQ ID NO:21). As ligações dissulfureto deverão envolver as cisteínas 2 com 5; 3 com 6 e 1 com 4. O significado funcional da similaridade do enrolamento sugere a formação de dímeros para a família IL-170. Em consequência disso, os dímeros IL-170 aproximariam dois receptores de superfície celular, através dos quais a transdução de sinal terá lugar.

Estas novas proteínas são designadas por proteínas relacionadas com CTLA-8 ou, de um modo geral, proteínas IL-170. As proteínas naturais deverão ser capazes de mediar várias respostas fisiológicas que conduziriam a respostas biológicas ou fisiológicas em células-alvo, por exemplo, aquelas respostas características da sinalização por citocinas. Estudos iniciais haviam localizado a mensagem que codifica esta proteína em várias linhas celulares de células hematopoiéticas. Os genes que codificam o antigénio CTLA-8 (IL-17) original foram mapeados no cromossoma 1A de ratinho e no cromossoma 2q31 humano. A proteína CTLA-8 murina foi originalmente clonada por Rouvier *et al.* (1993) *J. Immunol.* 150:5445-5456. A IL-173 humana foi mapeada no cromossoma 13q11. Sequências similares para proteínas em outras espécies de mamíferos deverão também estar disponíveis.

A proteína CTLA-8 purificada, quando cultivada com sinoviócitos, é capaz de induzir a secreção de IL-6 a partir destas células. Esta indução é revertida com a adição de um anticorpo neutralizante produzido contra a CTLA-8 humana. As células endoteliais, epiteliais, de fibroblastos e de carcinomas também exibem respostas ao tratamento com CTLA-8.

Estes resultados sugerem que CTLA-8 poderá estar implicada na fibrose inflamatória, por exemplo, na psoríase, na esclerodermia, na fibrose pulmonar ou na cirrose. A CTLA-8 também poderá causar a proliferação de carcinomas ou de outras células cancerosas, na medida em que IL-6 actua frequentemente como um factor de crescimento para essas células. Como tal, é provável que os membros acabados de encontrar desta família relacionada possuam actividades biológicas similares ou relacionadas.

As descrições abaixo estão orientadas, para fins exemplificativos, para as proteínas IL-174 murina ou humana, mas são igualmente aplicáveis a concretizações relacionadas de outras espécies.

II. Ácidos nucleicos

As Tabelas 1-6 descrevem as sequências nucleotídicas e de aminoácidos de várias sequências de novos membros da família IL-170. As sequências nucleotídicas descritas e os reagentes relacionados são vantajosos para a construção de clones de ADN adequados para estender os clones em ambas as direcções, com o objectivo de determinar a sequência de comprimento total ou as sequências de flanco; para a expressão de polipéptidos IL-170; ou, por exemplo, para o isolamento de um gene homólogo a partir de outra origem natural. Tipicamente, as sequências serão úteis no isolamento de outros genes, por exemplo, variantes alélicas de ratinho, e aplicar-se-ão procedimentos similares para isolar genes de outras espécies, por exemplo, animais de sangue quente, tais como pássaros e mamíferos. A hibridação cruzada permitirá o isolamento de genes de outras espécies. Um determinado número de diferentes abordagens deverá estar disponível para isolar com êxito um clone de ácido nucleico adequado de outras origens.

Tabela 1: Sequência nucleotídica que codifica o polipéptido IL-172 de um primata, por exemplo, o ser humano, e sequência de aminoácidos prevista. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. O local de clivagem previsto do péptido de sinal está indicado, mas poderá estar deslocado alguns resíduos para qualquer um dos lados; local de glicosilação putativo nos resíduos 55-57. SEQ ID NO:1 e 2.

ATG GAC TGG CCT CAC AAC CTG CTG TTT CTT CTT ACC ATT TCC ATC TTC Met Asp Trp Pro His Asn Leu Leu Phe Leu Leu Thr Ile Ser Ile Phe -20 -15 -10 -5	48
CTG GGG CTG GGC CAG CCC AGG AGC CCC AAA AGC AAG AGG AAG GGG CAA Leu Gly Leu Gly Gln Pro Arg Ser Pro Lys Ser Lys Arg Lys Gly Gln 1 5 10	96
GGG CGG CCT GGG CCC CTG GTC CCT GGC CCT CAC CAG GTG CCA CTG GAC Gly Arg Pro Gly Pro Leu Val Pro Gly Pro His Gln Val Pro Leu Asp 15 20 25	144
CTG GTG TCA CGG ATG AAA CCG TAT GCC CGC ATG GAG GAG TAT GAG AGG Leu Val Ser Arg Met Lys Pro Tyr Ala Arg Met Glu Glu Tyr Glu Arg 30 35 40	192
AAC ATC GAG GAG ATG GTG GCC CAG CTG AGG AAC AGC TCA GAG CTG GCC Asn Ile Glu Glu Met Val Ala Gln Leu Arg Asn Ser Ser Glu Leu Ala 45 50 55 60	240
CAG AGA AAG TGT GAG GTC AAC TTG CAG CTG TGG ATG TCC AAC AAG AGG Gln Arg Lys Cys Glu Val Asn Leu Gln Leu Trp Met Ser Asn Lys Arg 65 70 75	288
AGC CTG TCT CCC TGG GGC TAC AGC ATC AAC CAC GAC CCC AGC CGT ATC Ser Leu Ser Pro Trp Gly Tyr Ser Ile Asn His Asp Pro Ser Arg Ile 80 85 90	336
CCC GTG GAC CTG CCG GAG GCA CGG TGC CTG TGT CTG GGC TGT GTG AAC Pro Val Asp Leu Pro Glu Ala Arg Cys Leu Cys Leu Gly Cys Val Asn 95 100 105	384
CCC TTC ACC ATG CAG GAG GAC CGC AGC ATG GTG AGC GTG CCG GTG TTC Pro Phe Thr Met Gln Glu Asp Arg Ser Met Val Ser Val Pro Val Phe 110 115 120	432
AGC CAG GTT CCT GTG CGC CGC CGC CTC TGC CCG CCA CCG CCC CGC ACA Ser Gln Val Pro Val Arg Arg Arg Leu Cys Pro Pro Pro Pro Arg Thr 125 130 135 140	480
GGG CCT TGC CGC CAG CGC GCA GTC ATG GAG ACC ATC GCT GTG GGC TGC Gly Pro Cys Arg Gln Arg Ala Val Met Glu Thr Ile Ala Val Gly Cys 145 150 155	528
ACC TGC ATC TTC TGA Thr Cys Ile Phe 160	543

MDWPHNLLFLLTISIFLGLG QPRSPKSKRKGGQGRGPLVFGPHQVPLDLVSRMKPYARMEEYERN
IEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVDLPEARCLCLGCVNPFT
MQEDRSMSVSPVFSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Os segmentos particularmente interessantes incluem, por exemplo, aqueles que começam ou terminam com gln1, val19, pro20, pro22, lys34, pro35, leu78, ser79, glu98, ala99, phe110, thr111, cys143 ou arg144.

Sequência nucleotídica que codifica o polipéptido IL-172 de um roedor, por exemplo, o ratinho, e sequência de aminoácidos prevista. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. O local previsto de clivagem do péptido de sinal está indicado, mas poderá estar deslocado alguns resíduos para qualquer um dos lados; local de glicosilação putativo nos resíduos 53-55. SEQ ID NO:3 e 4.

ATG GAC TGG CCG CAC AGC CTG CTC TTC CTC CTG GCC ATC TCC ATC TTC	48
Met Asp Trp Pro His Ser Leu Leu Phe Leu Leu Ala Ile Ser Ile Phe	
-22 -20 -15 -10	
CTG GCG CCA AGC CAC CCC CGG AAC ACC AAA GGC AAA AGA AAA GGG CAA	96
Leu Ala Pro Ser His Pro Arg Asn Thr Lys Gly Lys Arg Lys Gly Gln	
-5 1 5 10	
GGG AGG CCC AGT CCC TTG GCC CCT GGG CCT CAT CAG GTG CCG CTG GAC	144
Gly Arg Pro Ser Pro Leu Ala Pro Gly Pro His Gln Val Pro Leu Asp	
15 20 25	
CTG GTG TCT CGA GTA AAG CCC TAC GCT CGA ATG GAA GAG TAT GAG CGG	192
Leu Val Ser Arg Val Lys Pro Tyr Ala Arg Met Glu Glu Tyr Glu Arg	
30 35 40	
AAC CTT GGG GAG ATG GTG GCC CAG CTG AGG AAC AGC TCC GAG CCA GCC	240
Asn Leu Gly Glu Met Val Ala Gln Leu Arg Asn Ser Ser Glu Pro Ala	
45 50 55	
AAG AAG AAA TGT GAA GTC AAT CTA CAG CTG TGG TTG TCC AAC AAG AGG	288
Lys Lys Lys Cys Glu Val Asn Leu Gln Leu Trp Leu Ser Asn Lys Arg	
60 65 70	
AGC CTG TCC CCA TGG GGC TAC AGC ATC AAC CAC GAC CCC AGC CGC ATC	336
Ser Leu Ser Pro Trp Gly Tyr Ser Ile Asn His Asp Pro Ser Arg Ile	
75 80 85 90	
CCT GCG GAC TTG CCC GAG GCG CGG TGC CTA TGT TTG GGT TGC GTG AAT	384
Pro Ala Asp Leu Pro Glu Ala Arg Cys Leu Cys Leu Gly Cys Val Asn	
95 100 105	
CCC TTC ACC ATG CAG GAG GAC CGT AGC ATG GTG AGC GTG CCA GTG TTC	432
Pro Phe Thr Met Gln Glu Asp Arg Ser Met Val Ser Val Pro Val Phe	
110 115 120	
AGC CAG GTG CCG GTG CGC CGC CGC CTC TGT CCT CAA CCT CCT CGC CCT	480
Ser Gln Val Pro Val Arg Arg Arg Leu Cys Pro Gln Pro Pro Arg Pro	
125 130 135	
GGG CCC TGC CGC CAG CGT GTC GTC ATG GAG ACC ATC GCT GTG GGT TGC	528
Gly Pro Cys Arg Gln Arg Val Val Met Glu Thr Ile Ala Val Gly Cys	
140 145 150	
ACC TGC ATC TTC TGA	543
Thr Cys Ile Phe	
155	
MDWPHSLFLLAISIFLAPSHP RNTKGKRKGQGRPSPLAPGPHQVPLDLVSRVKPYARMEEYERN	
LGEMVAQLRNSSEPAAKKCEVNLQLWLSNKRSLSPWGYSINHDPSRIPADLPEARCLCLGCVNPFT	
MQEDRSMVSVFVSQVPVRRRLCPQPPRPQPCRQVRVMMETIAVGCTCIF	

Os segmentos particularmente interessantes incluem, por exemplo, aqueles que começam ou terminam com arg1, ala17,

pro18, pro20, his21, lys32, pro33, leu76, ser77, glu96, ala97, phe108, thr109, cys141 ou arg142.

Tabela 2: Sequência nucleotídica que codifica o polipéptido IL-173 de um primata, por exemplo, o ser humano, e sequência de aminoácidos prevista. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:5 e 6.

TGC GCG GAC CGG CCG GAG GAG CTA CTG GAG CAG CTG TAC GGG CGC CTG	48
Cys Ala Asp Arg Pro Glu Glu Leu Leu Glu Gln Leu Tyr Gly Arg Leu	
1 5 10 15	
GCG GCC GGC GTG CTC AGT GCC TTC CAC CAC ACG CTG CAG CTG GGG CCG	96
Ala Ala Gly Val Leu Ser Ala Phe His Thr Leu Gln Leu Gly Pro	
20 25 30	
CGT GAG CAG GCG CGC AAC GCG AGC TGC CCG GCA GGG GGC AGG CCC GCC	144
Arg Glu Gln Ala Arg Asn Ala Ser Cys Pro Ala Gly Gly Arg Pro Ala	
35 40 45	
GAC CGC CGC TTC CGG ACG CCC ACC AAC CTG CGC AGC GTG TCG CCC TGG	192
Asp Arg Arg Phe Arg Thr Pro Thr Asn Leu Arg Ser Val Ser Pro Trp	
50 55 60	
GCC TAC AGA ATC TCC TAC GAC CCG GCG AGG TAC CCC AGG TAC CTG CCT	240
Ala Tyr Arg Ile Ser Tyr Asp Pro Ala Arg Tyr Pro Arg Tyr Leu Pro	
65 70 75 80	
GAA GCC TAC TGC CTG TGC CGG GGC TGC CTG ACC GGG CTG TTC GGC GAG	288
Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu Thr Gly Leu Phe Gly Glu	
85 90 95	
GAG GAC GTG CGC TTC CGC AGC GCC CCT GTC TAC ATG CCC ACC GTC GTC	336
Glu Asp Val Arg Phe Arg Ser Ala Pro Val Tyr Met Pro Thr Val Val	
100 105 110	
CTG CGC CGC ACC CCC GCC TGC GCC GGC GGC CGT TCC GTC TAC ACC GAG	384
Leu Arg Arg Thr Pro Ala Cys Ala Gly Gly Arg Ser Val Tyr Thr Glu	
115 120 125	
GCC TAC GTC ACC ATC CCC GTG GGC TGC ACC TGC GTC CCC GAG CCG GAG	432
Ala Tyr Val Thr Ile Pro Val Gly Cys Thr Cys Val Pro Glu Pro Glu	
130 135 140	
AAG GAC GCA GAC AGC ATC AAC T	454
Lys Asp Ala Asp Ser Ile Asn	
145 150	
CADRPEELLEQLYGRLAAGVLSAFHHTLQLGPREQARNASCPAGGRPADDRFRTPTNLRS	
VSPWAYRISYDPARYPRYLPEAYCLCRGCLTGLFGCEEDVRFRSAPVYMPTVVLRRTPACA	
GGRSVYTEAYVTIPVGCTCVPEPEKDADSIN	

Sequência nucleotídica suplementar que codifica o polipéptido IL-173 de um primata, por exemplo, o ser humano, e sequência de aminoácidos prevista. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:7 e 8.

```

gcccgggcag gtggcgacct cgctcagtcg gcttctcggt ccaagtcccc gggtctgg 58
atg ctg gta gcc ggc ttc ctg ctg gcg ctg ccg ccg agc tgg gcc gcg 106
Met Leu Val Ala Gly Phe Leu Leu Ala Leu Pro Pro Ser Trp Ala Ala
-15 -10 -5
ggc gcc ccg agg gcg ggc agg cgc ccc gcg ccg ccg ccg ggc tgc gcg 154
Gly Ala Pro Arg Ala Gly Arg Arg Pro Ala Arg Pro Arg Gly Cys Ala
-1 1 5 10 15
gac ccg ccg gag gag cta ctg gag cag ctg tac ggg cgc ctg gcg gcc 202
Asp Arg Pro Glu Glu Leu Leu Glu Gln Leu Tyr Gly Arg Leu Ala Ala
20 25 30
ggc gtg ctc agt gcc ttc cac cac acg ctg cag ctg ggg ccg cgt gag 250
Gly Val Leu Ser Ala Phe His His Thr Leu Gln Leu Gly Pro Arg Glu
35 40 45
cag gcg cgc aac gcg agc tgc ccg gca ggg ggc agg ccc gcc gac cgc 298
Gln Ala Arg Asn Ala Ser Cys Pro Ala Gly Gly Arg Pro Ala Asp Arg
50 55 60
cgc ttc ccg ccg ccc acc aac ctg cgc agc gtg tgc ccc tgg gcc tac 346
Arg Phe Arg Pro Pro Thr Asn Leu Arg Ser Val Ser Pro Trp Ala Tyr
65 70 75
aga atc tcc tac gac ccg gcg agg tac ccc agg tac ctg cct gaa gcc 394
Arg Ile Ser Tyr Asp Pro Ala Arg Tyr Pro Arg Tyr Leu Pro Glu Ala
80 85 90 95
tac tgc ctg tgc ccg ggc tgc ctg acc ggg ctg ttc ggc gag gag gac 442
Tyr Cys Leu Cys Arg Gly Cys Leu Thr Gly Leu Phe Gly Glu Glu Asp
100 105 110
gtg cgc ttc ccg agc gcc cct gtc tac atg ccc acc gtc gtc ctg cgc 490
Val Arg Phe Arg Ser Ala Pro Val Tyr Met Pro Thr Val Val Leu Arg
115 120 125

```

```

cgc acc ccc gcc tgc gcc ggc ggc cgt tcc gtc tac acc gag gcc tac 538
Arg Thr Pro Ala Cys Ala Gly Arg Ser Val Tyr Thr Glu Ala Tyr
      130      135      140

gtc acc atc ccc gtg ggc tgc acc tgc gtc ccc gag ccg gag aag gac 586
Val Thr Ile Pro Val Gly Cys Thr Cys Val Pro Glu Pro Glu Lys Asp
      145      150      155

gca gac agc atc aac tcc agc atc gac aaa cag ggc gcc aag ctc ctg 634
Ala Asp Ser Ile Asn Ser Ser Ile Asp Lys Gln Gly Ala Lys Leu Leu
      160      165      170      175

ctg ggc ccc aac gac gcg ccc gct ggc ccc tgaggccggt cctgccccgg 684
Leu Gly Pro Asn Asp Ala Pro Ala Gly Pro
      180      185

gaggtctccc cgccccgcat cccgaggcgc ccaagctgga gccgcctgga gggctcggtc 744

ggcgacctct gaagagagtg caccgagcaa accaagtgcc ggagcaccag cgccgctttt 804

ccatggagac tcgtaagcag cttcatctga caccgggcatc cctggcttgc ttttagctac 864

aagcaagcag cgtggctgga agctgatggg aaacgaccgg gcacgggcat cctgtgtgcg 924

gccccgcatgg agggtttggg aaagttcacg gaggtccct gaggagcctc tcagatcggc 984

tgctgcgggt gcagggcggt actcacgctt ggggtgcttgc caaagagata gggacgcata 1044

tgcttttttaa agcaatctaa aaataataat aagtatagcg actatatacc tactttttaa 1104

atcaactgtt ttgaatagag gcagagctat tttatattat caaatgagag ctactctgtt 1164

acatttctta acatataaac atcgtttttt acttcttctg gtagaatttt ttaaagcata 1224

attggaatcc ttggataaat tttgtagctg gtacactctg gcctgggtct ctgaattcag 1284

cctgtcaccg atggctgact gatgaaatgg acacgtctca tctgaccac tcttccttcc 1344

actgaaggtc ttcacgggcc tccaggcctc gtgccgaatt c 1385

MLVAGFLLALPPSWAAGAPRAGRPRPARPRGCADRPEELLEQLYGRLAGVLSAFHHTLQLGPREQARNA
SCPAGGRPADRRFRPPTNLRVSPWAYRISYDPARYPRYLPEAYCLRCGLTGGLFGEEDVFRFSAPVYM
PTVVLRRTPACAGGRSVYTEAYVTIPVGCTCVPEPEKDADSINSSIDKQAKLLLPNDAPAGP

```

Os motivos previstos importantes incluem, por exemplo, cAMP-PK nas posições 50-53, 66-69, 72-75 e 113-116; fosforilação Ca nas posições 82-84 e 166-168; locais miristolilo nas posições 57-61 e 164-166; e locais de fosforilação nas posições 50, 53, 72, 75, 80, 82, 113 e 116.

Sequência nucleotídica que codifica o polipéptido IL-173 de um roedor, por exemplo, o rato, e sequência de aminoácidos prevista. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:9 e 10.

```

TTT CCG AGA TAC CTG CCC GAA GCC TAC TGC CTG TGC CGA GGC TGT CTG 48
Phe Pro Arg Tyr Leu Pro Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu
  1      5      10      15

ACC GGG CTC TAC GGT GAG GAG GAC TTC CGC TTT CGC AGC GCA CCC GTC 96
Thr Gly Leu Tyr Gly Glu Glu Asp Phe Arg Phe Arg Ser Ala Pro Val
      20      25      30

TTC TCT CCG GCG GTG GTG CTG CGG CGC ACG GCG GCC T 133
Phe Ser Pro Ala Val Val Leu Arg Arg Thr Ala Ala
      35      40

FPRYLPEAYCLRCGLTGGLYGEEDFRFRSAPVFSAPVVLRRATA

```

Sequência nucleotídica suplementar que codifica o polipéptido IL-173 de um roedor, por exemplo, o ratinho, e sequência de aminoácidos prevista. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:11 e 12.

```

atg ttg ggg aca ctg gtc tgg atg ctc ctc gtc ggc ttc ctg ctg gca 48
Met Leu Gly Thr Leu Val Trp Met Leu Leu Val Gly Phe Leu Leu Ala
-20 -15 -10

ctg gcg cgc ggc cgc gcg gcg ggc gcg ctg agg acc ggg agg cgc ccg 96
Leu Ala Pro Gly Arg Ala Ala Gly Ala Leu Arg Thr Gly Arg Arg Pro
-5 -1 1 5

gcg cgc cgc cgc gac tgc gcg gac cgc cca gag gag ctc ctg gag cag 144
Ala Arg Pro Arg Asp Cys Ala Asp Arg Pro Glu Glu Leu Leu Glu Gln
10 15 20

ctg tac ggg cgc ctg gcg gcc ggc gtg ctc agc gcc ttc cac cac acg 192
Leu Tyr Gly Arg Leu Ala Ala Gly Val Leu Ser Ala Phe His His Thr
25 30 35 40

ctg cag ctc ggg ccg cgc gag cag gcg cgc aat gcc agc tgc ccg gcc 240
Leu Gln Leu Gly Pro Arg Glu Gln Ala Arg Asn Ala Ser Cys Pro Ala
45 50 55

ggg ggc agg gcc gcc gac cgc cgc ttc cgc cca ccc acc aac ctg cgc 288
Gly Gly Arg Ala Ala Asp Arg Arg Phe Arg Pro Pro Thr Asn Leu Arg
60 65 70

agc gtg tgc ccc tgg gcg tac agg att tcc tac gac cct gct cgc ttt 336
Ser Val Ser Pro Trp Ala Tyr Arg Ile Ser Tyr Asp Pro Ala Arg Phe
75 80 85

ccg agg tac ctg ccc gaa gcc tac tgc ctg tgc cga ggc tgc ctg acc 384
Pro Arg Tyr Leu Pro Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu Thr
90 95 100

ggg ctc tac ggg gag gag gac ttc cgc ttt cgc agc aca ccc gtc ttc 432
Gly Leu Tyr Gly Glu Glu Asp Phe Arg Phe Arg Ser Thr Pro Val Phe
105 110 115 120

tct cca gcc gtg gtg ctg cgc cgc aca gcg gcc tgc gcg ggc ggc cgc 480
Ser Pro Ala Val Val Leu Arg Arg Thr Ala Ala Cys Ala Gly Gly Arg
125 130 135

tct gtg tac gcc gaa cac tac atc acc atc ccg gtg ggc tgc acc tgc 528
Ser Val Tyr Ala Glu His Tyr Ile Thr Ile Pro Val Gly Cys Thr Cys
140 145 150

gtg ccc gag ccg gac aag tcc gcg gac agt gcg aac tcc agc atg gac 576
Val Pro Glu Pro Asp Lys Ser Ala Asp Ser Ala Asn Ser Ser Met Asp
155 160 165

aag ctg ctg ctg ggg ccc gcc gac agg cct gcg ggc cgc tgatgccggg 625
Lys Leu Leu Leu Gly Pro Ala Asp Arg Pro Ala Gly Arg
170 175 180

gactgccgc catggcccag ctctctgcat gcatcaggtc ccttgccct gacaaaaccc 685
accccatgat ccttgccgc tgccataattt ttccaaaagg acagctacat aagctttaa 745
tatatttttc aaagtagaca ctacatatct acaactattt tgaatagtgg cagaaactat 805
tttcatatta gtaatttaga gcaagcatgt tgttttttaa cttctttgat atacaagcac 865
atcacacaca tcccgttttc ctctagtagg attcttgagt gcataattgt agtgctcaga 925
tgaacttcct tctgctgcac tgtgccctgt cctgagtcct ctctgtggc ccaagcttac 985
taagggtgata atgagtgctc cggatctggg cacctaaggt ctccaggtcc ctggagaggg 1045

```

```
agggatgtgg gggggctagg aaccaagcgc ccctttgttc tttagcttat ggatggtctt 1105  
aacctttataa agattaaagt ttttgggtgtt attcttttc 1143
```

```
MLGTLVWMLLVGFLALAPGRAAGALRTGRRPARPRDCADRPEELLEQLYGRLAAGVLSAFHHTLQLGPRE  
QARNASCPAGGRAADRRFRPPTNLRSVSPWAYRISYDPARFPRYLPEAYCLCRGCLTGLYGEEDFRFRSTP  
VFSPAVVLRRTAACAGGRSVYAEHYITIPVGCTCVPEPDKSADSANSMDKLLGPADRPAGR.
```

Os motivos previstos importantes incluem, por exemplo, locais cAMP-PK nas posições 50-53, 66-69, 72-75 e 113-116; locais de fosforilação Ca nas posições 82-84, 159-161 e 166-168; locais miristolilo nas posições 57-61 e 101-105; locais N-glicosilo nas posições 51-53 e 164-166; locais de fosforilação nas posições 50, 53, 72, 75, 80, 82, 113 e 116; e locais de fosforilação PKC nas posições 4-6

Tabela 3: Sequência nucleotídica que codifica o polipeptídeo IL-174 de um primata, por exemplo, o ser humano, e sequência de aminoácidos prevista. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:13 e 14.

tgagtgtgca	gtgccagc	atg	tac	cag	gtg	gtt	gca	ttc	ttg	gca	atg	gtc	51
		Met	Tyr	Gln	Val	Val	Ala	Phe	Leu	Ala	Met	Val	
		-15						-10					
atg	gga	acc	cac	acc	tac	agc	cac	tgg	ccc	agc	tgc	tgc	99
Met	Gly	Thr	His	Thr	Tyr	Ser	His	Trp	Pro	Ser	Cys	Cys	
-5			-1	1				5				10	
ggg	cag	gac	acc	tct	gag	gag	ctg	ctg	agg	tgg	agc	act	147
Gly	Gln	Asp	Thr	Ser	Glu	Glu	Leu	Leu	Arg	Trp	Ser	Thr	
		15					20					25	
cct	ccc	cta	gag	cct	gct	agg	ccc	aac	cgc	cac	cca	gag	195
Pro	Pro	Leu	Glu	Pro	Ala	Arg	Pro	Asn	Arg	His	Pro	Glu	
		30					35				40	Ser	
gcc	agt	gaa	gat	gga	ccc	ctc	aac	agc	agg	gcc	atc	tcc	243
Ala	Ser	Glu	Asp	Gly	Pro	Leu	Asn	Ser	Arg	Ala	Ile	Ser	
	45					50					55	Pro	
tat	gag	ttg	gac	aga	gac	ttg	aac	cgg	ctc	ccc	cag	gac	291
Tyr	Glu	Leu	Asp	Arg	Asp	Leu	Asn	Arg	Leu	Pro	Gln	Asp	
60					65				70			Leu	
												Tyr	75
gcc	cgt	tgc	ctg	tgc	ccg	cac	tgc	gtc	agc	cta	cag	aca	339
Ala	Arg	Cys	Leu	Cys	Pro	His	Cys	Val	Ser	Leu	Gln	Thr	
				80					85			Gly	
												Ser	90
atg	gac	ccc	cgg	ggc	aac	tgc	gag	ctg	ctc	tac	cac	aac	387
Met	Asp	Pro	Arg	Gly	Asn	Ser	Glu	Leu	Leu	Tyr	His	Asn	
		95					100					Gln	
												Thr	105
ttc	tac	cgg	cgg	cca	tgc	cat	ggc	gag	aag	ggc	acc	cac	435
Phe	Tyr	Arg	Arg	Pro	Cys	His	Gly	Glu	Lys	Gly	Thr	His	
		110					115					Lys	
												Gly	120
												Tyr	
tgc	ctg	gag	cgc	agg	ctg	tac	cgt	gtt	tcc	tta	gct	tgt	483
Cys	Leu	Glu	Arg	Arg	Leu	Tyr	Arg	Val	Ser	Leu	Ala	Cys	
	125					130					135	Val	
												Cys	Val
cgg	ccc	cgt	gtg	atg	ggc	tag							504
Arg	Pro	Arg	Val	Met	Gly								
140					145								

MYQVVAFLAMVMGTHTYSHWPSCCPKGGDTSEELLRWSTVPVPPLEPARPNRHPESCRASEDGPL
NSRAISFPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNSELLYHNQTVFYRRPCHGE
KGTHKGYCLERRLYRVSLACVCVRPRVMG

Os motivos previstos importantes incluem, por exemplo, locais CAMP-PK nas posições 21-24, 53-56 e 95-98; locais de fosforilação Ca nas posições 15-17, 16-18 e 45-47; locais miristolilo nas posições 12-16, 115-119 e 118-122; um local N-glicosilo nas posições 104-107; locais de fosforilação nas posições 21, 23, 43, 53, 56, 95, 98 e 131; locais de fosforilação PKC nas posições 41-43 e 119-121; e um local de tirosina-quinase nas posições 95-102.

