

(22) Data de pedido: **2000.05.30**

(30) Prioridade(s): **1999.06.22 US 140650 P**
1999.06.23 US 141037 P
1999.07.20 US 144758 P
1999.10.29 US 162506 P
1999.12.09 US 170262 P

(43) Data de publicação do pedido: **2002.03.13**

(45) Data e BPI da concessão: **2007.03.21**
026/2007

(73) Titular(es):

GENENTECH, INC.

1 DNA WAY SOUTH SAN FRANCISCO, CA
94080-4990

US

(72) Inventor(es):

RHONA C. KABAKOFF

US

KEVIN P. BAKER

US

AUDREY GODDARD

US

AUSTIN L. GURNEY

US

COLIN K. WATANABE

US

(74) Mandatário:

ANTÓNIO JOÃO COIMBRA DA CUNHA FERREIRA
RUA DAS FLORES, Nº 74, 4º AND 1249-235 LISBOA

PT

(54) Epígrafe: **MÉTODOS E COMPOSIÇÕES PARA A INIBIÇÃO DO CRESCIMENTO DE CÉLULAS NEOPLÁSICAS.**

(57) Resumo:

MÉTODOS E COMPOSIÇÕES PARA A INIBIÇÃO DO CRESCIMENTO DE CÉLULAS NEOPLÁSICAS.

RESUMO

"Métodos e composições para a inibição do crescimento de células neoplásicas"

A presente invenção refere-se a métodos e composições para a inibição do crescimento de células neoplásicas. Em particular, a presente invenção refere-se a composições e métodos antitumorais para o tratamento de tumores. A invenção refere-se ainda a métodos de rastreio para a identificação de compostos inibidores do crescimento, *i.e.*, antitumorais. A presente invenção refere-se a novos polipéptidos e moléculas de ácido nucleico que codificam esses polipéptidos. São também aqui proporcionados vectores e células hospedeiras compreendendo essas sequências de ácido nucleico, moléculas polipeptídicas quiméricas compreendendo os polipéptidos da presente invenção fundidos com sequências polipeptídicas heterólogas, anticorpos que se ligam aos polipéptidos da presente invenção e a métodos para a produção dos polipéptidos da presente invenção.

DESCRIÇÃO

"Métodos e composições para a inibição do crescimento de células neoplásicas"

CAMPO DA INVENÇÃO

A presente invenção refere-se a métodos e composições para a inibição do crescimento de células neoplásicas. Em particular, a presente invenção refere-se a composições e métodos antitumorais para o tratamento de tumores. A invenção refere-se ainda a métodos de rastreio para a identificação de compostos inibidores do crescimento, *i.e.* antitumorais.

ANTECEDENTES DA INVENÇÃO

Os tumores malignos (cancros) são a segunda principal causa de morte nos Estados Unidos, depois das doenças cardíacas (Boring et al., CA Cancer J. Clin., 43:7 [1993]).

O cancro caracteriza-se por um aumento do número de células anómalas ou neoplásicas derivadas de um tecido normal, que proliferam formando uma massa tumoral, pela invasão de tecidos adjacentes por essas células neoplásicas tumorais e pela geração de células malignas que eventualmente se espalham através do sistema sanguíneo ou linfático para nódulos linfáticos regionais e para locais distantes (metástase). Num estado canceroso uma célula prolifera em condições sob as quais as células normais não cresceriam. O cancro manifesta-se numa grande variedade de formas caracterizadas por diferentes graus de invasão e agressividade.

Apesar dos avanços recentes na terapia do cancro, existe uma grande necessidade de novos agentes terapêuticos capazes de inibir o crescimento de células neoplásicas. Deste modo, é o objectivo da presente invenção identificar compostos capazes de inibir o crescimento de células neoplásicas, tais como as células cancerosas.

SUMÁRIO DA INVENÇÃO

A. Concretizações

A presente invenção refere-se a composições para a inibição do crescimento de células neoplásicas. Mais particularmente, a invenção refere-se a composições para o tratamento de tumores, incluindo cancros, tais como os cancros da mama, da próstata, do cólon, pulmonar, ovariano, renal e do SNC, leucemia, melanoma, etc., em pacientes mamíferos, preferivelmente seres humanos.

Num aspecto, a presente invenção refere-se a composições de matéria úteis para a inibição do crescimento de células neoplásicas compreendendo uma quantidade eficaz de um polipéptido PRO como aqui definido, em mistura com um transportador farmacêuticamente aceitável. Numa concretização preferida a composição de matéria compreende uma quantidade inibidora do crescimento de um polipéptido PRO. Noutra concretização preferida, a composição compreende uma quantidade citotóxica de um polipéptido PRO. Opcionalmente, as composições de matéria podem conter um ou mais agentes adicionais, inibidores do crescimento e/ou citotóxicos e/ou outros agentes quimioterapêuticos.

Noutro aspecto, a presente invenção refere-se a composições de matéria úteis para o tratamento de um tumor num mamífero compreendendo uma quantidade terapeuticamente eficaz de um polipéptido PRO como aqui definido. O tumor é preferivelmente um cancro.

As composições da invenção podem ser utilizadas num método para a inibição do crescimento de uma célula tumoral compreendendo a exposição da célula a uma quantidade eficaz de um polipéptido PRO como aqui definido. O método pode ser realizado *in vitro* ou *in vivo*.

Ainda noutra concretização, a invenção refere-se a um artigo de fabrico compreendendo:

- (a) um recipiente;
- (b) uma composição compreendendo um agente activo contido no recipiente; em que a composição é eficaz para a inibição

do crescimento de células neoplásicas, e.g., o crescimento de células tumorais, e o agente activo na composição é um polipéptido PRO como aqui definido; e

- (c) um rótulo afixado no referido recipiente, ou um folheto incluído no referido recipiente, indicando a utilização do referido polipéptido PRO para a inibição do crescimento de células neoplásicas.

Artigos de fabrico similares compreendendo um polipéptido PRO como aqui definido, numa quantidade que é terapeuticamente eficaz para o tratamento de tumores, estão também no âmbito da presente invenção. Estão também no âmbito da invenção artigos de fabrico compreendendo um polipéptido PRO como aqui definido e um outro agente inibidor do crescimento, um agente citotóxico ou um agente quimioterapêutico.

B. Concretizações adicionais

Noutras concretizações da presente invenção, a invenção proporciona uma molécula de ácido nucleico isolada compreendendo uma sequência nucleotídica que codifica um polipéptido PRO.

Num aspecto, a molécula de ácido nucleico isolada compreende uma sequência nucleotídica possuindo pelo menos cerca de 80% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 81% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 82% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 83% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 84% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 85% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 86% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 87% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 88% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 89% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 90% de identidade de sequência de ácido nucleico,

alternativamente pelo menos cerca de 91% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 92% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 93% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 94% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 95% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 96% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 97% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 98% de identidade de sequência de ácido nucleico e alternativamente pelo menos cerca de 99% de identidade de sequência de ácido nucleico, com (a) uma molécula de ADN que codifica um polipéptido PRO possuindo uma sequência de aminoácidos de comprimento completo como aqui se descreve, uma sequência de aminoácidos a que falta o péptido de sinal como aqui se descreve, um domínio extracelular de uma proteína transmembranar, com ou sem o péptido de sinal, como aqui se descreve ou qualquer outro fragmento especificamente definido da sequência de aminoácidos de comprimento completo como aqui se descreve, ou (b) o complemento da molécula de ADN de (a).

Noutros aspectos, a molécula de ácido nucleico isolada compreende uma sequência nucleotídica possuindo pelo menos cerca de 80% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 81% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 82% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 83% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 84% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 85% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 86% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 87% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 88% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 89% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 90% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 91% de identidade de

sequência de ácido nucleico, alternativamente pelo menos cerca de 92% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 93% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 94% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 95% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 96% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 97% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 98% de identidade de sequência de ácido nucleico e alternativamente pelo menos cerca de 99% de identidade de sequência de ácido nucleico, com (a) uma molécula de ADN compreendendo a sequência de codificação de um ADNc do polipéptido PRO de comprimento completo como aqui se descreve, a sequência de codificação de um polipéptido PRO a que falta o péptido de sinal como aqui se descreve, a sequência de codificação de um domínio extracelular de um polipéptido PRO transmembranar, com ou sem o péptido de sinal, como aqui se descreve, ou a sequência de codificação de qualquer outro fragmento especificamente definido da sequência de aminoácidos de comprimento completo como aqui se descreve, ou (b) o complemento da molécula de ADN de (a).

Num outro aspecto, a invenção refere-se a uma molécula de ácido nucleico isolada compreendendo uma sequência nucleotídica possuindo pelo menos cerca de 80% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 81% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 82% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 83% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 84% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 85% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 86% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 87% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 88% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 89% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 90% de identidade de

sequência de ácido nucleico, alternativamente pelo menos cerca de 91% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 92% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 93% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 94% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 95% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 96% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 97% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 98% de identidade de sequência de ácido nucleico e alternativamente pelo menos cerca de 99% de identidade de sequência de ácido nucleico, com (a) uma molécula de ADN que codifica o mesmo polipéptido maduro que é codificado por qualquer um dos ADNc de proteína humana depositados na ATCC como aqui se descreve, ou (b) o complemento da molécula de ADN de (a).

Outro aspecto da invenção proporciona uma molécula de ácido nucleico isolada incluindo uma sequência nucleotídica que codifica um polipéptido PRO que apresenta o domínio transmembranar eliminado, ou o domínio transmembranar inactivado, ou é complementar a essa sequência nucleotídica de codificação, em que o(s) domínio(s) transmembranar(es) destes polipéptidos são aqui revelados. Assim, encontram-se contemplados domínios extracelulares solúveis dos polipéptidos PRO aqui descritos.

Outra concretização refere-se a fragmentos de uma sequência de codificação de um polipéptido PRO, ou do seu complemento, que podem ser úteis, por exemplo, como sondas de hibridação, para fragmentos de codificação de um polipéptido PRO que podem codificar opcionalmente um polipéptido compreendendo um local de ligação para um anticorpo anti-PRO, ou como sondas oligonucleotídicas anti-sentido. Estes fragmentos de ácido nucleico têm usualmente pelo menos cerca de 20 nucleótidos de comprimento, alternativamente pelo menos cerca de 30 nucleótidos de comprimento, alternativamente pelo menos cerca de 40 nucleótidos de comprimento, alternativamente pelo menos cerca de 50 nucleótidos de comprimento, alternativamente pelo

menos cerca de 60 nucleótidos de comprimento, alternativamente pelo menos cerca de 70 nucleótidos de comprimento, alternativamente pelo menos cerca de 80 nucleótidos de comprimento, alternativamente pelo menos cerca de 90 nucleótidos de comprimento, alternativamente pelo menos cerca de 100 nucleótidos de comprimento, alternativamente pelo menos cerca de 110 nucleótidos de comprimento, alternativamente pelo menos cerca de 120 nucleótidos de comprimento, alternativamente pelo menos cerca de 130 nucleótidos de comprimento, alternativamente pelo menos cerca de 140 nucleótidos de comprimento, alternativamente pelo menos cerca de 150 nucleótidos de comprimento, alternativamente pelo menos cerca de 160 nucleótidos de comprimento, alternativamente pelo menos cerca de 170 nucleótidos de comprimento, alternativamente pelo menos cerca de 180 nucleótidos de comprimento, alternativamente pelo menos cerca de 190 nucleótidos de comprimento, alternativamente pelo menos cerca de 200 nucleótidos de comprimento, alternativamente pelo menos cerca de 250 nucleótidos de comprimento, alternativamente pelo menos cerca de 300 nucleótidos de comprimento, alternativamente pelo menos cerca de 350 nucleótidos de comprimento, alternativamente pelo menos cerca de 400 nucleótidos de comprimento, alternativamente pelo menos cerca de 450 nucleótidos de comprimento, alternativamente pelo menos cerca de 500 nucleótidos de comprimento, alternativamente pelo menos cerca de 600 nucleótidos de comprimento, alternativamente pelo menos cerca de 700 nucleótidos de comprimento, alternativamente pelo menos cerca de 800 nucleótidos de comprimento, alternativamente pelo menos cerca de 900 nucleótidos de comprimento e alternativamente pelo menos cerca de 1000 nucleótidos de comprimento, em que, neste contexto, o termo "cerca de" significa o comprimento da sequência nucleotídica referido, mais ou menos 10% desse comprimento referido. Note-se que podem ser determinados novos fragmentos de uma sequência nucleotídica que codifica o polipéptido PRO de forma rotineira efectuando o alinhamento da sequência nucleotídica que codifica o polipéptido PRO com outras sequências nucleotídicas conhecidas utilizando qualquer um dos vários de programas de alinhamento de sequências bem conhecidos e determinando qual(ais) o(s) fragmento(s) da sequência nucleotídica que codifica o polipéptido PRO são novos. Todas estas sequências nucleotídicas que codificam

polipéptidos PRO estão aqui contempladas. Estão também contemplados os fragmentos de polipéptidos PRO codificados por esses fragmentos de moléculas nucleotídicas, preferivelmente os fragmentos de polipéptidos PRO que compreendem um local de ligação para um anticorpo anti-PRO.

Noutra concretização, a invenção proporciona um polipéptido PRO isolado codificado por qualquer das sequências de ácido nucleico isoladas atrás identificadas.

Num determinado aspecto, a invenção refere-se a um polipéptido PRO isolado, compreendendo uma sequência de aminoácidos possuindo pelo menos cerca de 80% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 81% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 82% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 83% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 84% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 85% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 86% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 87% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 88% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 89% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 90% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 91% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 92% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 93% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 94% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 95% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 96% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 97% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 98% de identidade de sequência de aminoácidos e alternativamente pelo menos cerca de 99% de identidade de sequência de aminoácidos, com um

polipéptido PRO possuindo uma sequência de aminoácidos de comprimento completo como aqui se descreve, uma sequência de aminoácidos a que falta o péptido de sinal como aqui se descreve, um domínio extracelular de uma proteína transmembranar, com ou sem o péptido de sinal, como aqui se descreve, ou qualquer outro fragmento especificamente definido da sequência de aminoácidos de comprimento completo como aqui se descreve.

Num outro aspecto, a invenção refere-se a um polipéptido PRO isolado compreendendo uma sequência de aminoácidos possuindo pelo menos cerca de 80% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 81% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 82% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 83% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 84% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 85% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 86% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 87% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 88% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 89% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 90% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 91% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 92% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 93% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 94% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 95% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 96% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 97% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 98% de identidade de sequência de aminoácidos e alternativamente pelo menos cerca de 99% de identidade de sequência de aminoácidos, com uma

sequência de aminoácidos codificada por qualquer dos ADNc de proteína humana depositados na ATCC como aqui se descreve.

Num outro aspecto, a invenção refere-se a um polipéptido PRO isolado compreendendo uma sequência de aminoácidos com uma pontuação de pelo menos cerca de 80% positivos, alternativamente pelo menos cerca de 81% positivos, alternativamente pelo menos cerca de 82% positivos, alternativamente pelo menos cerca de 83% positivos, alternativamente pelo menos cerca de 84% positivos, alternativamente pelo menos cerca de 85% positivos, alternativamente pelo menos cerca de 86% positivos, alternativamente pelo menos cerca de 87% positivos, alternativamente pelo menos cerca de 88% positivos, alternativamente pelo menos cerca de 89% positivos, alternativamente pelo menos cerca de 90% positivos, alternativamente pelo menos cerca de 91% positivos, alternativamente pelo menos cerca de 92% positivos, alternativamente pelo menos cerca de 93% positivos, alternativamente pelo menos cerca de 94% positivos, alternativamente pelo menos cerca de 95% positivos, alternativamente pelo menos cerca de 96% positivos, alternativamente pelo menos cerca de 97% positivos, alternativamente pelo menos cerca de 98% positivos e alternativamente pelo menos cerca de 99% positivos, em comparação com a sequência de aminoácidos de um polipéptido PRO possuindo uma sequência de aminoácidos de comprimento completo como aqui se descreve, uma sequência de aminoácidos a que falta o péptido de sinal como aqui se descreve, um domínio extracelular de uma proteína transmembrana, com ou sem o péptido de sinal, como aqui se descreve, ou qualquer outro fragmento especificamente definido da sequência de aminoácidos de comprimento completo como aqui se descreve.

Num aspecto específico, a invenção proporciona um polipéptido PRO isolado sem a sequência de sinal N-terminal e/ou a metionina de iniciação e que é codificado por uma sequência nucleotídica que codifica uma tal sequência de aminoácidos, como aqui previamente descrito. Os processos para a sua produção são também aqui descritos, sendo que esses processos compreendem a cultura de uma célula hospedeira

compreendendo um vector que compreende a molécula de ácido nucleico codificante apropriada, em condições adequadas para expressão do polipéptido PRO, e a recuperação do polipéptido PRO a partir da cultura celular.

Outro aspecto da invenção proporciona um polipéptido PRO isolado que apresenta o domínio transmembranar eliminado ou o domínio transmembranar inactivado. São também aqui descritos processos para a sua produção, em que esses processos compreendem a cultura de uma célula hospedeira compreendendo um vector que compreende a molécula de ácido nucleico codificante apropriada, em condições adequadas para expressão do polipéptido PRO, e a recuperação do polipéptido PRO a partir da cultura celular.

Está também contemplado um método de identificação de agonistas de um polipéptido PRO, que compreende o contacto do polipéptido PRO com uma molécula candidata e a monitoração de uma actividade biológica mediada pelo referido polipéptido PRO. Preferivelmente, o polipéptido PRO é um polipéptido PRO nativo.

Ainda numa outra concretização, a invenção refere-se a uma composição de matéria compreendendo um polipéptido PRO em combinação com uma transportador. Opcionalmente, o transportador é um transportador farmacêuticamente aceitável.

Outra concretização da presente invenção refere-se à utilização de um polipéptido PRO para a preparação de um medicamento útil no tratamento de uma condição que é responsiva ao polipéptido PRO.

Em concretizações adicionais da presente invenção, a invenção proporciona vectores compreendendo ADN que codifica qualquer dos polipéptidos aqui descritos. Proporcionam-se igualmente células hospedeiras compreendendo qualquer destes vectores. Como exemplo, as células hospedeiras podem ser células CHO, células de *E. coli*, de levedura ou de insecto infectadas por baculovírus. Proporciona-se adicionalmente um processo para produzir qualquer dos polipéptidos aqui descritos e que compreendo a cultura das células hospedeiras em condições adequadas para a expressão do polipéptido

desejado, e a recuperação do polipéptido desejado a partir da cultura celular.

Noutras concretizações, a invenção proporciona moléculas quiméricas compreendendo qualquer dos polipéptidos aqui descritos fundido com um polipéptido ou uma sequência de aminoácidos heterólogos. Constituem exemplos de tais moléculas quiméricas quaisquer dos polipéptidos aqui descritos fundidos com uma sequência de marcação epitópica ou uma região Fc de uma imunoglobulina.

Está igualmente contemplado um anticorpo que se liga especificamente a qualquer dos polipéptidos descritos acima ou adiante. Opcionalmente, o anticorpo é um anticorpo monoclonal, um anticorpo humanizado, um fragmento de anticorpo ou um anticorpo de cadeia simples.

Ainda noutras concretizações, a invenção proporciona sondas oligonucleotídicas úteis para isolar sequências nucleotídica genómicas e de ADNc ou sondas anti-sentido, em que essas sondas podem ser derivadas de qualquer das sequências nucleotídicas descritas acima ou adiante.

DESCRIÇÃO SUCINTA DAS FIGURAS

A Figura 1 mostra uma sequência nucleotídica (SEQ ID NO:1) de um ADNc de PRO4400 de sequência nativa, em que SEQ ID NO:1 é um clone aqui designado "DNA87974-2609".

A Figura 2 mostra a sequência de aminoácidos (SEQ ID NO:2) derivada da sequência de codificação de SEQ ID NO:1 apresentada na Figura 1.