Sequência nucleotídica que codifica o polipeptídeo IL-174 de um roedor, por exemplo, o ratinho, e sequência de aminoácidos prevista. Também é possível utilizar sequências de

ácidos nucleicos complementares para muitos fins. SEQ ID NO:15 e 16.

CGG CAC AGG CGG CAC AAA GCC CGG AGA GTG GCT GAA GTG GAG CTC TGC Arg His Arg Arg His Lys Ala Arg Arg Val Ala Glu Val Glu Leu Cys 1 5 10 15	48
ATC TGT ATC CCC CCC AGA GCC TCT GAG CCA CAC CCA CCA CGC AGA ATC Ile Cys Ile Pro Pro Arg Ala Ser Glu Pro His Pro Pro Arg Arg Ile 20 25 30	96
CTG CAG GGC CAG CAA GGA TGG CCT CTC AAC AGC AGG GCC ATC TCT CCT Leu Gln Gly Gln Gln Gly Trp Pro Leu Asn Ser Arg Ala Ile Ser Pro 35 40 45	144
TGG AGC TAT GAG TTG GAC AGG GAC TTG AAT CGG GTC CCC CAG GAC TGG Trp Ser Tyr Glu Leu Asp Arg Asp Leu Asn Arg Val Pro Gln Asp Trp 50 55 60	192
TAC CAC GCT CGA TGC CTG TGC CCA CAC TGC GTC ACG CTA CAG ACA GGC Tyr His Ala Arg Cys Leu Cys Pro His Cys Val Thr Leu Gln Thr Gly 65 70 75 80	240
TCC CAC ATG GAC CCG CTG GGC AAC TCC GTC CCA CTT TAC CAC AAC CAG Ser His Met Asp Pro Leu Gly Asn Ser Val Pro Leu Tyr His Asn Gln 85 90 95	288
ACG GTC TTC TAC CGG CGG CCA TGC ATG GCG AGG AAG GTA CCC ATC GCC Thr Val Phe Tyr Arg Arg Pro Cys Met Ala Arg Lys Val Pro Ile Ala 100 105 110	336
GCT ACT GCT TGG AGC GCA GGT CTA CCG AGT CTC CTT GGC TTG TGT GTG Ala Thr Ala Trp Ser Ala Gly Leu Pro Ser Leu Leu Gly Leu Cys Val 115 120 125	384
TGT GCG GCC CCG GGT CAT GGC TTA GTC ATG CTC ACC ATC TGC CTG AGG Cys Ala Ala Pro Gly His Gly Leu Val Met Leu Thr Ile Cys Leu Arg 130 135 140	432
TGAATGCCGG GTGGGAGAGA GGGCCAGGTG TACATCACCT GCCAATGCCG GCCGGGTTCA	492
AGCCTGCAAA GCCTACCTGA AGCAGCAGGT CCCGGGACAG GATGGAGACT TGGGGAGAAA	552
TCTGACTTTT GCACTTTTGT GAGCATTTTG GGAAGAGCAG GTTCGCTTGT GCTGTAGAGA	612
TGCTGTTG	620
RHRRHKARRVAEVELCICIPPRASEPHPPRRILQGQGWPLNSRAISPWSYELDRDLNRVPQDWYHARC LCPHCVTLQTGSHMDPLGNSVFLYHNQTVFYRRPCMARKVPIAATAWSAGLPSSLGLCVCAAPGHGLVM LTICLR	

Sequência nucleotídica suplementar que codifica o polipéptido IL-174 de um roedor, por exemplo, o ratinho, e sequência de aminoácidos prevista. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:17 e 18.

```

atg tac cag gct gtt gca ttc ttg gca atg atc gtg gga acc cac acc 48
Met Tyr Gln Ala Val Ala Phe Leu Ala Met Ile Val Gly Thr His Thr
-15 -10 -5 -1

gtc agc ttg cgg atc cag gag ggc tgc agt cac ttg ccc agc tgc tgc 96
Val Ser Leu Arg Ile Gln Glu Gly Cys Ser His Leu Pro Ser Cys Cys
1 5 10 15

ccc agc aaa gag caa gaa ccc ccg gag gag tgg ctg aag tgg agc tct 144
Pro Ser Lys Glu Gln Glu Pro Pro Glu Glu Trp Leu Lys Trp Ser Ser
20 25 30

gca tct gtg tcc ccc cca gag cct ctg agc cac acc cac cac gca gaa 192
Ala Ser Val Ser Pro Pro Glu Pro Leu Ser His Thr His His Ala Glu
35 40 45

tcc tgc agg gcc agc aag gat ggc ccc ctc aac agc agg gcc atc tct 240
Ser Cys Arg Ala Ser Lys Asp Gly Pro Leu Asn Ser Arg Ala Ile Ser
50 55 60

cct tgg agc tat gag ttg gac agg gac ttg aat cgg gtc ccc cag gac 288
Pro Trp Ser Tyr Glu Leu Asp Arg Asp Leu Asn Arg Val Pro Gln Asp
65 70 75 80

ctg tac cac gct cga tgc ctg tgc cca cac tgc gtc agc cta cag aca 336
Leu Tyr His Ala Arg Cys Leu Cys Pro His Cys Val Ser Leu Gln Thr
85 90 95

ggc tcc cac atg gac ccg ctg ggc aac tcc gtc cca ctt tac cac aac 384
Gly Ser His Met Asp Pro Leu Gly Asn Ser Val Pro Leu Tyr His Asn
100 105 110

cag acg gtc ttc tac cgg cgg cca tgc cat ggt gag gaa ggt acc cat 432
Gln Thr Val Phe Tyr Arg Arg Pro Cys His Gly Glu Glu Gly Thr His
115 120 125

cgc cgc tac tgc ttg gag cgc agg ctc tac cga gtc tcc ttg gct tgt 480
Arg Arg Tyr Cys Leu Glu Arg Arg Leu Tyr Arg Val Ser Leu Ala Cys
130 135 140

gtg tgt gtg cgg ccc cgg gtc atg gct tagtcatgct caccacctgc 527
Val Cys Val Arg Pro Arg Val Met Ala
145 150

ctgaggctga tgcccgggttg ggagagaggg ccaggtgtac aatcaccttg ccaatgctggg 587
ccgggttcaa gccctccaaa gccctacctg aagcagcagg ctcccgggac aagatggagg 647
acttggggag aaactctgac ttttgacttt tttggaagca cttttgggaa ggagcagggtt 707
ccgcttgtgc tgctagagga tgctgttgtg gcattttctac tcaggaacgg actccaaagg 767
cctgctgacc ctggaagcca tactcctggc tcctttcccc tgaatcccc aactcctggc 827
acaggcactt tctccacctc tccccctttg ccttttgttg tgtttgtttg tgcattgcaa 887
ctctgcgtgc agccagggtgt aattgccttg aaggatggtt ctgagggtgaa agctgttatc 947
gaaagtgaag agatttatcc aaataaacat ctgtgttt 985

MYQAVAFAMIVGTHTVSLRIQEGCSHLPSCCPSKEQEPPEEWLKWSSASVSPPEPLSHTHAESCRAS
KDGPLNSRAISPWSYELDRDLNRVPQDLYHARCLCPHCVSLQTGSHMDPLGNSVPLYHNQTVFYRRPCH
GEEGTHRRYCLERRLYRVSLACVVRPRVMA

```

Os motivos previstos importantes incluem, por exemplo, locais cAMP-PK nas posições 29-32 e 61-64; locais de fosforilação Ca nas posições 18-20, 53-55 e 67-69; um local miristolilo nas posições 123-127; um local de N-glicosilação nas posições 112-114; locais de fosforilação nas posições 29, 31, 51, 53, 61, 64, 139 e 141; e locais de fosforilação PKC nas posições 2-4, 49-51 e 127-129.

Tabela 4: Sequência nucleotídica que codifica IL-171 de um primata, por exemplo, o ser humano, no código da IUPAC. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:19.

```
GACACGGATG AGGACCGCTA TCCACAGAAG CTGGCCTTCG CCGAGTGCCT GTGCAGAGGC      60
TGTATCGATG CACGGACGGG CCGCGAGACA GCTGCGCTCA ACTCCGTGCG GCTGCTCCAG      120
AGCCTGCTGG TGCTGCGCCG CCGGCCCTGC TCCCGCGACG GCTCGGGGCT CCCCACACCT      180
GGGGCCTTTG CCTTCCACAC CGAGTTCATC CACGTCCCCG TCGGCTGCAC CTGCGTGCTG      240
CCCCGTTCAA GTGTGACCGC CAAGGCCGTG GGGCCCTTAG NTGACACCGT GTGCTCCCCA      300
GAGGGACCCC TATTTATGGG AATTATGGTA TTATATGCTT CCCACATACT TGGGGCTGGC      360
ATCCCGNGCT GAGACAGCCC CCTGTTCTAT TCAGCTATAT GGGGAGAAGA GTAGACTTTC      420
AGCTAAGTGA AAAGTGNAAC GTGCTGACTG TCTGCTGTCT TNCTACTNAT GCTAGCCCGA      480
GTGTTCACTC TGAGCCTGTT AAATATAGGC GGTATGTAC C                          521
```

As SEQ ID NO:20 e 21 são a sequência de ADNc traduzível e a sequência polipeptídica no software PATENTIN.

```
GAC ACG GAT GAG GAC CGC TAT CCA CAG AAG CTG GCC TTC GCC GAG TGC      48
Asp Thr Asp Glu Asp Arg Tyr Pro Gln Lys Leu Ala Phe Ala Glu Cys
1      5      10      15

CTG TGC AGA GGC TGT ATC GAT GCA CGG ACG GGC CGC GAG ACA GCT GCG      96
Leu Cys Arg Gly Cys Ile Asp Ala Arg Thr Gly Arg Glu Thr Ala Ala
20     25     30

CTC AAC TCC GTG CGG CTG CTC CAG AGC CTG CTG GTG CTG CGC CGC CGG      144
Leu Asn Ser Val Arg Leu Leu Gln Ser Leu Leu Val Leu Arg Arg Arg
35     40     45

CCC TGC TCC CGC GAC GGC TCG GGG CTC CCC ACA CCT GGG GCC TTT GCC      192
Pro Cys Ser Arg Asp Gly Ser Gly Leu Pro Thr Pro Gly Ala Phe Ala
50     55     60

TTC CAC ACC GAG TTC ATC CAC GTC CCC GTC GGC TGC ACC TGC GTG CTG      240
Phe His Thr Glu Phe Ile His Val Pro Val Gly Cys Thr Cys Val Leu
65     70     75     80

CCC CGT TCA AGT GTG ACC GCC AAG GCC GTG GGG CCC TTA Gnt GAC ACC      288
Pro Arg Ser Ser Val Thr Ala Lys Ala Val Gly Pro Leu Xaa Asp Thr
85     90     95

GTG TGC TCC CCA GAG GGA CCC CTA TTT ATG GGA ATT ATG GTA TTA TAT      336
Val Cys Ser Pro Glu Gly Pro Leu Phe Met Gly Ile Met Val Leu Tyr
100    105    110

GCT TCC CAC ATA CTT GGG GCT GGC ATC CCG nGC TGAGACAGCC CCCTGTTCTA      389
Ala Ser His Ile Leu Gly Ala Gly Ile Pro Xaa
115    120

TTCAGCTATA TGGGGAGAAG AGTAGACTTT CAGCTAAGTG AAAAGTGCAA CGTGCTGACT      449

GTCTGCTGTC GTCCTACTCA TGCTAGCCCG AGTGTTCACT CTGAGCCTGT TAAATATAGG      509

CGGTTATGTA CC                          521

DTDED RYPQKLAFACLRCG CIDARTGRETAALNSVRLQLSLVLRRLRRPCSRDGSGLPTPGAFHFTEFI
HVPVGCTCVLPRSSVTAKAVGPLXDTVCSPGGLFMGIMVLYASHILGAGIPX
```

Sequência nucleotídica suplementar que codifica IL-171 de um primata, por exemplo, o ser humano. Também é possível

utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:22 e 23.

```

gtgtggcctc aggtataaga ggggctgctg ccagggtgcat ggccagggtgc acctgtggga 60
ttgccgccag gtgtgcaggc cgctccaagc ccagcctgcc ccgctgccgc cacc atg 117
Met
acg ctc ctc ccc ggc ctc ctg ttt ctg acc tgg ctg cac aca tgc ctg 165
Thr Leu Leu Pro Gly Leu Leu Phe Leu Thr Trp Leu His Thr Cys Leu
-15 -10 -5 -1
gcc cac cat gac ccc tcc ctc agg ggg cac ccc cac agt cac ggt acc 213
Ala His His Asp Pro Ser Leu Arg Gly His Pro His Ser His Gly Thr
1 5 10 15
cca cac tgc tac tcg gct gag gaa ctg ccc ctc ggc cag gcc ccc cca 261
Pro His Cys Tyr Ser Ala Glu Glu Leu Pro Leu Gly Gln Ala Pro Pro
20 25 30
cac ctg ctg gct cga ggt gcc aag tgg ggg cag gct ttg cct gta gcc 309
His Leu Leu Ala Arg Gly Ala Lys Trp Gly Gln Ala Leu Pro Val Ala
35 40 45
ctg gtg tcc agc ctg gag gca gca agc cac agg ggg agg cac gag agg 357
Leu Val Ser Ser Leu Glu Ala Ala Ser His Arg Gly Arg His Glu Arg
50 55 60
ccc tca gct acg acc cag tgc ccg gtg ctg cgg ccg gag gag gtg ttg 405
Pro Ser Ala Thr Thr Gln Cys Pro Val Leu Arg Pro Glu Glu Val Leu
65 70 75 80
gag gca gac acc cac cag cgc tcc atc tca ccc tgg aga tac cgt gtg 453
Glu Ala Asp Thr His Gln Arg Ser Ile Ser Pro Trp Arg Tyr Arg Val
85 90 95
gac acg gat gag gac cgc tat cca cag aag ctg gcc ttc gcc gag tgc 501
Asp Thr Asp Glu Asp Arg Tyr Pro Gln Lys Leu Ala Phe Ala Glu Cys
100 105 110
ctg tgc aga ggc tgt atc gat gca cgg acg ggc cgc gag aca gct gcg 549
Leu Cys Arg Gly Cys Ile Asp Ala Arg Thr Gly Arg Glu Thr Ala Ala
115 120 125
ctc aac tcc gtg cgg ctg ctc cag agc ctg ctg gtg ctg cgc cgc cgg 597
Leu Asn Ser Val Arg Leu Leu Gln Ser Leu Leu Val Leu Arg Arg Arg
130 135 140
ccc tgc tcc cgc gac ggc tcg ggg ctc ccc aca cct ggg gcc ttt gcc 645
Pro Cys Ser Arg Asp Gly Ser Gly Leu Pro Thr Pro Gly Ala Phe Ala
145 150 155 160
ttc cac acc gag ttc atc cac gtc ccc gtc ggc tgc acc tgc gtg ctg 693
Phe His Thr Glu Phe Ile His Val Pro Val Gly Cys Thr Cys Val Leu
165 170 175
ccc cgt tca gtg tgaccgccga ggccgtgggg cccctagact ggacacgtgt 745
Pro Arg Ser Val
180

```

```

gctccccaga gggcaccccc tatttatgtg tatttattgg tatttatatg cctcccccaa 805
cactaccctt ggggtctggg cattccccgt gtctggagga cagcccccca ctgttctcct 865
catctccagc ctcagtagtt gggggtagaa ggagctcagc acctcttcca gcccttaaag 925
ctgcagaaaa ggtgtcacac ggctgcctgt accttggtc cctgtcctgc tcccggttc 985
ccttacccta tcaactggcct caggcccccg caggctgcct ctccccaacc tccttgaag 1045
taccctgtt tcttaaacia ttatttaagt gtacgtgtat tattaaactg atgaacacat 1105
cc

```

1107

```

MTLLPGLLFLTWLHTCLAHHDPSLRGHPHSHGTPHCYSAEELPLGQAPPHLLARGAKWGQALPVALVSS
LEAASHRGRHERPSATTQCPVLRPEEVLEADTHQRSISPWRYRVDTEEDRYPOKLAFAECLCRGCIDAR
TGRETAALNSVRLQLSLLVLRRLPCSRDGSGLPTGAFAFHTEFIHVPVGCTCVLPRSV

```

Tabela 5: Sequência nucleotídica que codifica a sequência de IL-175 de um primata, por exemplo, o ser humano, no código da IUPAC. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:24.

```

GAGAAAGAGC TTCCTGCACA AAGTAAGCCA CCAGCGCAAC ATGACAGTGA AGACCCCTGCA      60
TGGCCCAGCC ATGGTCAAGT ACTTGCTGCT GTCGATATTG GGGCTTGCCT TTCTGAGTGA      120
GGCGGCAGCT CGGAAAATCC CCAAAGTAGG ACATACTTTT TTCCAAAAGC CTGAGAGTTG      180
CCCGCCTGTG CCAGGAGGTA GTATGAAGCT TGACATTGGC ATCATCAATG AAAACCAGCG      240
CGTTTCCATG TCACGTAACA TCGAGAGCCG CTCCACCTCC CCTGGAATT ACACTGTCAC      300
TTGGGACCCC AACCGGTACC CCTCGAAGTT GTACAGGCCC AAGTGTAGGA ACTTGGGCTG      360
TATCAATGCT CAAGGAAAGG AAGACATCTN CATGAATTCC GTC                          403

```

As SEQ ID NO:25 e 26 são a sequência de ADNc e a sequência polipeptídica traduzíveis no software PATENTIN. O local previsto de clivagem do péptido de sinal está indicado, mas poderá estar deslocado alguns resíduos para qualquer um dos lados; local de glicosilação putativo nos resíduos 53-55.

```

GAGAAAGAGC TTCCTGCACA AAGTAAGCCA CCAGCGCAAC ATGACAGTGA AGACCCTGCA      60
TGGCCCAGCC ATG GTC AAG TAC TTG CTG CTG TCG ATA TTG GGG CTT GCC      109
      Met Val Lys Tyr Leu Leu Leu Ser Ile Leu Gly Leu Ala
      -20              -15              -10

TTT CTG AGT GAG GCG GCA GCT CGG AAA ATC CCC AAA GTA GGA CAT ACT      157
Phe Leu Ser Glu Ala Ala Ala Arg Lys Ile Pro Lys Val Gly His Thr
      -5              1              5

TTT TTC CAA AAG CCT GAG AGT TGC CCG CCT GTG CCA GGA GGT AGT ATG      205
Phe Phe Gln Lys Pro Glu Ser Cys Pro Pro Val Pro Gly Gly Ser Met
      10              15              20              25

AAG CTT GAC ATT GGC ATC ATC AAT GAA AAC CAG CGC GTT TCC ATG TCA      253
Lys Leu Asp Ile Gly Ile Ile Asn Glu Asn Gln Arg Val Ser Met Ser
      30              35              40

CGT AAC ATC GAG AGC CGC TCC ACC TCC CCC TGG AAT TAC ACT GTC ACT      301
Arg Asn Ile Glu Ser Arg Ser Thr Ser Pro Trp Asn Tyr Thr Val Thr
      45              50              55

TGG GAC CCC AAC CGG TAC CCC TCG AAG TTG TAC AGG CCC AAG TGT AGG      349
Trp Asp Pro Asn Arg Tyr Pro Ser Lys Leu Tyr Arg Pro Lys Cys Arg
      60              65              70

AAC TTG GGC TGT ATC AAT GCT CAA GGA AAG GAA GAC ATC TCC ATG AAT      397
Asn Leu Gly Cys Ile Asn Ala Gln Gly Lys Glu Asp Ile Ser Met Asn
      75              80              85

TCC GTC
Ser Val
90      403

MVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINENQRVSMsrNIESRST
SPWNYTVTWDPNRYPSKLYRPKCRNLGCINAQGKEDIXMNSV

```

Os segmentos particularmente interessantes incluem, por exemplo, aqueles que começam ou terminam com arg1, cys17, pro18, pro19, val20, thr49, ser50, arg69, pro70 e o final da sequência disponível.

Tabela 6: Sequência nucleotídica que codifica IL-176 de um primata, por exemplo, o ser humano. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:27 e 28:

```

tc gtg ccg tat ctt ttt aaa aaa att att ctt cac ttt ttt gcc tcc 47
  Val Pro Tyr Leu Phe Lys Lys Ile Ile Leu His Phe Phe Ala Ser
    1          5          10          15

tat tac ttg tta ggg aga ccc aat ggt agt ttt att cct tgg gga tac 95
Tyr Tyr Leu Leu Gly Arg Pro Asn Gly Ser Phe Ile Pro Trp Gly Tyr
          20          25          30

ata gta aat act tca tta aag tcg agt aca gaa ttt gat gaa aag tgt 143
Ile Val Asn Thr Ser Leu Lys Ser Ser Thr Glu Phe Asp Glu Lys Cys
          35          40          45

gga tgt gtg gga tgt act gcc gcc ttc aga agt cca cac act gcc tgg 191
Gly Cys Val Gly Cys Thr Ala Ala Phe Arg Ser Pro His Thr Ala Trp
          50          55          60

agg gag aga act gct gtt tat tca ctg att aag cat ttg ctg tgt acc 239
Arg Glu Arg Thr Ala Val Tyr Ser Leu Ile Lys His Leu Leu Cys Thr
          65          70          75

aac tac ttt tca tgt ctt atc tta att ctc ata aca gtc att 281
Asn Tyr Phe Ser Cys Leu Ile Leu Ile Leu Thr Val Ile
          80          85          90

tgatatttta aaaaacccca gaaatctgag aaagagataa agtgggttgc tcaagggtat 341
agaacagact accatgtggt gtatttcaga ttttaattca tgtttgtctg attttaagtt 401
ttgttcgctt gccagggtac cccacaaaaa tgccaggcag ggcattttca tgatgcactt 461
gagatacctg aaatgacagg gtagcatcac acctgagagg ggtaaaggat gggaacctac 521
cttccatggc cgctgcttgg cagtctcttg ctgcatgcta gcagagccac tgtatatgtg 581
ccgaggctct gagaattaac tgcttaaaga actgccttct ggagggagaa gagcacaaga 641
tcacaattaa ccatatacac atcttactgt gcgaggtcat tgagcaatac aggagggatt 701
ttatacattt tagcaactat cttcaaaacc tgagctatag ttgtattctg ccccttctct 761
ctgggcaaaa gtgtaaaagt ttg 784

VPYLFKKIILHFFASYLLGRENGSFIPWGYIVNTSLKSSTEFDEKCGCVGCTAAFRSPHTAWRER
TAVYSLIKHLLCTNYFSCILILITVI

```

Sequência nucleotídica que codifica IL-177 de um primata, por exemplo, o ser humano. Também é possível utilizar sequências de ácidos nucleicos complementares para muitos fins. SEQ ID NO:29 e 30:

```

gtg act gta ttg tgg gga cag gaa gca caa att ccc atg tgg atc act 48
Val Thr Val Leu Trp Gly Gln Glu Ala Gln Ile Pro Met Trp Ile Thr
1 5 10 15

agg aga gat aat aag tgg ggt cat ttc acc cct tgg tcc cct gct tcc 96
Arg Arg Asp Asn Lys Trp Gly His Phe Thr Pro Trp Ser Pro Ala Ser
20 25 30

aga ccc aaa gag gcc tac atg gca ttg tgc ttc ctt ctt agt tgt agg 144
Arg Pro Lys Glu Ala Tyr Met Ala Leu Cys Phe Leu Leu Ser Cys Arg
35 40 45

agg tgt gag ata caa tca ttt gcc tct gac ttt gag ggt tgg tcc 189
Arg Cys Glu Ile Gln Ser Phe Ala Ser Asp Phe Glu Gly Trp Ser
50 55 60

tagcatgccc ctgaccagta gccccttaaa tacttcattg atatggaagg tctctgaatc 249

ttcgtgggct taatctacca ctctctgaag ttcttatgtc tttcaaaggc ctctaaaatc 309

tctgccatgt cttgctcadc cagttgttag catgatgtca ttgatacagt ggactttgga 369

atctaagtgg ggagacactg gtaagtgacc aattacttca cctgtgggtgt gcaagccaga 429

tcaggaagcc tctacctgca cgacaacaca t 460

VTVLWGQEAQIPMWITRRDNKWHGFTPWSPASRPKEAYMALCFLLSRRCEIQSFASDFEGWS

```

Tabela 7: Alinhamento de vários membros da família CTLA-8/IL-170. A sequência de CTLA-8 de rato é a SEQ ID NO:31 (ver GB L13839; 293329/30); a sequência de CTLA-8 de ratinho é a SEQ ID NO:32 (ver GB 1469917/8); a CTLA-8 humana é a SEQ ID NO:33 (ver GB U32659; 115222/3); e a ORF13 do vírus herpes saimiri é a SEQ ID NO:34 (ver GB Y13183; 2370235). Alinhamento de múltiplas sequências utilizando CLUSTAL X (1.64b).

```

IL-74_Mu -----MYQAVAFIAMIVGTHTVSLRI----QEGCSHLPSCCPSKEQEPPEEWLKWS
IL-74_Hu -----MYQVVAFLAMVMGTHTY-----S---HWPSCCPSKQDTSSEELLRWS
IL-72_Hu -----MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLVPGPHQVPLDLVSRMK
IL-72_Mu -----MDWPHSLLFLLAISIFLAPSHPRNTKGKRGQGRPSPLAPGPHQVPLDLVSRVK
IL-73_Mu --MLGTLVWMLLVGFLALALPGRAAGALRT--GRRP--ARPRDCADRPEELLEQLYGRLA
IL-73_Hu -----MLVAGFLLALPPSWAAGAPRA--GRRP--ARPRGCADRPEELLEQLYGRLA
IL-17_Hu --MTPGKTSLSVSLLLLSLEAIVKAGITIP-----RNPGCPSNEDKNFPRTVMVNL
IL-17_Hs --MTFRKTSLV-LLLLLSIDCIVKSEITSA-----QTPRCCLAANN-SFPRSVMTL
IL-17_Rt -----MCLMLLLLNLLEATVKA AVLIP-----QSSVCPNAEANNFLQNVKVN
IL-17_Mu -----MLLLLLSLAATVKA AAIIP-----QSSACPNTAKDFLQNVKVN
IL-75_Hu -----MVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIIN
IL-71_Hu MTLPLGLLFLTWLHTCLAHHDPSLRGHPHSHGTPHCYSAEELPLGQAPPHELLARGAKWGQ
*

IL-74_Mu S-----ASVSPP-EPLSHTHHAES---CRASKD-GPLNSRAISFWSYELDRDLNRV
IL-74_Hu T-----VPVPL-EPARPNRHPES---CRASED-GPLNSRAISFWRVYELDRDLNRL
IL-72_Hu P-YARMEEYERNIEEMVAQLRNSSELAQ-RKCEVNLQLWMSNKRSLSPWGYSINHDPRI
IL-72_Mu P-YARMEEYERNLGEVVAQLRNSSEPAK-KKCEVNLQLWLSNKRSLSPWGYSINHDPRI
IL-73_Mu AGVLSAFHHTLQLGPR-EQARNASCAGGRAADRRFR-PPTNLRSVSPWAYRISYDPAF
IL-73_Hu AGVLSAFHHTLQLGPR-EQARNASCAGGRPADDRFR-PPTNLRSVSPWAYRISYDPAF
IL-17_Hu N-----IHNRNTNTN-----P-KRSSDYNRSTSPWNLRNEDPERY
IL-17_Hs S-----IRNWTSS-----KRASDYNRSTSPWTLHRNEDQDRY
IL-17_Rt K-----VINSLSKA-----SSRRPSDYLRSTSPWTLRNEDPDRY
IL-17_Mu K-----VFNSLGAKV-----SSRRPSDYLRSTSPWTLHRNEDPDRY
IL-75_Hu E-----N---QRVSMS-----R---NIESRSTSPWNYTVTWDPNRY
IL-71_Hu ALPVALVSSLEAAASHRGRHERPSATTQCPLRPPEEVLEADTHQRSTSPWRYVDTDERY
* * *

```

```

IL-74_Mu      PQDLYHARCLCPHCVSLQTGSHMDPLGNSVPLYHNQTVFYRR--PCHGEEGTHRRYCLER
IL-74_Hu      PQDLYHARCLCPHCVSLQTGSHMDPRGNSSELLYHNQTVFYRR--PCHGEKGTHKGYCLER
IL-72_Hu      PVDLPEARCLCLGCVNPFTM-QEDRSMVSVPVFS-QVPVRRR--LCPPPP--RTGPCRQR
IL-72_Mu      PADLPEARCLCLGCVNPFTM-QEDRSMVSVPVFS-QVPVRRR--LCPQPP--RPGPCRQR
IL-73_Mu      PRYLPEAYCLCRGCLTGLYG-EEDFRFRSTPVFS-PAVVLRTAACAG-----GRSVYA
IL-73_Hu      PRYLPEAYCLCRGCLTGLFG-EEDVRFRSAPVYM-PTVVLRTTACAG-----GRSVYT
IL-17_Hu      PSVIWEAKCRHLGCINADGN--VDYHMNSVPIQQEILVLRREPPHCPN-----SFR
IL-17_Hs      PSVIWEAKCRHLGCINADGN--VDYHMNSVPIQQEILVLRREPPHCPN-----SFR
IL-17_Rt      PSVIWEAQRHQRVCVNAEGK--LDHMHNSVLIQQEILVLKREPEKCF-----TFR
IL-17_Mu      PSVIWEAQRHQRVCVNAEGK--LDHMHNSVLIQQEILVLKREPEKCF-----TFR
IL-75_Hu      PSEVVQAQCRNLGCINAQGG--EDISMNSVPIQQETLVVRRKHQGC SV-----SFQ
IL-71_Hu      POKLAFACLCRGCIDARTG-RETAALNSVRLQLSLLVLRRRRPCRSDGSGLPTPGAFAFH
* : * * * :

```



```

IL-74_Mu      RLYR-VSLACVCVRPRVMA-----
IL-74_Hu      RLYR-VSLACVCVRPRVMG-----
IL-72_Hu      AVMETIAGVCTCIF-----
IL-72_Mu      VVMETIAGVCTCIF-----
IL-73_Mu      EHYITIPVGCTCVPEPDKSADSANSSMDK----LLGPA DRPAGR
IL-73_Hu      EAYVTIPVGCTCVPEPEKDADSINSSIDKQAKLLGPNDA PAGP
IL-17_Hu      LEKILVSVGCTCVTPIVHHVA-----
IL-17_Hs      LEKMLVTVGCTCVTPIVHNVD-----
IL-17_Rt      VEKMLVGVGCTCVSSIVRHAS-----
IL-17_Mu      VEKMLVGVGCTCVASIVRQAA-----
IL-75_Hu      LEKVLVTVGCTCVTPVIHHVQ-----
IL-71_Hu      TEFIHVPVGCTCVLPRSV-----
: : * * :

```

Os segmentos particularmente interessantes incluem, por exemplo, aqueles que correspondem aos segmentos de IL-172 ou IL-175, indicados acima, com os outros membros da família.

A proteína ou os polipéptidos purificados são úteis para gerar anticorpos através de métodos padrão, como descrito acima. Os péptidos sintéticos ou a proteína purificada podem ser apresentados a um sistema imunitário para gerar uma composição de ligação específica, por exemplo, anticorpos monoclonais ou policlonais. Ver, por exemplo, Coligan (1991) *Current Protocols in Immunology*, Wiley/Greene; e Harlow & Lane (1989) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press.

Por exemplo, a composição de ligação específica poderia ser utilizada para rastrear uma biblioteca de expressão preparada a partir de uma linha celular que expressa uma proteína IL-174. O rastreio pode ser efectuado por coloração padrão da proteína expressa à superfície ou por "panning". O rastreio da expressão intracelular também pode ser efectuado por meio de vários procedimentos de coloração ou de imunofluorescência. As composições de ligação poderiam ser utilizadas para purificar por afinidade ou para separar as células que expressam a proteína.

Este invento contempla a utilização de ADN isolado ou de fragmentos para codificar uma proteína ou polipéptido IL-174

biologicamente activo correspondente. Adicionalmente, este invento abrange o ADN isolado ou recombinante que codifica uma proteína ou polipéptido biologicamente activo e que é capaz de hibridar, sob condições apropriadas, com as sequências de ADN de IL-174 aqui descritas. A referida proteína ou polipéptido biologicamente activo pode ser um antigénio intacto, ou um fragmento, e possuir uma sequência de aminoácidos tal como descrita na Tabela 3. O ADN isolado pode ter as respectivas sequências reguladoras nos flancos 5' e 3', por exemplo, promotores, amplificadores, sinais de adição poliA e outros.

Um ácido nucleico "isolado" é um ácido nucleico, por exemplo, um ARN, um ADN ou um polímero misto, que está substancialmente separado dos outros componentes que acompanham naturalmente uma sequência nativa, por exemplo, ribossomas, polimerases e sequências genómicas de flanco da espécie progenitora. O termo abrange uma sequência de ácido nucleico que foi removida do seu ambiente natural, e inclui isolados de ADN recombinante ou clonado e análogos quimicamente sintetizados ou análogos biologicamente sintetizados por sistemas heterólogos. Uma molécula substancialmente pura inclui formas isoladas da molécula. Em alternativa, uma espécie purificada poderá ser separada dos componentes do hospedeiro de um sistema de expressão recombinante. A magnitude da homologia de um ácido nucleico destes será tipicamente menor que os vectores grandes, por exemplo, menos de dezenas de kB, tipicamente menos de vários kB e, preferencialmente, na gama de 2-6 kB.

Um ácido nucleico isolado será geralmente uma composição homogénea de moléculas, mas, em algumas concretizações, conterá uma heterogeneidade secundária. Esta heterogeneidade é normalmente encontrada nas extremidades do polímero, ou em partes que não são críticas para uma função ou uma actividade biológica desejadas.

Um ácido nucleico "recombinante" é definido quer pelo seu método de produção quer pela sua estrutura. Relativamente ao seu método de produção, por exemplo, um produto produzido por um processo, o processo é a utilização de técnicas de ácidos nucleicos recombinantes, por exemplo, envolvendo uma intervenção humana na sequência nucleotídica, tipicamente na

selecção ou produção. Em alternativa, ele pode ser um ácido nucleico preparado pela criação de uma sequência que compreende a fusão de dois fragmentos que não são naturalmente contíguos um ao outro, mas que pretende excluir os produtos da natureza, por exemplo, os mutantes naturais. Assim, por exemplo, os produtos preparados por transformação de células com qualquer vector de ocorrência não natural estão abrangidos, assim como o estão os ácidos nucleicos que compreendem uma sequência derivada por meio de qualquer processo oligonucleotídico sintético. Isto é frequentemente executado para substituir um codão por um codão redundante que codifica o mesmo aminoácido ou um aminoácido conservativo, ao mesmo tempo que se introduz ou remove um local de reconhecimento de sequência. Em alternativa, o processo é realizado para unir segmentos de ácidos nucleicos de funções desejadas, com o intuito de gerar uma entidade genética única compreendendo uma combinação desejada de funções que não é encontrada nas formas naturais habitualmente disponíveis. Os locais de reconhecimento de enzimas de restrição são frequentemente o alvo destas manipulações artificiais, embora outros alvos com locais específicos, por exemplo, promotores, locais de replicação do ADN, sequências de regulação, sequências de controlo ou outras funções úteis, possam ser incorporados por concepção. Um conceito similar destina-se a um polipéptido recombinante, por exemplo, de fusão. Estão especificamente incluídos os ácidos nucleicos sintéticos que, por redundância do código genético, codificam polipéptidos similares a fragmentos destes antigénios e fusões de sequências provenientes de variantes de várias espécies diferentes.

Um "fragmento" significativo num contexto de ácidos nucleicos é um segmento contíguo de pelo menos cerca de 17 nucleótidos, geralmente pelo menos 20 nucleótidos, mais geralmente pelo menos 23 nucleótidos, usualmente pelo menos 26 nucleótidos, mais usualmente pelo menos 29 nucleótidos, frequentemente pelo menos 32 nucleótidos, mais frequentemente pelo menos 35 nucleótidos, tipicamente pelo menos 38 nucleótidos, mais tipicamente pelo menos 41 nucleótidos, habitualmente pelo menos 44 nucleótidos, mais habitualmente pelo menos 47 nucleótidos, preferencialmente pelo menos 50 nucleótidos, mais preferencialmente pelo menos

53 nucleótidos e, em concretizações particularmente preferidas, terá pelo menos 56 ou mais nucleótidos. Os referidos fragmentos poderão ter extremidades em qualquer posição, mas especialmente nas fronteiras entre os domínios estruturais.

Em outras concretizações, o invento proporciona polinucleótidos (ou polipéptidos) que compreendem uma pluralidade de segmentos distintos, por exemplo, não sobreponíveis, com o comprimento especificado. Tipicamente, a pluralidade serão pelo menos dois, mais habitualmente pelo menos três e, preferencialmente, 5, 7 ou mesmo mais. Embora sejam proporcionados os comprimentos mínimos, comprimentos maiores, de várias dimensões, poderão ser apropriados, por exemplo, um de comprimento 7 e dois de comprimento 12.

Um ADN que codifica uma proteína IL-174 será particularmente útil para identificar genes, ARNm e espécies de ADNc que codificam proteínas relacionadas ou homólogas, assim como os ADN que codificam proteínas homólogas de diferentes espécies. Há provavelmente homólogos em outras espécies, incluindo os primatas. Várias proteínas CTLA-8 deverão ser homólogas e estão aqui compreendidas. Contudo, mesmo as proteínas que possuem uma relação evolutiva mais distante com o antígeno podem ser facilmente isoladas, em condições apropriadas, utilizando estas sequências se forem suficientemente homólogas. As proteínas CTLA-8 de primatas são de particular interesse.

Estão aqui também descritas moléculas e fragmentos de ADN recombinante que possuem uma sequência de ADN idêntica ou fortemente homóloga dos ADN de IL-174 isolados aqui apresentados. Em particular, as sequências estarão frequentemente ligadas operativamente a segmentos de ADN que controlam a transcrição, a tradução e a replicação do ADN. Em alternativa, clones recombinantes derivados das sequências genómicas, por exemplo, contendo intrões, serão úteis para estudos transgênicos, incluindo, por exemplo, células e organismos transgênicos, e para terapia gênica. Ver, por exemplo, Goodnow (1992) "Transgenic Animals" in Roitt (ed.) Encyclopedia of Immunology, Academic Press, San Diego, pp. 1502-1504; Travis (1992) Science 256:1392-1394; Kuhn *et al.*

(1991) Science 254:707-710; Capecchi (1989) Science 244:1288; Robertson (ed. 1987) Teratocarcinomas and Embryonic Stem Cells: A Practical Approach, IRL Press, Oxford; Rosenberg (1992) J. Clinical Oncology 10:180-199; e Cournoyer & Caskey (1993) Ann. Rev. Immunol. 11:297-329.

As sequências de ácidos nucleicos homólogas, quando comparadas, apresentam uma similaridade significativa. Os padrões para a homologia em ácidos nucleicos são medidas da homologia geralmente utilizadas na arte, quer por comparação de sequências quer baseadas em condições de hibridação. As condições de hibridação estão descritas em maior detalhe abaixo.

Uma homologia substancial no contexto da comparação de sequências de ácidos nucleicos significa que os segmentos, ou as suas cadeias complementares, quando comparados, são idênticos quando estão alinhados de uma forma óptima, com as inserções ou as deleções de nucleótidos apropriadas, em pelo menos cerca de 50% dos nucleótidos, geralmente pelo menos 56%, mais geralmente pelo menos 59%, usualmente pelo menos 62%, mais usualmente pelo menos 65%, frequentemente pelo menos 68%, mais frequentemente pelo menos 71%, tipicamente pelo menos 74%, mais tipicamente pelo menos 77%, habitualmente pelo menos 80%, mais habitualmente pelo menos cerca de 85%, preferencialmente pelo menos cerca de 90%, mais preferencialmente pelo menos cerca de 95 a 98% ou mais e, em concretizações particulares, num valor tão elevado quanto cerca de 99% ou mais dos nucleótidos. Em alternativa, existe uma homologia substancial quando os segmentos hibridam, em condições de hibridação selectivas, com uma cadeia, ou o seu complemento, tipicamente utilizando uma sequência derivada da Tabela 2, 3 ou 6. Normalmente, a hibridação selectiva terá lugar quando existe uma homologia de pelo menos cerca de 55% ao longo de uma extensão de pelo menos 14 nucleótidos, preferencialmente pelo menos cerca de 65%, mais preferencialmente pelo menos cerca de 75% e, ainda mais preferencialmente, pelo menos cerca de 90%. Consultar Kanehisa (1984) Nuc. Acids Res. 12:203-213. O comprimento da comparação da homologia, tal como descrita, poderá ser ao longo de extensões mais longas e, em certas concretizações, será ao longo de uma extensão de pelo menos cerca de 17 nucleótidos,

usualmente pelo menos cerca de 20 nucleótidos, mais usualmente pelo menos cerca de 24 nucleótidos, tipicamente pelo menos cerca de 28 nucleótidos, mais tipicamente pelo menos cerca de 40 nucleótidos, preferencialmente pelo menos cerca de 50 nucleótidos e, mais preferencialmente, pelo menos cerca de 75 a 100 ou mais nucleótidos.

As condições rigorosas, relativamente a homologia no contexto da hibridação, serão condições rigorosas combinadas de sal, temperatura, solventes orgânicos e outros parâmetros, tipicamente os que são controlados em reacções de hibridação. As condições rigorosas de temperatura incluirão geralmente temperaturas superiores a cerca de 30°C, mais geralmente superiores a cerca de 37°C, tipicamente superiores a cerca de 45°C, mais tipicamente superiores a cerca de 55°C, preferencialmente superiores a cerca de 65°C e, mais preferencialmente, superiores a cerca de 70°C. As condições rigorosas de sal serão habitualmente uma concentração inferior a cerca de 1000 mM, habitualmente inferior a cerca de 500 mM, mais habitualmente inferior a cerca de 400 mM, tipicamente inferior a cerca de 300 mM, preferencialmente inferior a cerca de 200 mM e, mais preferencialmente, inferior a cerca de 150 mM. Contudo, a combinação de parâmetros é muito mais importante que a medida que qualquer parâmetro individual. Consultar, por exemplo, Wetmur & Davidson (1968) J. Mol. Biol. 31:349-370. A hibridação em condições rigorosas deverá proporcionar um sinal de pelo menos 2x sobre o fundo, preferencialmente pelo menos 3-5x ou mais.

Em alternativa, para a comparação de sequências, normalmente uma sequência actua como sequência de referência em relação à qual as sequências de teste são comparadas. Ao utilizar um algoritmo de comparação de sequências, as sequências de referência e de teste são introduzidas num computador, designam-se coordenadas para as subsequências, se necessário, e os parâmetros do programa do algoritmo para as sequências são designados. O algoritmo de comparação das sequências calcula, em seguida, a percentagem de identidade de sequência da(s) sequência(s) de teste em relação à sequência de referência, com base nos parâmetros do programa designados.

O alinhamento óptimo das sequências para efeitos de comparação pode ser efectuado, por exemplo, pelo algoritmo de homologia local de Smith & Waterman (1981) Adv. Appl. Math. 2:482, pelo algoritmo para alinhamento global de Needleman & Wunsch (1970) J. Mol. Biol. 48:443, pelo método de busca de similaridade de Pearson & Lipman (1988) Proc. Nat'l Acad. Sci. USA 85:2444, pelas implementações computadorizadas destes algoritmos (GAP, BESTFIT, FASTA e TFASTA no pacote de *software* de Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, WI) ou por inspecção visual (consultar, de uma forma geral, Ausubel *et al.*, *supra*).

Um exemplo de um algoritmo útil é o PILEUP. O PILEUP cria um alinhamento de múltiplas sequências a partir de um grupo de sequências relacionadas, utilizando alinhamentos progressivos par a par para mostrar a relação e a percentagem de identidade de sequência. Também representa graficamente uma árvore ou dendograma mostrando as relações de agrupamento ("clustering") utilizadas para criar o alinhamento. O PILEUP utiliza uma simplificação do método de alinhamento progressivo de Feng & Doolittle (1987) J. Mol. Evol. 35:351-360. O método utilizado é similar ao método descrito por Higgins & Sharp (1989) CABIOS 5:151-153. O programa pode alinhar até 300 sequências, cada uma delas com um comprimento máximo de 5 000 nucleótidos ou aminoácidos. O procedimento de alinhamento múltiplo começa com o alinhamento par a par das duas sequências mais similares, produzindo um agrupamento ("cluster") de duas sequências alinhadas. Este agrupamento é depois alinhado com a sequência ou agrupamento de sequências alinhadas seguinte mais relacionado. Dois agrupamentos de sequências são alinhados por uma simples extensão do alinhamento par a par de duas sequências individuais. O alinhamento final é obtido por meio de uma série de alinhamentos progressivos par a par. O programa é corrido, designando sequências específicas e as respectivas coordenadas de aminoácidos ou de nucleótidos para as regiões de comparação das sequências e designando os parâmetros do programa. Por exemplo, uma sequência de referência pode ser comparada com outras sequências de teste para determinar a relação percentual de identidade de sequência, utilizando os seguintes parâmetros: peso do hiato predefinido (3,00), peso do comprimento do hiato predefinido (0,10) e hiatos de extremidade ponderadas.

Outro exemplo de um algoritmo que é adequado para determinar a percentagem de identidade de sequência e de similaridade de sequência é o algoritmo BLAST, que é descrito por Altschul et al. (1990) J. Mol. Biol. 215:403-410. O software para realizar as análises BLAST está disponível ao público através do National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo envolve a identificação inicial de pares de sequências com uma pontuação alta (HSP de "high scoring pair") através da identificação de palavras curtas de comprimento W na sequência de busca ("query"), que correspondem ou satisfazem uma pontuação de limiar positiva T quando são alinhadas com uma palavra com o mesmo comprimento que está presente numa sequência da base de dados. T é designado por limiar de pontuação da palavra vizinha (Altschul et al., supra). Estas correspondências iniciais com a palavra vizinha actuam como sementes para iniciar buscas no sentido de encontrar HSP mais compridas que as contenham. As correspondências com a palavra são depois estendidas em ambas as direcções, ao longo de cada sequência, o mais longe que a pontuação cumulativa do alinhamento possa ser aumentada. A extensão das correspondências com a palavra em cada direcção é interrompida quando: a pontuação cumulativa do alinhamento diminui, na quantidade X, do valor máximo alcançado; a pontuação cumulativa atinge zero ou um valor inferior devido à acumulação de um ou mais alinhamentos de resíduos com uma pontuação negativa; ou é atingido o final de qualquer uma das sequências. Os parâmetros do algoritmo BLAST W, T e X determinam a sensibilidade e a velocidade do alinhamento. O programa BLAST utiliza como valores predefinidos um comprimento de palavra (W) de 11, a matriz de pontuação BLOSUM62 (ver Henikoff & Henikoff (1989) Proc. Nat'l Acad. Sci. USA 89:10915), alinhamentos (B) de 50, um valor de expectativa (E) de 10, M=5, N=4 e uma comparação de ambas as cadeias.

Além de calcular a percentagem de identidade de sequência, o algoritmo BLAST também efectua uma análise estatística da similaridade entre duas sequências (consultar, por exemplo, Karlin & Altschul (1993) Proc. Nat'l Acad. Sci. USA 90:5873-5787). Uma medida da similaridade proporcionada

pelo algoritmo BLAST é a probabilidade mais pequena de soma ($P(N)$), que proporciona uma indicação da probabilidade com que uma correspondência entre duas sequências nucleotídicas ou de aminoácidos ocorreria por acaso. Por exemplo, um ácido nucleico é considerado similar a uma sequência de referência se a probabilidade mais pequena de soma, numa comparação do ácido nucleico de teste com o ácido nucleico de referência, for inferior a cerca de 0,1; mais preferencialmente inferior a cerca de 0,01; ainda mais preferencialmente inferior a cerca de 0,001.

Uma indicação adicional de que duas sequências de ácidos nucleicos de polipéptidos são substancialmente idênticas é o facto do polipéptido codificado pelo primeiro ácido nucleico reagir imunologicamente de forma cruzada com o polipéptido codificado pelo segundo ácido nucleico, como descrito abaixo. Assim, normalmente, um polipéptido é substancialmente idêntico a um segundo polipéptido, por exemplo, quando os dois péptidos diferem apenas através de substituições conservativas. Outra indicação de que duas sequências de ácidos nucleicos são substancialmente idênticas é o facto das duas moléculas hibridarem uma com a outra em condições rigorosas, como descrito abaixo.

As proteínas semelhantes a CTLA-8 de outras espécies de mamíferos podem ser clonadas e isoladas por meio de hibridação entre espécies cruzadas fortemente aparentadas, por exemplo, o ser humano, tal como descrito nas Tabelas 1-7. A homologia poderá ser relativamente baixa entre espécies fracamente aparentadas, pelo que a hibridação de espécies fortemente aparentadas é aconselhável. Em alternativa, a produção de uma preparação de anticorpos que apresenta menos especificidade em termos de espécie poderá ser útil em abordagens de clonagem de expressão.

III. Proteína IL-174 purificada

A sequência prevista da sequência do polipéptido IL-173 de primata, por exemplo, o ser humano, e de roedor, por exemplo, o ratinho, está apresentada na Tabela 2. De forma idêntica, na Tabela 3 é proporcionada a sequência de IL-174 de primata, por exemplo, o ser humano, e é-lhe atribuída a

designação SEQ ID NO:14. Uma IL-174 de roedor, por exemplo, murina, está também descrita na Tabela 3. As sequências peptídicas permitem a preparação de péptidos para gerar anticorpos destinados ao reconhecimento destes segmentos.

Tal como aqui utilizados, os termos “proteína IL-174 de primata” e “proteína IL-174 de roedor” englobarão, quando utilizados num contexto de proteínas, uma proteína possuindo as sequências de aminoácidos designadas ilustradas na Tabela 3, ou um fragmento significativo de uma proteína destas. Eles também se referem a um polipéptido derivado de um primata ou de um roedor que apresenta uma função biológica similar ou interage com componentes de ligação específicos para a proteína IL-174. Estes componentes de ligação, por exemplo, anticorpos, normalmente ligam-se a uma proteína IL-174 com uma afinidade elevada, por exemplo, pelo menos cerca de 100 nM, usualmente melhor que cerca de 30 nM, preferencialmente melhor que cerca de 10 nM e, mais preferencialmente, melhor que cerca de 3 nM. Proteínas homólogas seriam encontradas em espécies de mamíferos que não o rato ou o ser humano, por exemplo, o ratinho e os primatas, e no genoma do vírus do herpes, por exemplo, ORF13. As espécies não mamíferas também devem possuir genes e proteínas estrutural ou funcionalmente relacionados.

O termo “polipéptido”, como aqui utilizado, inclui um fragmento ou segmento significativo e engloba uma extensão de resíduos de aminoácidos de pelo menos cerca de 8 aminoácidos, geralmente pelo menos 10 aminoácidos, mais geralmente pelo menos 12 aminoácidos, frequentemente pelo menos 14 aminoácidos, mais frequentemente pelo menos 16 aminoácidos, tipicamente pelo menos 18 aminoácidos, mais tipicamente pelo menos 20 aminoácidos, usualmente pelo menos 22 aminoácidos, mais usualmente pelo menos 24 aminoácidos, preferencialmente pelo menos 26 aminoácidos, mais preferencialmente pelo menos 28 aminoácidos e, em concretizações particularmente preferidas, pelo menos cerca de 30 ou mais aminoácidos. As extremidades específicas deste segmento encontrar-se-ão em quaisquer combinações dentro da proteína, preferencialmente englobando os domínios estruturais.