DESCRIÇÃO DETALHADA DA INVENÇÃO

As expressões "polipéptido PRO" e "PRO", como aqui utilizadas e quando imediatamente seguidas por uma designação numérica referem-se a vários polipéptidos em que a designação completa (*i.e.*, PRO/número) se refere a sequências polipeptídicas específicas tal como aqui descritas. As expressões "polipéptido PRO/número" e "PRO/número" em que o termo "número" é proporcionado como uma designação numérica real, tal como aqui utilizadas, abrangem polipéptidos de

sequência nativa e variantes dos polipéptidos (que são aqui adicionalmente definidos). Os polipéptidos PRO aqui descritos podem ser isolados de várias fontes, tal como de tipos de tecidos humanos ou de outra fonte, ou preparados por métodos recombinantes ou sintéticos.

Um “polipéptido PRO de sequência nativa” compreende um polipéptido apresentando a mesma sequência de aminoácidos que o polipéptido PRO correspondente derivado da natureza. Estes polipéptidos PRO de sequência nativa podem ser isolados da natureza ou podem ser produzidos por meios recombinantes ou sintéticos. A expressão “polipéptido PRO de sequência nativa” abrange especificamente formas de ocorrência natural truncadas ou excretadas, do polipéptido PRO específico (por exemplo, uma sequência de domínio extracelular), formas variantes de ocorrência natural (por exemplo, formas de *splicing* alternativo) e variantes alélicas de ocorrência natural, do polipéptido. Em várias concretizações da invenção, os polipéptidos PRO de sequência nativa aqui revelados são polipéptidos de sequência nativa maduros ou de comprimento completo compreendendo as sequências de aminoácidos de comprimento completo apresentadas nas figuras anexas. Os codões de início e de paragem são apresentados a negrito e sublinhados nas figuras. Contudo, apesar de se apresentarem os polipéptidos PRO revelados nas figuras anexas como começando com resíduos de metionina, designados aqui como a posição de aminoácido 1 nas figuras, é conceptível e possível que outros resíduos de metionina localizados quer a montante quer a jusante da posição de aminoácido 1 nas figuras possam ser empregues como o resíduo de aminoácido de partida para os polipéptidos PRO.

O “domínio extracelular” ou “ECD” do polipéptido PRO refere-se a uma forma do polipéptido PRO que se encontra essencialmente isenta dos domínios transmembranar e citoplasmático. Habitualmente, o ECD de um polipéptido PRO terá menos de 1% desses domínios transmembranar e/ou citoplasmático e preferivelmente terá menos que 0,5% desses domínios. Entender-se-á que quaisquer domínios transmembranares identificados para os polipéptidos PRO da presente invenção são identificados de acordo com os critérios rotineiramente empregues na especialidade para identificar

esse tipo de domínios hidrófobos. As fronteiras exactas de um domínio transmembranar podem variar mas muito provavelmente não variarão mais de cerca de 5 aminoácidos de qualquer das extremidades do domínio tal como aqui inicialmente identificado. Opcionalmente, e consequentemente, um domínio extracelular de um polipéptido PRO pode conter desde cerca de 5 ou menos aminoácidos de cada lado do limite do domínio transmembranar/domínio extracelular tal como identificado nos Exemplos ou na descrição e os polipéptidos com ou sem o péptido de sinal associado e os ácidos nucleicos que os codificam, estão contemplados na presente invenção.

A localização aproximada dos "péptidos de sinal" dos diferentes polipéptidos PRO aqui revelados apresentam-se no presente fascículo e/ou nas figuras anexas. Note-se, contudo, que o limite C-terminal de um péptido de sinal pode variar mas muito provavelmente não variará mais de cerca de 5 aminoácidos de cada lado do limite C-terminal do péptido de sinal tal como inicialmente aqui identificado, em que o limite C-terminal do péptido de sinal pode ser identificado de acordo com os critério rotineiramente empregues na especialidade para identificar esse tipo de elemento de sequência de aminoácidos (por exemplo, Nielsen et al., Prot. Eng., 10:1-6 (1997) e von Heinje et al., Nucl. Acids Res., 14:4683-4690 (1986)). Adicionalmente, reconhece-se igualmente que nalguns casos a clivagem de uma sequência de sinal de um polipéptido excretado não é inteiramente uniforme, resultando em mais do que uma espécie excretada. Os polipéptidos maduros em que o péptido de sinal é clivado num intervalo não superior a cerca de 5 aminoácidos de qualquer dos lados do limite C-terminal do péptido de sinal tal como aqui identificado e os polinucleótidos que os codificam, estão contemplados na presente invenção. "Variante de polipéptido PRO" significa um polipéptido PRO activo tal como definido atrás ou adiante possuindo pelo menos cerca de 80% de identidade de sequência de aminoácidos com uma sequência de um polipéptido PRO de sequência nativa de comprimento completo como aqui se descreve, uma sequência de um polipéptido PRO a que falta o péptido de sinal como aqui se descreve, um domínio extracelular de um polipéptido PRO, com ou sem o péptido de sinal, como aqui se descreve ou qualquer outro fragmento de uma sequência de um polipéptido PRO de comprimento completo

como aqui se descreve. Estas variantes de polipéptidos PRO incluem, por exemplo, polipéptidos PRO em que um ou mais resíduos de aminoácido são adicionados, ou eliminados, no terminal N ou C da sequência de aminoácidos nativa de comprimento completo. Vulgarmente, uma variante de polipéptido PRO terá pelo menos cerca de 80% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 81% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 82% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 83% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 84% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 85% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 86% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 87% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 88% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 89% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 90% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 91% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 92% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 93% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 94% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 95% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 96% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 97% de identidade de sequência de aminoácidos, alternativamente pelo menos cerca de 98% de identidade de sequência de aminoácidos e alternativamente pelo menos cerca de 99% de identidade de sequência de aminoácidos, com uma sequência de um polipéptido PRO de sequência nativa de comprimento completo como aqui se descreve, uma sequência de um polipéptido PRO a que falta o péptido de sinal como aqui se descreve, um domínio extracelular de um polipéptido PRO, com ou sem o péptido de sinal, como aqui se descreve ou qualquer outro fragmento especificamente definido de uma sequência de um polipéptido PRO de comprimento completo como aqui se

descreve. Vulgarmente, os polipéptidos PRO variantes têm pelo menos cerca de 10 aminoácidos de comprimento, alternativamente pelo menos cerca de 20 aminoácidos de comprimento, alternativamente pelo menos cerca de 30 aminoácidos de comprimento, alternativamente pelo menos cerca de 40 aminoácidos de comprimento, alternativamente pelo menos cerca de 50 aminoácidos de comprimento, alternativamente pelo menos cerca de 60 aminoácidos de comprimento, alternativamente pelo menos cerca de 70 aminoácidos de comprimento, alternativamente pelo menos cerca de 80 aminoácidos de comprimento, alternativamente pelo menos cerca de 90 aminoácidos de comprimento, alternativamente pelo menos cerca de 100 aminoácidos de comprimento, alternativamente pelo menos cerca de 150 aminoácidos de comprimento, alternativamente pelo menos cerca de 200 aminoácidos de comprimento, alternativamente pelo menos cerca de 300 aminoácidos de comprimento, ou mais. "Percentagem (%) de identidade de sequência de aminoácidos" em relação a sequências de polipéptido PRO aqui identificadas define-se como a percentagem de resíduos de aminoácido numa sequência candidata que são idênticos aos resíduos de aminoácido na sequência de PRO, após alinhamento das sequências e introdução de hiatos, se necessário, para se conseguir a máxima percentagem de identidade de sequências, e não considerando quaisquer substituições conservativas como parte da identidade de sequências. O alinhamento para fins da determinação da percentagem de identidade de sequências de aminoácidos pode ser realizado de várias maneiras que são conhecidas dos peritos na especialidade, por exemplo, utilizando *software* de computador disponível ao público tal como o *software* BLAST, BLAST-2, ALIGN, ALIGN-2 ou Megalign (DNASTAR). Os peritos na especialidade podem determinar os parâmetros apropriados para a medição do alinhamento, incluindo quaisquer algoritmos necessários para se conseguir o alinhamento máximo ao longo de todo o comprimento das sequências que se estão a comparar. Para os presentes fins, contudo, os valores de % de identidade de sequência de aminoácidos são obtidos como se descreve adiante utilizando o programa de computador para comparação de sequências ALIGN-2, sendo o código-fonte completo do programa ALIGN-2 proporcionado na Tabela 1. O programa de computador para comparação de sequências ALIGN-2 é da autoria da Genentech, Inc., e o código-fonte apresentado na Tabela 1 foi

apresentado com a documentação para o utilizador no U.S. Copyright Office, Washington D.C., 20559, onde está registado com o n.º de registo U.S. Copyright Registration TXU510087. O programa ALIGN-2 está disponível ao público através da Genentech, Inc., South San Francisco, California, ou pode ser compilado a partir do código-fonte proporcionado na Tabela 1. O programa ALIGN-2 deverá ser compilado para utilização num sistema operativo UNIX, preferivelmente UNIX V4.0D digital. Todos os parâmetros de comparação de sequências são estabelecidos pelo programa ALIGN-2 e não variam.

Para os presentes fins, a % de identidade de sequência de aminoácidos de uma determinada sequência de aminoácidos A relativamente a, com ou contra, uma determinada sequência de aminoácidos B (que pode ser alternativamente articulado como, uma determinada sequência de aminoácidos A que apresenta ou compreende uma determinada % de identidade de sequência de aminoácidos relativamente a, com ou contra, uma determinada sequência de aminoácidos B), calcula-se como se segue:

$$100 \text{ vezes a fracção } X/Y$$

em que X é o número de resíduos de aminoácidos pontuados como correspondências idênticas pelo programa de alinhamento de sequências ALIGN-2 no alinhamento efectuado pelo programa, de A e B, e em que Y é o número total de resíduos de aminoácidos em B. Notar-se-á que quando o comprimento da sequência de aminoácidos A não é igual ao comprimento da sequência de aminoácidos B, a % de identidade de sequência de aminoácidos de A relativamente a B não será igual à % de identidade de sequência de aminoácidos de B relativamente a A. Como exemplos de cálculos de % de identidade de sequência de aminoácidos, as Tabelas 2-3 demonstram como calcular a % de identidade de sequência de aminoácidos da sequência de aminoácidos designada "Proteína de comparação" relativamente à sequência de aminoácidos designada "PRO".

Salvo especificado em contrário, todos os valores de % de identidade de sequência de aminoácidos aqui utilizados são obtidos como descrito acima utilizando o programa de computador para comparação de sequências ALIGN-2. Contudo, a % de identidade de sequência de aminoácidos pode também ser

determinada utilizando o programa para comparação de sequências NCBI-BLAST2 (Altschul et al., Nucleic Acids Res., 25:3389-3402 (1997)). O programa para comparação de sequências NCBI-BLAST2 pode ser obtido em <http://www.ncbi.nlm.nih.gov>. ou obtido de outro modo no National Institute of Health, Bethesda, MD. O NCBI-BLAST2 utiliza vários parâmetros de busca, em que todos esses parâmetros de busca estão estabelecidos como valores por defeito incluindo, por exemplo, "unmask = yes", "strand = all", "expected occurrences = 10", "minimum low complexity length = 15/5", "multi-pass e-value = 0.01", "constant for multi-pass = 25", "dropoff for final gapped alignment = 25" e "scoring matrix = BLOSUM62".

Nas situações em que se utiliza o NCBI-BLAST2 para as comparações de sequências de aminoácidos, a % de identidade de sequência de aminoácidos de uma determinada sequência de aminoácidos A relativamente a, com ou contra, uma determinada sequência de aminoácidos B (que pode ser alternativamente articulado como uma determinada sequência de aminoácidos A que apresenta ou compreende uma determinada % de identidade de sequência de aminoácidos relativamente a, com ou contra, uma determinada sequência de aminoácidos B), calcula-se como se segue:

$$100 \text{ vezes a fracção } X/Y$$

em que X é o número de resíduos de aminoácidos pontuados como correspondências idênticas pelo programa de alinhamento de sequências NCBI-BLAST2 no alinhamento efectuado pelo programa, de A e B, e em que Y é o número total de resíduos de aminoácidos em B. Notar-se-á que quando o comprimento da sequência de aminoácidos A não é igual ao comprimento da sequência de aminoácidos B, a % de identidade de sequência de aminoácidos de A relativamente a B não será igual à % de identidade de sequência de aminoácidos de B relativamente a A.

Adicionalmente, a % de identidade de sequência de aminoácidos pode também ser determinada utilizando o programa de computador WU-BLAST-2 (Altschul et al., Methods in Enzymology, 266:460-480 (1996)). A maioria dos parâmetros de busca do WU-BLAST-2 são estabelecidos em valores por defeito. Os parâmetros não estabelecidos em valores por defeito, i.e.,

os parâmetros ajustáveis, são colocados nos seguintes valores: "overlap span = 1", "overlap fraction = 0.125", "word threshold (T) = 11", e "scoring matrix = BLOSUM62". Para os presentes fins, o valor da % de identidade de sequência de aminoácidos é determinado dividindo (a) o número de resíduos de aminoácidos com correspondência idêntica entre a sequência de aminoácidos do polipéptido PRO de interesse apresentando a sequência derivada do polipéptido PRO nativo e a sequência de aminoácidos de interesse de comparação (*i.e.*, a sequência contra a qual se está a comparar o polipéptido PRO de interesse que pode ser um polipéptido PRO variante) tal como determinado por WU-BLAST-2, por (b) o número total de resíduos de aminoácidos do polipéptido PRO de interesse. Por exemplo, na expressão "um polipéptido compreendendo uma sequência de aminoácidos A que apresenta, ou apresentando, pelo menos 80% de identidade de sequência de aminoácidos relativamente à sequência de aminoácidos B", a sequência de aminoácidos A é a sequência de aminoácidos de comparação de interesse e a sequência de aminoácidos B é a sequência de aminoácidos do polipéptido PRO de interesse.

"Polinucleótido de PRO variante" ou "sequência de ácido nucleico de PRO variante" significam uma molécula de ácido nucleico que codifica um polipéptido PRO activo tal como se define adiante e que possui pelo menos cerca de 80% de identidade de sequência de ácido nucleico com uma sequência de ácido nucleico que codifica uma sequência de um polipéptido PRO de sequência nativa de comprimento completo como aqui se descreve, uma sequência de um polipéptido PRO de sequência nativa de comprimento completo a que falta o péptido de sinal como aqui se descreve, um domínio extracelular de um polipéptido PRO, com ou sem o péptido de sinal, como aqui se descreve ou qualquer outro fragmento de uma sequência de um polipéptido PRO de comprimento completo como aqui se descreve. Vulgarmente, um polinucleótido de PRO variante terá pelo menos cerca de 80% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 81% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 82% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 83% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 84% de identidade de sequência de ácido nucleico,

alternativamente pelo menos cerca de 85% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 86% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 87% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 88% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 89% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 90% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 91% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 92% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 93% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 94% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 95% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 96% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 97% de identidade de sequência de ácido nucleico, alternativamente pelo menos cerca de 98% de identidade de sequência de ácido nucleico e ainda alternativamente pelo menos cerca de 99% de identidade de sequência de ácido nucleico, com uma sequência de ácido nucleico que codifica uma sequência de um polipéptido PRO de sequência nativa de comprimento completo como aqui se descreve, uma sequência de um polipéptido PRO de sequência nativa de comprimento completo a que falta o péptido de sinal como aqui se descreve, um domínio extracelular de um polipéptido PRO, com ou sem a sequência de sinal, como aqui se descreve ou qualquer outro fragmento de uma sequência de um polipéptido PRO de comprimento completo como aqui se descreve. As variantes não abrangem a sequência nucleotídica nativa.

Vulgarmente, os polinucleótidos de PRO variantes têm pelo menos cerca de 30 nucleótidos de comprimento, alternativamente pelo menos cerca de 60 nucleótidos de comprimento, alternativamente pelo menos cerca de 90 nucleótidos de comprimento, alternativamente pelo menos cerca de 120 nucleótidos de comprimento, alternativamente pelo menos cerca de 150 nucleótidos de comprimento, alternativamente pelo menos cerca de 180 nucleótidos de comprimento, alternativamente pelo menos cerca de 210 nucleótidos de

comprimento, alternativamente pelo menos cerca de 240 nucleótidos de comprimento, alternativamente pelo menos cerca de 270 nucleótidos de comprimento, alternativamente pelo menos cerca de 300 nucleótidos de comprimento, alternativamente pelo menos cerca de 450 nucleótidos de comprimento, alternativamente pelo menos cerca de 600 nucleótidos de comprimento, alternativamente pelo menos cerca de 900 nucleótidos de comprimento, ou mais. A "percentagem (%) de identidade de sequência de ácido nucleico" relativamente às sequências de ácido nucleico que codificam o polipéptido PRO aqui identificadas, define-se como a percentagem de nucleótidos numa sequência candidata que são idênticos aos nucleótidos numa sequência de ácido nucleico que codifica um polipéptido PRO após alinhamento das sequências e introdução de hiatos, caso necessário, para se conseguir a máxima percentagem de identidade de sequências. O alinhamento para fins da determinação da percentagem de identidade de sequências de ácido nucleico pode ser realizado de várias maneiras que são conhecidas dos peritos na especialidade, por exemplo, utilizando *software* de computador disponível ao público tal como o *software* BLAST, BLAST-2, ALIGN, ALIGN-2 ou Megalign (DNASTAR). Os peritos na especialidade podem determinar os parâmetros apropriados para a medição do alinhamento, incluindo quaisquer algoritmos necessários para se conseguir o alinhamento máximo ao longo de todo o comprimento das sequências que se estão a comparar. Para os presentes fins, contudo, os valores de % de identidade de sequência de ácido nucleico são obtidos como se descreve adiante utilizando o programa de computador para comparação de sequências ALIGN-2, sendo o código-fonte completo do programa ALIGN-2 proporcionado na Tabela 1. O programa de computador para comparação de sequências ALIGN-2 é da autoria da Genentech, Inc., e o código-fonte apresentado na Tabela 1 foi apresentado com a documentação para o utilizador no U.S. Copyright Office, Washington D.C., 20559, onde está registado com o n.º de registo US Copyright Registration TXU510087. O programa ALIGN-2 está disponível ao público através da Genentech, Inc., South San Francisco, California ou pode ser compilado a partir do código-fonte proporcionado na Tabela 1. O programa ALIGN-2 deverá ser compilado para utilização num sistema operativo UNIX, preferivelmente UNIX V4.0D digital.

Todos os parâmetros de comparação de sequências são estabelecidos pelo programa ALIGN-2 e não variam.

Para os presentes fins, a % de identidade de sequência de ácido nucleico de uma determinada sequência de ácido nucleico C relativamente a, com ou contra, uma determinada sequência de ácido nucleico D (que pode ser alternativamente articulado como uma determinada sequência de ácido nucleico C que apresenta ou compreende uma determinada % de identidade de sequência de ácido nucleico relativamente a, com ou contra, uma determinada sequência de ácido nucleico D), calcula-se como se segue:

$$100 \text{ vezes a fracção } W/Z$$

em que W é o número de nucleótidos pontuados como correspondências idênticas pelo programa de alinhamento de sequências ALIGN-2 no alinhamento efectuado pelo programa, de C e D, e em que Z é o número total de nucleótidos em D. Notar-se-á que quando o comprimento da sequência de ácido nucleico C não é igual ao comprimento da sequência de ácido nucleico D, a % de identidade de sequência de ácido nucleico de C relativamente a D não será igual à % de identidade de sequência de ácido nucleico de D relativamente a C. Como exemplos de cálculo da % de identidade de sequência de ácido nucleico, as tabelas 4-5 demonstram como calcular a % de identidade de sequência de ácido nucleico da sequência de ácido nucleico designada como "ADN de Comparação" relativamente à sequência de ácido nucleico designada como "PRO-ADN".