O termo "composição de ligação" refere-se a moléculas que se ligam com especificidade à proteína IL-174, por exemplo, uma forma do tipo ligando-receptor ou uma interacção anticorpo-antigénio, ou a compostos, por exemplo, proteínas que se associam especificamente à proteína IL-174, por exemplo, numa interacção natural proteína-proteína fisiologicamente relevante, covalente ou não covalente. A molécula poderá ser um polímero ou um reagente químico. Não existe nenhuma implicação quanto à proteína IL-174 ser o ligando ou o receptor de uma interacção ligando-receptor, tirando o facto da interacção apresentar uma especificidade similar, por exemplo, afinidade específica. Um análogo funcional poderá ser uma proteína com modificações estruturais ou poderá ser uma molécula totalmente não relacionada, por exemplo, que possui uma forma molecular que interage com os determinantes de ligação apropriados. As proteínas poderão actuar como agonistas ou antagonistas de um receptor; ver, por exemplo, Goodman *et al.* (eds. 1990) Goodman & Gilman's: The Pharmacological Bases of Therapeutics 8th ed.), Pergamon Press.

A solubilidade de um polipéptido ou fragmento depende do ambiente e do polipéptido. Muitos parâmetros afectam a solubilidade dos polipéptidos, incluindo a temperatura, o ambiente electrolítico, o tamanho e as características moleculares do polipéptido e a natureza do solvente. Tipicamente, a temperatura à qual o polipéptido é utilizado varia entre cerca de 4°C e cerca de 65°C. Usualmente, a temperatura em uso é superior a cerca de 18°C e, mais usualmente, é superior a cerca de 22°C. Para fins de diagnóstico, a temperatura será normalmente próxima da temperatura ambiente ou mais quente, mas inferior à temperatura de desnaturação dos componentes no ensaio. Para fins terapêuticos, a temperatura será normalmente a temperatura corporal, tipicamente cerca de 37°C para os seres humanos, embora em certas situações a temperatura possa ser aumentada ou diminuída *in situ* ou *in vitro*.

Os electrólitos aproximar-se-ão normalmente das condições fisiológicas *in situ*, mas poderão ser modificados para uma força iónica superior ou inferior quando tal for vantajoso. Os iões efectivos poderão ser modificados, por exemplo, para

estarem em conformidade com os tampões padrão utilizados em contextos fisiológicos ou analíticos.

O tamanho e a estrutura do polipéptido deverão geralmente encontrar-se num estado substancialmente estável e não num estado desnaturado. O polipéptido poderá estar associado a outros polipéptidos numa estrutura quaternária, por exemplo, para conferir solubilidade, ou associado a lípidos ou detergentes de uma forma que se aproxima às interações de bicamada lipídica natural.

O solvente será normalmente um tampão biologicamente compatível, de um tipo utilizado para a conservação das actividades biológicas, e normalmente aproximar-se-á de um solvente fisiológico. O solvente terá um pH neutro, geralmente compreendido entre cerca de 5 e 10, preferencialmente cerca de 7,5. Em algumas ocasiões, adicionar-se-á um detergente, geralmente um detergente suave e não desnaturante, por exemplo, CHS ou CHAPS, ou uma concentração suficientemente baixa para evitar uma desintegração significativa das propriedades estruturais ou fisiológicas do antigénio.

A solubilidade é reflectida pela sedimentação medida em unidades Svedberg, que são uma medida da velocidade de sedimentação de uma molécula em condições particulares. A determinação da velocidade de sedimentação era classicamente efectuada numa ultracentrífuga analítica, mas agora é normalmente realizada numa ultracentrífuga padrão. Ver Freifelder (1982) *Physical Biochemistry* (2nd ed.), W.H. Freeman; e Cantor & Schimmel (1980) *Biophysical Chemistry*, parts 1-3, W.H. Freeman & Co., San Francisco. Como uma determinação rudimentar, uma amostra contendo um polipéptido putativamente solúvel é centrifugada numa ultracentrífuga padrão de tamanho normal a cerca de 50 000 rpm durante cerca de 10 minutos, e as moléculas solúveis permanecerão no sobrenadante. Uma partícula ou polipéptido solúvel é normalmente menos de cerca de 30S, mais tipicamente menos de cerca de 15S, usualmente menos de cerca de 10S, mais usualmente menos de cerca de 6S e, em concretizações particulares, preferencialmente menos de cerca de 4S e, mais preferencialmente, menos de cerca de 3S.

IV. Produção da proteína IL-174. Espécies miméticas

O ADN que codifica a proteína IL-174 ou os seus fragmentos pode ser obtido por síntese química, rastreio de bibliotecas de ADNc ou rastreio de bibliotecas genómicas preparadas a partir de uma grande variedade de linhas celulares ou amostras de tecidos.

Este ADN pode ser expresso numa grande variedade de células hospedeiras para a síntese de uma proteína completa ou de fragmentos que podem, pelo seu lado, ser utilizados para gerar anticorpos monoclonais ou policlonais, para estudos de ligação, para a construção e a expressão de moléculas modificadas e para estudos de estrutura/função. Cada antigénio, ou os seus fragmentos, pode ser expresso em células hospedeiras que estão transformadas ou transfectadas com vectores de expressão apropriados. Estas moléculas podem ser substancialmente purificadas para ficarem isentas de contaminantes proteicos ou celulares, além daqueles derivados do hospedeiro recombinante, sendo por conseguinte particularmente úteis em composições farmacêuticas quando combinadas com um veículo e/ou diluente farmacêuticamente aceitável. O antigénio, ou suas porções, poderá ser expresso sob a forma de fusões com outras proteínas.

Os vectores de expressão são normalmente construções de ADN ou ARN de replicação autónoma contendo o gene do antigénio desejado, ou seus fragmentos, normalmente ligado operativamente a elementos de controlo genético adequados, que são reconhecidos numa célula hospedeira apropriada. Estes elementos de controlo são capazes de promover a expressão num hospedeiro adequado. O tipo específico de elementos de controlo necessários para realizar a expressão dependerá da eventual célula hospedeira utilizada. Geralmente, os elementos de controlo genético podem incluir um sistema promotor procariótico ou um sistema de controlo da expressão de um promotor eucariótico e normalmente incluem um promotor da transcrição, um operador opcional para controlar o início da transcrição, amplificadores da transcrição para aumentar o nível da expressão de ARNm, uma sequência que codifica um local de ligação dos ribossomas apropriado e sequências que terminam a transcrição e a tradução. Os vectores de expressão

normalmente também contêm uma origem de replicação que permite que o vector se replique independentemente da célula hospedeira. Métodos de amplificação do número de cópias do vector também são conhecidos; ver, por exemplo, Kaufman *et al.* (1985) *Molec. and Cell. Biol.* 5:1750-1759.

Os vectores deste invento contêm ADN que codifica uma proteína IL-174 ou um seu fragmento, normalmente que codifica um polipéptido biologicamente activo. O ADN pode estar sob o controlo de um promotor viral e pode codificar uma marca de selecção. Este invento contempla adicionalmente a utilização destes vectores de expressão, que são capazes de expressar o ADNc eucariótico codificando uma proteína IL-174, num hospedeiro procariótico ou eucariótico, em que o vector é compatível com o hospedeiro e em que o ADNc eucariótico que codifica o antigénio é inserido no vector de tal forma que o crescimento do hospedeiro contendo o vector expressa o ADNc em questão. Normalmente, os vectores de expressão são concebidos para uma replicação estável nas suas células hospedeiras ou para amplificação de modo a aumentar grandemente o número total de cópias do gene desejável por célula. Nem sempre é necessário exigir que um vector de expressão se replique numa célula hospedeira, por exemplo, é possível realizar a expressão transitória do antigénio ou seus fragmentos em vários hospedeiros, utilizando vectores que não contêm uma origem de replicação que seja reconhecida pela célula hospedeira. Também é possível utilizar vectores que provocam a integração de um gene da proteína IL-174, ou seus fragmentos, no ADN do hospedeiro por recombinação, ou integrar um promotor que controla a expressão de um gene endógeno.

Os vectores, como aqui utilizados, compreendem plasmídeos, vírus, bacteriófagos, fragmentos de ADN integráveis e outros veículos que permitem a integração de fragmentos de ADN no genoma do hospedeiro. Os vectores de expressão são vectores especializados contendo elementos de controlo genético que realizam a expressão de genes ligados operativamente. Os plasmídeos são a forma de vector mais habitualmente utilizada, mas todas as outras formas de vectores que desempenham uma função equivalente e que são, ou se tornam, conhecidos na arte são adequados para utilização neste invento. Ver, por exemplo, Pouwels *et al.* (1985 &

Supplements) Cloning Vectors: A Laboratory Manual, Elsevier, N.Y.; e Rodriguez et al. (eds. 1988) Vectors: A Survey of Molecular Cloning Vectors and Their Uses, Butterworth, Boston, MA.

As células transformadas incluem células, preferencialmente de mamífero, que foram transformadas ou transfectadas com vectores contendo um gene de IL-174, normalmente construído utilizando técnicas de ADN recombinante. As células hospedeiras transformadas geralmente expressam o antigénio, ou seus fragmentos, mas para fins de clonagem, amplificação e manipulação do seu ADN não necessitam de expressar a proteína. Este invento contempla ainda cultivar as células transformadas num meio nutriente, permitindo, assim, que a proteína se acumule na cultura. A proteína pode ser recuperada, quer a partir da cultura quer do meio de cultura.

Para os objectivos deste invento, as sequências de ADN estão ligadas operativamente quando estão funcionalmente relacionadas umas com as outras. Por exemplo, o ADN de uma pré-sequência ou sequência de comando de secreção está ligado operativamente a um polipéptido caso seja expresso como uma pré-proteína ou participe na orientação do polipéptido para a membrana celular ou na secreção do polipéptido. Um promotor está ligado operativamente a uma sequência codificante se controlar a transcrição do polipéptido; um local de ligação ao ribossoma está ligado operativamente a uma sequência codificante caso esteja posicionado para permitir a tradução. Normalmente, ligado operativamente significa contíguo e em enquadramento (em fase). Contudo, determinados elementos genéticos, como os genes repressores, não estão ligados em posição contígua, embora continuem a ligar-se a sequências operadoras que, pelo seu lado, controlam a expressão.

As células hospedeiras adequadas incluem os procariotas, os eucariotas inferiores e os eucariotas superiores. Os procariotas incluem ambos os organismos Gram-negativos e Gram-positivos, por exemplo, *E. coli* e *B. subtilis*. Os eucariotas inferiores incluem as leveduras, por exemplo, *S. cerevisiae* e *Pichia*, e espécies do género *Dictyostelium*. Os eucariotas superiores incluem linhas celulares estabelecidas de cultura

de tecidos de células animais, tanto de origem não mamífera, por exemplo, células de insecto e pássaros, como de origem mamífera, por exemplo, humana, primatas e roedores.

Os sistemas vector-hospedeiro procariótico incluem uma grande variedade de vectores para muitas espécies diferentes. Tal como aqui utilizados, *E. coli* e os seus vectores serão utilizados genericamente para incluir os vectores equivalentes utilizados em outros procariotas. Um vector representativo para amplificar o ADN é pBR322 ou muitos dos seus derivados. Os vectores que podem ser utilizados para expressar a proteína IL-174 ou os seus fragmentos incluem, embora não estejam limitados a estes, vectores como aqueles que contêm o promotor lac (série pUC); o promotor trp (pBR322-trp); o promotor lpp (série pIN); os promotores pP ou pR de lambda (pOTS); ou promotores híbridos como ptac (pDR540). Ver Brosius *et al.* (1988) "Expression Vectors Employing Lambda-, trp-, lac- and lpp-derived Promoters", in Rodriguez and Denhardt (eds.) *Vectors: A Survey of Molecular Cloning Vectors and Their Uses*, Butterworth, Boston, Chapter 10, pp. 205-236.

Os eucariotas inferiores, por exemplo, as leveduras e *Dictyostelium*, poderão ser transformados com vectores que codificam proteínas IL-174. Para os objectivos deste invento, o hospedeiro eucariota inferior mais comum é a levedura de padeiro, *Saccharomyces cerevisiae*. Este será utilizado para representar genericamente os eucariotas inferiores, embora um certo número de outras estirpes e espécies esteja igualmente disponível. Os vectores para leveduras consistem normalmente numa origem de replicação (excepto se forem do tipo integrante), um gene de selecção, um promotor, o ADN que codifica a proteína desejada, ou seus fragmentos, e sequências para terminação da tradução, poliadenilação e terminação da transcrição. Os vectores de expressão adequados para levedura incluem promotores constitutivos como o promotor da 3-fosfoglicerato-quinase e vários outros promotores de genes de enzimas glicolíticas, ou promotores indutíveis como o promotor da álcool-desidrogenase 2 ou o promotor da metalotionina. Os vectores adequados incluem derivados dos seguintes tipos: replicação autónoma e baixo número de cópias (como a série YRp), replicação autónoma e número elevado de cópias (como a

série YEp); tipos integrantes (como a série YIp) ou minicromossomas (como a série YCp).

As células de cultura de tecidos de eucariotas superiores são as células hospedeiras preferidas para a expressão da proteína IL-174 funcionalmente activa. Em princípio, muitas linhas de células de cultura de tecidos de eucariotas superiores são viáveis, por exemplo, sistemas de expressão baculovírus-insecto, sejam elas de uma fonte invertebrada ou de uma fonte vertebrada. Contudo, as células de mamíferos são preferidas devido ao processamento, tanto co-tradução como pós-tradução. A transformação ou transfecção e a propagação destas células tornou-se um procedimento de rotina. Os exemplos de linhas celulares úteis incluem as células HeLa, as linhas de células de ovário de *hamster* chinês (CHO), as linhas de células de rim de rato bebé (BRK), as linhas de células de insecto, as linhas de células de pássaro e as linhas de células de macaco (COS). Os vectores de expressão para estas linhas de células geralmente incluem uma origem de replicação, um promotor, um local de iniciação da tradução, locais de excisão-reparação ("splicing") do ARN (caso se utilize ADN genómico), um local de poliadenilação e um local de terminação da transcrição. Estes vectores normalmente também contêm um gene de selecção ou um gene de amplificação. Os vectores de expressão adequados poderão ser plasmídeos, vírus ou retrovírus que transportam promotores derivados, por exemplo, de fontes como adenovírus, SV40, parvovírus, o vírus *vaccinia* ou citomegalovírus. Os exemplos representativos de vectores de expressão adequados incluem pCDNA1; pCD, ver Okayama *et al.* (1985) *Mol. Cell Biol.* 5:1136-1142; pMC1neo Poly-A, ver Thomas *et al.* (1987) *Cell* 51:503-512; e um vector de baculovírus como pAC 373 ou pAC 610, ver O'Reilly *et al.* (1992) *Baculovirus Expression Vectors: A Laboratory Manual*, Freeman and Co., CRC Press, Boca Raton, Fl.

Frequentemente, será desejável expressar um polipéptido da proteína IL-174 num sistema que proporcione um padrão de glicosilação específico ou definido. Neste caso, o padrão habitual será aquele proporcionado naturalmente pelo sistema de expressão. Contudo, o padrão é modificável por exposição do polipéptido, por exemplo, uma forma não glicosilada, a proteínas de glicosilação apropriadas introduzidas num sistema

de expressão heterólogo. Por exemplo, o gene da proteína IL-174 poderá ser co-transformado com um ou mais genes que codificam enzimas de glicosilação de um mamífero ou outras enzimas de glicosilação. Utilizando esta abordagem, determinados padrões de glicosilação de mamíferos serão alcançáveis ou aproximados em células de procariotas ou outras.

A proteína IL-174, ou um seu fragmento, poderá ser manipulada para ser ligada por via de fosfatidilinositol (PI) a uma membrana celular, podendo ser removida das membranas por tratamento com uma enzima de clivagem do fosfatidilinositol, por exemplo, a fosfatidilinositol-fosfolipase C. Isto liberta o antigénio numa forma biologicamente activa e permite a purificação por meio de procedimentos padrão da química de proteínas. Ver, por exemplo, Low (1989) *Biochim. Biophys. Acta* 988:427-454; Tse et al. (1985) *Science* 230:1003-1008; e Brunner et al. (1991) *J. Cell Biol.* 114:1275-1283.

Agora que a proteína IL-174 foi caracterizada, é possível preparar fragmentos ou seus derivados através de processos convencionais de síntese de péptidos. Estes incluem processos como os descritos em Stewart & Young (1984) *Solid Phase Peptide Synthesis*, Pierce Chemical Co., Rockford, IL; Bodanszky & Bodanszky (1984) *The Practice of Peptide Synthesis*, Springer-Verlag, New York; e Bodanszky (1984) *The Principles of Peptide Synthesis*, Springer-Verlag, New York. Por exemplo, é possível utilizar um método de azida, um método de cloreto de ácido, um método de anidrido de ácido, um método de anidrido misto, um método de éster activo (por exemplo, o éster p-nitrofenílico, o éster N-hidroxissuccinimídico ou o éster cianometílico), um método de carbodiimidazole, um método de oxirredução ou um método de diciclo-hexilcarbodiimida (DCCD)/aditivo. As sínteses em fase sólida e em fase de solução são ambas aplicáveis aos processos anteriores.

A proteína IL-174, os fragmentos ou os derivados são adequadamente preparados de acordo com os processos acima referidos normalmente empregues na síntese de péptidos, quer por um denominado processo passo a passo, que compreende condensar um aminoácido com o aminoácido terminal, um a um em sequência, quer por acoplamento de fragmentos peptídicos ao

aminoácido terminal. Os grupos amino que não estão a ser utilizados na reacção de acoplamento estão normalmente protegidos para evitar o acoplamento num local incorrecto.

Caso seja adoptada uma síntese em fase sólida, o aminoácido C-terminal está preso a um transportador ou um suporte insolúvel através do seu grupo carboxilo. O transportador insolúvel não está particularmente limitado, desde que possua uma capacidade de ligação a um grupo carboxilo reactivo. Os exemplos destes transportadores insolúveis incluem resinas de halometilo, como seja a resina de clorometilo ou a resina de bromometilo, resinas de hidroximetilo, resinas de fenol, resinas de terc-alquiloxycarbonilo tratadas com hidrazida e compostos similares.

Um aminoácido protegido ao nível do grupo amino é ligado em sequência, através de condensação do seu grupo carboxilo activado com o grupo amino reactivo do péptido ou da cadeia previamente formado, para sintetizar o péptido passo a passo. Após sintetizar a sequência completa, o péptido é separado do transportador insolúvel para produzir o péptido. Esta abordagem em fase sólida é descrita por Merrifield *et al.* (1963) in *J. Am. Chem. Soc.* 85:2149-2156.

A proteína preparada e os seus fragmentos podem ser isolados e purificados a partir da mistura reaccional por meio de separação de péptidos, por exemplo, por extracção, precipitação, electroforese, várias formas de cromatografia e métodos similares. As proteínas IL-174 deste invento podem ser obtidas em graus variáveis de pureza dependendo da utilização desejada. A purificação pode ser obtida por utilização das técnicas de purificação de proteínas aqui descritas, ou por utilização dos anticorpos aqui descritos em cromatografia de afinidade em imunoabsorvente. Esta cromatografia de afinidade em imunoabsorvente é executada ligando primeiro os anticorpos a um suporte sólido e contactando depois os anticorpos ligados com lisados solubilizados de células originais apropriadas, lisados de outras células que expressam a proteína ou lisados ou sobrenadantes de células que produzem a proteína IL-174 em resultado de técnicas de ADN; ver abaixo.

V. Variantes físicas

Encontram-se igualmente aqui descritas proteínas ou péptidos que possuem uma homologia de sequência de aminoácidos substancial com a sequência de aminoácidos da proteína IL-174. As variantes incluem variantes de espécies ou variantes alélicas.

A homologia de sequência de aminoácidos, ou identidade de sequência, é determinada através de otimização das correspondências de resíduos, se necessário, introduzindo hiatos conforme requerido. Isto altera-se quando se consideram as substituições conservativas como correspondências. As substituições conservativas tipicamente incluem substituições dentro dos seguintes grupos: glicina, alanina; valina, isoleucina, leucina; ácido aspártico, ácido glutâmico; asparagina, glutamina; serina, treonina; lisina, arginina; e fenilalanina, tirosina. As sequências de aminoácidos homólogas destinam-se tipicamente a incluir as variações alélicas e inter-espécies naturais em cada sequência da proteína respectiva. As proteínas ou péptidos homólogos típicos terão uma homologia entre 25-100% (se for possível introduzir hiatos) e 50-100% (se as substituições conservativas estiveram incluídas) com a sequência de aminoácidos da proteína IL-170. As medidas de homologia serão pelo menos cerca de 35%, geralmente pelo menos 40%, mais geralmente pelo menos 45%, frequentemente pelo menos 50%, mais frequentemente pelo menos 55%, tipicamente pelo menos 60%, mais tipicamente pelo menos 65%, usualmente pelo menos 70%, mais usualmente pelo menos 75%, preferencialmente pelo menos 80%, mais preferencialmente pelo menos 80% e, em concretizações particularmente preferidas, pelo menos 85% ou mais. Ver também Needleham *et al.* (1970) *J. Mol. Biol.* 48:443-453; Sankoff *et al.* (1983) Chapter One in *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, Addison-Wesley, Reading, MA; e os pacotes de *software* de IntelliGenetics, Mountain View, CA e de University of Wisconsin Genetics Computer Group, Madison, WI. O ADN isolado que codifica uma proteína IL-174 pode ser facilmente modificado através de substituições nucleotídicas, deleções nucleotídicas, inserções nucleotídicas e inversões de extensões nucleotídicas. Estas modificações resultam em novas

sequências de ADN que codificam estes antígenos, os seus derivados ou proteínas que possuem uma actividade fisiológica, imunogénica ou antigénica similar. Estas sequências modificadas podem ser utilizadas para produzir antígenos mutantes ou para melhorar a expressão. A expressão melhorada poderá envolver a amplificação do gene, um aumento da transcrição, um aumento da tradução e outros mecanismos. Estes derivados mutantes da proteína IL-174 incluem mutações predeterminadas ou dirigidas da respectiva proteína ou seus fragmentos. A "proteína IL-174 mutante" engloba um polipéptido que de outro modo cai dentro da definição de homologia da proteína IL-174 murina ou IL-174 humana como apresentada acima, mas que possui uma sequência de aminoácidos que difere da sequência da proteína IL-174 como encontrada na natureza, seja por meio de deleção, de substituição ou de inserção. Em particular, a "proteína IL-174 mutante obtida por mutação dirigida" inclui geralmente proteínas que possuem uma homologia significativa com a sequência da Tabela 3 e que partilham várias actividades biológicas, por exemplo, antigénica ou imunogénica, com aquelas sequências, e que em concretizações preferidas contêm a maior parte das sequências descritas. Conceitos similares aplicam-se a diferentes proteínas IL-174, particularmente aquelas encontradas em vários animais de sangue quente, por exemplo, mamíferos e pássaros. Como anteriormente afirmado, salienta-se que as descrições pretendem geralmente englobar todas as proteínas IL-174, não estando limitadas à concretização de ratinho especificamente discutida.

Embora os locais de mutação dirigida sejam predeterminados, os mutantes não precisam de ser mutantes dirigidos. A mutagénese da proteína IL-174 pode ser realizada fazendo inserções ou deleções de aminoácidos. É possível gerar substituições, deleções, inserções ou quaisquer combinações delas para chegar a uma construção final. As inserções incluem fusões amino-terminais ou carboxi-terminais. A mutagénese aleatória pode ser realizada num codão-alvo, e os mutantes expressos podem depois ser rastreados quanto à actividade desejada. Os métodos para fazer mutações de substituição em locais predeterminados do ADN que possui uma sequência conhecida são bem conhecidos na arte, por exemplo, por mutagénese com o iniciador M13 ou técnicas de reacção da

polimerase em cadeia (PCR). Ver também Sambrook *et al.* (1989) & Ausubel *et al.* (1987 & Supplements).

As mutações do ADN não devem normalmente colocar as sequências codificantes fora do quadro de leitura (fora de fase) e, de preferência, não criarão regiões complementares que possam hibridar para produzir uma estrutura de ARNm secundária como voltas ou ganchos de cabelo.

O presente invento também proporciona proteínas recombinantes, por exemplo, proteínas de fusão heterólogas utilizando segmentos provenientes destas proteínas. Uma proteína de fusão heteróloga é uma fusão de proteínas ou segmentos que normalmente não estão fundidos da mesma maneira. Assim, o produto de fusão de uma imunoglobulina com um polipéptido IL-174 é uma molécula de proteína contínua possuindo sequências fundidas numa ligação peptídica típica, normalmente constituída por um único produto de tradução e que apresenta propriedades derivadas de cada péptido original. Um conceito similar aplica-se às sequências de ácidos nucleicos heterólogas.

Além disso, é possível produzir construções novas mediante a combinação de domínios funcionais similares de outras proteínas. Por exemplo, segmentos de ligação ao antigénio ou outros segmentos poderão ser "trocados" entre fragmentos ou polipéptidos de fusão novos e diferentes. Consultar, por exemplo, Cunningham *et al.* (1989) *Science* 243: 1330-1336; e O'Dowd *et al.* (1988) *J. Biol. Chem.* 263:15985-15992. Por conseguinte, novos polipéptidos quiméricos que exibem novas combinações de especificidades resultarão da ligação funcional de domínios biologicamente relevantes e outros domínios funcionais.

O método do fosforamidito descrito por Beaucage & Carruthers (1981) *Tetra. Letts.* 22:1859-1862 produzirá fragmentos de ADN sintéticos adequados. Um fragmento de cadeia dupla será frequentemente obtido por síntese da cadeia complementar e hibridação da cadeia em condições apropriadas, ou por adição da cadeia complementar utilizando a ADN-polimerase com uma sequência iniciadora adequada, por exemplo, técnicas de PCR.

VI. Variantes funcionais

O bloqueio da resposta fisiológica às proteínas IL-174 poderá resultar da inibição da ligação do antigénio ao seu parceiro de ligação natural, por exemplo, através de inibição competitiva. Assim, os ensaios *in vitro* do presente invento utilizarão frequentemente a proteína isolada, membranas de células que expressam uma proteína IL-174 recombinante associada à membrana, fragmentos solúveis compreendendo fragmentos de ligação ou fragmentos ligados a substratos em fase sólida. Estes ensaios também permitirão a determinação de diagnóstico dos efeitos de mutações e modificações no segmento de ligação ou de mutações e modificações na proteína, por exemplo, análogos.

Este invento também contempla a utilização de ensaios competitivos de rastreio de fármacos, por exemplo, em que anticorpos neutralizantes para o antigénio ou fragmentos do parceiro de ligação competem com um composto de teste pela ligação à proteína. Desta forma, os anticorpos podem ser utilizados para detectar a presença de qualquer polipéptido que partilhe um ou mais locais de ligação antigénicos da proteína e também podem ser utilizados para ocupar locais de ligação na proteína que, caso contrário, podiam interagir com um parceiro de ligação.

Adicionalmente, é possível utilizar anticorpos neutralizantes contra a proteína IL-174 e fragmentos solúveis do antigénio que contêm um local de ligação ao receptor de afinidade elevada, para inibir a função do antigénio nos tecidos, por exemplo, em tecidos que experimentam uma fisiologia anormal.

Os “derivados” dos antigénios IL-174 incluem mutantes da sequência de aminoácidos, variantes de glicosilação e conjugados covalentes ou agregados com outros grupos químicos. Os derivados covalentes podem ser preparados mediante ligação de certas funcionalidades a grupos que são encontrados nas cadeias laterais dos aminoácidos de IL-174 ou nas extremidades N-terminal ou C-terminal, através de meios que são bem conhecidos na arte. Estes derivados podem incluir, sem

limitação, amidas ou ésteres alifáticos da extremidade carboxilo ou de resíduos que contêm cadeias laterais carboxilo, derivados O-acílicos de resíduos que contêm grupos hidroxilo e derivados N-acílicos do aminoácido na posição amino-terminal ou de resíduos que contêm grupos amino, por exemplo, lisina ou arginina. Os grupos acilo são seleccionados entre o conjunto de grupos alquilo que incluem os grupos alquilo normais em C₃-C₁₈, formando assim espécies alcanoílo ou aroílo. A ligação covalente a proteínas de transporte poderá ser importante quando os grupos imunogénicos são haptenos.

Em particular, estão incluídas as alterações da glicosilação, por exemplo, efectuadas por modificação dos padrões de glicosilação de um polipéptido durante a sua síntese e processamento ou em passos de processamento adicionais. Os meios particularmente preferidos para realizar isto são a exposição do polipéptido a enzimas de glicosilação derivadas de células que normalmente proporcionam este processamento, por exemplo, enzimas de glicosilação de mamíferos. As enzimas de desglicosilação também estão contempladas. Igualmente abrangidas estão versões da mesma sequência de aminoácidos primária que possuem outras modificações secundárias, incluindo resíduos de aminoácidos fosforilados, por exemplo, fosfotirosina, fosfosserina ou fosfotreonina.

Um grupo principal de derivados consiste nos conjugados covalentes da proteína IL-174 ou seus fragmentos com outras proteínas ou polipéptidos. Estes derivados podem ser sintetizados em cultura recombinante, tal como fusões N-terminais ou C-terminais, ou através da utilização de agentes conhecidos na arte pela sua utilidade na reticulação de proteínas através dos grupos laterais reactivos. Os locais de derivatização preferidos do antigénio com agentes de reticulação são ao nível dos grupos amino livres, grupos hidrato de carbono e resíduos de cisteína.

Polipéptidos de fusão entre as proteínas IL-174 e outras proteínas homólogas ou heterólogas também são proporcionados. Os polipéptidos homólogos poderão ser fusões entre diferentes marcadores de superfície, resultando, por exemplo, em uma proteína híbrida que apresenta especificidade de ligação ao

receptor. De forma similar, é possível construir fusões heterólogas que apresentariam uma combinação de propriedades ou actividades das proteínas das quais derivam. Os exemplos típicos são fusões de um polipéptido repórter, por exemplo, a luciferase, com um segmento ou domínio de um antigénio, por exemplo, um segmento de ligação ao receptor, pelo que a presença ou localização do antigénio fundido poderá ser facilmente determinada. Ver, por exemplo, Dull *et al.*, Patente U.S. 4,859,609. Outros parceiros para a fusão génica incluem a β -galactosidase bacteriana, a *trpE*, a proteína A, a β -lactamase, a alfa-amilase, a álcool-desidrogenase e o factor alfa de conjugação de levedura. Consultar, por exemplo, Godowski *et al.* (1988) *Science* 241:812-816.

O método do fosforamidito descrito por Beaucage & Carruthers (1981) *Tetra. Letts.* 22:1859-1862 produzirá fragmentos de ADN sintéticos adequados. Um fragmento de cadeia dupla será frequentemente obtido por síntese da cadeia complementar e hibridação da cadeia em condições apropriadas, ou por adição da cadeia complementar utilizando a ADN-polimerase com uma sequência iniciadora adequada.

Estes polipéptidos também poderão ter resíduos de aminoácidos que foram quimicamente modificados por fosforilação, sulfonação, biotinilação ou por adição ou remoção de outros grupos, particularmente aqueles que possuem formas moleculares semelhantes aos grupos fosfato. Em algumas concretizações, as modificações serão reagentes de marcação úteis ou actuarão como alvos de purificação, por exemplo, ligandos de afinidade.

As proteínas de fusão serão tipicamente preparadas por métodos de ácidos nucleicos recombinantes ou por métodos de polipéptidos sintéticos. As técnicas para a manipulação e expressão de ácidos nucleicos estão descritas de uma forma geral, por exemplo, em Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed.), Vols. 1-3, Cold Spring Harbor Laboratory. As técnicas para a síntese de polipéptidos estão descritas, por exemplo, em Merrifield (1963) *J. Amer. Chem. Soc.* 85:2149-2156; Merrifield (1986) *Science* 232:341-347; e Atherton *et al.* (1989) *Solid Phase Peptide Synthesis: A Practical Approach*, IRL Press, Oxford.

O invento também contempla a utilização de derivados das proteínas IL-174 que são distintos das variações ao nível da sequência de aminoácidos ou ao nível da glicosilação. Estes derivados poderão envolver a associação covalente ou agregativa com grupos químicos. Estes derivados geralmente caem nas três classes: (1) sais, (2) modificações covalentes nas cadeias laterais e no resíduo terminal e (3) complexos de adsorção, por exemplo, com membranas celulares. Estes derivados covalentes ou agregativos são úteis como imunogénios, como reagentes em imunoensaios ou em métodos de purificação, como seja a purificação por afinidade de antigénios ou de outras proteínas de ligação. Por exemplo, um antigénio IL-174 pode ser imobilizado por ligação covalente num suporte sólido, como a Sepharose activada com brometo de cianogénio, através de métodos que são bem conhecidos na arte, ou adsorvido em superfícies de poliolefinas, com ou sem reticulação com glutaraldeído, para utilização no ensaio ou na purificação de anticorpos anti-proteína IL-174 ou do seu receptor ou de outro parceiro de ligação. Os antigénios IL-174 também podem ser marcados com um grupo detectável, por exemplo, radioiodados pelo procedimento com cloramina T, ligados covalentemente a quelatos de terras raras ou conjugados com outro grupo fluorescente para utilizar em ensaios de diagnóstico. A purificação da proteína IL-170 poderá ser realizada por meio de anticorpos ou parceiros de ligação imobilizados. Um antigénio ou fragmento IL-174 solubilizado deste invento pode ser utilizado como um imunogénio para a produção de anti-soros ou de anticorpos específicos para a proteína ou seus fragmentos. O antigénio purificado pode ser utilizado para rastrear anticorpos monoclonais ou fragmentos de ligação preparados por imunização com várias formas de preparações impuras contendo a proteína. Em particular, o termo "anticorpos" também compreende fragmentos de ligação ao antigénio de anticorpos naturais. As proteínas IL-170 purificadas também podem ser utilizadas como um reagente para detectar quaisquer anticorpos gerados em resposta à presença de níveis elevados da proteína ou de fragmentos celulares contendo o antigénio, em que ambos poderão ser um diagnóstico de uma condição fisiológica anormal ou específica ou de uma condição de doença. Adicionalmente, os fragmentos do antigénio também poderão actuar como imunogénios

para produzir os anticorpos do presente invento, como descrito imediatamente abaixo. Por exemplo, este invento contempla anticorpos produzidos contra sequências de aminoácidos codificadas pela sequência nucleotídica apresentada na Tabela 3, ou fragmentos de proteínas contendo esta sequência. Em particular, este invento contempla anticorpos que possuem afinidade de ligação para, ou são produzidos contra, fragmentos específicos que se prevê localizarem-se fora da bicamada lipídica.

O presente invento contempla o isolamento de variantes de espécies adicionais fortemente aparentadas. A análise de transferência de Southern estabeleceu que existem entidades genéticas similares em outros mamíferos, por exemplo, o rato e o ser humano. É provável que as proteínas IL-174 estejam disseminadas em variantes de espécies, por exemplo, roedores, lagomorfos, carnívoros, artiodáctilos, perissodáctilos e primatas.

O invento também proporciona meios para isolar um grupo de antigénios relacionados, que apresentam tanto disparidades como similaridades em termos de estrutura, expressão e função. A elucidação de muitos dos efeitos fisiológicos dos antigénios será grandemente acelerada pelo isolamento e caracterização de diferentes variantes de espécies. Em particular, o presente invento proporciona sondas úteis para identificar entidades genéticas homólogas adicionais em diferentes espécies.

Os genes isolados permitirão a transformação de células desprovidas de expressão da proteína IL-174 correspondente, por exemplo, tipos ou células de espécies que não possuem os antigénios correspondentes e que deveriam exibir uma actividade de fundo negativa. A expressão dos genes transformados permitirá o isolamento de linhas celulares antígenicamente puras, com variantes de espécie definidas ou únicas. Esta abordagem permitirá uma detecção e discriminação mais sensíveis dos efeitos fisiológicos das proteínas IL-174. Fragmentos subcelulares, por exemplo, citoplastos ou fragmentos membranares, podem ser isolados e utilizados.

A dissecação dos elementos estruturais críticos que realizam as várias funções fisiológicas ou de diferenciação

proporcionadas pelas proteínas é possível utilizando técnicas padrão da biologia molecular moderna, particularmente na comparação de membros da classe relacionada. Ver, por exemplo, a técnica de mutagénese de varrimento de homólogos descrita em Cunningham *et al.* (1989) *Science* 243:1339-1336; e as abordagens utilizadas em O'Dowd *et al.* (1988) *J. Biol. Chem.* 263:15985-15992; e Lechleiter *et al.* (1990) *EMBO J.* 9: 4381-4390.

Em particular, os domínios ou segmentos funcionais podem ser substituídos entre variantes de espécies para determinar que características estruturais são importantes tanto na afinidade como na especificidade para o parceiro de ligação, assim como na transdução de sinal. Um conjunto de diferentes variantes será utilizado para rastrear as moléculas que exibem propriedades combinadas de interacção com diferentes variantes de espécies de parceiros de ligação.

A internalização do antigénio poderá ocorrer em determinadas circunstâncias, e poderá ocorrer uma interacção entre os componentes intracelulares e os segmentos "extracelulares" das proteínas envolvidas nas interacções. Os segmentos específicos da interacção da proteína IL-174 com outros componentes intracelulares poderão ser identificados por mutagénese ou por meios bioquímicos directos, por exemplo, métodos de reticulação ou afinidade. A análise estrutural por métodos cristalográficos ou outros métodos físicos também será aplicável. A investigação adicional do mecanismo da função biológica incluirá o estudo de componentes associados, que poderão ser isolados por métodos de afinidade ou por meios genéticos, por exemplo, a análise de complementação de mutantes.

O estudo adicional da expressão e controlo da proteína IL-174 será prosseguido. Os elementos de controlo associados aos antigénios poderão apresentar padrões de expressão de desenvolvimento diferencial, padrões de expressão específicos de tecidos ou outros padrões de expressão. As regiões genéticas a montante ou a jusante, por exemplo, os elementos de controlo, são importantes.

Os estudos estruturais do antigénio conduzirão à concepção de novas variantes, particularmente de análogos que

exibem propriedades agonistas ou antagonistas sobre os parceiros de ligação. Isto pode ser combinado com métodos de rastreio previamente descritos para isolar variantes que exibem espectros desejados de actividades.

A expressão em outros tipos de células resultará frequentemente em diferenças de glicosilação num antigénio particular. Várias variantes de espécies poderão exibir funções distintas com base em diferenças estruturais que não a sequência de aminoácidos. As modificações diferenciais poderão ser responsáveis pela função diferencial, e a elucidação dos efeitos é agora possível.

Assim, o presente invento proporciona reagentes importantes relacionados com a interacção antigénio-parceiro de ligação. Embora a descrição anterior se tenha centrado essencialmente na proteína IL-174 murina e IL-174 humana, os peritos na arte reconhecerão imediatamente que o invento abrange outros antigénios, por exemplo, variantes de ratinho e de outras espécies de mamíferos ou variantes alélicas, assim como as suas variantes.

VII. *Kits*

Este invento também contempla a utilização de proteínas IL-174, seus fragmentos, péptidos e os seus produtos de fusão numa variedade de *kits* e métodos de diagnóstico para detectar a presença de uma composição de ligação. Tipicamente, o *kit* possuirá um compartimento contendo um péptido ou um segmento génico definido de IL-174 ou um reagente que reconhece um ou o outro, por exemplo, fragmentos do antigénio ou anticorpos.

Um *kit* para determinar a afinidade de ligação de um composto de teste a uma proteína IL-174 compreenderia tipicamente um composto de teste; um composto marcado, por exemplo, um anticorpo possuindo uma afinidade de ligação conhecida para o antigénio; uma fonte de proteína IL-174 (natural ou recombinante) e um meio para separar o composto marcado ligado do composto marcado livre, tal como uma fase sólida para imobilizar o antigénio. Uma vez os compostos rastreados, aqueles que possuem uma afinidade de ligação adequada para o antigénio podem ser avaliados em ensaios

biológicos apropriados, como são bem conhecidos na arte, para determinar se exibem ou não actividades biológicas similares ao antigénio natural. A disponibilidade dos polipéptidos recombinantes da proteína IL-174 também proporciona padrões bem definidos para calibrar esses ensaios.

Um *kit* preferido para determinar a concentração, por exemplo, de uma proteína IL-174 numa amostra compreenderia tipicamente um composto marcado, por exemplo, um anticorpo, possuindo uma afinidade de ligação conhecida para o antigénio, uma fonte de antigénio (natural ou recombinante) e um meio para separar o composto marcado ligado do composto marcado livre, por exemplo, uma fase sólida para imobilizar a proteína IL-174. Compartimentos contendo os reagentes e instruções serão normalmente proporcionados.

Um método para determinar a concentração de proteína IL-174 numa amostra compreenderia tipicamente os passos de: (1) preparação das membranas a partir de uma amostra constituída por uma fonte de proteína IL-174 ligada à membrana; (2) lavagem das membranas e suspensão das mesmas num tampão; (3) solubilização do antigénio através de incubação das membranas num meio de cultura ao qual foi adicionado um detergente adequado; (4) ajuste da concentração de detergente do antigénio solubilizado; (5) contacto e incubação da referida diluição com anticorpo radiomarcado para formar complexos; (6) recuperação dos complexos, por exemplo, por filtração através de filtros tratados com polietilenoimina; e (7) medição da radioactividade dos complexos recuperados.

Os anticorpos, incluindo os fragmentos de ligação ao antigénio, específicos para a proteína IL-174 ou seus fragmentos são úteis em aplicações de diagnóstico para detectar a presença de níveis elevados de proteína IL-174 e/ou seus fragmentos. Estes ensaios de diagnóstico podem utilizar lisados, células vivas, células fixadas, imunofluorescência, culturas de células, fluidos corporais e podem ainda envolver a detecção de antigénios relacionados com a proteína no soro ou em fluidos similares. Os ensaios de diagnóstico poderão ser homogéneos (sem um passo de separação entre o reagente livre e o complexo proteína-proteína) ou heterogéneos (com um passo de separação). Existem vários ensaios comerciais, tais como o

radioimunoensaio (RIA), o ensaio imunossorvente ligado a enzima (ELISA), o imunoensaio enzimático (EIA), a técnica de imunoensaio enzimático de multiplicação (EMIT), o imunoensaio fluorescente com substrato marcado (SLFIA) e ensaios idênticos. Por exemplo, é possível utilizar anticorpos não marcados mediante o recurso a um segundo anticorpo que está marcado e que reconhece o anticorpo para uma proteína IL-174 ou para um seu fragmento particular. Ensaio similares foram também extensamente discutidos na literatura. Ver, por exemplo, Harlow & Lane (1988) *Antibodies: A Laboratory Manual*, CSH.

Os anticorpos anti-idiotípicos poderão ter uma utilização similar para diagnosticar a presença de anticorpos contra uma proteína IL-174, dado que isso poderá ser um diagnóstico de vários estados anormais. Por exemplo, a sobreprodução de proteína IL-174 poderá resultar na produção de várias reacções imunológicas que poderão ser um diagnóstico de estados fisiológicos anormais, particularmente em condições celulares proliferativas como o cancro ou a diferenciação anormal.

Frequentemente, os reagentes para os ensaios de diagnóstico são fornecidos em *kits*, de modo a otimizar a sensibilidade do ensaio. Para o presente invento, dependendo da natureza do ensaio, é proporcionado o protocolo e o marcador - o anticorpo marcado ou não marcado, ou a proteína IL-174 marcada - geralmente em conjunção com outros aditivos, como tampões, estabilizadores, materiais necessários para a produção de sinal, como sejam os substratos para as enzimas, e espécies similares. Preferencialmente, o *kit* também conterá instruções para uma correcta utilização e eliminação do conteúdo após o uso. Tipicamente, o *kit* possui compartimentos para cada reagente útil. Desejavelmente, os reagentes são proporcionados como um pó liofilizado seco, em que os reagentes poderão ser reconstituídos num meio aquoso, proporcionando concentrações apropriadas dos mesmos para realizar o ensaio.

Quaisquer dos constituintes acima mencionados dos ensaios de diagnóstico e de rastreio de fármacos poderão ser utilizados sem modificação ou poderão ser modificados de uma variedade de formas. Por exemplo, a marcação poderá ser obtida

por união covalente ou não covalente de um grupo que proporciona, directa ou indirectamente, um sinal detectável. Em qualquer um destes ensaios, o antigénio, o composto de teste, a proteína IL-174 ou os anticorpos contra ela podem ser marcados directa ou indirectamente. As possibilidades para a marcação directa incluem grupos de marcação: radiomarcadores como ^{125}I , enzimas (Patente U.S. 3,645,090) como a peroxidase e a fosfatase alcalina e marcadores fluorescentes (Patente U.S. 3,940,475) capazes de monitorizar a alteração da intensidade de fluorescência, o desvio de comprimento de onda ou a polarização de fluorescência. As possibilidades para a marcação indirecta incluem a biotinilação de um constituinte, seguida de ligação a avidina acoplada a um dos grupos de marcação acima referidos.

Há também numerosos métodos para separar o antigénio ligado do antigénio livre ou, em alternativa, o composto de teste ligado do composto de teste livre. A proteína IL-174 pode ser imobilizada em várias matrizes, seguida de lavagem. As matrizes apropriadas incluem o plástico, como seja uma placa de ELISA, os filtros e as esferas. Os métodos de imobilização da proteína IL-174 numa matriz incluem, sem limitação, a adesão directa ao plástico, a utilização de um anticorpo de captura, o acoplamento químico e a biotina-avidina. O último passo nesta abordagem envolve a precipitação do complexo proteína-proteína por meio de um de vários métodos, incluindo aqueles que utilizam, por exemplo, um solvente orgânico como o polietilenoglicol ou um sal como o sulfato de amónio. Outras técnicas de separação adequadas incluem, sem limitação, o método de partículas magnetizáveis e anticorpo marcado com fluoresceína descrito em Rattle *et al.* (1984) Clin. Chem. 30: 1457-1461, e a separação com partículas magnéticas e anticorpos duplos como descrita na Patente U.S. 4,659,678.

Os métodos de ligação das proteínas ou seus fragmentos aos vários marcadores têm sido amplamente referidos na literatura e não necessitam de uma discussão detalhada aqui. Muitas das técnicas envolvem o uso de grupos carboxilo activados através da utilização de carbodiimida ou de ésteres activos para formar ligações peptídicas, a formação de tioéteres por reacção de um grupo mercapto com um halogéneo

activado como cloroacetilo, ou uma olefina activada como a maleimida para ligação. As proteínas de fusão também encontram uso nestas aplicações.

Outro aspecto de diagnóstico aqui descrito envolve a utilização de sequências oligonucleotídicas ou polinucleotídicas retiradas da sequência de uma proteína IL-174. Estas sequências podem ser utilizadas como sondas para detectar níveis de mensagem do antigénio em amostras de doentes suspeitos de terem uma condição anormal, por exemplo, cancro ou um problema de desenvolvimento. A preparação de ambas as sequências nucleotídicas de ARN e ADN, a marcação das sequências e o tamanho preferido das sequências recebeu ampla descrição e discussão na literatura. Normalmente, uma sonda oligonucleotídica deve ter pelo menos cerca de 14 nucleótidos, geralmente pelo menos cerca de 18 nucleótidos, e as sondas polinucleotídicas poderão ter até várias quilobases. É possível utilizar vários marcadores, mais habitualmente radionuclidos, particularmente ^{32}P . Contudo, outras técnicas também poderão ser empregues, tal como a utilização de nucleótidos modificados com biotina para introdução num polinucleótido. A biotina depois actua como o local para a ligação a avidina ou aos anticorpos, que poderão estar marcados com uma grande variedade de marcadores, tais como radionuclidos, agentes fluorescentes, enzimas ou espécies similares. Em alternativa, é possível utilizar anticorpos que podem reconhecer cópias específicas, incluindo cópias de ADN, cópias de ARN, cópias híbridas ADN-ARN ou cópias ADN-proteína. Os anticorpos, pelo seu lado, poderão ser marcados e o ensaio efectuado com o dúplex ligado a uma superfície, de modo que ao dar-se a formação do dúplex na superfície, a presença do anticorpo ligado ao dúplex pode ser detectada. A utilização de sondas para o novo ARN anti-sentido poderá ser aplicada em quaisquer técnicas convencionais tais como a hibridação de ácidos nucleicos, o rastreio positivo e negativo, a utilização de sondas recombinatórias, a técnica de "hybrid-released translation" (HRT) e a técnica de "hybrid-arrested translation" (HART). Isto também inclui técnicas de amplificação como a reacção em cadeia pela polimerase (PCR). Outra abordagem utiliza, por exemplo, ácidos nucleicos anti-sentido, incluindo a introdução de ARN de cadeia dupla (ARNcd) para interferir geneticamente com a função génica como

descrito, por exemplo, em Misquitta *et al.* (1999) *Proc. Nat'l. Acad. Sci. USA* 96:1451-1456, e/ou de ribozimas para bloquear a tradução de um ARNm de IL-70 específico. A utilização de métodos anti-sentido para inibir a tradução de genes *in vitro* é bem conhecida na arte. Marcus-Sakura (1988) *Anal. Biochem.* 172:289; Akhtar (ed. 1995) *Delivery Strategies for Antisense Oligonucleotide Therapeutics*, CRC Press, Inc.

Os *kits* de diagnóstico que também testam a presença qualitativa ou quantitativa de outros marcadores estão igualmente contemplados. O diagnóstico ou prognóstico poderá depender da combinação de múltiplas indicações utilizadas como marcadores. Assim, os *kits* poderão testar combinações de marcadores. Ver, por exemplo, Viallet *et al.* (1989) *Progress in Growth Factor Res.* 1:89-97.

O âmbito alargado deste invento será melhor compreendido por meio dos exemplos que se seguem, que não se destinam a limitar o invento a concretizações específicas.

EXEMPLOS

I. Métodos gerais

Alguns dos métodos padrão estão descritos ou referenciados, por exemplo, em Maniatis *et al.* (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed.), vols. 1-3, CSH Press, NY; Ausubel *et al.*, *Biology*, Greene Publishing Associates, Brooklyn, NY; ou Ausubel *et al.* (1987 & Supplements) *Current Protocols in Molecular Biology*, Greene/Wiley, New York; Innis *et al.* (eds. 1990) *PCR Protocols: A Guide to Methods and Applications*, Academic Press, N.Y.; e Kohler *et al.* (1995) *Quantitation of mRNA by Polymerase Chain Reaction*, Springer-Verlag, Berlin. Os métodos para a purificação de proteínas incluem métodos como a precipitação com sulfato de amónio, a cromatografia em coluna, a electroforese, a centrifugação, a cristalização e outros. Consultar, por exemplo, Ausubel *et al.* (1987 e suplementos periódicos); Deutscher (1990) "Guide to Protein Purification" in *Methods in Enzymology*, vol. 182 e outros volumes nesta série; e a literatura do fabricante sobre utilização de

produtos para purificação de proteínas, por exemplo, Pharmacia, Piscataway, N.J., ou Bio-Rad, Richmond, CA. A combinação com técnicas recombinantes permite a fusão com segmentos apropriados, por exemplo, com uma sequência FLAG ou uma sequência equivalente que pode ser fundida via uma sequência removível por proteases. Ver, por exemplo, Hochuli (1989) *Chemische Industrie* 12:69-70; Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) *Genetic Engineering, Principle and Methods* 12:87-98, Plenum Press, N.Y.; e Crowe *et al.* (1992) *QIAexpress: The High Level Expression & Protein Purification System* QUIAGEN, Inc., Chatsworth, CA.

As técnicas imunológicas padrão estão descritas, por exemplo, em Hertzzenberg *et al.* (eds. 1996) *Weir's Handbook of Experimental Immunology* vols. 1-4, Blackwell Science; Coligan (1991) *Current Protocols in Immunology*, Wiley/Greene, NY; e *Methods in Enzymology* vols. 70, 73, 74, 84, 92, 93, 108, 116, 121, 132, 150, 162 & 163. Os ensaios de citocinas estão descritos, por exemplo, em Thomson (ed. 1998) *The Cytokine Handbook* (3d ed.) Academic Press, San Diego; Mire-Sluis & Thorpe (1998) *Cytokines*, Academic Press, San Diego; Metcalf & Nicola (1995) *The Hematopoietic Colony Stimulating Factors*, Cambridge University Press; e Aggarwal & Gutterman (1991) *Human Cytokines*, Blackwell Pub.

Os ensaios para as actividades biológicas vasculares são bem conhecidos na arte. Eles abrangerão as actividades angiogénica e angiostática em tumores ou outros tecidos, por exemplo, a proliferação do músculo liso arterial (ver, por exemplo, Koyoma *et al.* (1996) *Cell* 87:1069-1078), a adesão de monócitos ao epitélio vascular (ver McEvoy *et al.* (1997) *J. Exp. Med.* 185:2069-2077), etc. Ver também Ross (1993) *Nature* 362:801-809; Rekhter & Gordon (1995) *Am. J. Pathol.* 147:668-677; Thyberg *et al.* (1990) *Atherosclerosis* 10:966-990; e Gumbiner (1996) *Cell* 84:345-357.