Salvo especificado em contrário, todos os valores de % de identidade de sequência de ácido nucleico aqui utilizados são obtidos como descrito acima utilizando o programa de computador para comparação de sequências ALIGN-2. Contudo, a % de identidade de sequências de ácido nucleico pode também ser determinada utilizando o programa para comparação de sequências NCBI-BLAST2 (Altschul et al., *Nucleic Acids Res.*, 25:3389-3402 (1997)). O programa para comparação de sequências NCBI-BLAST2 pode ser obtido de <http://www.ncbi.nlm.nih.gov>, ou obtido de outro modo do National Institute of Health, Bethesda, MD. O NCBI-BLAST2 utiliza vários parâmetros de

busca, em que todos esses parâmetros de busca estão estabelecidos em valores por defeito incluindo, por exemplo, "unmask = yes", "strand = all", "expected occurrences =10", "minimum low complexity length =15/5", "multi-pass e-value = 0.01", "constant for multi-pass = 25", "dropoff for final gapped alignment = 25" e "scoring matrix = BLOSUM62".

Nas situações em que se emprega o NCBI-BLAST2 para as comparações de sequências, a % de identidade de sequência de ácido nucleico de uma dada sequência de ácido nucleico C relativamente a, com ou contra, uma determinada sequência de ácido nucleico D (que pode ser alternativamente articulado como uma determinada sequência de ácido nucleico C que apresenta ou compreende uma determinada % de identidade de ácidos nucleicos relativamente a, com ou contra, uma determinada sequência de ácido nucleico D) calcula-se como se segue:

$$100 \text{ vezes a fracção } W/Z$$

em que W é o número de nucleótidos pontuados como correspondências idênticas pelo programa de alinhamento de sequências NCBI-BLAST2 no alinhamento efectuado pelo programa, de C e D, e em que Z é o número total de nucleótidos em D. Notar-se-á que quando o comprimento da sequência de ácido nucleico C não é igual ao comprimento da sequência de ácido nucleico D, a % de identidade de ácidos nucleicos de C relativamente a D não será igual à % de identidade de sequência de ácido nucleico de D relativamente a C.

Adicionalmente, os valores de % de identidade de sequência de ácido nucleico podem também ser gerados utilizando o programa de computador WU-BLAST-2 (Altschul et al., Methods in Enzymology, 266:460-480 (1996)). A maioria dos parâmetros de busca do WU-BLAST-2 são estabelecidos em valores por defeito. Os parâmetros não colocados em valores por defeito, i.e., os parâmetros ajustáveis, são estabelecidos nos seguintes valores: "overlap span = 1", "overlap fraction = 0.125", "word threshold (T) = 11" e "scoring matrix = BLOSUM62". Para os presentes fins, o valor da % de identidade de sequência de ácido nucleico é determinado dividindo (a) o número de nucleótidos com correspondência idêntica entre a

sequência de ácido nucleico da molécula de ácido nucleico que codifica o polipéptido PRO de interesse apresentando uma sequência derivada do ácido nucleico que codifica a sequência do polipéptido PRO de sequência nativa e a molécula de ácido nucleico de comparação de interesse (*i.e.*, a sequência contra a qual se está a comparar a molécula de ácido nucleico que codifica o polipéptido PRO de interesse que pode ser um polinucleótido de PRO variante) tal como determinado por WU-BLAST-2, por (b) o número total de nucleótidos da molécula de ácido nucleico que codifica o polipéptido PRO de interesse. Por exemplo, na expressão “uma molécula de ácido nucleico isolada compreendendo uma sequência de ácido nucleico A que apresenta ou apresentando pelo menos 80% de identidade de sequência de ácido nucleico com a sequência de ácido nucleico B”, a sequência de ácido nucleico A é a molécula de ácido nucleico de interesse de comparação e a sequência de ácido nucleico B é a sequência de ácido nucleico da molécula de ácido nucleico que codifica o polipéptido PRO de interesse.

Noutras concretizações, os polinucleótidos de PRO variantes são moléculas de ácido nucleico que codificam um polipéptido PRO activo e que são capazes de hibridar, preferivelmente sob condições de hibridação e lavagem rigorosas, com sequências nucleotídicas que codificam o polipéptido PRO de comprimento completo apresentado nas figuras aqui anexas. Os polipéptidos PRO variantes podem ser aqueles que são codificados por um polinucleótido de PRO variante.

A expressão “positivos”, no contexto de comparações de identidade de sequências de aminoácidos efectuadas como descrito acima, inclui não apenas os resíduos de aminoácidos nas sequências comparadas que são idênticos, mas também aqueles que apresentam propriedades semelhantes. Os resíduos de aminoácido que pontuam um valor positivo com um resíduo de aminoácido de interesse são aqueles que, ou são idênticos ao resíduo de aminoácido de interesse, ou são uma substituição preferida (tal como definido na Tabela 6 adiante) do resíduo de aminoácido de interesse.

Para os presentes fins, o valor de % de positivos de uma determinada sequência de aminoácidos A relativamente a, com ou

contra, uma determinada sequência de aminoácidos B (que pode ser alternativamente articulado como uma determinada sequência de aminoácidos A que apresenta ou compreende uma determinada % de positivos relativamente, com ou contra uma determinada sequência de aminoácidos B), calcula-se como se segue:

$$100 \text{ vezes a fracção } X/Y$$

em que X é o número de resíduos de aminoácidos pontuados como um valor positivo pelo programa de alinhamento de sequências ALIGN-2 no alinhamento efectuado pelo programa, de A e B, e em que Y é o número total de resíduos de aminoácidos em B. Notar-se-á que quando o comprimento da sequência de aminoácidos A não é igual ao comprimento da sequência de aminoácidos B, a % de positivos de A relativamente a B não será igual à % de positivos de B relativamente a A.

“Isolados”, quando utilizada para descrever os diferentes polipéptidos aqui revelados, significa um polipéptido que foi identificado e separado e/ou recuperado de uma componente do seu ambiente natural. Preferivelmente, o polipéptido isolado está isento de associação com todos os componentes com os quais está naturalmente associado. Os componentes contaminantes do seu ambiente natural são materiais que interfeririam tipicamente com utilizações em diagnóstico ou terapêuticas do polipéptido e podem incluir enzimas, hormonas e outros solutos proteínáceos ou não proteínáceos. Em concretizações preferidas, o polipéptido será purificado (1) até um grau suficiente para se obterem pelo menos 15 resíduos da sequência de aminoácidos N-terminais ou internos por utilização de um sequenciador de copo rotativo, ou (2) até à homogeneidade por SDS-PAGE sob condições não redutoras ou redutoras utilizando coloração com azul de Coomassie ou, preferivelmente, com prata. Os polipéptidos isolados incluem polipéptidos *in situ* no interior de células recombinantes, dado que não estará presente pelo menos um componente do ambiente natural do polipéptido PRO. Habitualmente, contudo, os polipéptidos isolados serão preparados através de pelo menos um passo de purificação.

Uma molécula de ácido nucleico “isolada” que codifica um polipéptido PRO ou uma molécula de ácido nucleico “isolada”

que codifica um anticorpo anti-PRO, é uma molécula de ácido nucleico que é identificada e separada de pelo menos uma molécula de ácido nucleico contaminante com a qual está habitualmente associada na fonte natural do ácido nucleico que codifica PRO ou na fonte natural do ácido nucleico que codifica o anti-PRO. Preferivelmente, o ácido nucleico isolado encontra-se isento de associação com todos os componentes com os quais está naturalmente associado. Uma molécula de ácido nucleico que codifica PRO isolada ou uma molécula de ácido nucleico que codifica um anti-PRO isolada encontram-se numa forma ou conformação distintos dos encontrados na natureza. As moléculas de ácido nucleico isoladas distinguem-se portanto da molécula de ácido nucleico que codifica PRO ou da molécula de ácido nucleico que codifica anti-PRO tal como existem em células naturais. Contudo, uma molécula de ácido nucleico isolada que codifica um polipéptido PRO ou uma molécula de ácido nucleico isolada que codifica um anticorpo anti-PRO incluem moléculas de ácido nucleico de PRO ou moléculas de ácido nucleico de anti-PRO contidas em células que habitualmente expressam polipéptidos PRO ou anticorpos anti-PRO em que, por exemplo, a molécula de ácido nucleico está numa localização cromossómica diferente da das células naturais.

A expressão "sequências de controlo" refere-se a sequências de ADN necessárias para a expressão de uma sequência de codificação operativamente ligada num organismo hospedeiro particular. As sequências de controlo que são adequadas para procariotas incluem, por exemplo, um promotor, opcionalmente uma sequência operadora, e um local de ligação ao ribossoma. Sabe-se que as células eucariotas utilizam promotores, sinais de poliadenilação e potenciadores.

O ácido nucleico encontra-se "operativamente ligado" quando está colocado numa relação funcional com outra sequência de ácido nucleico. Por exemplo, o ADN para uma pré-sequência ou um comando de secreção encontra-se operativamente ligado ao ADN para um polipéptido PRO se é expresso como uma pré-proteína que participa na secreção do polipéptido; um promotor ou um potenciador encontram-se operativamente ligados a uma sequência de codificação se afectam a transcrição da sequência; ou um local de ligação ao ribossoma encontra-se

operativamente ligado a uma sequência de codificação se está posicionado de modo a facilitar a tradução. Em geral, "operativamente ligado" significa que as sequências de ADN ligadas são contíguas e no caso de um comando de secreção, contíguas e em fase de leitura. Contudo, os potenciadores não têm que ser contíguos. A ligação consegue-se por ligação em locais de restrição convenientes. Se estes locais não existirem, utilizam-se adaptadores ou ligantes oligonucleotídicos sintéticos de acordo com a prática convencional.

A expressão "anticorpo" utiliza-se no seu sentido mais abrangente e cobre especificamente, por exemplo, anticorpos monoclonais anti-PRO simples (incluindo anticorpos agonistas), composições de anticorpos anti-PRO com especificidade poliepitópica, anticorpos anti-PRO de cadeia simples e fragmentos de anticorpos anti-PRO (ver adiante). A expressão "anticorpo monoclonal" tal como aqui utilizada refere-se a um anticorpo obtido de uma população de anticorpos substancialmente homogêneos, *i.e.*, os anticorpos individuais que constituem a população são idênticos excepto relativamente a possíveis mutações de ocorrência natural que podem estar presentes em quantidades mínimas.

O "rigor" das reacções de hibridação é facilmente determinável por uma pessoa competente na matéria e geralmente é um cálculo empírico dependente do comprimento da sonda, da temperatura de lavagem e da concentração salina. Em geral, sondas mais longas requerem temperaturas mais elevadas para uma hibridação correcta, enquanto que sondas mais curtas necessitam de temperaturas mais baixas. A hibridação geralmente depende da capacidade do ADN desnaturado tornar a hibridar quando estão presentes cadeias complementares num ambiente abaixo da sua temperatura de fusão. Quanto maior o grau de homologia pretendido entre a sonda e a sequência hibridável, maior a temperatura relativa que pode ser utilizada. Em resultado, temperaturas relativas mais elevadas tenderão a tornar as condições reaccionais mais rigorosas, enquanto que temperaturas mais baixas, menos. Para detalhes adicionais e explicação sobre o rigor das reacções de hibridação, veja-se Ausubel *et al.*, Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

“Condições rigorosas” ou “condições altamente rigorosas”, tal como aqui definido, podem ser identificadas como aquelas que: (1) empregam uma força iônica baixa e uma temperatura elevada para lavagem, por exemplo cloreto de sódio 0,015 M/citrato de sódio 0,0015 M/dodecilssulfato de sódio a 0,1% a 50°C; (2) empregam durante a hibridação um agente desnaturante, tal como formamida, por exemplo, formamida a 50% (v/v) com albumina sérica bovina a 0,1%/Ficoll a 0,1%/polivinilpirrolidona a 0,1%/tampão de fosfato de sódio 50 mM a pH 6,5 com cloreto de sódio 750 mM, citrato de sódio 75 mM a 42°C; ou (3) empregam formamida a 50%, SSC 5x (NaCl 0,75 M, citrato de sódio 0,075 M), fosfato de sódio 50 mM (pH 6,8), pirofosfato de sódio a 0,1%, solução de Denhardt 5x, ADN de esperma de salmão tratado com ultra-sons (50 µg/ml), SDS a 0,1% e sulfato de dextrano a 10% a 42°C, com lavagens a 42°C com SSC 0,2x (cloreto de sódio/citrato de sódio) e formamida a 50% a 55°C, seguido por uma lavagem de rigor elevado consistindo em SSC 0,1x contendo EDTA a 55°C.

“Condições moderadamente rigorosas” podem ser identificadas como descrito por Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989 e incluem a utilização de uma solução de lavagem e de condições de hibridação (e.g., temperatura, força iônica e % de SDS) menos rigorosas do que as descritas acima. Constitui um exemplo de condições moderadamente rigorosas a incubação durante a noite a 37°C numa solução compreendendo: formamida a 20%, SSC 5x (NaCl 150 mM, citrato trissódico 15 mM), fosfato de sódio 50 mM (pH 7,6), solução de Denhardt 5x, sulfato de dextrano a 10% e ADN de esperma de salmão cortado e desnaturado a 20 mg/ml, seguida de lavagem dos filtros com SSC 1x a cerca de 37°C-50°C. O perito na especialidade saberá como ajustar a temperatura, a força iônica, etc. conforme necessário para acomodar factores tais como o comprimento da sonda e semelhantes.

A expressão “marcado com epítipo” quando aqui utilizada refere-se a um polipéptido quimérico compreendendo um polipéptido PRO fundido com um “polipéptido marcador”. O polipéptido marcador tem resíduos suficientes para proporcionar um epítipo contra o qual pode ser desenvolvido um anticorpo, mas é suficientemente curto de modo a não

interferir com a actividade do polipéptido com o qual se encontra fundido. O polipéptido marcador é preferivelmente também fracamente único de modo a que o anticorpo não apresente reactividade cruzada substancial com outros epítomos. Geralmente os polipéptidos marcadores adequados têm pelo menos seis resíduos de aminoácidos e têm usualmente entre cerca de 8 e 50 resíduos de aminoácidos (preferivelmente, entre cerca de 10 e 20 resíduos de aminoácido).

Tal como aqui utilizada, a expressão "imunoadesina" designa moléculas do tipo anticorpo que combinam a especificidade de ligação de uma proteína heteróloga (uma "adesina") com as funções efectoras dos domínios constantes das imunoglobulinas. Estruturalmente, as imunoadesinas compreendem uma fusão de uma sequência de aminoácidos com a especificidade de ligação desejada que é diferente do local de reconhecimento e ligação ao antigénio de um anticorpo (*i.e.*, é "heteróloga") e de uma sequência de domínio constante de imunoglobulina. A parte adesina de uma molécula de imunoadesina é tipicamente uma sequência de aminoácidos contíguos incluindo pelo menos o local de ligação de um receptor ou de um ligando. A sequência de domínio constante de imunoglobulina da imunoadesina pode ser obtida de qualquer imunoglobulina tal como os subtipos IgG-1, IgG-2, IgG-3 ou IgG-4, IgA (incluindo IgA-1 e IgA-2), IgE, IgD ou IgM.

"Activo" ou "actividade", para os presentes fins, referem-se a forma(s) dos polipéptidos PRO que retêm uma actividade biológica e/ou imunológica dos polipéptidos PRO nativos ou de ocorrência natural, em que actividade "biológica" se refere a uma função biológica (de inibição ou de estimulação) causada por um polipéptido PRO nativo ou de ocorrência natural, diferente da capacidade para induzir a produção de um anticorpo contra um epítomo antigénico apresentado por um polipéptido PRO nativo ou de ocorrência natural, e uma actividade "imunológica" refere-se à capacidade de induzir a produção de um anticorpo contra um epítomo antigénico apresentado por um polipéptido PRO nativo ou de ocorrência natural.

"Actividade biológica" no contexto de um anticorpo ou de outro agonista que possa ser identificado pelos ensaios de

rastreio aqui revelados (e.g., uma molécula orgânica ou inorgânica pequena, um péptido, etc.) utiliza-se para referir a capacidade dessas moléculas para invocar um ou mais dos efeitos aqui enumerados em ligação à definição de uma "quantidade terapeuticamente eficaz". Numa concretização específica, "actividade biológica" é a capacidade de inibir o crescimento ou a proliferação de células neoplásicas. Uma actividade biológica preferida é a inibição, incluindo retardamento ou paragem completa, do crescimento de uma célula tumoral alvo (e.g., de cancro). Outra actividade biológica preferida é a actividade citotóxica que resulta na morte da célula tumoral alvo (e.g., de cancro). Ainda outra actividade biológica preferida é a indução de apoptose de uma célula tumoral alvo (e.g., de cancro).

A expressão "actividade imunológica" significa reactividade imunológica cruzada com pelo menos um epítipo de um polipéptido PRO. "Reactividade imunológica cruzada", tal como aqui se utiliza, significa que o polipéptido candidato é capaz de inibir competitivamente a actividade biológica qualitativa de um polipéptido PRO que apresenta esta actividade, com anti-soros policlonais criados contra o polipéptido PRO activo conhecido. Estes anti-soros preparam-se, de maneira convencional, por injeção de cabras ou coelhos, por exemplo, por via subcutânea, com o análogo activo conhecido em adjuvante completo de Freund seguida de injeção intraperitoneal ou subcutânea de reforço em meio incompleto de Freund. A reactividade imunológica cruzada é preferivelmente "específica", o que significa que a afinidade de ligação da molécula que apresenta reactividade imunológica cruzada (e.g., um anticorpo) identificada, para com o polipéptido PRO correspondente, é significativamente superior (preferivelmente pelo menos cerca de 2 vezes, mais preferivelmente pelo menos cerca de 4 vezes, ainda mais preferivelmente pelo menos cerca de 6 vezes, o mais preferivelmente pelo menos cerca de 8 vezes mais elevada) à afinidade de ligação dessa molécula para com qualquer outro polipéptido nativo conhecido.

"Tumor", como aqui se utiliza, refere-se a todo o crescimento e proliferação de células neoplásicas, quer

malignas quer benignas, e a todas as células e tecidos pré-cancerosos e cancerosos.

Os termos "cancro" e "canceroso" referem-se a, ou descrevem, a condição fisiológica em mamíferos que é tipicamente caracterizada por um crescimento celular desregulado. Os exemplos de cancros incluem, mas não se lhes limitam, carcinoma, linfoma, blastoma, sarcoma e leucemia. Exemplos mais particulares destes cancros incluem cancro da mama, cancro da próstata, cancro do cólon, cancro de células escamosas, cancro do pulmão de células pequenas, cancro do pulmão de células não pequenas, cancro dos ovários, cancro cervical, cancro gastrointestinal, cancro pancreático, glioblastoma, cancro do fígado, cancro da bexiga, hepatoma, cancro colorrectal, carcinoma do endométrio, carcinoma das glândulas salivares, cancro dos rins, cancro vulvar, cancro da tiróide, carcinoma hepático e vários tipos de cancro da cabeça e pescoço.

"Tratamento" consiste numa intervenção realizada com a intenção de prevenir o desenvolvimento ou a alterar a patologia de uma desordem. Deste modo, "tratamento" refere-se tanto ao tratamento terapêutico como a medidas profilácticas ou preventivas. Os que necessitam do tratamento incluem aqueles que já têm a desordem assim como aqueles em quem a desordem se pretende prevenir. No tratamento de tumores (e.g., cancro), um agente terapêutico pode diminuir directamente a patologia de células tumorais, ou tornar as células tumorais mais susceptíveis ao tratamento com outros agentes terapêuticos, e.g., radiação e/ou quimioterapia.

A "patologia" do cancro inclui todos os fenómenos que comprometem o bem estar do paciente. Estes incluem, sem limitação, crescimento celular anómalo ou incontrollável, metástase, interferência com o normal funcionamento de células vizinhas, libertação de citocinas ou outros produtos de secreção em níveis anómalos, supressão ou agravamento da resposta inflamatória ou imunológica, etc.

Uma "quantidade eficaz" de um polipéptido aqui descrito ou de um seu agonista, em referência a inibição do crescimento de células neoplásicas, é uma quantidade capaz de inibir, em

alguma extensão, o crescimento de células alvo. A expressão inclui uma quantidade capaz de invocar um efeito inibidor do crescimento, citostático e/ou citotóxico e/ou apoptose das células alvo. Uma "quantidade eficaz" de um polipéptido PRO ou de um seu agonista para fins de inibição do crescimento de células neoplásicas, pode ser determinada empiricamente e de uma maneira rotineira.