Os ensaios para as actividades biológicas das células neurais estão descritos, por exemplo, em Wouterlood (ed. 1995) *Neuroscience Protocols* modules 10, Elsevier; *Methods in Neurosciences*, Academic Press; e *Neuromethods* Humana Press, Totowa, NJ. A metodologia dos sistemas do desenvolvimento está

descrita, por exemplo, em Meisami (ed.) *Handbook of Human Growth and Developmental Biology*, CRC Press; e Chrispeels (ed.) *Molecular Techniques and Approches in Developmental Biology*, Interscience.

A análise computacional das sequências é efectuada, por exemplo, utilizando programas de *software* disponíveis, incluindo aqueles do GCG (U. Wisconsin) e as fontes do GenBank. As bases de dados públicas de sequências também foram utilizadas, por exemplo, GenBank e outras.

Muitas técnicas aplicáveis a IL-170 poderão ser aplicadas a estas novas entidades, como descrito, por exemplo, em USSN.

As análises de FACS estão descritas em Melamed *et al.* (1990) *Flow Cytometry and Sorting*, Wiley-Liss, Inc., New York, NY; Shapiro (1988) *Practical Flow Cytometry*, Liss, New York, NY; e Robinson *et al.* (1993) *Handbook of Flow Cytometry Methods*, Wiley-Liss, New York, NY.

II. Isolamento de um clone de ADN que codifica a proteína IL-170

O isolamento de CTLA-8 murina está descrito em Rouvier *et al.* (1993) *J. Immunol.* 150:5445-5456. Estão disponíveis métodos similares para isolar homólogos entre espécies das proteínas IL-173, IL-174, IL-176 e IL-177, juntamente com as proteínas IL-171, IL-172 e IL-175.

Fonte das mensagens de IL-170

Várias linhas celulares são rastreadas com uma sonda apropriada quanto a uma expressão de nível elevado do mensageiro. As linhas celulares apropriadas são seleccionadas com base nos níveis de expressão da mensagem de IL-170 adequada.

Isolamento de um clone que codifica IL-170

Utilizam-se técnicas de PCR padrão para amplificar uma sequência génica de IL-170 a partir de uma biblioteca genómica ou de ADNc, ou a partir de ARNm. Uma biblioteca genómica ou de ADNc humana é obtida e rastreada com uma sonda sintética ou de ADNc apropriada. Iniciadores de PCR poderão ser preparados. Os

iniciadores adequados são seleccionados, por exemplo, entre as sequências proporcionadas, e um clone de comprimento total é isolado. Várias combinações de iniciadores, com vários comprimentos e eventualmente com diferenças ao nível da sequência, poderão ser preparadas. O clone de comprimento total pode ser utilizado como sonda de hibridação para pesquisar outros genes homólogos, utilizando condições de hibridação rigorosas ou menos rigorosas.

Em outro método, utilizam-se oligonucleótidos para rastrear uma biblioteca. Em combinação com técnicas de reacção da polimerase em cadeia (PCR), utilizam-se oligonucleótidos sintéticos, em orientações apropriadas, como iniciadores para seleccionar os clones correctos a partir de uma biblioteca.

III. Caracterização bioquímica de proteínas IL-170

Uma proteína IL-170 é expressa em células heterólogas, por exemplo, a forma nativa ou uma forma recombinante que apresenta o péptido FLAG na extremidade carboxilo. Ver, por exemplo, Crowe *et al.* (1992) OIAexpress: The High Level Expression and Protein Purification System QIAGEN, Inc. Chatsworth, CA; e Hopp *et al.* (1988) Bio/Technology 6:1204-1210. Estas duas formas são introduzidas em vectores de expressão, por exemplo, pME18S ou pEE12, e subsequentemente transfectadas em células apropriadas, por exemplo, células COS-7 ou NSO, respectivamente. As células electroporadas são cultivadas, por exemplo, durante 48 horas em meio RPMI suplementado com soro fetal de bovino a 10%. As células são depois incubadas com ³⁵S-Met e ³⁵S-Cys, de forma a marcar as proteínas celulares. A comparação das proteínas sob condições redutoras em SDS-PAGE deverá mostrar que as células transfectadas com os clones de comprimento total segregam um polipéptido com o tamanho apropriado, por exemplo, cerca de 15 000 daltons. O tratamento com endoglicosidasas demonstrará se há ou não formas N-glicosiladas.

IV. Produção e purificação em larga escala de proteínas IL-170

Para os ensaios biológicos, a proteína IL-170 de mamífero é produzida em grandes quantidades, por exemplo, com células COS-7 transfectadas e crescidas em meio RPMI suplementado com Nutridoma HU a 1% (Boehringer Mannheim, Mannheim, Alemanha),

sendo subsequentemente purificada. A purificação poderá utilizar a cromatografia de afinidade com anticorpos ou técnicas de purificação de proteínas, por exemplo, recorrendo a anticorpos para determinar as propriedades de separação.

De modo a produzir quantidades maiores de proteínas nativas, é possível preparar transformantes estáveis de células NSO de acordo com a metodologia desenvolvida por Celltech (Slough, Berkshire, Reino Unido; Pedidos de patente internacional W086/05807, W087/04462, W089/01036 e W089/10404).

Tipicamente, 1 litro de sobrenadante contendo IL-173 ou IL-173-FLAG humana é passado, por exemplo, por uma coluna de 60 ml de iões Zn^{2+} enxertados numa matriz quelante de Sepharose Fast Flow (Pharmacia, Upsalla, Suécia). Após lavagem com 10 volumes de tampão de ligação (Kit His-Bind Buffer, Novagen, Madison, WI), as proteínas retidas pelos iões metálicos são eluídas com um gradiente de imidazole 20-100 mM. O teor da proteína IL-173-FLAG humana nas fracções eluídas é determinado por "dot blot", utilizando o anticorpo monoclonal anti-FLAG M2 (Eastman Kodak, New Haven, CT), ao passo que o teor de IL-173 humana é avaliado, por exemplo, por coloração de um gel de SDS-PAGE não redutor com prata. As fracções contendo IL-170 são depois combinadas e dialisadas contra PBS, sendo subsequentemente utilizadas em ensaios biológicos ou adicionalmente purificadas, por exemplo, por HPLC de permuta aniónica numa coluna de DEAE. Um terceiro passo de cromatografia de filtração em gel poderá ser efectuado numa coluna SUPERDEX G-75 HRD30 (Pharmacia, Uppsala, Suécia). A purificação poderá ser avaliada, por exemplo, por SDS-PAGE corado com prata.

V. Preparação de anticorpos contra IL-173

Ratinhos Balb/c consanguíneos são imunizados intraperitonealmente, por exemplo, com 1 ml de IL-173-FLAG humana purificada, emulsionada em adjuvante completo de Freund no dia 0, e em adjuvante incompleto de Freund nos dias 15 e 22. Os ratinhos são reforçados com 0,5 ml de IL-173 humana purificada, administrada intravenosamente.

O anti-soro policlonal é recolhido. O soro pode ser purificado em relação aos anticorpos. Os anticorpos podem ser subsequentemente processados, por exemplo, para Fab, Fab2, Fv ou fragmentos similares.

Criam-se hibridomas utilizando, por exemplo, a linha de células de mieloma não secretoras SP2/0-Ag8 e polietilenoglicol 1000 (Sigma, St. Louis, MO) como agente de fusão. As células do hibridoma são colocadas numa placa de cultura de tecidos Falcon de 96 poços (Becton Dickinson, NJ) e alimentadas com DMEM F12 (Gibco, Gaithersburg, MD) suplementado com gentamicina 80 µg/ml, glutamina 2 mM, soro de cavalo a 10% (Gibco, Gaithersburg, MD), ADCM a 1% (CRTS, Lyon, França), azasserina 10^{-5} M (Sigma, St. Louis, MO) e hipoxantina 5×10^{-5} M. Os sobrenadantes do hibridoma são pesquisados quanto a produção do anticorpo contra a proteína IL-173 humana por imunocitoquímica, utilizando células COS-7 transfectadas com IL-173 humana e fixadas com acetona, e por ELISA, utilizando IL-173-FLAG humana purificada a partir dos sobrenadantes de células COS-7 como um antigénio de revestimento. Alíquotas de clones de células positivas são expandidas durante 6 dias e criopreservadas, bem como propagadas no líquido ascítico de ratinhos Balb/c tratados com pristano (2,6,10,14-tetrametilpentadecano, Sigma, St. Louis, MO), que haviam recebido uma injeção intraperitoneal de pristano 15 dias antes. Tipicamente, administram-se intraperitonealmente cerca de 10^5 células de hibridoma em 1 ml de PBS e, 10 dias mais tarde, o líquido ascítico é recolhido de cada ratinho.

Após centrifugação do líquido ascítico, a fracção de anticorpos é isolada por precipitação com sulfato de amónio e cromatografia de permuta aniónica numa coluna de silício Zephyr-D (IBF Sepracor) equilibrada com Tris 20 mM, pH 8,0. As proteínas são eluídas com um gradiente de NaCl (no intervalo de NaCl 0 a 1 M). Recolhem-se fracções de 2 ml, que são testadas por ELISA quanto à presença de anticorpo anti-IL-173. As fracções contendo actividade específica anti-IL-173 são combinadas, dialisadas e congeladas. Alíquotas dos anticorpos monoclonais purificados poderão ser marcadas com peroxidase.

As preparações dos anticorpos, policlonais ou monoclonais, poderão ser absorvidas de forma cruzada, esgotadas ou combinadas para criar reagentes que apresentam as combinações desejadas de selectividades e especificidades. Antigénios específicos definidos podem ser imobilizados numa matriz sólida e utilizados para esgotar ou escolher selectivamente as capacidades de ligação desejadas.

VI. Quantificação de IL-173 humana

Entre os anticorpos específicos para IL-173, seleccionam-se os isolados clonais apropriados para quantificar os níveis de IL-173 humana recorrendo a um ensaio de tipo sanduíche. Os anticorpos purificados são diluídos, por exemplo, para 2 µg/ml em tampão de revestimento (tampão carbonato, pH 9,6; Na₂CO₃ 15 mM; NaHCO₃ 35 mM). Esta solução diluída é revestida sobre os poços de uma placa ELISA de 96 poços (Immunoplate Maxisorp F96 certificada, NUNC, Dinamarca) durante a noite à temperatura ambiente. As placas são depois lavadas manualmente, por exemplo, com um tampão de lavagem que consiste em solução salina tamponada com fosfato e Tween 20 a 0,05% (Technicon Diagnostics, EUA). Adicionam-se a cada poço 110 µl de CTLA-8 humana purificada, diluída em tampão TBS-B-T [Tris 20 mM, NaCl 150 mM, BSA a 1% (Sigma, St. Louis, MO) e Tween 20 a 0,05%]. Após 3 horas de incubação a 37°C, as placas são lavadas uma vez. Adicionam-se 100 µl de Ac marcado com peroxidase, diluído para 5 µg/ml em tampão TBS-B-T, a cada poço e incuba-se durante 2 horas a 37°C. Os poços são depois lavados três vezes com tampão de lavagem. Adicionam-se a cada poço 100 µl do substrato da peroxidase, 2,2'-azino-bis(ácido 3-etilbenzotiazoino-6-sulfónico) (ABTS), diluído para 1 mg/ml em tampão citrato/fosfato, e a reacção colorimétrica é lida a 405 nm.

VII. Distribuição dos genes de IL-170

A proteína IL-173 humana foi identificada a partir de uma sequência derivada de uma biblioteca de ADNc do córtex frontal do cérebro de um epiléptico. A proteína IL-173 de rato foi derivada de uma biblioteca de ADNc de cóclea, cérebro, cerebelo, olho, pulmão e rim. Mais uma vez, os genes parecem ser bastante raros, o que sugere que as distribuições da expressão seriam extremamente restringidas.

A proteína IL-174 de ratinho foi identificada a partir de uma sequência obtida de uma biblioteca de ADNc derivada de um embrião de ratinho. O gene parece ser bastante raro, o que sugere que a distribuição da expressão seria extremamente restringida.

A proteína IL-171 humana foi identificada a partir de uma sequência derivada de um linfócito T apoptótico. O gene parece ser bastante raro, o que sugere que a distribuição da expressão seria extremamente restringida.

A proteína IL-172 humana foi identificada a partir de sequências derivadas do coração, fígado e baço, timo e tumor do timo de feto humano e do feto total. A proteína de ratinho foi obtida de sequências derivadas de embrião, glândula mamária e órgãos combinados de ratinho. Ambos os genes parecem ser bastante raros, o que sugere que a sua distribuição da expressão seria extremamente restringida.

A proteína IL-175 humana foi identificada a partir de uma sequência derivada de um linfócito T activado durante 12 horas com tiouridina. O gene parece ser bastante raro, o que sugere que a distribuição da expressão seria extremamente restringida.

VIII. Mapeamento cromossómico dos genes de IL-170

Utiliza-se um ADNc isolado que codifica o gene de IL-170 apropriado. O mapeamento cromossómico é uma técnica padrão. Ver, por exemplo, BIOS Laboratories (New Haven, CT) e os métodos para utilização de um painel de híbridos de células somáticas de ratinho com PCR.

O gene de IL-173 humana localiza-se no cromossoma humano 13q11.

IX. Isolamento de homólogos de IL-170

Uma composição de ligação, por exemplo, um anticorpo, é utilizada para rastrear uma biblioteca de expressão preparada a partir de uma linha celular que expressa uma proteína IL-170. Utilizam-se técnicas de coloração padrão para detectar

ou separar o antigénio intracelular ou expresso à superfície, ou as células transformadas que expressam à superfície são rastreadas mediante ensaios de "panning". O rastreio da expressão intracelular é efectuado por vários procedimentos de coloração ou imunofluorescência. Ver também McMahan *et al.* (1991) EMBO J. 10:2821-2832.

Métodos idênticos são aplicáveis para isolar as variantes de espécies ou as variantes alélicas. As variantes de espécies são isoladas utilizando técnicas de hibridação de espécies cruzadas baseadas num isolado ou fragmento de comprimento total de uma das espécies como sonda, ou espécies apropriadas.

X. Isolamento de receptores para IL-170

Estão disponíveis métodos para rastrear uma biblioteca de expressão preparada a partir de uma linha celular que expressa potenciais receptores de IL-170. Um ligando IL-170 marcado é produzido como descrito acima. Utilizam-se técnicas de coloração padrão para detectar ou separar o receptor expresso à superfície, ou as células transformadas que expressam à superfície são rastreadas mediante ensaios de "panning". Ver também McMahan *et al.* (1991) EMBO J. 10:2821-2832.

Por exemplo, no dia 0, efectue um pré-revestimento de lâminas Permanox de 2 câmaras com 1 ml de fibronectina 10 ng/ml em PBS por câmara, durante 30 min à temperatura ambiente. Enxagúe uma vez com PBS. Em seguida, plaqueie as células COS, numa quantidade de $2-3 \times 10^5$ células por câmara, em 1,5 ml de meio de crescimento. Incube durante a noite a 37°C.

No dia 1, para cada amostra, prepare 0,5 ml de uma solução de DEAE-dextrano 66 µg/ml, cloroquina 66 µM e 4 µg de ADN em DME isento de soro. Para cada conjunto, prepare um controlo positivo, por exemplo, de ADNc de IL-170-FLAG humana numa diluição de 1 e 1/200, e uma prova negativa. Enxagúe as células com DME isento de soro. Adicione a solução de ADN e incube durante 5 h a 37°C. Retire o meio e adicione 0,5 ml de DMSO a 10% em DME durante 2,5 min. Retire e lave uma vez com DME. Adicione 1,5 ml de meio de crescimento e incube durante a noite.

No dia 2, mude o meio. Nos dias 3 ou 4, as células são fixadas e coradas. Enxagúe as células duas vezes com solução salina tamponada de Hank (HBSS) e fixe-as em paraformaldeído (PFA) a 4%/glucose durante 5 min. Lave 3x com HBSS. As lâminas poderão ser armazenadas a -80°C após remoção de todo o líquido. Para cada câmara, efectuem-se incubações de 0,5 ml da seguinte forma. Adicione HBSS/saponina (0,1%) com 32 μl /ml de NaN_3 1M durante 20 min. As células são depois lavadas com HBSS/saponina 1x. Adicione anticorpo solúvel às células e incube durante 30 min. Lave as células duas vezes com HBSS/saponina. Adicione o segundo anticorpo, por exemplo, anticorpo anti-ratinho Vector, numa diluição de 1/200, e incube durante 30 min. Prepare a solução ELISA, por exemplo, solução de peroxidase de rábano Vector Elite ABC e pré-incube durante 30 min. Utilize, por exemplo, 1 gota de solução A (avidina) e 1 gota de solução (B) (biotina) por 2,5 ml de HBSS/saponina. Lave as células duas vezes com HBSS/saponina. Adicione a solução HRP ABC e incube durante 30 min. Lave as células duas vezes com HBSS, faça uma segunda lavagem durante 2 min, o que fecha as células. Em seguida, adicione ácido diaminobenzóico Vector (DAB) durante 5 a 10 min. Utilize 2 gotas de tampão mais 4 gotas de DAB mais 2 gotas de H_2O_2 por 5 ml de água destilada em vidro. Retire cuidadosamente a câmara e enxagúe a lâmina com água. Seque ao ar durante alguns minutos, adicione 1 gota de Crystal Mount e uma lamela. Coza durante 5 min a $85-90^{\circ}\text{C}$.

Em alternativa, o ligando marcado é utilizado para purificar ou separar por afinidade as células que expressam o receptor. Ver, por exemplo, Sambrook *et al.* ou Ausubel *et al.*

É possível efectuar muitas modificações e variações deste invento. As concretizações específicas aqui descritas são oferecidas apenas a título de exemplo, e o invento será limitado apenas pelos termos das reivindicações anexas.

LISTAGEM DE SEQUÊNCIAS

A SEQ ID NO:1 é a sequência de ácido nucleico de IL-172 de primata.

A SEQ ID NO:2 é a sequência polipeptídica de IL-172 de primata.

A SEQ ID NO:3 é a sequência de ácido nucleico de IL-172 murina.

A SEQ ID NO:4 é a sequência polipeptídica de IL-172 murina.

A SEQ ID NO:5 é a sequência de ácido nucleico de IL-173 de primata.

A SEQ ID NO:6 é a sequência polipeptídica de IL-173 de primata.

A SEQ ID NO:7 é uma sequência de ácido nucleico suplementar de IL-173 de primata.

A SEQ ID NO:8 é uma sequência polipeptídica suplementar de IL-173 de primata.

A SEQ ID NO:9 é a sequência de ácido nucleico de IL-173 murina.

A SEQ ID NO:10 é a sequência polipeptídica de IL-173 murina.

A SEQ ID NO:11 é uma sequência de ácido nucleico suplementar de IL-173 murina.

A SEQ ID NO:12 é uma sequência polipeptídica suplementar de IL-173 murina.

A SEQ ID NO:13 é a sequência de ácido nucleico de IL-174 de primata.

A SEQ ID NO:14 é a sequência polipeptídica de IL-174 de primata.

A SEQ ID NO:15 é a sequência de ácido nucleico de IL-174 murina.

A SEQ ID NO:16 é a sequência polipeptídica de IL-174 murina.

A SEQ ID NO:17 é uma sequência de ácido nucleico suplementar de IL-174 murina.

A SEQ ID NO:18 é uma sequência polipeptídica suplementar de IL-174 murina.

A SEQ ID NO:19 é a sequência de ácido nucleico IUPAC de IL-171 de primata.

A SEQ ID NO:20 é a sequência de ácido nucleico de IL-171 de primata.

A SEQ ID NO:21 é a sequência polipeptídica de IL-171 de primata.

A SEQ ID NO:22 é uma sequência de ácido nucleico suplementar de IL-171 de primata.

A SEQ ID NO:23 é uma sequência polipeptídica suplementar de IL-171 de primata.

A SEQ ID NO:24 é a sequência de ácido nucleico IUPAC de IL-175 de primata.

A SEQ ID NO:25 é a sequência de ácido nucleico de IL-175 de primata.

A SEQ ID NO:26 é a sequência polipeptídica de IL-175 de primata.

A SEQ ID NO:27 é a sequência de ácido nucleico de IL-176 de primata.

A SEQ ID NO:28 é a sequência polipeptídica de IL-176 de primata.

A SEQ ID NO:29 é a sequência de ácido nucleico de IL-177 de primata.

A SEQ ID NO:30 é a sequência polipeptídica de IL-177 de primata.

A SEQ ID NO:31 é a sequência polipeptídica de CTLA-8 de rato.

A SEQ ID NO:32 é a sequência polipeptídica de CTLA-8 de ratinho.

A SEQ ID NO:33 é a sequência polipeptídica de CTLA-8 de primata.

A SEQ ID NO:34 é a sequência polipeptídica de CTLA-8 viral.

<110> Schering Corporation

<120> Citoquinas de Mamífero Purificadas; Reagentes e Métodos
Relacionados

<130> DX0917K

<140>

<141>

<150> USSN 09/228,822

<151> 1999-01-11

<160> 34

<170> PatentIn Ver. 2.1

<210> 1

<211> 543

<212> ADN

<213> primata

<220>

<221> CDS

<222> (1)..(540)

<220>

<221> mat_peptide

<222> (61)..(540)

<400> 1

```

atg gac tgg cct cac aac ctg ctg ttt ctt ctt acc att tcc atc ttc 48
Met Asp Trp Pro His Asn Leu Leu Phe Leu Leu Thr Ile Ser Ile Phe
-20 -15 -10 -5

ctg ggg ctg ggc cag ccc agg agc ccc aaa agc aag agg aag ggg caa 96
Leu Gly Leu Gly Gln Pro Arg Ser Pro Lys Ser Lys Arg Lys Gly Gln
-1 1 5 10

ggg cgg cct ggg ccc ctg gtc cct ggc cct cac cag gtg cca ctg gac 144
Gly Arg Pro Gly Pro Leu Val Pro Gly Pro His Gln Val Pro Leu Asp
15 20 25

ctg gtg tca cgg atg aaa ccg tat gcc cgc atg gag gag tat gag agg 192
Leu Val Ser Arg Met Lys Pro Tyr Ala Arg Met Glu Glu Tyr Glu Arg
30 35 40

aac atc gag gag atg gtg gcc cag ctg agg aac agc tca gag ctg gcc 240
Asn Ile Glu Glu Met Val Ala Gln Leu Arg Asn Ser Ser Glu Leu Ala
45 50 55 60

cag aga aag tgt gag gtc aac ttg cag ctg tgg atg tcc aac aag agg 288
Gln Arg Lys Cys Glu Val Asn Leu Gln Leu Trp Met Ser Asn Lys Arg
65 70 75

agc ctg tct ccc tgg ggc tac agc atc aac cac gac ccc agc cgt atc 336
Ser Leu Ser Pro Trp Gly Tyr Ser Ile Asn His Asp Pro Ser Arg Ile
80 85 90

ccc gtg gac ctg ccg gag gca cgg tgc ctg tgt ctg ggc tgt gtg aac 384
Pro Val Asp Leu Pro Glu Ala Arg Cys Leu Cys Leu Gly Cys Val Asn
95 100 105

ccc ttc acc atg cag gag gac cgc agc atg gtg agc gtg ccg gtg ttc 432
Pro Phe Thr Met Gln Glu Asp Arg Ser Met Val Ser Val Pro Val Phe
110 115 120

agc cag gtt cct gtg cgc cgc cgc ctc tgc ccg cca ccg ccc cgc aca 480
Ser Gln Val Pro Val Arg Arg Arg Leu Cys Pro Pro Pro Arg Thr
125 130 135 140

ggg cct tgc cgc cag cgc gca gtc atg gag acc atc gct gtg ggc tgc 528
Gly Pro Cys Arg Gln Arg Ala Val Met Glu Thr Ile Ala Val Gly Cys
145 150 155

acc tgc atc ttc tga 543
Thr Cys Ile Phe
160

```

<210> 2

<211> 180

<212> PRT

<213> primata

<400> 2

```

Met Asp Trp Pro His Asn Leu Leu Phe Leu Leu Thr Ile Ser Ile Phe
-20          -15          -10          -5

Leu Gly Leu Gly Gln Pro Arg Ser Pro Lys Ser Lys Arg Lys Gly Gln
      -1  1          5          10

Gly Arg Pro Gly Pro Leu Val Pro Gly Pro His Glh Val Pro Leu Asp
      15          20          25

Leu Val Ser Arg Met Lys Pro Tyr Ala Arg Met Glu Glu Tyr Glu Arg
      30          35          40

Asn Ile Glu Glu Met Val Ala Gln Leu Arg Asn Ser Ser Glu Leu Ala
      45          50          55          60

Gln Arg Lys Cys Glu Val Asn Leu Gln Leu Trp Met Ser Asn Lys Arg
      65          70          75

Ser Leu Ser Pro Trp Gly Tyr Ser Ile Asn His Asp Pro Ser Arg Ile
      80          85          90

Pro Val Asp Leu Pro Glu Ala Arg Cys Leu Cys Leu Gly Cys Val Asn
      95          100          105

Pro Phe Thr Met Gln Glu Asp Arg Ser Met Val Ser Val Pro Val Phe
      110          115          120

Ser Gln Val Pro Val Arg Arg Arg Leu Cys Pro Pro Pro Pro Arg Thr
      125          130          135          140

Gly Pro Cys Arg Gln Arg Ala Val Met Glu Thr Ile Ala Val Gly Cys
      145          150          155

Thr Cys Ile Phe
      160

```

<210> 3

<211> 543

<212> ADN

<213> roedor

<220>

<221> CDS

<222> (1)..(540)

<220>

<221> mat_peptide

<222> (67)..(540)

<400> 3

```

atg gac tgg ccg cac agc ctg ctc ttc ctc ctg gcc atc tcc atc ttc   48
Met Asp Trp Pro His Ser Leu Leu Phe Leu Leu Ala Ile Ser Ile Phe
      -20          -15          -10

ctg gcg cca agc cac ccc cgg aac acc aaa ggc aaa aga aaa ggg caa   96
Leu Ala Pro Ser His Pro Arg Asn Thr Lys Gly Lys Arg Lys Gly Gln
      -5          -1  1          5          10