Uma "quantidade terapeuticamente eficaz", em referência ao tratamento de tumores, refere-se a uma quantidade capaz de invocar um ou mais dos seguintes efeitos: (1) inibição, em alguma extensão, do crescimento do tumor, incluindo, retardamento e paragem completa do crescimento; (2) redução do número de células tumorais; (3) redução da dimensão do tumor; (4) inibição (*i.e.*, redução, retardamento ou paragem completa) da infiltração de células tumorais em órgãos periféricos; (5) inibição (*i.e.*, redução, retardamento ou paragem completa) de metástases; (6) aumento da resposta imunitária antitumoral, a qual pode, mas não tem que, resultar na regressão ou rejeição do tumor; e/ou (7) alívio, em alguma extensão, de um ou mais sintomas associados à desordem. Uma "quantidade terapeuticamente eficaz" de um polipéptido PRO ou de um seu agonista, para fins de tratamento de tumores, pode ser determinada empiricamente e de uma maneira rotineira.

Uma "quantidade inibidora do crescimento" de um polipéptido PRO ou de um seu agonista, é uma quantidade capaz de inibir o crescimento de uma célula, especialmente uma célula tumoral, *e.g.*, de cancro, quer *in vitro* quer *in vivo*. Uma "quantidade inibidora do crescimento" de um polipéptido PRO ou de um seu agonista, para fins de inibição do crescimento de células neoplásicas, pode ser determinada empiricamente e de uma maneira rotineira.

Uma "quantidade citotóxica" de um polipéptido PRO ou de um seu agonista, é uma quantidade capaz de causar a destruição de uma célula, especialmente de uma célula tumoral, *e.g.*, de cancro, quer *in vitro* quer *in vivo*. Uma "quantidade citotóxica" de um polipéptido PRO ou de um seu agonista, para fins de inibição do crescimento de células neoplásicas, pode ser determinada empiricamente e de uma maneira rotineira.

A expressão "agente citotóxico", como aqui se utiliza, refere-se a uma substância que inibe ou previne a função de células e/ou causa a destruição de células. Pretende-se incluir na expressão isótopos radioactivos (e.g., I^{131} , I^{125} , Y^{90} e Re^{186}), agentes quimioterapêuticos e toxinas tais como toxinas enzimaticamente activas de origem bacteriana, fúngica, vegetal ou animal, ou seus fragmentos.

Um "agente quimioterapêutico" é um composto químico útil no tratamento de tumores, e.g., cancro. Os exemplos de agentes quimioterapêuticos incluem adriamicina, doxorubicina, epirubicina, 5-fluorouracilo, arabinósido de citosina ("Ara-C"), ciclofosfamida, tiotepa, busulfano, citoxina, taxóides, e.g., paclitaxel (Taxol, Bristol-Myers Squibb Oncology, Princeton, NJ), e doxetaxel (Taxotere, Rhône-Poulenc Rorer, Antony, França), toxotere, metotrexato, cisplatina, melfalano, vinblastina, bleomicina, etoposido, ifosfamida, mitomicina C, mitoxantrona, vincristina, vinorelbina, carboplatina, teniposido, daunomicina, carminomicina, aminopterina, dactinomicina, mitomicinas, esperamicinas (veja-se, Patente U.S. 4 675 187), melfalano e outras mostardas de azoto relacionadas. Estão também incluídos nesta definição os agentes hormonais que actuam para regular ou inibir a acção de hormonas sobre tumores tais como tamoxifeno e onapristona.

Um "agente inibidor do crescimento" quando aqui utilizado, refere-se a um composto ou uma composição que inibem o crescimento de uma célula, especialmente uma célula tumoral, e.g., de cancro, quer *in vitro* quer *in vivo*. Assim, o agente inibidor do crescimento é um que reduz significativamente a percentagem das células alvo em fase S. Os exemplos de agentes inibidores do crescimento incluem agentes que bloqueiam a progressão do ciclo celular (num sítio que não a fase S), tais como os agentes que induzem paragem G1 e paragem em fase M. Os bloqueadores de fase M clássicos incluem as vincas (vincristina e vinblastina), o taxol e os inibidores de topo II tais como doxorubicina, epirubicina, daunorubicina, etoposido e bleomicina. Os agentes que param G1 também espalham paragem em fase S, por exemplo, agentes alquilantes do ADN tais como tamoxifeno, prednisona, dacarbazina, mecloretamina, cisplatina, metotrexato,

5-fluorouracilo e ara-C. Pode-se encontrar informação adicional em The Molecular Basis of Cancer, Mendelsohn and Israel, eds., Capítulo 1, intitulado "Cell cycle regulation, oncogens, and antineoplastic drugs" de Murakami *et al.*, (WB Saunders: Philadelphia, 1995), especialmente p. 13.

O termo "citoquina" é um termo genérico para proteínas libertadas por uma população celular que actuam sobre outra célula como mediadores intercelulares. Os exemplos destas citoquinas são linfoquinas, monoquinas e hormonas polipeptídicas tradicionais. Estão incluídas entre as citoquinas a hormona do crescimento tal como a hormona do crescimento humano, N-metionil-hormona do crescimento humana, e hormona do crescimento bovina; hormona paratiróide; tiroxina; insulina; pró-insulina; relaxina; pro-relaxina; hormonas glicoproteínicas tais como hormona estimulante dos folículos (FSH), hormona estimulante da tiróide (TSH) e hormona luteinizante (LH); factor de crescimento hepático; factor de crescimento de fibroblastos; prolactina; lactogénio placentário; factor de necrose tumoral α e β ; substância inibidora da substância mulleriana; péptido associado à gonadotropina de ratinho; inibina; activina; factor de crescimento endotelial vascular; integrina; trombopoietina (TPO); factores de crescimento de nervos tais como NGF- β ; factor de crescimento de plaquetas; factores de crescimento de transformação (TGF) tais como TGF- α e TGF- β ; factor de crescimento semelhante da insulina I e II; eritropoietina (EPO); factores osteoindutores; interferões tais como interferão- α , - β e - γ ; factores estimulantes de colónias (CSF) tais como CSF de macrófagos (M-CSF); CSF de granulócitos-macrófagos (GM-CSF); e CSF de granulócitos (G-CSF); interleucinas (IL) tais como IL-1, IL-1 α , IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-11, IL-12, um factor de necrose tumoral tal como TNF- α ou TNF- β ; e outros factores polipeptídicos incluindo LIF e ligando *kit* (KL). Como aqui se utiliza, o termo citoquina inclui proteínas de fontes naturais ou de cultura de células recombinantes e equivalentes biologicamente activas das citoquinas de sequência nativa.

O termo "pró-fármaco", como utilizado no presente pedido, refere-se a uma forma precursora ou derivada de uma substância farmacologicamente activa que é menos citotóxica para células

tumorais em comparação com o fármaco progenitor e é susceptível de ser enzimaticamente activada ou convertida na forma progenitora, mais activa. Veja-se, e.g., Wilman, "Prodrugs in Cancer Chemotherapy", Biochemical Society Transactions, 14, pp. 375-382, 615th Meeting Belfast (1986) and Stella et al., "Prodrugs: A Chemical Approach to Targeted Drug Delivery," Directed Drug Delivery, Borchardt et al., (ed.), pp. 247-267, Humana Press (1985). Os pró-fármacos da presente invenção incluem, mas não se lhes limitam, pró-fármacos contendo fosfato, pró-fármacos contendo tiofosfato, pró-fármacos glicosilados ou pró-fármacos contendo fenilacetamida opcionalmente substituídos, pró-fármacos de 5-fluorocitosina e de outra 5-fluorouridina, que podem ser derivados numa forma de pró-fármaco para utilização na presente invenção incluem, mas não se lhes limitam, os agentes quimioterapêuticos acima descritos.

O termo "agonista" é utilizado no seu sentido mais amplo e inclui qualquer molécula que imite uma actividade biológica de um polipéptido PRO nativo aqui descrito. As moléculas agonistas adequadas incluem especificamente anticorpos agonistas ou fragmentos de anticorpos, fragmentos ou variantes de sequências de aminoácidos de polipéptidos PRO nativos, péptidos, moléculas orgânicas pequenas, etc. Os métodos para a identificação de agonistas de um polipéptido PRO podem compreender o contacto de uma célula tumoral com uma molécula agonista candidata e medição da inibição do crescimento das células tumorais.

Administração "crónica" refere-se à administração do(s) agente(s) num modo contínuo, em oposição a um modo agudo, de modo a manter o efeito terapêutico inicial (actividade) durante um período de tempo prolongado. Administração "intermitente" é o tratamento que não é feito consecutivamente sem interrupção, mas é, em vez disso, de natureza cíclica.

"Mamífero", para fins de tratamento, refere-se a qualquer animal classificado como mamífero, incluindo seres humanos, animais domésticos e de quinta, e animais de zoológico, desporto, ou animais de estimação, tais como cães, gatos, gado vacum, cavalos, ovelhas, porcos, cabras, coelhos, etc. Preferivelmente, o mamífero é o ser humano.

Administração "em combinação com" um ou mais outros agentes terapêuticos inclui a administração simultânea (concorrente) e consecutiva, por qualquer ordem.

"Transportadores", como aqui se utiliza, incluem transportadores, excipientes ou estabilizantes, farmacologicamente aceitáveis, que não são tóxicos para a célula ou para o mamífero a eles exposto nas dosagens e concentrações empregues. Frequentemente, o transportador fisiologicamente aceitável é uma solução aquosa de pH tamponado. Os exemplos de transportadores fisiologicamente aceitáveis incluem tampões tais como fosfato, citrato e outros ácidos orgânicos; antioxidantes incluindo ácido ascórbico; polipéptidos de baixo peso molecular (menos de cerca de 10 resíduos); proteínas, tais como albumina sérica, gelatina ou imunoglobulinas; polímeros hidrófilos tais como polivinilpirrolidona; aminoácidos tais como glicina, glutamina, asparagina, arginina ou lisina; monossacáridos, dissacáridos e outros hidratos de carbono, incluindo glucose, manose ou dextrinas; agentes quelantes tais como EDTA; aldóis tais como manitol ou sorbitol; contra-íões formadores de sais tais como sódio; e/ou tensioactivos não iónicos tais como TWEEN[™], polietilenoglicol (PEG) e PLURONICS[™].

"Anticorpos nativos" e "imunoglobulinas nativas" são habitualmente glicoproteínas heterotetraméricas com cerca de 150 000 dalton, compostas por duas cadeias leves (L) idênticas e duas cadeias pesadas (H) idênticas. Cada cadeia leve encontra-se ligada a uma cadeia pesada através de uma ligação dissulfureto covalente, embora o número de ligações dissulfureto varie entre as cadeias pesadas de diferentes isotipos de imunoglobulinas. Cada cadeia pesada e leve apresenta igualmente pontes de dissulfureto regularmente espaçadas. Cada cadeia pesada apresenta numa das extremidades um domínio variável (V_H) seguido por vários domínios constantes. Cada cadeia leve apresenta um domínio variável (V_L) numa das extremidades e um domínio constante na sua outra extremidade; o domínio constante da cadeia leve está alinhado com o primeiro domínio constante da cadeia pesada e o domínio variável da cadeia leve está alinhado com o domínio variável da cadeia pesada. Crê-se que determinados resíduos de

aminoácidos formam uma interface entre os domínios variáveis de cadeia leve e de cadeia pesada.

A expressão “variável” refere-se ao facto de certas porções dos domínios variáveis diferirem extensamente na sequência entre anticorpos e serem utilizadas na ligação e especificidade de cada anticorpo particular relativamente ao seu antigénio particular. Contudo, a variabilidade não se encontra uniformemente distribuída ao longo dos domínios variáveis dos anticorpos. Encontra-se concentrada em três segmentos denominados regiões determinantes de complementaridade (CDR) ou regiões hipervariáveis tanto nos domínios variáveis de cadeia leve como de cadeia pesada. As porções mais altamente conservadas dos domínios variáveis denominam-se regiões de esqueleto (FR). Cada um dos domínios variáveis das cadeias pesada e leve nativas compreende quatro regiões FR, que adoptam largamente uma configuração em folha β ligadas por três CDR que formam ansas que ligam a estrutura em folha β e nalguns casos fazem parte dessa estrutura. As CDR em cada cadeia são mantidas juntas em estreita proximidade pelas regiões FR e, juntamente com as CDR da outra cadeia, contribuem para a formação do local de ligação ao antigénio dos anticorpos (veja-se Kabat et al., NIH Publ. No. 91-3242, Vol. I, páginas 647-669 (1991)). Os domínios constantes não estão directamente envolvidos na ligação de um anticorpo a um antigénio mas apresentam diferentes funções efectoras tais como a participação do anticorpo na toxicidade celular dependente de anticorpos.

A expressão “região hipervariável”, quando aqui utilizada, refere-se aos resíduos de aminoácidos de um anticorpo que são responsáveis pela ligação ao antigénio. A região hipervariável compreende resíduos de aminoácido de uma “região determinante de complementaridade” ou “CDR” (i.e., os resíduos 24-34 (L1), 50-56 (L2) e 89-97 (L3) no domínio variável da cadeia leve e 31-35 (H1), 50-65 (H2) e 95-102 (H3) no domínio variável da cadeia pesada; Kabat et al., Sequences of Proteins of Immunological Interest, 5ª Ed. Public Health Service, National Institute of Health, Bethesda, Maryland [1991]) e/ou os resíduos de uma “ansa hipervariável” (i.e., os resíduos 26-32 (L1), 50-52 (L2) e 91-96 (L3) no domínio variável de cadeia leve e 26-32 (H1), 53-55 (H2) e 96-101 (H3)

no domínio variável de cadeia pesada; Clothia e Lesk, J. Mol. Biol., 196:901-917 [1987]). Os resíduos de "esqueleto" ou "FR" são os resíduos do domínio variável que não os resíduos da região hipervariável tal como aqui definido.

Os "fragmentos de anticorpos" compreendem uma porção de um anticorpo intacto, preferivelmente a região de ligação ao antígeno ou variável do anticorpo intacto. Os exemplos de fragmentos de anticorpos incluem os fragmentos Fab, Fab', F(ab')₂ e Fv; diacorpos; anticorpos lineares (Zapata et al., Protein Eng., 8(10): 1057-1062 [1995]); moléculas de anticorpo de cadeia simples; e anticorpos multiespecíficos formados a partir de fragmentos de anticorpos.

A digestão de anticorpos com papaína produz dois fragmentos de ligação ao antígeno idênticos, denominados fragmentos "Fab", cada um com um único local de ligação ao antígeno e um fragmento "Fc" residual, uma designação que reflecte a capacidade para cristalizar facilmente. O tratamento com pepsina origina um fragmento F(ab')₂ que possui dois locais de combinação com antígeno e é ainda capaz de ligar cruzadamente o antígeno.

"Fv" é o fragmento de anticorpo mínimo que contém um local de reconhecimento e ligação ao antígeno, completo. Esta região consiste num dímero de um domínio variável de cadeia pesada e um de cadeia leve em forte associação não covalente. É nesta configuração que as três CDR de cada domínio variável interagem para definir um local de ligação ao antígeno na superfície do dímero V_H-V_L. Colectivamente, as seis CDR conferem ao anticorpo especificidade de ligação ao antígeno. Contudo, mesmo apenas um domínio variável (ou metade de um Fv compreendendo apenas três CDR específicas para um antígeno) tem a capacidade de reconhecer e ligar o antígeno, apesar de o fazer com menor afinidade do que o local de ligação inteiro.

O fragmento Fab contém também o domínio constante da cadeia leve e o primeiro domínio constante (CH1) da cadeia pesada. Os fragmentos Fab diferem de fragmentos Fab' por adição de alguns resíduos no terminal carboxi do domínio de cadeia pesada CH1 incluindo uma ou mais cisteínas da região de charneira do anticorpo. Fab'-SH é a designação aqui utilizada

para Fab' em que o(s) resíduo(s) de cisteína dos domínios constantes possuem um grupo tiol livre. Produziram-se originalmente fragmentos de anticorpo F(ab')₂ como pares de fragmentos Fab' que apresentam cisteínas de charneira entre si. São igualmente conhecidos outros acoplamentos químicos de fragmentos de anticorpos.

As "cadeias leves" de anticorpos (imunoglobulinas) de quaisquer espécies de vertebrados podem ser atribuídas a um ou dois tipos claramente distintos denominados κ e λ , com base nas sequências de aminoácido dos seus domínios constantes.

Dependendo da sequência de aminoácidos do domínio constante das suas cadeias pesadas, as imunoglobulinas podem ser atribuídas a diferentes classes. Existem cinco classes principais de imunoglobulinas: IgA, IgD, IgE, IgG e IgM e várias destas podem ser adicionalmente subdivididas em subclasses (isotipos), e.g., IgG1, IgG2, IgG3, IgG4, IgA e IgA2.

A expressão "anticorpo monoclonal" tal como aqui utilizada, refere-se a um anticorpo obtido de uma população de anticorpos substancialmente homogêneos, i.e., os anticorpos individuais que constituem a população são idênticos excepto relativamente a possíveis mutações de ocorrência natural que possam estar presentes em quantidades mínimas. Os anticorpos monoclonais são altamente específicos, sendo dirigidos contra um único local antigénico. Adicionalmente, contrariamente às preparações de anticorpos convencionais (policlonais) que incluem tipicamente anticorpos diferentes dirigidos contra determinantes (epítomos) diferentes, cada anticorpo monoclonal é dirigido contra um único determinante no antigénio. Além da sua especificidade, os anticorpos monoclonais são vantajosos por serem sintetizados por cultura de hibridomas, não contaminados por outras imunoglobulinas. O adjectivo "monoclonal" indica o carácter do anticorpo como sendo obtido de uma população substancialmente homogênea de anticorpos e não deve ser entendido como requerendo a produção do anticorpo por qualquer método em particular. Por exemplo, os anticorpos monoclonais a utilizar de acordo com a presente invenção podem ser produzidos pelo método do hibridoma, pela primeira vez

descrito por Kohler et al., Nature, 256:495 [1975], ou podem ser produzidos por métodos de ADN recombinante (veja-se, por exemplo, a Patente U.S. 4 816 567). Os “anticorpos monoclonais” podem ser também isolados de bibliotecas de anticorpos sobre fagos utilizando as técnicas descritas em Clackson et al., Nature, 352:624-628 [1991] e Marks et al., J. Mol. Biol., 222:581-597 (1991), por exemplo.

Os presentes anticorpos monoclonais incluem especificamente anticorpos “quiméricos” (imunoglobulinas) nos quais uma porção da cadeia pesada e/ou leve é idêntica ou homóloga às sequências correspondentes em anticorpos derivados de uma espécie particular ou pertencentes a uma classe ou subclasse particular de anticorpo, enquanto que a(s) cadeia(s) restante(s) são idênticas ou homólogas às sequências correspondentes em anticorpos derivados de outra espécie ou pertencentes a outra classe ou subclasse de anticorpos, assim como fragmentos de tais anticorpos, desde que apresentem a actividade biológica desejada (Patente U.S. 4 816 567; Morrison et al., Proc. Natl. Acad. Sci. USA, 81:6851-6855 [1984]).