```

```

ggg agg ccc agt ccc ttg gcc cct ggg cct cat cag gtg ccg ctg gac 144
Gly Arg Pro Ser Pro Leu Ala Pro Gly Pro His Gln Val Pro Leu Asp
                      15                      20                      25

ctg gtg tct cga gta aag ccc tac gct cga atg gaa gag tat gag cgg 192
Leu Val Ser Arg Val Lys Pro Tyr Ala Arg Met Glu Glu Tyr Glu Arg
                      30                      35                      40

aac ctt ggg gag atg gtg gcc cag ctg agg aac agc tcc gag cca gcc 240
Asn Leu Gly Glu Met Val Ala Gln Leu Arg Asn Ser Ser Glu Pro Ala
                      45                      50                      55

aag aag aaa tgt gaa gtc aat cta cag ctg tgg ttg tcc aac aag agg 288
Lys Lys Lys Cys Glu Val Asn Leu Gln Leu Trp Leu Ser Asn Lys Arg
                      60                      65                      70

agc ctg tcc cca tgg ggc tac agc atc aac cac gac ccc agc cgc atc 336
Ser Leu Ser Pro Trp Gly Tyr Ser Ile Asn His Asp Pro Ser Arg Ile
                      75                      80                      85                      90

cct gcg gac ttg ccc gag gcg cgg tgc cta tgt ttg ggt tgc gtg aat 384
Pro Ala Asp Leu Pro Glu Ala Arg Cys Leu Cys Leu Gly Cys Val Asn
                      95                      100                      105

ccc ttc acc atg cag gag gac cgt agc atg gtg agc gtg cca gtg ttc 432
Pro Phe Thr Met Gln Glu Asp Arg Ser Met Val Ser Val Pro Val Phe
                      110                      115                      120

agc cag gtg ccg gtg cgc cgc cgc ctc tgt cct caa cct cct cgc cct 480
Ser Gln Val Pro Val Arg Arg Arg Leu Cys Pro Gln Pro Pro Arg Pro
                      125                      130                      135

ggg ccc tgc cgc cag cgt gtc gtc atg gag acc atc gct gtg ggt tgc 528
Gly Pro Cys Arg Gln Arg Val Val Met Glu Thr Ile Ala Val Gly Cys
                      140                      145                      150

acc tgc atc ttc tga 543
Thr Cys Ile Phe
155

```

<210> 4

<211> 180

<212> PRT

<213> roedor

<400> 4

```

Met Asp Trp Pro His Ser Leu Leu Phe Leu Leu Ala Ile Ser Ile Phe
-20                      -15                      -10

```

```

Leu Ala Pro Ser His Pro Arg Asn Thr Lys Gly Lys Arg Lys Gly Gln
-5                      -1 1                      5                      10

```

```

Gly Arg Pro Ser Pro Leu Ala Pro Gly Pro His Gln Val Pro Leu Asp
.15                      20                      25

```

```

Leu Val Ser Arg Val Lys Pro Tyr Ala Arg Met Glu Glu Tyr Glu Arg
30                      35                      40

```

```

Asn Leu Gly Glu Met Val Ala Gln Leu Arg Asn Ser Ser Glu Pro Ala

```

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

<210> 5
 <211> 454
 <212> ADN
 <213> primata
 <220>
 <221> CDS
 <222> (1)..(453)

<400> 5	
tgc gcg gac ccg ccg gag gag cta ctg gag cag ctg tac ggg cgc ctg	48
Cys Ala Asp Arg Pro Glu Glu Leu Leu Gln Leu Tyr Gly Arg Leu	
1 5 10 15	
gcg gcc ggc gtg ctc agt gcc ttc cac cac acg ctg cag ctg ggg ccg	96
Ala Ala Gly Val Leu Ser Ala Phe His His Thr Leu Gln Leu Gly Pro	
20 25 30	
cgt gag cag gcg cgc aac gcg agc tgc ccg gca ggg ggc agg ccc gcc	144
Arg Glu Gln Ala Arg Asn Ala Ser Cys Pro Ala Gly Gly Arg Pro Ala	
35 40 45	
gac cgc cgc ttc ccg acg ccc acc aac ctg cgc agc gtg tgc ccc tgg	192
Asp Arg Arg Phe Arg Thr Pro Thr Asn Leu Arg Ser Val Ser Pro Trp	
50 55 60	
gcc tac aga atc tcc tac gac ccg gcg agg tac ccc agg tac ctg cct	240
Ala Tyr Arg Ile Ser Tyr Asp Pro Ala Arg Tyr Pro Arg Tyr Leu Pro	
65 70 75 80	
gaa gcc tac tgc ctg tgc ccg ggc tgc ctg acc ggg ctg ttc ggc gag	288
Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu Thr Gly Leu Phe Gly Glu	
85 90 95	
gag gac gtg cgc ttc cgc agc gcc cct gtc tac atg ccc acc gtc gtc	336
Glu Asp Val Arg Phe Arg Ser Ala Pro Val Tyr Met Pro Thr Val Val	
100 105 110	

```

ctg cgc cgc acc ccc gcc tgc gcc gcc gcc cgt tcc gtc tac acc gag 384
Leu Arg Arg Thr Pro Ala Cys Ala Gly Gly Arg Ser Val Tyr Thr Glu
      115          120          125

gcc tac gtc acc atc ccc gtg gcc tgc acc tgc gtc ccc gag ccg gag 432
Ala Tyr Val Thr Ile Pro Val Gly Cys Thr Cys Val Pro Glu Pro Glu
      130          135          140

aag gac gca gac agc atc aac t. 454
Lys Asp Ala Asp Ser Ile Asn
145          150

```

<210> 6
 <211> 151
 <212> PRT
 <213> primata

```

<400> 6
Cys Ala Asp Arg Pro Glu Glu Leu Leu Glu Gln Leu Tyr Gly Arg Leu
  1          5          10          15

Ala Ala Gly Val Leu Ser Ala Phe His His Thr Leu Gln Leu Gly Pro
      20          25          30

Arg Glu Gln Ala Arg Asn Ala Ser Cys Pro Ala Gly Gly Arg Pro Ala
      35          40          45

Asp Arg Arg Phe Arg Thr Pro Thr Asn Leu Arg Ser Val Ser Pro Trp
      50          55          60

Ala Tyr Arg Ile Ser Tyr Asp Pro Ala Arg Tyr Pro Arg Tyr Leu Pro
      65          70          75          80

Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu Thr Gly Leu Phe Gly Glu
      85          90          95

Glu Asp Val Arg Phe Arg Ser Ala Pro Val Tyr Met Pro Thr Val Val
      100          105          110

Leu Arg Arg Thr Pro Ala Cys Ala Gly Gly Arg Ser Val Tyr Thr Glu
      115          120          125

Ala Tyr Val Thr Ile Pro Val Gly Cys Thr Cys Val Pro Glu Pro Glu
      130          135          140

Lys Asp Ala Asp Ser Ile Asn
145          150

```

<210> 7
 <211> 1385
 <212> ADN
 <213> primata

 <220>
 <221> CDS
 <222> (59)..(664)

 <220>
 <221> mat_peptide
 <222> (110)..(664)

<400> 7

gcccgggcag gtggcgacct cgctcagtcg gcttctcggt ccaagtcgcc gggtctgg 58
 atg ctg gta gcc ggc ttc ctg ctg gcg ctg ccg ccg agc tgg gcc gcg 106
 Met Leu Val Ala Gly Phe Leu Leu Ala Leu Pro Pro Ser Trp Ala Ala
 -15 -10 -5
 ggc gcc ccg agg gcg ggc agg cgc ccc gcg cgg ccg cgg ggc tgc gcg 154
 Gly Ala Pro Arg Ala Gly Arg Arg Pro Ala Arg Pro Arg Gly Cys Ala
 -1 1 5 10 15
 gac cgg ccg gag gag cta ctg gag cag ctg tac ggg cgc ctg gcg gcc 202
 Asp Arg Pro Glu Glu Leu Leu Glu Gln Leu Tyr Gly Arg Leu Ala Ala
 20 25 30
 ggc gtg ctc agt gcc ttc cac cac acg ctg cag ctg ggg ccg cgt gag 250
 Gly Val Leu Ser Ala Phe His His Thr Leu Gln Leu Gly Pro Arg Glu
 35 40 45
 cag gcg cgc aac gcg agc tgc ccg gca ggg ggc agg ccc gcc gac cgc 298
 Gln Ala Arg Asn Ala Ser Cys Pro Ala Gly Gly Arg Pro Ala Asp Arg
 50 55 60
 cgc ttc cgg ccg ccc acc aac ctg cgc agc gtg tgc ccc tgg gcc tac 346
 Arg Phe Arg Pro Pro Thr Asn Leu Arg Ser Val Ser Pro Trp Ala Tyr
 65 70 75
 aga atc tcc tac gac ccg gcg agg tac ccc agg tac ctg cct gaa gcc 394
 Arg Ile Ser Tyr Asp Pro Ala Arg Tyr Pro Arg Tyr Leu Pro Glu Ala
 80 85 90 95
 tac tgc ctg tgc cgg ggc tgc ctg acc ggg ctg ttc ggc gag gag gac 442
 Tyr Cys Leu Cys Arg Gly Cys Leu Thr Gly Leu Phe Gly Glu Glu Asp
 100 105 110
 gtg cgc ttc cgc agc gcc cct gtc tac atg ccc acc gtc gtc ctg cgc 490
 Val Arg Phe Arg Ser Ala Pro Val Tyr Met Pro Thr Val Val Leu Arg
 115 120 125
 cgc acc ccc gcc tgc gcc ggc ggc cgt tcc gtc tac acc gag gcc tac 538
 Arg Thr Pro Ala Cys Ala Gly Gly Arg Ser Val Tyr Thr Glu Ala Tyr
 130 135 140
 gtc acc atc ccc gtg ggc tgc acc tgc gtc ccc gag ccg gag aag gac 586
 Val Thr Ile Pro Val Gly Cys Thr Cys Val Pro Glu Pro Glu Lys Asp
 145 150 155
 gca gac agc atc aac tcc agc atc gac aaa cag ggc gcc aag ctc ctg 634
 Ala Asp Ser Ile Asn Ser Ser Ile Asp Lys Gln Gly Ala Lys Leu Leu
 160 165 170 175
 ctg ggc ccc aac gac gcg ccc gct ggc ccc tgaggccggt cctgccccgg 684
 Leu Gly Pro Asn Asp Ala Pro Ala Gly Pro
 180 185
 gaggtctccc cggcccgcat cccgaggcgc ccaagctgga gccgcctgga gggctcggtc 744
 ggcgacctct gaagagagtg caccgagcaa accaagtgcc ggagcaccag cgcgccttt 804

ccatggagac tcgtaagcag cttcatctga cacgggcac cctggcttgc ttttagctac 864
 aagcaagcag cgtggctgga agctgatggg aaacgacccg gcacgggcat cctgtgtgcg 924
 gcccgcatgg aggggtttgga aaagttcacg gaggtccctc gaggagcctc tcagatcggc 984
 tgctgcgggt gcagggcgtg actcaccgct ggggtgcttgc caaagagata gggacgcata 1044
 tgctttttta agcaatctaa aaataataat aagtatagcg actatatacc tactttttaa 1104
 atcaactggt ttgaatagag gcagagctat tttatattat caaatgagag ctactctgtt 1164
 acatttttta acatataaac atcgtttttt actttctctg gtagaatttt ttaaagcata 1224
 attggaatcc ttggataaat tttgtagctg gtacactctg gcctgggtct ctgaattcag 1284
 cctgtcaccg atggctgact gatgaaatgg acacgtctca totgacccac tcttccttcc 1344
 actgaaggtc ttcacgggcc tccaggcctc gtgccgaatt c 1385

<210> 8

<211> 202

<212> PRT

<213> primata

<400> 8

Met Leu Val Ala Gly Phe Leu Leu Ala Leu Pro Pro Ser Trp Ala Ala
 -15 -10 -5

Gly Ala Pro Arg Ala Gly Arg Arg Pro Ala Arg Pro Arg Gly Cys Ala
 -1 1 5 10 15

Asp Arg Pro Glu Glu Leu Leu Glu Gln Leu Tyr Gly Arg Leu Ala Ala
 20 25 30

Gly Val Leu Ser Ala Phe His His Thr Leu Gln Leu Gly Pro Arg Glu
 35 40 45

Gln Ala Arg Asn Ala Ser Cys Pro Ala Gly Gly Arg Pro Ala Asp Arg
 50 55 60

Arg Phe Arg Pro Pro Thr Asn Leu Arg Ser Val Ser Pro Trp Ala Tyr
 65 70 75

Arg Ile Ser Tyr Asp Pro Ala Arg Tyr Pro Arg Tyr Leu Pro Glu Ala
 80 85 90 95

Tyr Cys Leu Cys Arg Gly Cys Leu Thr Gly Leu Phe Gly Glu Glu Asp
 100 105 110

Val Arg Phe Arg Ser Ala Pro Val Tyr Met Pro Thr Val Val Leu Arg
 115 120 125

Arg Thr Pro Ala Cys Ala Gly Gly Arg Ser Val Tyr Thr Glu Ala Tyr
 130 135 140

Val Thr Ile Pro Val Gly Cys Thr Cys Val Pro Glu Pro Glu Lys Asp
 145 150 155

Ala Asp Ser Ile Asn Ser Ser Ile Asp Lys Gln Gly Ala Lys Leu Leu
 160 165 170 175

Leu Gly Pro Asn Asp Ala Pro Ala Gly Pro
 180 185

<210> 9
 <211> 133
 <212> ADN
 <213> roedor
 <220>
 <221> CDS
 <222> (1)..(132)

<400> 9
 ttt ccg aga tac ctg ccc gaa gcc tac tgc ctg tgc cga ggc tgt ctg 48
 Phe Pro Arg Tyr Leu Pro Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu
 1 5 10 15
 acc ggg ctc tac ggt gag gag gac ttc cgc ttt cgc agc gca ccc gtc 96
 Thr Gly Leu Tyr Gly Glu Glu Asp Phe Arg Phe Arg Ser Ala Pro Val
 20 25 30
 ttc tct ccg gcg gtg gtg ctg cgg cgc acg gcg gcc t 133
 Phe Ser Pro Ala Val Val Leu Arg Arg Thr Ala Ala
 35 40

<210> 10
 <211> 44
 <212> PRT
 <213> roedor

<400> 10
 Phe Pro Arg Tyr Leu Pro Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu
 1 5 10 15
 Thr Gly Leu Tyr Gly Glu Glu Asp Phe Arg Phe Arg Ser Ala Pro Val
 20 25 30
 Phe Ser Pro Ala Val Val Leu Arg Arg Thr Ala Ala
 35 40

<210> 11
 <211> 1143
 <212> ADN
 <213> roedor
 <220>
 <221> CDS
 <222> (1)..(615)
 <220>
 <221> mat_peptide
 <222> (73)..(615)

<400> 11
 atg ttg ggg aca ctg gtc tgg atg ctc ctc gtc ggc ttc ctg ctg gca 48
 Met Leu Gly Thr Leu Val Trp Met Leu Leu Val Gly Phe Leu Leu Ala

-20	-15	-10	
ctg gcg ccg ggc cgc gcg gcg ggc gcg	ctg agg acc ggg agg cgc ccg	96	
Leu Ala Pro Gly Arg Ala Ala Gly Ala	Leu Arg Thr Gly Arg Arg Pro		
-5	-1 1 5		
gcg cgg ccg cgg gac tgc gcg gac cgg cca gag gag ctc ctg gag cag	144		
Ala Arg Pro Arg Asp Cys Ala Asp Arg Pro Glu Glu Leu Leu Glu Gln			
10	15 20		
ctg tac ggg cgg ctg gcg gcc ggc gtg ctc agc gcc ttc cac cac acg	192		
Leu Tyr Gly Arg Leu Ala Ala Gly Val Leu Ser Ala Phe His His Thr			
25	30 35 40		
ctg cag ctc ggg ccg cgc gag cag gcg cgc aat gcc agc tgc ccg gcc	240		
Leu Gln Leu Gly Pro Arg Glu Gln Ala Arg Asn Ala Ser Cys Pro Ala			
45	50 55		
ggg gcc agg gcc gcc gac cgc cgc ttc cgg cca ccc acc aac ctg cgc	288		
Gly Gly Arg Ala Ala Asp Arg Arg Phe Arg Pro Pro Thr Asn Leu Arg			
60	65 70		
agc gtg tgc ccc tgg gcg tac agg att tcc tac gac cct gct cgc ttt	336		
Ser Val Ser Pro Trp Ala Tyr Arg Ile Ser Tyr Asp Pro Ala Arg Phe			
75	80 85		
ccg agg tac ctg ccc gaa gcc tac tgc ctg tgc cga gcc tgc ctg acc	384		
Pro Arg Tyr Leu Pro Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu Thr			
90	95 100		
ggg ctc tac ggg gag gag gac ttc cgc ttt cgc agc aca ccc gtc ttc	432		
Gly Leu Tyr Gly Glu Glu Asp Phe Arg Phe Arg Ser Thr Pro Val Phe			
105	110 115 120		
tct cca gcc gtg gtg ctg cgg cgc aca gcg gcc tgc gcg gcc gcc cgc	480		
Ser Pro Ala Val Val Leu Arg Arg Thr Ala Ala Cys Ala Gly Gly Arg			
125	130 135		
tct gtg tac gcc gaa cac tac atc acc atc cgg gtg gcc tgc acc tgc	528		
Ser Val Tyr Ala Glu His Tyr Ile Thr Ile Pro Val Gly Cys Thr Cys			
140	145 150		
gtg ccc gag ccg gac aag tcc gcg gac agt gcg aac tcc agc atg gac	576		
Val Pro Glu Pro Asp Lys Ser Ala Asp Ser Ala Asn Ser Ser Met Asp			
155	160 165		
aag ctg ctg ctg ggg ccc gcc gac agg cct gcg ggg cgc tgatgccggg	625		
Lys Leu Leu Leu Gly Pro Ala Asp Arg Pro Ala Gly Arg			
170	175 180		
gactgccgcg catggcccag cttcctgcat gcatcaggtc ccctggccct gacaaaaacc 685			
accccatgat ccctggccgc tgccctaattt ttccaaaagg acagctacat aagctttaaa 745			
tatatttttc aaagtagaca ctacatatct acaactattt tgaatagtgg cagaaactat 805			
tttcatatta gtaatttaga gcaagcatgt tgtttttaaa cttctttgat atacaagcac 865			
atcacacaca tcccgttttc ctctagtagg attcttgagt gcataattgt agtgctcaga 925			
tgaacttccct tctgctgcac tgtgccctgt ccctgagttct ctctgtggc ccaagcttac 985			
taagggtgata atgagtgtc cggatctggg cacctaaggt ctccaggtcc ctggagaggg 1045			
agggatgtgg gggggctagg aaccaagcgc ccctttgttc tttagcttat ggatgtctt 1105			
aactttataa agattaaagt ttttggtgtt attctttc 1143			

<210> 12

<211> 205

<212> PRT

<213> roedor

<400> 12

Met Leu Gly Thr Leu Val Trp Met Leu Leu Val Gly Phe Leu Leu Ala
 -20 -15 -10

Leu Ala Pro Gly Arg Ala Ala Gly Ala Leu Arg Thr Gly Arg Arg Pro
 -5 -1 1 5

Ala Arg Pro Arg Asp Cys Ala Asp Arg Pro Glu Glu Leu Leu Glu Gln
 10 15 20

Leu Tyr Gly Arg Leu Ala Ala Gly Val Leu Ser Ala Phe His His Thr
 25 30 35 40

Leu Gln Leu Gly Pro Arg Glu Gln Ala Arg Asn Ala Ser Cys Pro Ala
 45 50 55

Gly Gly Arg Ala Ala Asp Arg Arg Phe Arg Pro Pro Thr Asn Leu Arg
 60 65 70

Ser Val Ser Pro Trp Ala Tyr Arg Ile Ser Tyr Asp Pro Ala Arg Phe
 75 80 85

Pro Arg Tyr Leu Pro Glu Ala Tyr Cys Leu Cys Arg Gly Cys Leu Thr
 90 95 100

Gly Leu Tyr Gly Glu Glu Asp Phe Arg Phe Arg Ser Thr Pro Val Phe
 105 110 115 120

Ser Pro Ala Val Val Leu Arg Arg Thr Ala Ala Cys Ala Gly Gly Arg
 125 130 135

Ser Val Tyr Ala Glu His Tyr Ile Thr Ile Pro Val Gly Cys Thr Cys
 140 145 150

Val Pro Glu Pro Asp Lys Ser Ala Asp Ser Ala Asn Ser Ser Met Asp
 155 160 165

Lys Leu Leu Leu Gly Pro Ala Asp Arg Pro Ala Gly Arg
 170 175 180

<210> 13

<211> 504

<212> ADN

<213> primata

<220>

<221> CDS

<222> (19)..(501)

<220>

<221> mat_peptide

<222> (67)..(501)

<400> 13

```

tgagtgtgca gtgccagc atg tac cag gtg gtt gca ttc ttg gca atg gtc 51
      Met Tyr Gln Val Val Ala Phe Leu Ala Met Val
      -15                -10

atg gga acc cac acc tac agc cac tgg ccc agc tgc tgc ccc agc aaa 99
Met Gly Thr His Thr Tyr Ser His Trp Pro Ser Cys Cys Pro Ser Lys
-5          -1 1          5          10

ggg cag gac acc tct gag gag ctg ctg agg tgg agc act gtg cct gtg 147
Gly Gln Asp Thr Ser Glu Glu Leu Arg Trp Ser Thr Val Pro Val
15          20          25

cct ccc cta gag cct gct agg ccc aac cgc cac cca gag tcc tgt agg 195
Pro Pro Leu Glu Pro Ala Arg Pro Asn Arg His Pro Glu Ser Cys Arg
30          35          40

gcc agt gaa gat gga ccc ctc aac agc agg gcc atc tcc ccc tgg aga 243
Ala Ser Glu Asp Gly Pro Leu Asn Ser Arg Ala Ile Ser Pro Trp Arg
45          50          55

tat gag ttg gac aga gac ttg aac cgg ctc ccc cag gac ctg tac cac 291
Tyr Glu Leu Asp Arg Asp Leu Asn Arg Leu Pro Gln Asp Leu Tyr His
60          65          70          75

gcc cgt tgc ctg tgc ccg cac tgc gtc agc cta cag aca ggc tcc cac 339
Ala Arg Cys Leu Cys Pro His Cys Val Ser Leu Gln Thr Gly Ser His
80          85          90

atg gac ccc cgg ggc aac tcg gag ctg ctc tac cac aac cag act gtc 387
Met Asp Pro Arg Gly Asn Ser Glu Leu Leu Tyr His Asn Gln Thr Val
95          100          105

ttc tac cgg cgg cca tgc cat ggc gag aag ggc acc cac aag ggc tac 435
Phe Tyr Arg Arg Pro Cys His Gly Glu Lys Gly Thr His Lys Gly Tyr
110          115          120

tgc ctg gag cgc agg ctg tac cgt gtt tcc tta gct tgt gtg tgt gtg 483
Cys Leu Glu Arg Arg Leu Tyr Arg Val Ser Leu Ala Cys Val Cys Val
125          130          135

cgg ccc cgt gtg atg ggc tag 504
Arg Pro Arg Val Met Gly
140          145

```

<210> 14

<211> 161

<212> PRT

<213> primata

<400> 14

```

Met Tyr Gln Val Val Ala Phe Leu Ala Met Val Met Gly Thr His Thr
-15          -10          -5          -1

```

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

<210> 15

<211> 620

<212> ADN

<213> roedor

<220>

<221> CDS

<222> (1)..(432)

<400> 15

CGG CAC AGG CGG CAC AAA GCC CGG AGA GTG GCT GAA GTG GAG CTC TGC	48
Arg His Arg Arg His Lys Ala Arg Arg Val Ala Glu Val Glu Leu Cys	
1 5 10 15	
ATC TGT ATC CCC CCC AGA GCC TCT GAG CCA CAC CCA CCA CGC AGA ATC	96
Ile Cys Ile Pro Pro Arg Ala Ser Glu Pro His Pro Pro Arg Arg Ile	
20 25 30	
CTG CAG GGC CAG CAA GGA TGG CCT CTC AAC AGC AGG GCC ATC TCT CCT	144
Leu Gln Gly Gln Gln Gly Trp Pro Leu Asn Ser Arg Ala Ile Ser Pro	
35 40 45	
TGG AGC TAT GAG TTG GAC AGG GAC TTG AAT CGG GTC CCC CAG GAC TGG	192
Trp Ser Tyr Glu Leu Asp Arg Asp Leu Asn Arg Val Pro Gln Asp Trp	
50 55 60	
TAC CAC GCT CGA TGC CTG TGC CCA CAC TGC GTC ACG CTA CAG ACA GGC	240
Tyr His Ala Arg Cys Leu Cys Pro His Cys Val Thr Leu Gln Thr Gly	
65 70 75 80	

TCC CAC ATG GAC CCG CTG GGC AAC TCC GTC CCA CTT TAC CAC AAC CAG 288
 Ser His Met Asp Pro Leu Gly Asn Ser Val Pro Leu Tyr His Asn Gln
 85 90 95
 ACG GTC TTC TAC CGG CGG CCA TGC ATG GCG AGG AAG GTA CCC ATC GCC 336
 Thr Val Phe Tyr Arg Arg Pro Cys Met Ala Arg Lys Val Pro Ile Ala
 100 105 110
 GCT ACT GCT TGG AGC GCA GGT CTA CCG AGT CTC CTT GGC TTG TGT GTG 384
 Ala Thr Ala Trp Ser Ala Gly Leu Pro Ser Leu Leu Gly Leu Cys Val
 115 120 125
 TGT GCG GCC CCG GGT CAT GGC TTA GTC ATG CTC ACC ATC TGC CTG AGG 432
 Cys Ala Ala Pro Gly His Gly Leu Val Met Leu Thr Ile Cys Leu Arg
 130 135 140
 TGAATGCCGG GTGGGAGAGA GGGCCAGGTG TACATCACCT GCCAATGCCG GCCGGGTTCA 492
 AGCCTGCAAA GCCTACCTGA AGCAGCAGGT CCCGGGACAG GATGGAGACT TGGGGAGAAA 552
 TCTGACTTTT GCACTTTTTG GAGCATTTTG GGAAGAGCAG GTTCGCTTGT GCTGTAGAGA 612
 TGCTGTTG 620

<210> 16

<211> 144

<212> PRT

<213> roedor

<400> 16

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

<210> 17

<211> 985

<212> ADN

<213> roedor

<220>

<221> CDS

<222> (1)..(507)

<220>

<221> mat_peptide

<222> (49)..(507)