Formas “humanizadas” de anticorpos não humanos (e.g., murinos) são imunoglobulinas quiméricas, cadeias de imunoglobulina ou seus fragmentos (tais como Fv, Fab, Fab', F(ab')₂ ou outras subsequências de ligação ao antigénio de anticorpos) que contêm uma sequência mínima derivada de imunoglobulina não humana. Na sua maioria, os anticorpos humanizados são imunoglobulinas humanas (anticorpo receptor) em que resíduos de uma CDR do receptor são substituídos por resíduos de uma CDR de uma espécie não humana (anticorpo dador) tal como de ratinho, rato ou coelho, apresentando a especificidade, afinidade e capacidade desejadas. Nalguns casos, substituem-se resíduos FR de Fv de imunoglobulina humana pelos correspondentes resíduos não humanos. Adicionalmente, os anticorpos humanizados podem compreender resíduos que não existem nem no anticorpo receptor nem nas sequências de CDR ou de esqueleto importadas. Estas modificações são efectuadas para refinar e maximizar adicionalmente o desempenho do anticorpo. Em geral, o anticorpo humanizado compreenderá substancialmente todos de pelo menos um, e tipicamente dois, domínios variáveis, nos

quais todas ou substancialmente todas as regiões CDR correspondem às de uma imunoglobulina não humana e todas ou substancialmente todas as regiões FR são as de uma sequência de imunoglobulina humana. O anticorpo humanizado optimamente também compreenderá pelo menos uma porção de uma região constante de imunoglobulina (Fc), tipicamente a de uma imunoglobulina humana. Para mais detalhes veja-se Jones et al., Nature, 321:522-525 (1986); Reichmann et al., Nature, 332:323-329 [1988]; e Presta, Curr. Op. Struct. Biol., 2:593-596 (1992). O anticorpo humanizado inclui um anticorpo PRIMATIZED™ em que a região de ligação ao antígeno do anticorpo é derivada de um anticorpo produzido por imunização de macacos *Macaque* com o antígeno de interesse.

Fragmentos de anticorpo "Fv de cadeia única" ou "sFv" incluem os domínios V_H e V_L de anticorpos em que estes domínios estão presentes numa única cadeia polipeptídica. Preferivelmente, o polipéptido Fv compreende adicionalmente um ligante polipeptídico entre os domínios V_H e V_L que permite que sFv forme a estrutura desejada para ligação ao antígeno. Para uma revisão sobre os sFv veja-se Pluckthun em The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg e Moore ed., Springer-Verlag, New York, p. 269-315 (1994).

A expressão "diacorpos" refere-se a fragmentos de anticorpo pequenos com dois locais de ligação ao antígeno em que os fragmentos compreendem um domínio variável de cadeia pesada (V_H) ligado a um domínio variável de cadeia leve (V_L) na mesma cadeia polipeptídica (V_H-V_L). Utilizando um ligante que é demasiado curto para permitir emparelhamento entre os dois domínios da mesma cadeia, forçam-se os domínios a emparelhar com os domínios complementares de outra cadeia e a criar dois locais de ligação ao antígeno. Descrevem-se diacorpos com mais detalhes em, por exemplo, EP 404 097; WO 93/11161; e Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

Um anticorpo "isolado" é um que foi identificado e separado e/ou recuperado de um componente do seu ambiente natural. Os componentes contaminantes do seu ambiente natural são materiais que interferiram com utilizações do anticorpo em diagnóstico ou terapêutica e podem incluir enzimas, hormonas e

outros solutos proteínáceos ou não proteínáceos. Em concretizações preferidas, o anticorpo será purificado (1) até mais de 95% em peso de anticorpo tal como determinado pelo método de Lowry e mais preferivelmente até mais de 99% em peso, (2) até um grau suficiente para se obterem pelo menos 15 resíduos da sequência de aminoácidos N-terminais ou internos por utilização de um sequenciador de copo rotativo, ou (3) até à homogeneidade por SDS-PAGE em condições redutoras ou não redutoras utilizando coloração com azul de Coomassie ou, preferivelmente, com prata. O anticorpo isolado inclui o anticorpo *in situ* no interior de células recombinantes dado que não estará presente pelo menos um componente do ambiente natural do anticorpo. Habitualmente, contudo, o anticorpo isolado será preparado através de pelo menos um passo de purificação.

A palavra "marcador", quando aqui utilizada, refere-se a um composto ou uma composição detectáveis que são directa ou indirectamente conjugados com o anticorpo de modo a gerar um anticorpo "marcado". O marcador pode ser detectável por si (e.g., marcadores radioisotópicos ou marcadores fluorescentes) ou, no caso de um marcador enzimático, pode catalisar uma alteração química de um composto ou composição substrato que é detectável. O marcador pode também ser uma entidade não detectável tal como uma toxina.

"Fase sólida" significa uma matriz não aquosa à qual o anticorpo da presente invenção pode aderir. Os exemplos de fases sólidas aqui abrangidas incluem as que são formadas parcial ou inteiramente por vidro (e.g., vidro de porosidade controlada), polissacáridos (e.g., agarose), poliacrilamidas, poliestireno, poli(álcool vinílico) e silicones. Em determinadas concretizações, dependendo do contexto, a fase sólida pode compreender o poço de uma placa de ensaio; noutras é uma coluna de purificação (e.g., uma coluna de cromatografia de afinidade). Esta expressão também inclui uma fase sólida descontínua de partículas discretas, tais como as descritas na patente U.S. 4 275 149.

Um "lipossoma" é uma vesícula pequena composta por vários tipos de lípidos, fosfolípidos e/ou tensioactivos, que é útil para entregar um fármaco (tal como um polipéptido PRO ou um

seu anticorpo) a um mamífero. Os componentes do lipossoma são habitualmente dispostos numa conformação em bicamada, semelhante ao arranjo dos lípidos das membranas biológicas.

Uma "molécula pequena" define-se aqui como possuindo um peso molecular inferior a cerca de 500 Daltons.

Tal como apresentado adiante, a Tabela 1 proporciona o código-fonte completo para o programa de computador para comparação de sequências, ALIGN-2. Este código-fonte pode ser rotineiramente compilado para utilização num sistema operativo UNIX para proporcionar o programa de computador para comparação de sequências, ALIGN-2.

Adicionalmente, as Tabelas 2-5 apresentam exemplos hipotéticos para utilização do método descrito adiante para determinar a % de identidade de sequência de aminoácidos (Tabelas 2-3) e a % de identidade de sequências de ácidos nucleicos (Tabelas 4-5) utilizando o programa de computador para comparação de sequências ALIGN-2, em que "PRO" representa a sequência de aminoácidos de um polipéptido PRO hipotético de interesse, "Proteína de Comparação" representa a sequência de aminoácidos de um polipéptido contra a qual se compara o polipéptido "PRO" de interesse, "PRO-ADN" representa uma sequência de ácido nucleico que codifica um PRO hipotética de interesse, "ADN de Comparação" representa a sequência nucleotídica de uma molécula de ácido nucleico contra a qual se compara a molécula de ácido nucleico "PRO-ADN" de interesse, "X", "Y" e "Z" representam, cada um, resíduos de aminoácido hipotéticos diferentes e "N", "L" e "V" representam, cada um, nucleótidos hipotéticos diferentes.

Tabela 1

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 / * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int _day[26][26] = {
/*   A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */ { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ { -2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ { -4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ { -1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ { -1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ { -1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ { -2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ { -1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ { _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, 0, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ { -2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ { -6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ { -3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

Tabela 1 (cont.)

```

/*
*/
#include <stdio.h>
#include <ctype.h>

#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
#define MX           4      /* save if there's at least MX-1 bases since last jmp */

#define DMAT         3      /* value of matching bases */
#define DMIS         0      /* penalty for mismatched bases */
#define DINS0         8      /* penalty for a gap */
#define DINS1         1      /* penalty per base */
#define PINS0         8      /* penalty for a gap */
#define PINS1         4      /* penalty per residue */

struct jmp {
    short          n[MAXJMP];    /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP];    /* base no. of jmp in seq x */
};
/* limits seq to 2^16 -1 */

struct diag {
    int            score;        /* score at last jmp */
    long           offset;       /* offset of prev block */
    short          ijmp;         /* current jmp index */
    struct jmp     jp;           /* list of jmps */
};

struct path {
    int            spc;          /* number of leading spaces */
    short          n[JMPS];      /* size of jmp (gap) */
    int            x[JMPS];      /* loc of jmp (last elem before gap) */
};

char             *ofile;        /* output file name */
char             *namex[2];     /* seq names: getseqs() */
char             *prog;         /* prog name for err msgs */
char             *seqx[2];      /* seqs: getseqs() */
int              dmax;          /* best diag: nw() */
int              dmax0;         /* final diag */
int              dna;           /* set if dna: main() */
int              endgaps;       /* set if penalizing end gaps */
int              gapx, gapy;    /* total gaps in seqs */
int              len0, len1;    /* seq lens */
int              ngapx, ngapy;  /* total size of gaps */
int              smax;          /* max score: nw() */
int              *xbm;          /* bitmap for matching */
long             offset;        /* current offset in jmp file */
struct diag      *dx;           /* holds diagonals */
struct path      pp[2];         /* holds path for seqs */

char             *calloc(), *malloc(), *index(), *strcpy();
char             *getseq(), *g_calloc();

```

Tabela 1 (cont.)

```

/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 * where file1 and file2 are two dna or two protein sequences.
 * The sequences can be in upper- or lower-case and may contain ambiguity
 * Any lines beginning with ';', '>' or '<' are ignored
 * Max file length is 65535 (limited by unsigned short x in the jmp struct)
 * A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 * Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)
    int    ac;
    char   *av[];
{
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';', '>' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0;                /* 1 to penalize endgaps */
    ofile = "align.out";        /* output file */

    nw();                       /* fill in the matrix, get the possible jmps */
    readjmps();                 /* get the actual jmps */
    print();                     /* print stats, alignment */

    cleanup(0);                 /* unlink any tmp files */
}

```

main

Tabela 1 (cont.)

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()
{
    char          *px, *py;          /* seqs and ptrs */
    int            *ndely, *dely;     /* keep track of dely */
    int            ndelx, delx;       /* keep track of delx */
    int            *tmp;              /* for swapping row0, row1 */
    int            mis;               /* score for each type */
    int            ins0, ins1;        /* insertion penalties */
    register       id;                /* diagonal index */
    register       ij;               /* jmp index */
    register       *col0, *col1;      /* score for curr, last row */
    register       xx, yy;            /* index into seqs */

    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;          /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

    /* fill in match matrix
     */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
         */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }
}

```

nw

Tabela 1 (cont.)

...nw

```

for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
     * favor new del over ongoing del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0+ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongoing del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0+ins1) >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
     * mis over any del and delx over dely
     */

```

Tabela 1 (cont.)

...nw

```

id = xx * yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    coll[yy] = mis;
else if (delx >= dely[yy]) {
    coll[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writeimps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
        dx[id].jp.n[ij] = ndelx;
        dx[id].jp.x[ij] = xx;
        dx[id].score = delx;
    }
} else {
    coll[yy] = dely[yy];
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writeimps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
        dx[id].jp.n[ij] = -ndely[yy];
        dx[id].jp.x[ij] = xx;
        dx[id].score = dely[yy];
    }
}
if (xx == len0 && yy < len1) {
    /* last col
    */
    if (endgaps)
        coll[yy] -= ins0 + ins1 * (len1 - yy);
    if (coll[yy] > smax) {
        smax = coll[yy];
        dmax = id;
    }
}
}
if (endgaps && xx < len0)
    coll[yy-1] -= ins0 + ins1 * (len0 - xx);
if (coll[yy-1] > smax) {
    smax = coll[yy-1];
    dmax = id;
}
tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)col1);
}

```

Tabela 1 (cont.)

```

/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC 3
#define P_LINE 256 /* maximum output line */
#define P_SPC 3 /* space between name or num and seq */

extern _day[26][26];
int olen; /* set output line length */
FILE *fx; /* output file */

print()
{
    int lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

```

print

Tabela 1 (cont.)

```

/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)
    int      lx, ly;          /* "core" (minus endgaps) */
    int      firstgap, lastgap; /* leading trailing overlap */
{
    int      nm, i0, i1, siz0, siz1;
    char     outx[32];
    double   pct;
    register n0, n1;
    register char *p0, *p1;

    /* get total matches, score
     */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
        else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
    }

    /* pct homology:
     * if penalizing endgaps, base is the shorter seq
     * else, knock off overhangs and take shorter core
     */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
    pct = 100.*((double)nm)/((double)lx);
    fprintf(fx, "\n");
    fprintf(fx, "< %d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);

```

getmat

Tabela 1 (cont.)

```

fprintf(fx, "< gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapx, (dna)? "base": "residue", (ngapx == 1)? "": "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
            ngapy, (dna)? "base": "residue", (ngapy == 1)? "": "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "\n< score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINS0, DINS1);
    else
        fprintf(fx,
            "\n< score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            smax, PINS0, PINS1);
    if (endgaps)
        fprintf(fx,
            "< endgaps penalized, left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base": "residue", (firstgap == 1)? "": "s",
            lastgap, (dna)? "base": "residue", (lastgap == 1)? "": "s");
    else
        fprintf(fx, "< endgaps not penalized\n");
}

static      nm;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];       /* jmp index for a path */
static      nc[2];       /* number at start of current line */
static      ni[2];       /* current elem number -- for gapping */
static      siz[2];
static char *ps[2];      /* ptr to current element */
static char *po[2];      /* ptr to next output char slot */
static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
    int      nm;          /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nm = stripname(namex[i]);
        if (nm > lmax)
            lmax = nm;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}

```

...getmat

pr_align

Tabela 1 (cont.)

```

for (nn = nm = 0, more = 1; more;) {
    for (i = more = 0; i < 2; i++) {
        /*
         * do we have more of this sequence?
         */
        if (!*ps[i])
            continue;

        more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
        }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
        }
        else { /* we're putting a seq element
            */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
            po[i]++;
            ps[i]++;

            /*
             * are we at next gap for this seq?
             */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                siz[i] = pp[i].n[ij[i] + 1];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i] + 1];
            }
            ni[i]++;
        }
    }
    if (++nn == olen || !more && nm) {
        dumpblock();
        for (i = 0; i < 2; i++)
            po[i] = out[i];
        nm = 0;
    }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()
{
    register i;

    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';
}

```

...pr_align

dumpblock

Tabela 1 (cont.)

...dumpblock

```

(void) putc('\n', fx);
for (i = 0; i < 2; i++) {
    if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
        if (i == 0)
            nums(i);
        if (i == 0 && *out[1])
            stars();
        putline(i);
        if (i == 0 && *out[1])
            fprintf(fx, star);
        if (i == 1)
            nums(i);
    }
}

```

```

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)
int ix; /* index in out[] holding seq line */
{
    char nline[P_LINE];
    register i, j;
    register char *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax + P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = *py;
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

```

nums

```

/*
 * put out a line (name, {num}, seq, {num}): dumpblock()
 */
static
putline(ix)
int ix;
{

```

putline

Tabela 1 (cont.)

...putline

```

int          i;
register char *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
    (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
    (void) putc(' ', fx);

/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
for (px = out[ix]; *px; px++)
    (void) putc(*px&0x7F, fx);
(void) putc('\n', fx);
}

```

```

/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()
{
    int          i;
    register char *p0, *p1, cx, *px;

    if (!*out[0] || (*out[0] == ' ' && *(p0[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(p0[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                cx = '.';
            else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';
}

```

stars

Tabela 1 (cont.)

```
/*
 * strip path or prefix from pn. return len: pr_align()
 */
static
stripname(pn)
    char    *pn;    /* file name (may be path) */
{
    register char    *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}
```

stripname

Tabela 1 (cont.)

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

char *jname = "/tmp/homgXXXXXX"; /* tmp file for jumps */
FILE *fj;

int cleanup(); /* cleanup tmp file */
long lseek();

/*
 * remove any tmp file if we blow
 */
cleanup(i)
int i;
{
    if (fj)
        (void) unlink(jname);
    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char *
getseq(file, len)
char *file; /* file name */
int *len; /* seq len */
{
    char line[1024], *pseq;
    register char *px, *py;
    int natgc, tlen;
    FILE *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';

```

cleanup

getseq

Tabela 1 (cont.)

```

py = pseq + 4;
*len = tlen;
rewind(fp);

while (fgets(line, 1024, fp)) {
    if (*line == ';' || *line == '<' || *line == '>')
        continue;
    for (px = line; *px != '\n'; px++) {
        if (isupper(*px))
            *py++ = *px;
        else if (islower(*px))
            *py++ = toupper(*px);
        if (index("ATGCU", *(py-1)))
            natgc++;
    }
    *py++ = '\0';
    *py = '\0';
    (void) fclose(fp);
    dna = natgc > (tlen/3);
    return(pseq+4);
}

char *
g_alloc(msg, nx, sz)
char *msg;          /* program, calling routine */
int nx, sz;          /* number and size of elements */
{
    char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
            fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
}

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()
{
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

    if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open() %s\n", prog, jname);
            cleanup(1);
        }
    }
    for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; i++) {
        while (1) {
            for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;

```

...getseq

g_alloc

readjmps

Tabela 1 (cont.)

...readjumps

```

        if (j < 0 && dx[dmax].offset && fj) {
            (void) lseek(fd, dx[dmax].offset, 0);
            (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
            (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
            dx[dmax].ijmp = MAXJMP-1;
        }
        else
            break;
    }
    if (i >= JMPS) {
        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
        cleanup(1);
    }
    if (j >= 0) {
        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) {
            /* gap in second seq */
            pp[1].n[i1] = -siz;
            xx += siz;
            /* id = xx - yy + len1 - 1
             */
            pp[1].x[i1] = xx - dmax + len1 - 1;
            gapy++;
            ngapy -= siz;
        }
        /* ignore MAXGAP when doing endgaps */
        siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
        i1++;
    }
    else if (siz > 0) { /* gap in first seq */
        pp[0].n[i0] = siz;
        pp[0].x[i0] = xx;
        gapx++;
        ngapx += siz;
        /* ignore MAXGAP when doing endgaps */
        siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
        i0++;
    }
    }
    else
        break;
}

/* reverse the order of jumps
 */
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
}

```

Tabela 1 (cont.)

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejumps(ix)                                writejumps
{
    int ix;

    char *mktemp();

    if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
            cleanup(1);
        }
        if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
    }
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

```

Tabela 2

PRO XXXXXXXXXXXXXXXX (Comprimento = 15 aminoácidos)

Proteína de Comparação XXXXXYYYYYYYY (Comprimento = 12 aminoácidos)

% de identidade de sequência de aminoácidos = (o número de resíduos de aminoácidos idênticos e na mesma posição entre as duas sequências de polipéptidos tal como determinado por ALIGN-2) a dividir por (o número total de resíduos de aminoácidos do polipéptido PRO) =

5 a dividir por 15 = 33,3%

Tabela 3

PRO XXXXXXXXXXXX (Comprimento = 10 aminoácidos)

Proteína de Comparação XXXXXYYYYYYZZYZ (Comprimento = 15 aminoácidos)

% de identidade de sequência de aminoácidos = (o número de resíduos de aminoácidos idênticos e na mesma posição entre as duas sequências de polipéptidos tal como determinado por ALIGN-2) a dividir por (o número total de resíduos de aminoácidos do polipéptido PRO)

= 5 a dividir por 10 = 50%

Tabela 4

PRO-ADN	NNNNNNNNNNNNNN	(Comprimento = 14 nucleótidos)
ADN de Comparação	NNNNNNLLLLLLLLLL -	(Comprimento = 16 nucleótidos)
% de identidade da sequência de ácido nucleico = (o número de nucleótidos idênticos e na mesma posição entre as duas sequências de ácido nucleico tal como determinado por ALIGN-2) a dividir por (o número total de nucleótidos da sequência de ácido nucleico PRO-ADN) = 6 a dividir por 14 = 42,9%		

Tabela 5

PRO-ADN	NNNNNNNNNNNN	(Comprimento = 12 nucleótidos)
ADN de Comparação	NNNNLLLV	(Comprimento = 9 nucleótidos)
% de identidade de sequência de ácido nucleico = (o número de nucleótidos idênticos e na mesma posição entre as duas sequências de ácido nucleico tal como determinado por ALIGN-2) a dividir por (o número total de nucleótidos da sequência de ácido nucleico PRO-ADN) = 4 a dividir por 12 = 33,3%		

II. Composições e Métodos da Invenção

A. Polipéptidos PRO de Comprimento Completo

A presente invenção proporciona recém-identificadas e isoladas sequências nucleotídica que codificam polipéptidos referidos no presente pedido como polipéptidos PRO. Em particular, identificaram-se e isolaram-se ADNc que codificam o polipéptido PRO, como descrito com mais detalhes nos Exemplos que se seguem.

Tal como revelado nos Exemplos adiante, clones de ADNc que codificam polipéptidos PRO foram depositados na ATCC. As sequências nucleotídica reais dos clones podem ser facilmente determinadas pelo perito na especialidade por sequenciação dos clones depositados utilizando métodos de rotina na especialidade. As sequências de aminoácidos previstas podem ser determinadas a partir das sequências nucleotídicas utilizando competências correntes. Para os polipéptidos PRO e ácidos nucleicos codificantes aqui descritos, a Requerente identificou o que se crê ser o enquadramento de leitura mais bem identificável a partir da informação sobre a sequência então disponível.

B. Variantes de PRO

Além dos polipéptidos PRO de sequência nativa de comprimento completo aqui descritos, está contemplado que se possam preparar variantes de PRO. As variantes de PRO podem ser preparadas por introdução das alterações apropriadas de nucleótidos no ADN de PRO e/ou por síntese do polipéptido PRO desejado. Os peritos na especialidade saberão que alterações de aminoácidos podem alterar processos pós-tradução do polipéptido PRO, tal como a alteração do número ou da posição de locais de glicosilação ou a alteração das características de ancoragem à membrana.

Podem efectuar-se variações no polipéptido PRO de sequência nativa de comprimento completo ou em vários domínios do polipéptido PRO aqui descritos, por exemplo, utilizando qualquer das técnicas e orientações para mutações conservativas e não conservativas apresentadas, por exemplo, na patente U.S. 5 364 934. As variações podem ser uma substituição, uma deleção ou uma inserção de um ou mais codões que codificam o polipéptido PRO e que resultam numa alteração da sequência de aminoácidos do polipéptido PRO comparativamente com o polipéptido PRO de sequência nativa. Opcionalmente a variação é por substituição de pelo menos um aminoácido por qualquer outro aminoácido num ou mais domínios do polipéptido PRO. Podem encontrar-se orientações relativas à determinação de qual o resíduo de aminoácido que pode ser inserido, substituído ou eliminado sem afectar adversamente a actividade desejada, comparando a sequência do polipéptido PRO com a de moléculas de proteínas homólogas conhecidas e minimizando o número de alterações na sequência de aminoácidos efectuadas em regiões de homologia elevada. As substituições de aminoácidos podem ser o resultado da substituição de um aminoácido por outro aminoácido apresentando propriedades estruturais e/ou químicas semelhantes, tal como a substituição de uma leucina por uma serina, *i.e.*, substituições de aminoácidos conservativas. As inserções ou as deleções podem opcionalmente ser na gama entre cerca de 1 a 5 aminoácidos. As variações permitidas podem ser determinadas fazendo sistematicamente inserções, deleções ou substituições de aminoácidos na sequência e testando as variantes resultantes

quanto à actividade apresentada pela sequência nativa madura ou de comprimento completo.

Proporcionam-se aqui fragmentos de polipéptidos PRO. Tais fragmentos podem ser truncados no terminal N ou no terminal C ou podem apresentar falta de resíduos internos, por exemplo, quando comparados com uma proteína nativa de comprimento completo. Certos fragmentos apresentam falta de resíduos de aminoácidos que não são essenciais para a actividade biológica desejada do polipéptido PRO.

Os fragmentos de PRO podem ser preparados por qualquer de várias técnicas convencionais. Os fragmentos peptídicos desejados podem ser sintetizados quimicamente. Uma abordagem alternativa envolve a geração de fragmentos de PRO por digestão enzimática, *e.g.*, por tratamento da proteína com uma enzima que se sabe que cliva proteínas em locais definidos por resíduos de aminoácidos particulares ou por digestão do ADN com enzimas de restrição adequadas e isolamento do fragmento desejado. Ainda outra técnica adequada envolve o isolamento e a amplificação de um fragmento de ADN que codifica um fragmento polipeptídico desejado através da reacção em cadeia com polimerase (PCR). Nos iniciadores 5' e 3' da PCR empregam-se oligonucleótidos que definem os terminais desejados do fragmento de ADN. Preferivelmente, os fragmentos de polipéptido PRO partilham pelo menos uma actividade biológica e/ou imunológica com o polipéptido PRO nativo apresentado nas figuras anexas.

Em concretizações particulares, apresentam-se substituições conservativas de interesse na Tabela 6 sob o título de substituições preferidas. Se tais substituições resultarem numa alteração da actividade biológica, então introduzem-se alterações mais substanciais, denominadas substituições exemplificativas na Tabela 6 ou tal como adicionalmente descrito adiante em referência a classes de aminoácidos, e rastreiam-se os produtos.

Tabela 6

Resíduo Original	Substituições Exemplificativas	Substituições preferidas
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucina	leu
Leu (L)	norleucina; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucina	leu

Obtêm-se modificações substanciais de função ou de identidade imunológica do polipéptido PRO por selecção de substituições que diferem significativamente no seu efeito sobre a manutenção (a) da estrutura do esqueleto do polipéptido na área da substituição, por exemplo, como uma conformação em folha ou em hélice, (b) da carga ou hidrofobicidade da molécula no local alvo, ou (c) do volume da cadeia lateral. Os resíduos de ocorrência natural dividem-se em grupos com base em propriedades comuns das cadeias laterais:

- (1) hidrófobos: norleucina, met, ala, val, leu, ile;
- (2) hidrófilos neutros: cys, ser, thr;

- (3) ácidos: asp, glu;
- (4) básicos: asn, gln, his, lys, arg;
- (5) resíduos que influenciam a orientação da cadeia: gly, pro;
e
- (6) aromáticos: trp, tyr, phe.

As substituições não conservativas implicarão trocar um membro de uma destas classes por outro de outra classe. Estes resíduos substituídos podem também ser introduzidos nos locais de substituição conservativa ou, mais preferivelmente, nos restantes locais (não conservados).

As variações podem ser efectuadas utilizando métodos conhecidos na especialidade tais como mutagénese mediada por oligonucleótidos (dirigida ao local), varrimento de alaninas e mutagénese por PCR. Pode efectuar-se mutagénese dirigida ao local [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], mutagénese com cassetes [Wells et al., Gene, 34:315 (1985)], mutagénese com selecção por restrição [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] ou outras técnicas conhecidas, sobre o ADN clonado, para produzir o ADN variante de PRO.

Pode também utilizar-se análise de varrimento de aminoácidos para identificar um ou mais aminoácidos ao longo de uma sequência contígua. Entre os aminoácidos preferidos para varrimento estão os aminoácidos neutros e relativamente pequenos. Tais aminoácidos incluem alanina, glicina, serina e cisteína. A alanina é tipicamente um aminoácido preferido para varrimento neste grupo porque elimina a cadeia lateral para além do carbono beta e é menos provável que altere a conformação da cadeia principal da variante [Cunningham e Wells, Science, 244: 1081-1085 (1989)]. A alanina é também tipicamente preferida porque é o aminoácido mais comum. Adicionalmente, encontra-se frequentemente tanto em posições enterradas como expostas [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. Se a substituição com alanina não proporcionar quantidades adequadas de variante, pode utilizar-se um aminoácido isotérico.

C. Modificações de Polipéptidos PRO

As modificações covalentes de polipéptidos PRO estão incluídas no âmbito da presente invenção. Um tipo de modificação covalente inclui a reacção de resíduos de aminoácidos alvo de um polipéptido PRO com um agente de derivatização orgânico que é capaz de reagir com cadeias laterais seleccionadas ou com os resíduos N- ou C-terminais do polipéptido PRO. A derivatização com agentes bifuncionais é útil, por exemplo, para reticulação de polipéptidos PRO com uma matriz ou uma superfície de suporte insolúveis em água para utilização no método de purificação de anticorpos anti-PRO e *vice-versa*. Os agentes de reticulação habitualmente utilizados incluem, por exemplo, 1,1-bis(diazoacetil)-2-feniletano, glutaraldeído, ésteres de N-hidroxissuccinimida, por exemplo, ésteres de ácido 4-azidossalicílico, imidoésteres homobifuncionais, incluindo ésteres de dissuccinimidilo tais como 3,3'-ditiobis(succinimidil-propionato), maleimidas bifuncionais tais como bis-N-maleimido-1,8-octano e agentes tais como metil-3-[(p-azidofenil)ditio]propioimidato.

Outras modificações incluem a desamidação de resíduos glutaminilo e asparaginilo nos resíduos glutamilo e aspartilo correspondentes, respectivamente, a hidroxilação de prolina e lisina, a fosforilação de grupos hidroxilo de resíduos serilo treonilo, a metilação de grupos α -amino de cadeias laterais de lisina, arginina e histidina [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, p. 79-86 (1983)], a acetilação da amina N-terminal e a amidação de qualquer grupo carboxilo C-terminal.

Outro tipo de modificação covalente do polipéptido PRO incluída no âmbito da presente invenção compreende alterar o padrão de glicosilação nativo do polipéptido. Para os presentes fins, entende-se que "alterar o padrão de glicosilação nativo", significa eliminar uma ou mais porções hidrato de carbono presentes em polipéptidos PRO de sequência nativa (quer por remoção do local de glicosilação subjacente quer por deleção da glicosilação por meios químicos e/ou enzimáticos) e/ou adicionar um ou mais locais de glicosilação que não se encontram presentes no polipéptido PRO de sequência nativa. Adicionalmente, a expressão inclui alterações

qualitativas na glicosilação das proteínas nativas, envolvendo uma alteração da natureza e das proporções das várias porções hidrato de carbono presentes.

A adição de locais de glicosilação ao polipéptido PRO pode obter-se por alteração da sequência de aminoácidos. A alteração pode ser efectuada, por exemplo, por adição de, ou substituição com, um ou mais resíduos de serina ou treonina no polipéptido PRO de sequência nativa (para locais de glicosilação ligada a O). A sequência de aminoácidos do polipéptido PRO pode ser opcionalmente alterada através de alterações ao nível do ADN, nomeadamente por mutação do ADN que codifica o polipéptido PRO em bases pré-seleccionadas de modo a que sejam gerados codões que se traduzirão nos aminoácidos desejados.

Outra maneira de aumentar o número de porções hidrato de carbono no polipéptido PRO é por acoplamento químico ou enzimático de glicósidos ao polipéptido. Tais métodos são descritos na especialidade, por exemplo, em WO 87/05330 publicado em 11 de Setembro de 1987 e em Aplin e Wriston, CRC Crit Rev. Biochem., p. 259-306 (1981).

A remoção de porções hidrato de carbono presentes no polipéptido PRO pode ser conseguida química ou enzimaticamente ou por substituição por mutação de codões que codificam resíduos de aminoácidos que servem como alvos para glicosilação. São conhecidas na especialidade técnicas de desglicosilação química e descrevem-se, por exemplo, em Hakimuddin, *et al.*, Arch. Biochem. Biophys., 259:52 (1987) e em Edge *et al.*, Anal. Biochem., 118:131 (1981). A clivagem enzimática de porções hidrato de carbono em polipéptidos pode ser conseguida utilizando várias endoglicosidases e exoglicosidases tal como descrito por Thotakura *et al.*, Meth. Enzymol., 138:350 (1987).

Outro tipo de modificação covalente de polipéptidos PRO inclui ligar o polipéptido PRO a um de entre vários polímeros não proteínáceos, *e.g.*, polietilenoglicol (PEG), polipropilenoglicol ou polioxialquilenos do modo estabelecido nas patentes U.S. 4 640 835; 4 496 689; 4 301 144; 4 670 417; 4 791 192 ou 4 179 337.

O polipéptido PRO da presente invenção pode também ser modificado de modo a formar uma molécula quimérica compreendendo um polipéptido PRO fundido a outro polipéptido ou a uma sequência de aminoácidos heterólogos.

Numa concretização, uma tal molécula quimérica compreende uma fusão do polipéptido PRO com um polipéptido marcador que proporciona um epítipo ao qual se pode ligar selectivamente um anticorpo anti-marcador. O marcador epitópico é geralmente colocado no terminal amino ou carboxilo do polipéptido PRO. A presença destas formas do polipéptido PRO marcadas com epítipo pode ser detectada utilizando um anticorpo contra o polipéptido marcador. Igualmente, a dotação com o marcador epitópico permite que o polipéptido PRO seja facilmente purificado por purificação de afinidade utilizando um anticorpo anti-marcador ou outro tipo de matriz de afinidade que se liga ao marcador epitópico. São bem conhecidos na especialidade vários polipéptidos marcadores e seus respectivos anticorpos. Os exemplos incluem marcadores poli-histidina (poli-His) ou poli-histidina-glicina (poli-His-gly); o polipéptido marcador HA flu e o seu anticorpo 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; o marcador c-myc e os seus anticorpos 8F9, 3C7, 6E10, G4, B7 e 9E10 [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; e o marcador glicoproteína D (gD) do vírus Herpes *Simplex* e o seu anticorpo [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Outros polipéptidos marcadores incluem o péptido Flag [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; o péptido epitópico KT3 [Martin et al., Science, 255:192-194 (1992)]; um péptido epitópico de α -tubulina [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; e o marcador peptídico de proteína do gene 10 de T7 [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

Numa concretização alternativa, a molécula quimérica pode compreender uma fusão do polipéptido PRO com uma imunoglobulina ou com uma região particular de uma imunoglobulina. Para uma forma bivalente da molécula quimérica (também referida como uma "imunoadesina"), tal fusão poderia ser com a região Fc de uma molécula de IgG. As fusões de Ig incluem preferivelmente a substituição de uma forma solúvel (domínio transmembranar eliminado ou inactivado) de um

polipéptido PRO no lugar de pelo menos uma região variável no interior de uma molécula de Ig. Numa concretização especialmente preferida, a fusão de imunoglobulina inclui as regiões de charneira, CH2 e CH3 ou as regiões de charneira, CH1, CH2 e CH3 de uma molécula de IgG1. Para a produção de fusões de imunoglobulina veja-se igualmente a patente US 5 428 130 concedida em 27 de Junho de 1995.

D. Preparação de Polipéptidos PRO

A descrição que se segue refere-se principalmente à produção de polipéptidos PRO por cultura de células transformadas ou transfectadas com um vector contendo ácido nucleico de polipéptido PRO. Contempla-se obviamente que se possam utilizar métodos alternativos, que são conhecidos na especialidade, para preparar polipéptidos PRO. Por exemplo, a sequência do polipéptido PRO, ou suas porções, podem ser produzidas por síntese peptídica directa utilizando técnicas em fase sólida [veja-se, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. Pode efectuar-se a síntese de proteínas *in vitro* utilizando técnicas manuais ou automáticas. A síntese automática pode ser conseguida, por exemplo, utilizando um sintetizador de péptidos da Applied Biosystems (Foster City, CA) utilizando as instruções do fabricante. Várias porções do polipéptido PRO podem ser sintetizadas quimicamente, separadamente e combinadas, utilizando métodos químicos ou enzimáticos para produzir o polipéptido PRO de comprimento completo.

1. Isolamento de ADN que Codifica Polipéptidos PRO

Pode obter-se ADN que codifica polipéptidos PRO a partir de uma biblioteca de ADNc preparada a partir de tecido que se crê possuir ARNm de PRO e expressá-lo a um nível detectável. Assim, pode obter-se convenientemente ADN de PRO humano a partir de uma biblioteca de ADNc preparada a partir de tecido humano, tal como descrito nos Exemplos. Pode também obter-se o gene que codifica PRO a partir de uma biblioteca genómica ou por procedimentos de síntese conhecidos (e.g. síntese automática de ácido nucleico).

As bibliotecas podem ser rastreadas com sondas (tais como anticorpos contra o polipéptido PRO ou oligonucleótidos com pelo menos cerca de 20-80 bases) concebidas para identificar o gene de interesse ou a proteína por ele codificada. O rastreio da biblioteca de ADNc ou genómica com a sonda seleccionada pode ser conduzido utilizando procedimentos *standard*, tais como os descritos em Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). A utilização da metodologia de PCR é um modo alternativo de isolar o gene que codifica o polipéptido PRO [Sambrook et al., *supra*; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

Os Exemplos que se seguem descrevem técnicas para rastreio de uma biblioteca de ADNc. As sequências oligonucleotídicas seleccionadas como sondas devem apresentar um comprimento suficiente e devem ser suficientemente não ambíguas de modo a minimizar falsos positivos. O oligonucleótido é preferivelmente marcado de modo a poder ser detectado por hibridação com ADN da biblioteca que está a ser rastreada. São bem conhecidos na especialidade métodos de marcação e incluem a utilização de marcadores radioactivos tais como ATP marcado com ^{32}P , biotinilação ou marcação enzimática. Proporcionam-se as condições de hibridação, incluindo rigor moderado e rigor elevado, em Sambrook et al., *supra*.

As sequências identificadas nestes métodos de rastreio de bibliotecas podem ser comparadas e alinhadas com outras sequências conhecidas depositadas e disponíveis em bases de dados públicas tais como GenBank ou outras bases de dados de sequências privadas. Pode determinar-se a identidade de sequência (ao nível dos aminoácidos ou dos nucleótidos) no interior de regiões definidas da molécula ou ao longo da sequência de comprimento completo, utilizando métodos conhecidos na especialidade e tal como aqui descrito.

Pode obter-se ácido nucleico possuindo sequências de codificação de proteínas por rastreio de bibliotecas de ADNc ou genómicas seleccionadas utilizando a sequência de aminoácidos deduzida aqui revelada pela primeira vez, e se

necessário, utilizando procedimentos convencionais de alongamento de iniciadores tal como descrito em Sambrook et al., *supra*, para detectar precursores e intermediários de processamento de ARNm que pode não ter sofrido transcrição inversa em ADNc.

2. Seleção e Transformação de Células Hospedeiras

Transfectam-se ou transformam-se células hospedeiras com vectores de expressão ou de clonagem aqui descritos para a produção de polipéptido PRO e cultivam-se em meios nutrientes convencionais modificados conforme apropriado para induzir promotores, seleccionar transformantes ou amplificar os genes que codificam as sequências desejadas. As condições de cultura tais como meio, temperatura, pH e semelhantes, podem ser seleccionadas pelo perito na especialidade sem experiências desnecessárias. Em geral, podem encontrar-se princípios, protocolos e técnicas práticas para maximizar a produtividade das culturas celulares em Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) e Sambrook et al., *supra*.