<400> 17

```

atg tac cag gct gtt gca ttc ttg gca atg atc gtg gga acc cac acc 48
Met Tyr Gln Ala Val Ala Phe Leu Ala Met Ile Val Gly Thr His Thr
-15 -10 -5 -1

gtc agc ttg cgg atc cag gag ggc tgc agt cac ttg ccc agc tgc tgc 96
Val Ser Leu Arg Ile Gln Glu Gly Cys Ser His Leu Pro Ser Cys Cys
1 5 10 15

ccc agc aaa gag caa gaa ccc ccg gag gag tgg ctg aag tgg agc tct 144
Pro Ser Lys Glu Gln Glu Pro Pro Glu Glu Trp Leu Lys Trp Ser Ser
20 25 30

gca tct gtg tcc ccc cca gag cct ctg agc cac acc cac cac gca gaa 192
Ala Ser Val Ser Pro Pro Glu Pro Leu Ser His Thr His His Ala Glu
35 40 45

tcc tgc agg gcc agc aag gat ggc ccc ctc aac agc agg gcc atc tct 240
Ser Cys Arg Ala Ser Lys Asp Gly Pro Leu Asn Ser Arg Ala Ile Ser
50 55 60

cct tgg agc tat gag ttg gac agg gac ttg aat cgg gtc ccc cag gac 288
Pro Trp Ser Tyr Glu Leu Asp Arg Asp Leu Asn Arg Val Pro Gln Asp
65 70 75 80

ctg tac cac gct cga tgc ctg tgc cca cac tgc gtc agc cta cag aca 336
Leu Tyr His Ala Arg Cys Leu Cys Pro His Cys Val Ser Leu Gln Thr
85 90 95

ggc tcc cac atg gac ccg ctg ggc aac tcc gtc cca ctt tac cac aac 384
Gly Ser His Met Asp Pro Leu Gly Asn Ser Val Pro Leu Tyr His Asn
100 105 110

cag acg gtc ttc tac cgg cgg cca tgc cat ggt gag gaa ggt acc cat 432
Gln Thr Val Phe Tyr Arg Arg Pro Cys His Gly Glu Glu Gly Thr His
115 120 125

cgc cgc tac tgc ttg gag cgc agg ctc tac cga gtc tcc ttg gct tgt 480
Arg Arg Tyr Cys Leu Glu Arg Arg Leu Tyr Arg Val Ser Leu Ala Cys
130 135 140

gtg tgt gtg cgg ccc cgg gtc atg gct tagtcatgct caccacctgc 527
Val Cys Val Arg Pro Arg Val Met Ala
145 150

ctgaggctga tgcccgttg ggagagaggg ccaggtgtac aatcaccttg ccaatgcggg 587

ccgggttcaa gccctcaaaa gccctacctg aagcagcagg ctcccgggac aagatggagg 647

acttggggag aaactctgac ttttgcactt tttggaagca cttttgggaa ggagcaggtt 707

```

ccgcttgtgc tgctagagga tgctgttggtg gcatttctac tcaggaacgg actccaaagg. 767
 cctgttgacc ctggaagcca tactcctggc tcctttcccc tgaatcccc aactcctggc 827
 acaggcaactt tctccacctc tccccctttg ccttttgttg tgtttgtttg tgcattgcaa 887
 ctctgcgtgc agccaggtgt aattgccttg aaggatggtt ctgaggtgaa agctgttatc 947
 gaaagtgaag agatttatcc aaataaacat ctgtgttt 985

<210> 18

<211> 169

<212> PRT

<213> roedor

<400> 18

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

<210> 19

<211> 521

<212> ADN

<213> primata

<220>

<221> misc_feature

<222> (1)..(521)

<223> nota= "n pode ser a, c, g ou t"

<400> 19

```

gacacggatg aggaccgcta tccacagaag ctggccttcg ccgagtgcct gtgcagaggg 60
tgtatcgatg cacggacggg ccgcgagaca gctgcgctca actccgtgcg gctgctccag 120
agcctgctgg tgctgcgccg ccggccctgc tcccgcgacg gtcgggggct cccacacct 180
ggggcctttg ccttcacacac cgagttcatc cacgtcccg tcggctgcac ctgctgtctg 240
ccccgttcaa gtgtgaccgc caaggccgtg gggcccttag ntgacaccgt gtgctcccca 300
gagggaaccc tatttatggg aattatggta ttatatgctt cccacatact tggggctggc 360
atcccgngct gagacagccc cctgttctat tcagctatat ggggagaaga gtagactttc 420
agctaagtga aaagtgnaac gtgctgactg tctgctgtcg tinctactnat gctagcccga 480
gtgttcactc tgagcctgtt aaatataggc ggttatgtac c 521

```

<210> 20

<211> 521

<212> ADN

<213> primata

<220>

<221> CDS

<222> (1)..(369)

<220>

<221> misc_feature

<222> (281)

<223> nota= "os nucleótidos 281, 367, 437, 462 e 468 estão indicados c; cada um pode alternativamente ser a, g, ou t; o aminoácido traduzido depende do código genético"

<400> 20

```

gac acg gat gag gac cgc tat cca cag aag ctg gcc ttc gcc gag tgc 48
Asp Thr Asp Glu Asp Arg Tyr Pro Gln Lys Leu Ala Phe Ala Glu Cys
 1          5          10          15

ctg tgc aga ggc tgt atc gat gca cgg acg ggc cgc gag aca gct gcg 96
Leu Cys Arg Gly Cys Ile Asp Ala Arg Thr Gly Arg Glu Thr Ala Ala
 20          25          30

ctc aac tcc gtg cgg ctg ctc cag agc ctg ctg gtg ctg cgc cgc cgg 144
Leu Asn Ser Val Arg Leu Leu Gln Ser Leu Leu Val Leu Arg Arg Arg
 35          40          45

ccc tgc tcc cgc gac ggc tcg ggg ctc ccc aca cct ggg gcc ttt gcc 192
Pro Cys Ser Arg Asp Gly Ser Gly Leu Pro Thr Pro Gly Ala Phe Ala
 50          55          60

ttc cac acc gag ttc atc cac gtc ccc gtc ggc tgc acc tgc gtg ctg 240
Phe His Thr Glu Phe Ile His Val Pro Val Gly Cys Thr Cys Val Leu
 65          70          75          80

ccc cgt tca agt gtg acc gcc aag gcc gtg ggg ccc tta gct gac acc 288
Pro Arg Ser Ser Val Thr Ala Lys Ala Val Gly Pro Leu Ala Asp Thr
 85          90          95

gtg tgc tcc cca gag gga ccc cta ttt atg gga att atg gta tta tat 336
Val Cys Ser Pro Glu Gly Pro Leu Phe Met Gly Ile Met Val Leu Tyr
100          105          110

gct tcc cac ata ctt ggg gct ggc atc ccg cgc tgagacagcc ccctgttcta 389
Ala Ser His Ile Leu Gly Ala Gly Ile Pro Arg
115          120

```

ttcagctata tggggagaag agtagacttt cagctaagtg aaaagtgcaa cgtgctgact 449
 gtctgctgtc gtcctactca tgctagcccg agtggttcaact ctgagcctgt taaatatagg 509
 cggttatgta cc 521

<210> 21

<211> 123

<212> PRT

<213> primata

<400> 21

Asp Thr Asp Glu Asp Arg Tyr Pro Gln Lys Leu Ala Phe Ala Glu Cys
 1 5 10 15

Leu Cys Arg Gly Cys Ile Asp Ala Arg Thr Gly Arg Glu Thr Ala Ala
 20 25 30

Leu Asn Ser Val Arg Leu Leu Gln Ser Leu Leu Val Leu Arg Arg Arg
 35 40 45

Pro Cys Ser Arg Asp Gly Ser Gly Leu Pro Thr Pro Gly Ala Phe Ala
 50 55 60

Phe His Thr Glu Phe Ile His Val Pro Val Gly Cys Thr Cys Val Leu
 65 70 75 80

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

Val Cys Ser Pro Glu Gly Pro Leu Phe Met Gly Ile Met Val Leu Tyr
 100 105 110

Ala Ser His Ile Leu Gly Ala Gly Ile Pro Arg
 115 120

<210> 22

<211> 1107

<212> ADN

<213> primata

<220>

<221> CDS

<222> (115)..(705)

<220>

<221> mat_peptide

<222> (166)..(705)

<400> 22

gtgtggcctc aggtataaga gcggtgctg ccaggtgcat ggccaggtgc acctgtggga 60

ttgccgccag gtgtgcaggc cgctccaagc ccagcctgcc ccgctgccgc cacc atg 117
 Met

acg ctc ctc ccc gcc ctc ctg ttt ctg acc tgg ctg cac aca tgc ctg 165
 Thr Leu Leu Pro Gly Leu Phe Leu Thr Trp Leu His Thr Cys Leu
 -15 -10 -5 -1

```

gcc cac cat gac ccc tcc ctc agg ggg cac ccc cac agt cac ggt acc 213
Ala His His Asp Pro Ser Leu Arg Gly His Pro His Ser His Gly Thr
  1             5             10             15

cca cac tgc tac tcg gct gag gaa ctg ccc ctc ggc cag gcc ccc cca 261
Pro His Cys Tyr Ser Ala Glu Glu Leu Pro Leu Gly Gln Ala Pro Pro
      20             25             30

cac ctg ctg gct cga ggt gcc aag tgg ggg cag gct ttg cct gta gcc 309
His Leu Leu Ala Arg Gly Ala Lys Trp Gly Gln Ala Leu Pro Val Ala
      35             40             45

ctg gtg tcc agc ctg gag gca gca agc cac agg ggg agg cac gag agg 357
Leu Val Ser Ser Leu Glu Ala Ala Ser His Arg Gly Arg His Glu Arg
      50             55             60

ccc tca gct acg acc cag tgc ccg gtg ctg cgg ccg gag gag gtg ttg 405
Pro Ser Ala Thr Thr Gln Cys Pro Val Leu Arg Pro Glu Glu Val Leu
      65             70             75             80

gag gca gac acc cac cag cgc tcc atc tca ccc tgg aga tac cgt gtg 453
Glu Ala Asp Thr His Gln Arg Ser Ile Ser Pro Trp Arg Tyr Arg Val
      85             90             95

gac acg gat gag gac cgc tat cca cag aag ctg gcc ttc gcc gag tgc 501
Asp Thr Asp Glu Asp Arg Tyr Pro Gln Lys Leu Ala Phe Ala Glu Cys
      100             105             110

ctg tgc aga ggc tgt atc gat gca cgg acg ggc cgc gag aca gct gcg 549
Leu Cys Arg Gly Cys Ile Asp Ala Arg Thr Gly Arg Glu Thr Ala Ala
      115             120             125

ctc aac tcc gtg cgg ctg ctc cag agc ctg ctg gtg ctg cgc cgc cgg 597
Leu Asn Ser Val Arg Leu Leu Gln Ser Leu Leu Val Leu Arg Arg Arg
      130             135             140

ccc tgc tcc cgc gac ggc tog ggg ctc ccc aca cct ggg gcc ttt gcc 645
Pro Cys Ser Arg Asp Gly Ser Gly Leu Pro Thr Pro Gly Ala Phe Ala
      145             150             155             160

ttc cac acc gag ttc atc cac gtc ccc gtc ggc tgc acc tgc gtg ctg 693
Phe His Thr Glu Phe Ile His Val Pro Val Gly Cys Thr Cys Val Leu
      165             170             175

ccc cgt tca gtg tgaccgccga ggccgtgggg cccctagact ggacacgtgt 745
Pro Arg Ser Val
      180

gtccccaga gggcaccccc tatttatgtg tatttattgg tatttatatg cctcccccaa 805

cactaccctt ggggtctggg cattccccgt gtctggagga cagcccccca ctgttctcct 865

catctccagc ctcagtagtt gggggtagaa ggagctcagc acctcttcca gcccttaaag 925

ctgcagaaaa ggtgtcacac ggctgcctgt accttggtc cctgtcctgc tcccggcttc 985

ccttaacccta tcaactggcct caggcccccg caggetgcct cttcccaacc tcttggaag 1045

taccctgtt tcttaaacia ttatttaagt gtacgtgtat tattaactg atgaacacat 1105

```

cc

1107

<210> 23

<211> 197

<212> PRT

<213> primata

<400> 23

```

Met Thr Leu Leu Pro Gly Leu Leu Phe Leu Thr Trp Leu His Thr Cys
  -15                -10                -5

Leu Ala His His Asp Pro Ser Leu Arg Gly His Pro His Ser His Gly
-1  1              5              10              15

Thr Pro His Cys Tyr Ser Ala Glu Glu Leu Pro Leu Gly Gln Ala Pro
      20              25              30

Pro His Leu Leu Ala Arg Gly Ala Lys Trp Gly Gln Ala Leu Pro Val
      35              40              45

Ala Leu Val Ser Ser Leu Glu Ala Ala Ser His Arg Gly Arg His Glu
      50              55              60

Arg Pro Ser Ala Thr Thr Gln Cys Pro Val Leu Arg Pro Glu Glu Val
      65              70              75

Leu Glu Ala Asp Thr His Gln Arg Ser Ile Ser Pro Trp Arg Tyr Arg
  80              85              90              95

Val Asp Thr Asp Glu Asp Arg Tyr Pro Gln Lys Leu Ala Phe Ala Glu
      100             105             110

Cys Leu Cys Arg Gly Cys Ile Asp Ala Arg Thr Gly Arg Glu Thr Ala
      115             120             125

Ala Leu Asn Ser Val Arg Leu Leu Gln Ser Leu Leu Val Leu Arg Arg
      130             135             140

Arg Pro Cys Ser Arg Asp Gly Ser Gly Leu Pro Thr Pro Gly Ala Phe
      145             150             155

Ala Phe His Thr Glu Phe Ile His Val Pro Val Gly Cys Thr Cys Val
      160             165             170             175

Leu Pro Arg Ser Val
      180

```

<210> 24

<211> 403

<212> ADN

<213> primata

<220>

<221> misc_feature

<222> (1)..(403)

<223> nota= "n pode ser a, c, g, ou t"

<400> 24

gagaaagagc ttcttgacaca aagtaagcca ccagcgcaac atgacagtga agaccctgca 60

```

tggeccagcc atggtcaagt acttgctgct gtcgatattg gggcttgccct ttctgagtga 120
ggcggcagct cggaaaatcc ccaaagtagg acatactttt ttccaaaagc ctgagagttg 180
cccgctgtgt ccaggaggta gtatgaagct tgacattggc atcatcaatg aaaaccagcg 240
cgtttccatg tcaçgtaaca tcgagagccg ctccacctcc ccttggaatt acactgtcac 300
ttggggacccc aaccggtacc cctcgaagtt gtacaggccc aagtgtagga acttgggctg 360
tatcaatgct caaggaaagg aagacatctn catgaattcc gtc 403

```

<210> 25
 <211> 403
 <212> ADN
 <213> primata

 <220>
 <221> CDS
 <222> (71)..(403)

 <220>
 <221> mat_peptide
 <222> (131)..(403)

 <220>
 <221> misc feature
 <222> (1)..(403)
 <223> nota= "n pode ser a, c, g, ou t; o aminoácido traduzido depende do código genético"

<400> 25
 gagaaaagagc ttctgcaca aagtaagcca ccagcgcaac atgacagtga agaccctgca 60

 tggcccagcc atg gtc aag tac ttg ctg ctg tcg ata ttg ggg ctt gcc 109
 Met Val Lys Tyr Leu Leu Leu Ser Ile Leu Gly Leu Ala
 -20 -15 -10

 ttt ctg agt gag gcg gca gct cgg aaa atc ccc aaa gta gga cat act 157
 Phe Leu Ser Glu Ala Ala Ala Arg Lys Ile Pro Lys Val Gly His Thr
 -5 -1 1 5

 ttt ttc caa aag cct gag agt tgc ccg cct gtg cca gga ggt agt atg 205
 Phe Phe Gln Lys Pro Glu Ser Cys Pro Pro Val Pro Gly Gly Ser Met
 10 15 20 25

 aag ctt gac att ggc atc atc aat gaa aac cag cgc gtt tcc atg tca 253
 Lys Leu Asp Ile Gly Ile Ile Asn Glu Asn Gln Arg Val Ser Met Ser
 30 35 40

 cgt aac atc gag agc cgc tcc acc tcc ccc tgg aat tac act gtc act 301
 Arg Asn Ile Glu Ser Arg Ser Thr Ser Pro Trp Asn Tyr Thr Val Thr
 45 50 55

 tgg gac ccc aac cgg tac ccc tcg aag ttg tac agg ccc aag tgt agg 349
 Trp Asp Pro Asn Arg Tyr Pro Ser Lys Leu Tyr Arg Pro Lys Cys Arg
 60 65 70

 aac ttg ggc tgt atc aat gct caa gga aag gaa gac atc tnc atg aat 397
 Asn Leu Gly Cys Ile Asn Ala Gln Gly Lys Glu Asp Ile Xaa Met Asn
 75 80 85

 tcc gtc 403
 Ser Val
 90

<210> 26
 <211> 111
 <212> PRT
 <213> primata

<400> 26

Met Val Lys Tyr Leu Leu Leu Ser Ile Leu Gly Leu Ala Phe Leu Ser
 -20 -15 -10 -5
 Glu Ala Ala Ala Arg Lys Ile Pro Lys Val Gly His Thr Phe Phe Gln
 -1 1 5 10
 Lys Pro Glu Ser Cys Pro Pro Val Pro Gly Gly Ser Met Lys Leu Asp
 15 20 25
 Ile Gly Ile Ile Asn Glu Asn Gln Arg Val Ser Met Ser Arg Asn Ile
 30 35 40
 Glu Ser Arg Ser Thr Ser Pro Trp Asn Tyr Thr Val Thr Trp Asp Pro
 45 50 55 60
 Asn Arg Tyr Pro Ser Lys Leu Tyr Arg Pro Lys Cys Arg Asn Leu Gly
 65 70 75
 Cys Ile Asn Ala Gln Gly Lys Glu Asp Ile Xaa Met Asn Ser Val
 80 85 90

<210> 27

<211> 784

<212> ADN

<213> primata

<220>

<221> CDS

<222> (3)..(281)

<400> 27

tc gtg ccg tat ctt ttt aaa aaa att att ctt cac ttt ttt gcc tcc 47
 Val Pro Tyr Leu Phe Lys Lys Ile Ile Leu His Phe Phe Ala Ser
 1 5 10 15
 tat tac ttg tta ggg aga ccc aat ggt agt ttt att cct tgg gga tac 95
 Tyr Tyr Leu Leu Gly Arg Pro Asn Gly Ser Phe Ile Pro Trp Gly Tyr
 20 25 30
 ata gta aat act tca tta aag tcg agt aca gaa ttt gat gaa aag tgt 143
 Ile Val Asn Thr Ser Leu Lys Ser Ser Thr Glu Phe Asp Glu Lys Cys
 35 40 45
 gga tgt gtg gga tgt act gcc gcc ttc aga agt cca cac act gcc tgg 191
 Gly Cys Val Gly Cys Thr Ala Ala Phe Arg Ser Pro His Thr Ala Trp
 50 55 60
 agg gag aga act gct gtt tat tca ctg att aag cat ttg ctg tgt acc 239
 Arg Glu Arg Thr Ala Val Tyr Ser Leu Ile Lys His Leu Leu Cys Thr
 65 70 75
 aac tac ttt tca tgt ctt atc tta att ctc ata aca gtc att 281

Asn Tyr Phe Ser Cys Leu Ile Leu Ile Leu Ile Thr Val Ile
 80 85 90

tgatatttta aaaaaccca gaaatctgag aaagagataa agtggtttgc tcaaggttat 341
 agaacagact accatgtgtt gtatttcaga ttttaattca tgtttgtctg attttaagtt 401
 ttgttcgctt gccagggtac cccacaaaaa tgccaggcag ggcattttca tgatgcactt 461
 gagatacctg aaatgacagg gtagcatcac acctgagagg ggtaaaggat gggaacctac 521
 cttccatggc cgctgcttgg cagtctcttg ctgcatgcta gcagagccac tgtatatgtg 581
 ccgaggctct gagaattaac tgcttaaaga actgccttct ggagggagaa ggcacaaga 641
 tcacaattaa ccatatacac atcttactgt gcgaggtcac tgagcaatac aggagggatt 701
 ttatacatth tagcaactat cttcaaaacc tgagctatag ttgtattctg ccccttcct 761
 ctgggcaaaa gtgtaaaagt ttg 784

<210> 28

<211> 93

<212> PRT

<213> primata

<400> 28

Val Pro Tyr Leu Phe Lys Lys Ile Ile Leu His Phe Phe Ala Ser Tyr
 1 5 10 15

Tyr Leu Leu Gly Arg Pro Asn Gly Ser Phe Ile Pro Trp Gly Tyr Ile
 20 25 30

Val Asn Thr Ser Leu Lys Ser Ser Thr Glu Phe Asp Glu Lys Cys Gly
 35 40 45

Cys Val Gly Cys Thr Ala Ala Phe Arg Ser Pro His Thr Ala Trp Arg
 50 55 60

Glu Arg Thr Ala Val Tyr Ser Leu Ile Lys His Leu Leu Cys Thr Asn
 65 70 75 80

Tyr Phe Ser Cys Leu Ile Leu Ile Leu Ile Thr Val Ile
 85 90

<210> 29

<211> 460

<212> ADN

<213> primata

<220>

<221> CDS

<222> (1)..(189)

<400> 29

gtg act gta ttg tgg gga cag gaa gca caa att ccc atg tgg atc act 48
 Val Thr Val Leu Trp Gly Gln Glu Ala Gln Ile Pro Met Trp Ile Thr
 1 5 10 15

```

agg aga gat aat aag tgg ggt cat ttc acc cct tgg tcc cct gct tcc 96
Arg Arg Asp Asn Lys Trp Gly His Phe Thr Pro Trp Ser Pro Ala Ser
          20                25                30

aga ccc aaa gag gcc tac atg gca ttg tgc ttc ctt ctt agt tgt agg 144
Arg Pro Lys Glu Ala Tyr Met Ala Leu Cys Phe Leu Leu Ser Cys Arg
          35                40                45

agg tgt gag ata caa tca ttt gcc tct gac ttt gag ggt tgg tcc 189
Arg Cys Glu Ile Gln Ser Phe Ala Ser Asp Phe Glu Gly Trp Ser
          50                55                60

tagcatgccc ctgaccagta gcccttaaa tacttcattg atatggaagg tctctgaatc 249

ttcgtgggct taatctacca ctctctgaag ttcttatgtc tttcaaaggc ctctaaaatc 309

tctgccatgt cttgctcatc cagttgttag catgatgtca ttgatacagt ggactttgga 369

atctaagtgg ggagacactg gtaagtgacc aattacttca cctgtggtgt gcaagccaga 429

tcaggaagcc tctacctgca cgacaacaca t 460

```

<210> 30
 <211> 63
 <212> PRT
 <213> primata

```

<400> 30
Val Thr Val Leu Trp Gly Gln Glu Ala Gln Ile Pro Met Trp Ile Thr
  1                5                10                15

Arg Arg Asp Asn Lys Trp Gly His Phe Thr Pro Trp Ser Pro Ala Ser
          20                25                30

Arg Pro Lys Glu Ala Tyr Met Ala Leu Cys Phe Leu Leu Ser Cys Arg
          35                40                45

Arg Cys Glu Ile Gln Ser Phe Ala Ser Asp Phe Glu Gly Trp Ser
          50                55                60

```

<210> 31
 <211> 150
 <212> PRT
 <213> roedor

```

<400> 31
Met Cys Leu Met Leu Leu Leu Leu Asn Leu Glu Ala Thr Val Lys
  1                5                10                15

Ala Ala Val Leu Ile Pro Gln Ser Ser Val Cys Pro Asn Ala Glu Ala
          20                25                30

Asn Asn Phe Leu Gln Asn Val Lys Val Asn Leu Lys Val Ile Asn Ser
          35                40                45

Leu Ser Ser Lys Ala Ser Ser Arg Arg Pro Ser Asp Tyr Leu Asn Arg
          50                55                60

Ser Thr Ser Pro Trp Thr Leu Ser Arg Asn Glu Asp Pro Asp Arg Tyr

```

65		70		75		80
Pro Ser Val Ile Trp Glu Ala Gln Cys Arg His Gln Arg Cys Val Asn						
	85			90		95
Ala Glu Gly Lys Leu Asp His His Met Asn Ser Val Leu Ile Gln Gln						
	100		105			110
Glu Ile Leu Val Leu Lys Arg Glu Pro Glu Lys Cys Pro Phe Thr Phe						
	115		120			125
Arg Val Glu Lys Met Leu Val Gly Val Gly Cys Thr Cys Val Ser Ser						
	130		135			140
Ile Val Arg His Ala Ser						
	145		150			

<210> 32

<211> 147

<212> PRT

<213> roedor

<400> 32

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

<210> 33

<211> 155

<212> PRT

<213> primata

<400> 33

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

<210> 34

<211> 151

<212> PRT

<213> viral

<400> 34

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

Phe Arg Leu Glu Lys Met Leu Val Thr Val Gly Cys Thr Cys Val Thr
130 135 140
Pro Ile Val His Asn Val Asp
145 150

Lisboa, 2009-12-29

REIVINDICAÇÕES

1. Polinucleótido compreendendo uma sequência que codifica o polipéptido IL-174 maduro compreendido na SEQ ID NO:14.

2. Polinucleótido de acordo com a reivindicação 1, compreendendo a porção codificante madura completa da SEQ ID NO:13.

3. Polinucleótido codificando pelo menos 16 aminoácidos contíguos do polipéptido IL-174 maduro compreendido na SEQ ID NO:14.

4. Polinucleótido compreendendo pelo menos 33 nucleótidos contíguos da SEQ ID NO:13.

5. Vector de expressão compreendendo um polinucleótido de acordo com a reivindicação 1.

6. Célula hospedeira contendo um vector de expressão de acordo com a reivindicação 5, em que a célula é:

- (a) uma célula procariótica;
- (b) uma célula eucariótica;
- (c) uma célula bacteriana;
- (d) uma célula de levedura;
- (e) uma célula de insecto;
- (f) uma célula de mamífero;
- (g) uma célula de ratinho;
- (h) uma célula de primata; ou
- (i) uma célula humana.

7. Método de preparação de um polipéptido IL-174, compreendendo a cultura da célula hospedeira de acordo com a reivindicação 6 em condições adequadas para a expressão do polipéptido.

8. Polipéptido compreendendo o polipéptido IL-174 maduro compreendido na SEQ ID NO:14.

9. Polipéptido compreendendo pelo menos 16 aminoácidos contíguos do polipéptido IL-174 maduro compreendido na SEQ ID NO:14.

10. Polipéptido de acordo com a reivindicação 9 que:

- (a) está numa composição estéril;
- (b) não é glicosilado;
- (c) está desnaturado;
- (d) é um polipéptido sintético;
- (e) está ligado a um substrato sólido; ou
- (f) é uma proteína de fusão com uma marca de detecção ou de purificação.

11. Polipéptido de acordo com a reivindicação 8 que está radiomarcado.

12. Método utilizando o polipéptido IL-174 de acordo com a reivindicação 8 para identificar um composto que se liga selectivamente ao polipéptido, que compreende a incubação do composto com o polipéptido em condições apropriadas; provocando desse modo a ligação do composto ao polipéptido.

13. *Kit* compreendendo o polipéptido IL-174 de acordo com a reivindicação 8 ou o polinucleótido de acordo com a reivindicação 1 ou 2.

Lisboa, 2009-12-29

RESUMO

**"Citoquinas de mamíferos relacionadas com a interleucina-17.
Polinucleótidos que as codificam. Utilizações."**

Antigénios relacionados com a CTLA-8 oriundos de mamíferos, reagentes relacionados incluindo proteínas purificadas, anticorpos específicos e ácidos nucleicos que codificam os referidos antigénios. São igualmente proporcionados métodos de utilização dos referidos reagentes e *kits* de diagnóstico.