Os métodos de transfecção de células eucariotas e de transformação de células procariotas, são conhecidos dos peritos na especialidade, por exemplo, CaCl_2 , CaPO_4 , mediada por lipossomas e electroporação. Dependendo da célula hospedeira utilizada, efectua-se a transformação utilizando técnicas *standard* apropriadas para as células. Utiliza-se geralmente o tratamento com cálcio que emprega cloreto de cálcio, tal como descrito em Sambrook et al., *supra*, ou a electroporação, para procariotas. Utiliza-se infecção com *Agrobacterium tumefaciens* para a transformação de determinadas células vegetais, tal como descrito por Shaw et al., Gene, 23:315 (1983) e em WO 89/05859 publicado em 29 de Junho de 1989. Para células de mamífero sem as paredes celulares, pode utilizar-se o método de precipitação com fosfato de cálcio de Graham e van der Eb, Virology, 52:456-457 (1978). Descreveram-se aspectos gerais de transfecções de sistemas hospedeiros de células de mamífero na patente U.S. 4 399 216. As transformações em levedura são tipicamente efectuadas de acordo com o método de Van Solingen et al., J. Bact., 130:946 (1977) e Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829

(1979). Contudo, podem também utilizar-se outros métodos para introduzir ADN em células, tais como micro-injecção nuclear, electroporação, fusão de protoplastos bacterianos com células intactas, ou policações, e.g., polibreno, poliornitina. Para várias técnicas de transformação de células de mamífero, veja-se, Keown et al., Methods in Enzymology, 185:527-537 (1990) e Mansour et al., Nature, 336:348-352 (1988).

As células hospedeiras adequadas para clonagem ou expressão do ADN nos vectores aqui referidos incluem células procariotas, de levedura ou eucariotas superiores. Os procariotas adequados incluem, mas não se lhes limitam, as eubactérias, tais como organismos Gram-negativos ou Gram-positivos, por exemplo, *Enterobacteriaceae* tais como *E. coli*. Várias estirpes de *E. coli* encontram-se disponíveis ao público, tais como *E. coli* K12 estirpe MM294 (ATCC 31 446); *E. coli* X1776 (ATCC 31 537); *E. coli* estirpe W3110 (ATCC 27 325) e K5 772 (ATCC 53 635). Outras células hospedeiras procariotas adequadas incluem *Enterobacteriaceae* tais como *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescens*, e *Shigella*, assim como *Bacilli* tais como *B. subtilis* e *B. licheniformis* (e.g., *B. licheniformis* 41P, revelado em DD 266 710 publicado em 12 de Abril de 1989), *Pseudomonas* tais como *P. aeruginosa* e *Streptomyces*. Estes exemplos são ilustrativos e não limitativos. A estirpe W3110 constitui um hospedeiro ou hospedeiro progenitor particularmente preferido porque é uma estirpe hospedeira comum para fermentações de produtos de ADN recombinante. Preferivelmente, a célula hospedeira excreta quantidades mínimas de enzimas proteolíticas. Por exemplo, a estirpe W3110 pode ser modificada para efectuar uma mutação genética nos genes que codificam proteínas endógenas para o hospedeiro, incluindo os exemplos de tais hospedeiros, *E. coli* W3110 estirpe 1A2, que possui o genótipo completo *tonA*; *E. coli* W3110 estirpe 9E4, que apresenta o genótipo completo *tonA ptr3*; *E. coli* W3110 estirpe 27C7 (ATCC 55 244), que possui o genótipo completo *tonA ptr3 phoA E15 (argF-lac)Ib9 degP empT kan^r*; *E. coli* W3110 estirpe 37D6, que possui o genótipo completo *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan^r*; *E. coli* W3110 estirpe 40B-1, que é uma estirpe 37D6 com uma mutação de deleção de *degP* sem resistência à canamicina; e

uma estirpe de *E. coli* possuindo protease periplasmática mutante revelada na patente U.S. 4 946 783 concedida em 7 de Agosto de 1990. Alternativamente, são adequados métodos de clonagem *in vitro*, e.g., PCR ou outras reacções com ácido nucleico-polimerases.

Além dos procariotas, os micróbios eucariotas, tais como fungos filamentosos ou leveduras, constituem hospedeiros de clonagem ou de expressão adequados para vectores que codificam PRO. A *Saccharomyces cerevisiae* é um microrganismo hospedeiro eucariota inferior habitualmente utilizado. Outros incluem *Schizosaccharomyces pombe* (Beach e Nurse, Nature, 290: 140 [1981]; EP 139 383 publicado em 2 de Maio de 1985); hospedeiros de *Kluyveromyces* (patente U.S. 4 943 529; Fleer et al., Bio/Technology, 9:968-975 (1991)) tal como, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 737 [1983]). *K. fragilis* (ATCC 12 424), *K. bulgaricus* (ATCC 16 045), *K. wickerhamii* (ATCC 24 178), *K. waltii* (ATCC 56 500), *K. drosophilum* (ATCC 36 906; Vanden Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans* e *K. marxianus*; *Yarrowia* (EP 402 226); *Pichia pastoris* (EP 183 070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244 234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* tal como *Schwanniomyces occidentalis* (EP 394 538 publicado em 31 de Outubro de 1990); e fungos filamentosos tais como, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 publicado em 10 de Janeiro de 1991), e hospedeiros *Aspergillus* tais como *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81:1470-1474 (1984)) e *A. niger* (Kelly e Hynes, EMBO J., 4:475-479 [1985]). São aqui adequadas leveduras metilotróficas e incluem, não se lhes limitando, leveduras capazes de crescer em metanol seleccionadas dos géneros que consistem em *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis* e *Rhodotorula*. Pode encontrar-se uma listagem de espécies específicas que constituem exemplos desta classe de leveduras em C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Células hospedeiras adequadas para a expressão de polipéptidos PRO glicosilados são derivadas a partir de organismos pluricelulares. Os exemplos de células de invertebrados incluem células de insectos tais como *Drosophila* S2 e *Spodoptera* Sf9, assim como células vegetais. Os exemplos de linhas celulares hospedeiras de mamífero úteis incluem células de ovário de *hamster* chinês (CHO) e células COS. Exemplos mais específicos incluem a linha CV1 de rim de macaco transformada com SV40 (COS-7, ATCC CRL 1651); a linha de rim embrionário humano (células 293 ou 293 subclonadas para crescimento em cultura em suspensão, Graham et al., J. Gen. Virol., 36:59 (1977)); as células de ovário de *hamster* chinês/-DHFR (CHO), Urlaub e Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); as células de Sertoli de ratinho (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); as células de pulmão humano (W138, ATCC CCL 75); as células de fígado humano (Hep G2, HB 8065); e tumor da mama de ratinho (MMT 060562, ATCC CCL51). Considera-se que a selecção da célula hospedeira apropriada se encontra abrangida pela perícia na especialidade.

3. Seleccção e Utilização de um Vector Replicável

O ácido nucleico (e.g., ADNc ou ADN genómico) que codifica polipéptidos PRO pode ser inserido num vector replicável para clonagem (amplificação do ADN) ou para expressão. Encontram-se disponíveis ao público vários vectores. O vector pode, por exemplo, estar sob a forma de um plasmídeo, um cosmídeo, uma partícula viral ou um fago. A sequência de ácido nucleico apropriada pode ser inserida no vector através de vários procedimentos. Em geral, insere-se o ADN em local(ais) de endonucleases de restrição apropriado(s) utilizando técnicas conhecidas na especialidade. Os componentes do vector incluem geralmente, mas não se lhes limitam, uma ou mais sequências de sinal, uma origem de replicação, um ou mais genes marcadores, um elemento potenciador, um promotor e uma sequência de terminação da transcrição. A construção de vectores adequados contendo um ou mais destes componentes emprega técnicas de ligação *standard* que são conhecidas do perito na especialidade.

O polipéptido PRO pode ser produzido de modo recombinante não apenas directamente mas também na forma de um polipéptido de fusão com um polipéptido heterólogo, que pode ser uma sequência de sinal ou outro polipéptido possuindo um local de clivagem específico no terminal N da proteína madura ou do polipéptido. Em geral, a sequência de sinal pode ser um componente do vector ou pode constituir uma parte do ADN que codifica PRO que é inserido no vector. A sequência de sinal pode ser uma sequência de sinal procariota seleccionada, por exemplo, do grupo de comandos da fosfatase alcalina, penicilinase, Ipp ou da enterotoxina II estável ao calor. Para secreção em levedura, a sequência de sinal pode ser, e.g., o comando da invertase de levedura, o comando de factor alfa (incluindo os comandos de factor α de *Saccharomyces* e *Kluyveromyces*, o último descrito na patente U.S. 5 010 182) ou o comando da fosfatase ácida, o comando da glucoamilase de *C. albicans* (EP 362 179 publicado em 4 de Abril de 1990) ou a sequência de sinal descrita em WO 90/13646 publicado em 15 de Novembro de 1990. Para a expressão em células de mamífero, podem utilizar-se sequências de sinal de mamífero para dirigir a secreção da proteína, tais como sequências de sinal de polipéptidos excretados da mesma espécie ou de espécies relacionadas, assim como comandos de secreção virais.

Ambos os vectores, de expressão e de clonagem, contêm uma sequência de ácido nucleico que permite que o vector replique numa ou mais células hospedeiras seleccionadas. Tais sequências são bem conhecidas para várias bactérias, leveduras e vírus. A origem de replicação do plasmídeo pBR322 é adequada para a maioria das bactérias Gram-negativas, a origem do plasmídeo 2 μ é adequada para leveduras e várias origens de virais (SV40, polioma, adenovírus, VSV ou BPV) são úteis para vectores de clonagem de células de mamífero.

Os vectores de expressão e de clonagem conterão tipicamente um gene de selecção, também denominado um marcador seleccionável. Os genes de selecção típicos codificam proteínas que (a) conferem resistência a antibióticos ou outras toxinas, e.g., ampicilina, neomicina, metotrexato ou tetraciclina, (b) complementam deficiências auxotróficas, ou (c) fornecem nutrientes críticos não disponíveis a partir de

meios complexos, e.g., o gene que codifica D-alanina-racemase para bacilos.

Os exemplos de marcadores de selecção adequados para células de mamífero, são os que permitem a identificação de células competentes para assimilação do ácido nucleico que codifica PRO, tais como DHFR ou timidina-quinase. Constitui uma célula hospedeira apropriada, quando se utiliza DHFR de tipo selvagem, a linha celular CHO deficiente em actividade de DHFR, preparada e propagada como descrito por Urlaub *et al.*, Proc. Natl. Acad. Sci. USA, 77:4216 (1980). Constitui um gene de selecção adequado para utilização em leveduras, o gene *trp1* presente no plasmídeo de levedura YRp7 [Stinchcomb *et al.*, Nature, 282:39 (1979); Kingsman *et al.*, Gene, 7:141 (1979); Tschemper *et al.*, Gene, 10:157 (1980)]. O gene *trp* proporciona um marcador de selecção para uma estirpe mutante de levedura que não apresenta a capacidade de crescer em triptofano, por exemplo, ATCC n.º 44076 ou PEP4-1 [Jones, Genetics, 85:12 (1977)].

Os vectores de expressão e de clonagem contêm habitualmente um promotor operativamente ligado à sequência de ácido nucleico que codifica PRO para dirigir a síntese do ARNm. São bem conhecidos promotores reconhecidos por várias células hospedeiras potenciais. Os promotores adequados para utilização com hospedeiros procariotas incluem os sistemas promotores da β -lactamase e da lactose [Chang *et al.*, Nature, 275:615(1978); Goeddel *et al.*, Nature, 281:544 (1979)], da fosfatase alcalina, um sistema promotor do triptofano (*trp*) [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36 776] e promotores híbridos tais como o promotor tac [deBoer *et al.*, Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Os promotores para utilização em sistemas bacterianos conterão também uma sequência de Shine-Dalgarno (S.D.) ligada operativamente ao ADN que codifica o polipéptido PRO.

Os exemplos de sequências promotoras adequadas para utilização com hospedeiros de levedura incluem os promotores para 3-fosfoglicerato-quinase [Hitzeman *et al.*, J. Biol. Chem., 255:2073 (1980)] ou outras enzimas glicolíticas [Hess *et al.*, J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], tais como enolase,

gliceraldeído-3-fosfato-desidrogenase, hexoquinase, piruvato-descarboxilase, fosfofrutoquinase, glucose-6-fosfato-isomerase, 3-fosfoglicerato-mutase, piruvato-quinase, triosefosfato-isomerase, fosfoglucoase-isomerase e glucoquinase.

Outros promotores de levedura, que são promotores indutíveis apresentando a vantagem adicional da transcrição ser controlada pelas condições de crescimento, são as regiões promotoras para a álcool-desidrogenase 2, o isocitocromo C, a fosfatase ácida, enzimas de degradação associadas com o metabolismo do azoto, a metalotioneína, a gliceraldeído-3-fosfato-desidrogenase e enzimas responsáveis pela utilização de maltose e galactose. Descrevem-se adicionalmente vectores e promotores adequados para utilização em expressão em levedura em EP 73 657.

A transcrição de PRO a partir de vectores em células hospedeiras de mamífero é controlada, por exemplo, por promotores obtidos a partir do genoma de vírus tais como do poliomavírus, poxvírus aviário (UK 2 211 504 publicado em 5 de Julho de 1989), adenovírus (tal como Adenovírus 2), papilomavírus bovino, vírus do sarcoma aviário, citomegalovírus, um retrovírus, vírus da hepatite B e vírus símio 40 (SV40), de promotores de mamífero heterólogos, e.g., o promotor da actina ou um promotor de imunoglobulina e de promotores de choque térmico, desde que esses promotores sejam compatíveis com os sistemas da célula hospedeira.

A transcrição de um ADN que codifica o polipéptido PRO por eucariotas superiores pode ser aumentada por inserção no vector de uma sequência potenciadora. Os potenciadores são elementos de ADN actuantes em *cis* habitualmente com cerca de 10 a 300 pb, que actuam sobre um promotor para aumentar a sua transcrição. São actualmente conhecidas muitas sequências potenciadoras de genes de mamífero (globina, elastase, albumina, α -fetoproteína e insulina). Tipicamente, contudo, utilizar-se-á um potenciador de um vírus de célula eucariota. Os exemplos incluem o potenciador de SV40 do lado tardio da origem de replicação (pb 100-270), o potenciador do promotor precoce de citomegalovírus, o potenciador de polioma do lado tardio da origem de replicação e potenciadores de adenovírus.

O potenciador pode ser submetido a *splicing* no vector numa posição 5' ou 3' relativamente à sequência de codificação do polipéptido PRO, mas localiza-se preferivelmente num local a 5' do promotor.

Os vectores de expressão utilizados em células hospedeiras eucariotas (células de leveduras, fungos, insectos, plantas, animais, seres humanos ou células nucleadas de outros organismos pluricelulares) conterão igualmente sequências necessárias para a terminação da transcrição e para a estabilização do ARNm. Estas sequências estão habitualmente disponíveis a partir das regiões não traduzidas a 5' e ocasionalmente a 3' de ADN ou ADNc eucariotas ou virais. Estas regiões contêm segmentos nucleotídicos transcritos como fragmentos poliadenilados na porção não traduzida do ARNm que codifica o polipéptido PRO.

Descrevem-se ainda outros métodos, vectores e células hospedeiras adequados para adaptação à síntese de polipéptidos PRO em cultura de células de vertebrado recombinantes em Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117 060; e EP 117 058.

4. Detecção de Amplificação/Expressão Génicas

A amplificação e/ou expressão génicas podem ser medidas numa amostra directamente, por exemplo, por *Southern blot* convencional, *Northern blot* para quantificar a transcrição do ARNm [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], *dot blot* (análise de ADN), ou hibridação *in situ* utilizando uma sonda apropriadamente marcada, baseada nas sequências aqui proporcionadas. Alternativamente, podem utilizar-se anticorpos que podem reconhecer cópias específicos, incluindo cópias de ADN, cópias de ARN e cópias híbridos ADN-ARN ou cópias ADN-proteína. Os anticorpos podem por sua vez ser marcados e o ensaio pode ser realizado no local onde o dúplex se encontra ligado a uma superfície, de modo a que após formação de um dúplex na superfície se possa detectar a presença de anticorpo ligado ao dúplex.

Alternativamente, a expressão génica pode ser medida por métodos imunológicos, tal como coloração imuno-histoquímica de células ou secções de tecidos e ensaio de cultura celular ou fluidos corporais, para quantificar directamente a expressão do produto génico. Os anticorpos úteis para coloração imuno-histoquímica e/ou para ensaio de fluidos de amostra podem ser monoclonais ou policlonais e podem ser preparados em qualquer mamífero. Convenientemente, os anticorpos podem ser preparados contra um polipéptido PRO de sequência nativa ou contra um péptido sintético baseado nas sequências de ADN aqui proporcionadas ou contra uma sequência exógena fundida com o ADN de PRO e codificando para um epítipo de anticorpo específico.

5. Purificação de Polipéptidos PRO

Podem recuperar-se formas de polipéptidos PRO do meio de cultura ou de lisados da célula hospedeira. Se estiverem ligadas à membrana podem ser libertadas da membrana utilizando uma solução de detergente adequada (e.g., Triton-X 100) ou por clivagem enzimática. As células empregues na expressão de polipéptidos PRO podem ser rompidas por vários meios, físicos ou químicos, tais como ciclos de congelação-descongelação, tratamento com ultra-sons, ruptura mecânica ou agentes de lise celular.

Pode ser desejável purificar polipéptidos PRO de proteínas ou polipéptidos celulares recombinantes. Os seguintes procedimentos constituem exemplos de procedimentos de purificação adequados: fraccionamento numa coluna de permuta iónica; precipitação com etanol; HPLC de fase inversa; cromatografia em sílica ou numa resina de permuta catiónica tal como DEAE; cromatofocagem; SDS-PAGE; precipitação com sulfato de amónio; filtração em gel utilizando, por exemplo, Sephadex G-75; colunas de proteína A-Sepharose para remover contaminantes tais como IgG; e colunas quelantes de metais para ligar formas marcadas com epítipo do polipéptido PRO. Podem utilizar-se vários métodos de purificação de proteínas e esses métodos são conhecidos na especialidade e estão descritos por exemplo em Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). O passo ou passos de

purificação seleccionados dependerão, por exemplo, da natureza do processo de produção utilizado e do polipéptido PRO particular produzido.

E. Anticorpos

Alguns candidatos a fármacos para utilização nas composições e métodos da presente invenção são anticorpos e fragmentos de anticorpos que imitam a actividade biológica de um polipéptido PRO.

1. Anticorpos Policlonais

São conhecidos dos peritos na especialidade métodos de preparação de anticorpos policlonais. Os anticorpos policlonais podem ser criados num mamífero, por exemplo, através de uma ou mais injeções de um agente imunizante e, se pretendido, um adjuvante. Tipicamente, o agente imunizante e/ou o adjuvante serão injectados no mamífero por múltiplas injeções subcutâneas ou intraperitoneais. O agente imunizante pode incluir o polipéptido PRO ou uma sua proteína de fusão. Pode ser útil conjugar o agente imunizante a uma proteína que se sabe ser imunogénica no mamífero a imunizar. Os exemplos de tais proteínas imunogénicas incluem, mas não se lhes limitam, hemocianina de lapa *Fissurella*, albumina sérica, tiroglobulina bovina e inibidor de tripsina de soja. Os exemplos de adjuvantes que podem ser empregues incluem adjuvante completo de Freund e adjuvante MPL-TDM (monofosforil-Lípido A, dicorinomicolato de trealose sintético). O protocolo de imunização pode ser seleccionado por um perito na especialidade sem experiências desnecessárias.

2. Anticorpos Monoclonais

Os anticorpos podem, alternativamente, ser anticorpos monoclonais. Os anticorpos monoclonais podem ser preparados utilizando métodos de hibridomas, tais como os descritos por Kohler e Milstein, Nature, 256:495 (1975). Num método de hibridoma imuniza-se tipicamente um ratinho, um *hamster* ou outro animal hospedeiro apropriado, com um agente imunizante, para eliciar linfócitos que produzem ou são capazes de produzir, anticorpos que se ligarão especificamente ao agente

imunizante. Alternativamente, os linfócitos podem ser imunizados *in vitro*.

O agente imunizante incluirá tipicamente o polipéptido PRO, ou uma sua proteína de fusão. Em geral, utilizam-se linfócitos de sangue periférico ("PBL") caso se pretendam células de origem humana, ou células de baço ou células de nódulos linfáticos caso se pretendam fontes de mamífero não humano. Os linfócitos são então fundidos com uma linha celular imortalizada utilizando um agente de fusão adequado, tal como polietilenoglicol, para formar uma célula de hibridoma [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) p. 59-103]. As linhas celulares imortalizadas são usualmente células de mamífero transformadas, nomeadamente células de mieloma de origem murídea, bovina e humana. Usualmente, empregam-se linhas celulares de mieloma de rato ou ratinho. As células de hibridoma podem ser cultivadas num meio de cultura adequado que preferivelmente contém uma ou mais substâncias que inibem o crescimento ou a sobrevivência das células imortalizadas não fundidas. Por exemplo, se as células progenitoras não contiverem a enzima hipoxantina-guanina-fosforribosil-transferase (HGPRT ou HPRT), o meio de cultura para os hibridomas incluirá tipicamente hipoxantina, aminopterina e timidina ("meio HAT"), substâncias que impedem o crescimento de células deficientes em HGPRT.

As linhas celulares imortalizadas preferidas são as que se fundem eficientemente, suportam expressão estável em níveis elevados de anticorpo pelas células produtoras de anticorpo seleccionadas e são sensíveis a um meio, tal como meio HAT. Constituem linhas celulares imortalizadas mais preferidas as linhas de mieloma murino, que podem ser obtidas, por exemplo, de Salk Institute Cell Distribution Center, San Diego, California e de American Type Culture Collection, Manassas, Virginia. Foram também descritas linhas celulares de mieloma humano e de heteromieloma de ratinho-humano para a produção de anticorpos monoclonais humanos [Kozbor, J. Immunol., 133:3001 (1984); Brodeur *et al.*, Monoclonal Antibodies Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) p. 51-63].

O meio de cultura no qual se cultivam as células de hibridoma pode então ser ensaiado quanto à presença de anticorpos monoclonais dirigidos contra o polipéptido PRO. Preferivelmente, determina-se a especificidade de ligação de anticorpos monoclonais produzidos pelas células de hibridoma por imunoprecipitação ou por um ensaio de ligação *in vitro*, tal como um radioimunoensaio (RIA) ou um ensaio de imunossorção com enzimas ligadas (ELISA). Estas técnicas e ensaios são conhecidos na especialidade. A afinidade de ligação do anticorpo monoclonal pode, por exemplo, ser determinada pela análise Scatchard de Munson e Pollard, Anal. Biochem., 107:220 (1980).

Após terem sido identificadas as células de hibridoma desejadas, os clones podem ser subclonados por procedimentos de diluição limitante e crescidas por métodos *standard* [Goding, *supra*]. Os meios de cultura adequados para este fim incluem, por exemplo, meio de Eagle modificado por Dulbecco e meio RPMI-1640. Alternativamente, as células de hibridoma podem ser crescidas *in vivo* na forma de ascites num mamífero.

Os anticorpos monoclonais excretados pelos subclones podem ser isolados ou purificados do meio de cultura ou do fluido ascítico por procedimentos de purificação de imunoglobulinas convencionais tais como, por exemplo, cromatografia em proteína A-Sepharose, em hidroxilapatite, electroforese em gel, diálise ou cromatografia de afinidade.

Os anticorpos monoclonais podem também ser preparados por métodos de ADN recombinante, tais como os descritos na Patente U.S. 4 816 567. O ADN que codifica os anticorpos monoclonais da invenção pode ser facilmente isolado e sequenciado utilizando procedimentos convencionais (e.g., utilizando sondas oligonucleotídicas que são capazes de se ligar especificamente a genes que codificam as cadeias pesada e leve de anticorpos murinos). As células de hibridoma da invenção servem como fonte preferida de um tal ADN. Uma vez isolado, o ADN pode ser colocado em vectores de expressão, que são então transfectados para células hospedeiras tais como células COS de símio, células de ovário de *hamster* chinês (CHO) ou células de mieloma que de outra forma não produzem proteína imunoglobulina, para obter a síntese de anticorpos monoclonais

nas células hospedeiras recombinantes. Os ADN podem também ser modificados, por exemplo, substituindo pela sequência de codificação para domínios constantes de cadeia pesada e leve humanos as sequências homólogas murinas [Patente U.S. 4 816 567; Morrison et al., *supra*] ou por ligação covalente à sequência de codificação da imunoglobulina de toda ou parte da sequência de codificação para um péptido que não imunoglobulina. Tal péptido que não imunoglobulina pode substituir os domínios constantes de um anticorpo da invenção ou pode substituir os domínios variáveis de um local de combinação com o antigénio de um anticorpo da invenção para criar um anticorpo quimérico bivalente.

Os anticorpos podem ser anticorpos monovalentes. São bem conhecidos na especialidade métodos para preparar anticorpos monovalentes. Por exemplo, um método envolve a expressão recombinante da cadeia leve e da cadeia pesada modificadas de imunoglobulina. A cadeia pesada é, em geral, truncada em qualquer ponto na região Fc de modo a evitar a ligação cruzada da cadeia pesada. Alternativamente, resíduos de cisteína relevantes são substituídos por outro resíduo de aminoácido ou são eliminados para evitar a ligação cruzada.

São também adequados métodos *in vitro* para preparar anticorpos monovalentes. Pode efectuar-se a digestão de anticorpos para produzir os seus fragmentos, nomeadamente fragmentos Fab, utilizando técnicas de rotina conhecidas na especialidade.

3. Anticorpos Humanos e Humanizados

Os anticorpos da invenção podem adicionalmente compreender anticorpos humanizados ou anticorpos humanos. As formas humanizadas de anticorpos não humanos (*e.g.*, murinos) são imunoglobulinas quiméricas, cadeias de imunoglobulina ou seus fragmentos (tais como Fv, Fab, Fab', F(ab')₂ ou outras subsequências de ligação ao antigénio de anticorpos) que contêm uma sequência mínima derivada de imunoglobulina não humana. Os anticorpos humanizados incluem imunoglobulinas humanas (anticorpo receptor) em que resíduos de uma região determinante de complementaridade (CDR) do receptor são substituídos pelos resíduos de uma CDR de uma espécie não

humana (anticorpo dador) tal como de ratinho, rato ou coelho possuindo a especificidade, afinidade e capacidade desejadas. Em alguns casos, substituem-se resíduos do esqueleto de Fv da imunoglobulina humana pelos resíduos não humanos correspondentes. Os anticorpos humanizados podem também compreender resíduos que não se encontram nem no anticorpo receptor, nem nas sequências CDR ou de esqueleto importadas. Em geral, o anticorpo humanizado compreenderá substancialmente todos de pelo menos um, e tipicamente dois, domínios variáveis, em que todas ou substancialmente todas as regiões CDR correspondem às da imunoglobulina não humana e todas ou substancialmente todas as regiões FR são as de uma sequência de consenso de uma imunoglobulina humana. O anticorpo humanizado compreenderá também optimamente pelo menos uma porção de uma região constante de imunoglobulina (Fc), tipicamente a de uma imunoglobulina humana [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); e Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

São bem conhecidos na especialidade métodos para humanizar anticorpos não humanos. Em geral, um anticorpo humanizado tem introduzidos um ou mais resíduos de aminoácidos de uma fonte que não é humana. Estes resíduos de aminoácidos não humanos são frequentemente referidos como resíduos de "importação", que são tipicamente obtidos de um domínio variável de "importação". A humanização pode ser essencialmente efectuada seguindo o método de Winter e colaboradores [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)], substituindo por CDR ou sequências de CDR de roedor as sequências correspondentes de um anticorpo humano. Assim, estes anticorpos "humanizados" são anticorpos quiméricos (Patente U.S. 4 816 567), em que substancialmente menos de um domínio variável intacto foi substituído pela sequência correspondente de uma espécie não humana. Na prática, os anticorpos humanizados são tipicamente anticorpos humanos em que alguns resíduos de CDR e possivelmente alguns resíduos de FR são substituídos por resíduos de locais análogos em anticorpos de roedor.

Os anticorpos humanos podem também ser produzidos utilizando várias técnicas conhecidas na especialidade, incluindo bibliotecas de exibição sobre fagos [Hoogenboom e Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. As técnicas de Cole et al. e Boerner et al. estão também disponíveis para a preparação de anticorpos monoclonais humanos (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) e Boerner et al., J. Immunol., 147(1):86-95 (1991)]. De igual modo, podem ser preparados anticorpos humanos por introdução de *loci* de imunoglobulina humana em animais transgênicos, e.g., ratinhos nos quais os gene de imunoglobulina endógenos foram parcial ou completamente inativados. Por provocação, observa-se produção do anticorpo humano, que é muito semelhante ao observado em seres humanos em todos os sentidos, incluindo no rearranjo de genes, montagem e repertório do anticorpo. Esta abordagem é descrita, por exemplo, nas Patentes U.S. 5 545 807; 5 545 806; 5 569 825; 5 625 126; 5 633 425; 5 661 016 e nas seguintes publicações científicas: Marks et al., Bio/Technology, 10:779-783 (1992); Lonberg et al., Nature, 368:856-859 (1994); Morrison, Nature, 368:812-13 (1994); Fishwild et al., Nature Biotechnology, 14:845-51 (1996); Neuberger, Nature Biotechnology, 14:826 (1996); Lonberg e Huszar, Intern. Rev. Immunol., 13:65-93 (1995).

4. Anticorpos Biespecíficos

Os anticorpos biespecíficos são anticorpos monoclonais, preferivelmente humanos ou humanizados, que possuem especificidades de ligação para pelo menos dois antigénios diferentes. No presente caso, uma das especificidades de ligação é para com o polipéptido PRO e a outra é para com qualquer outro antigénio e preferivelmente para com uma proteína ou um receptor ou subunidade de receptor da superfície celular.

São conhecidos na especialidade métodos para preparar anticorpos biespecíficos. Tradicionalmente, a produção recombinante de anticorpos biespecíficos baseia-se na co-expressão de dois pares cadeia pesada/cadeia leve de imunoglobulina, em que as duas cadeias pesadas têm especificidades diferentes [Milstein e Cuello, Nature,

305:537-539 (1983)]. Dada a distribuição aleatória das cadeias pesada e leve de imunoglobulina, estes hibridomas (quadromas) produzem uma mistura potencial de dez moléculas de anticorpo diferentes, das quais só uma tem a estrutura biespecífica correcta. A purificação da molécula correcta é usualmente conseguida por passos de cromatografia de afinidade. Revelam-se procedimentos semelhantes em WO 93/08829, publicado em 13 de Maio de 1993 e em Traunecker et al., EMBO J., 10:3655-3659 (1991).

Os domínios variáveis de anticorpos com as especificidades de ligação desejadas (locais de combinação anticorpo-antigénio) podem ser fundidos com sequências de domínios constantes de imunoglobulina. A fusão é, preferivelmente, com um domínio constante de cadeia pesada de imunoglobulina, compreendendo pelo menos parte das regiões de charneira, CH2 e CH3. Prefere-se ter a primeira região constante de cadeia pesada (CH1) contendo o local necessário para a ligação à cadeia leve presente em pelo menos uma das fusões. Os ADN que codificam as fusões de cadeia pesada de imunoglobulina e, caso se pretenda, a cadeia leve de imunoglobulina, são inseridos em vectores de expressão separados que são co-transfectados num organismo hospedeiro adequado. Para mais detalhes sobre a geração de anticorpos biespecíficos veja-se, por exemplo, Suresh et al., Methods in Enzymology, 121:210 (1986).

De acordo com outra abordagem descrita em WO 96/27011, a interface entre um par de moléculas de anticorpo pode ser manipulada para maximizar a percentagem de heterodímeros que se recuperam da cultura de células recombinantes. A interface preferida compreende pelo menos uma parte da região CH3 de um domínio constante de anticorpo. Neste método, substituem-se uma ou mais cadeias laterais pequenas de aminoácidos da interface da primeira molécula de anticorpo por cadeias laterais maiores (e.g., tirosina ou triptofano). Criam-se "cavidades" de compensação de tamanho idêntico ou semelhante à(s) cadeia(s) lateral(is) grande(s) na interface da segunda molécula de anticorpo substituindo cadeias laterais grandes de aminoácidos por mais pequenas (e.g., alanina ou treonina). Isto proporciona um mecanismo para aumentar o rendimento do

heterodímero relativamente a outros produtos finais indesejáveis, tais como homodímeros.

Podem preparar-se anticorpos biespecíficos na forma de anticorpos de comprimento completo ou fragmentos de anticorpo (e.g., anticorpos biespecíficos $F(ab')_2$). Foram descritas na literatura técnicas para gerar anticorpos biespecíficos a partir de fragmentos de anticorpo. Por exemplo, podem preparar-se anticorpos biespecíficos utilizando ligação química. Brennan et al., Science, 229:81 (1985) descrevem um procedimento em que se clivam proteoliticamente anticorpos intactos para gerar fragmentos $F(ab')_2$. Estes fragmentos são reduzidos na presença do agente de complexação de ditiol, arsenieto de sódio, para estabilizar ditióis vicinais e impedir a formação de dissulfureto intermolecular. Os fragmentos Fab' gerados são então convertidos nos derivados tionitrobenzoato (TNB). Um dos derivados Fab' -TNB é então reconvertido no Fab' -tiol por redução com mercaptoetilamina e é misturado com uma quantidade equimolar do outro derivado Fab' -TNB para formar o anticorpo biespecífico. Os anticorpos biespecíficos produzidos podem ser utilizados como agentes para a imobilização selectiva de enzimas.

Os fragmentos Fab' podem ser recuperados directamente de *E. coli* e acoplados quimicamente para formar anticorpos biespecíficos. Shalaby et al., J. Exp. Med., 175:217-225 (1992) descrevem a produção de uma molécula de anticorpo biespecífico $F(ab')_2$ completamente humanizado. Cada fragmento Fab' foi excretado separadamente de *E. coli* e sujeito a acoplamento químico directo *in vitro* para formar o anticorpo biespecífico. O anticorpo biespecífico assim formado era capaz de se ligar a células que super-expressam o receptor ErbB2 e a células T humanas normais, bem como desencadear a actividade lítica de linfócitos citotóxicos humanos contra alvos de tumor da mama humanos.

Foram também descritas várias técnicas para preparar e isolar fragmentos de anticorpos biespecíficos directamente a partir da cultura de células recombinantes. Por exemplo, produziram-se anticorpos biespecíficos utilizando "fechos-de-correr" de leucinas. Kostelny et al., J. Immunol., 148(5):1547-1553 (1992). Ligaram-se os péptidos de "fecho-de-

correr" de leucinas das proteínas Fos e Jun a porções Fab' de dois anticorpos diferentes através de fusão génica. Reduziram-se os homodímeros de anticorpo na região de charneira para formar monómeros e seguidamente re-oxidaram-se para formar os heterodímeros de anticorpo. Este método pode também ser utilizado para a produção de homodímeros de anticorpo. A tecnologia dos "diacorpos" descrita por Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993) proporcionou um mecanismo alternativo para preparar fragmentos de anticorpo biespecíficos. Os fragmentos compreendem um domínio variável de cadeia pesada (V_H) ligado a um domínio variável de cadeia leve (V_L) através de um ligante que é demasiado curto para permitir emparelhamento entre os dois domínios na mesma cadeia. Assim, os domínios V_H e V_L de um fragmento são forçados a emparelhar com os domínios V_L e V_H complementares de outro fragmento, formando assim dois locais de ligação ao antigénio. Foi também relatada outra estratégia para preparar fragmentos de anticorpo biespecíficos por utilização de dímeros Fv de cadeia única (sFv). Veja-se, Gruber et al., J. Immunol., 152:5368 (1994).

Estão contemplados anticorpos com mais de duas valências. Por exemplo, podem preparar-se anticorpos triespecíficos. Tutt et al., J. Immunol., 147:60 (1991).

Podem ligar-se anticorpos biespecíficos exemplificativos a dois epítomos diferentes num determinado polipéptido PRO. Alternativamente, pode combinar-se um braço anti-polipéptido PRO com um braço que se liga a uma molécula desencadeadora num leucócito, tal como uma molécula receptora de células T (e.g., CD2, CD3, CD28 ou B7) ou receptores Fc para IgG (FcγR), tais como FcγRI (CD64), FcγRII (CD32) e FcγRIII (CD16) de modo a focar os mecanismos de defesa celular na célula que expressa o polipéptido PRO particular. Os anticorpos biespecíficos podem ser também utilizados para localizar agentes citotóxicos em células que expressam um determinado polipéptido PRO. Estes anticorpos possuem um braço de ligação ao polipéptido PRO e um braço que se liga a um agente citotóxico ou um quelante de radionuclidos, tal como EOTUBE, DPTA, DOTA ou TETA. Outro anticorpo biespecífico de interesse liga o polipéptido PRO e liga adicionalmente o factor tissular (TF).

5. Anticorpos Heteroconjugados

Os anticorpos heteroconjugados estão também no âmbito da presente invenção. Os anticorpos heteroconjugados são constituídos por dois anticorpos ligados covalentemente. Estes anticorpos foram, por exemplo, propostos para direccionar células do sistema imunitário para células indesejadas [Patente U.S. 4 676 980] e para tratamento de infecção por VIH [WO 91/00360; WO 92/200373; EP 03089]. Está contemplado que os anticorpos possam ser preparados *in vitro* utilizando métodos conhecidos em química de síntese de proteínas, incluindo os que envolvem agentes de reticulação. Por exemplo, podem construir-se imunotoxinas utilizando uma reacção de permuta de dissulfureto ou por formação de uma ligação tioéter. Os exemplos de reagentes adequados para este fim incluem iminotiolato e metil-4-mercaptobutirimidato e os revelados, por exemplo, na Patente U.S. 4 676 980.

6. Manipulação da Função Efectora

Pode ser desejável modificar o anticorpo da invenção relativamente à função efectora, de modo a aumentar, e.g., a eficácia do anticorpo no tratamento do cancro. Por exemplo, pode(m) introduzir-se resíduo(s) de cisteína na região Fc, permitindo assim a formação de ligações dissulfureto intercadeia nesta região. O anticorpo homodimérico assim gerado pode ter capacidade de internalização melhorada e/ou morte celular mediada por complemento e citotoxicidade celular dependente de anticorpo (ADCC) melhoradas. Veja-se, Caron et al., J. Exp. Med., 176:1191-1195 (1992) e Shopes, J. Immunol., 148:2918-2922 (1992). Podem também preparar-se anticorpos homodiméricos com actividade antitumoral aumentada utilizando agentes de reticulação heterobifuncionais, tal como descrito em Wolff et al., Cancer Research 53:2560-2565 (1993). Alternativamente, pode manipular-se um anticorpo que tem regiões Fc duplas e pode portanto apresentar capacidades de lise de complemento e ADCC melhoradas. Veja-se, Stevenson et al., Anti-Cancer Drug Design, 3:219-230 (1989).

7. Imunoconjugados

A invenção refere-se também a imunoconjugados incluindo um anticorpo conjugado com um agente citotóxico, tal como um agente quimioterapêutico, uma toxina (e.g., uma toxina enzimaticamente activa de origem bacteriana, fúngica, vegetal ou animal ou seus fragmentos) ou um isótopo radioactivo (i.e., um radioconjugado).

Foram descritos acima agentes quimioterapêuticos úteis na geração destes imunoconjugados. As toxinas enzimaticamente activas e seus fragmentos que podem ser utilizados incluem cadeia A da difteria, fragmentos activos não ligantes da toxina da difteria, cadeia A da exotoxina (de *Pseudomonas aeruginosa*), cadeia A da ricina, cadeia A da abrina, cadeia A da modecina, alfa-sarcina, proteínas de *Aleurites fordii*, proteínas de diantina, proteínas de *Phytolaca americana* (PAPI, PAPII e PAP-S), inibidor de *Momordica charantia*, curcina, crotina, inibidor de *Sapaonaria officinalis*, gelonina, mitogelina, restrictocina, fenomicina, enomicina e os tricotecenos. Encontram-se disponíveis vários radionuclídeos para a produção de anticorpos radioconjugados. Os exemplos incluem ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y e ^{186}Re .

Os conjugados de anticorpo e agente citotóxico são preparados utilizando vários agentes de acoplamento de proteínas bifuncionais, tais como N-succinimidil-3-(2-piridilditiol)propionato (SPDP), iminotiolano (IT), derivados bifuncionais de imidoésteres (tal como adipimidato de dimetilo HCL), ésteres activos (tal como suberato de dissuccinimidilo), aldeídos (tal como glutaraldeído), compostos bis-azido (tal como bis(p-azidobenzoíl)hexanodiamina), derivados bis-diazónio (tal como bis-(p-diazónio-benzoíl)etilenodiantina), diisocianatos (tal como 2,6-diisocianato de tolueno) compostos de flúor bis-activos (tal como 1,5-difluoro-2,4-dinitrobenzeno). Por exemplo, pode ser preparada uma imunotoxina de ricina tal como descrito em Vitetta et al., Science, 238:1098 (1987). Constitui um exemplo de agente quelante para a conjugação de radionucleótidos ao anticorpo, o ácido 1-isotiocianatobenzil-3-metildietilenotriaminopenta-acético (MX-DTPA) marcado com carbono-14. Veja-se WO94/11026.

Noutra concretização, o anticorpo pode ser conjugado com um "receptor" (tal como estreptavidina) para utilização no pré-direccionamento a tumores em que o conjugado anticorpo-receptor é administrado ao paciente, seguido de remoção do conjugado não ligado da circulação utilizando um agente de depuração e seguidamente administra-se um "ligando" (e.g., avidina) que está conjugada com um agente citotóxico (por exemplo, um radionucleótido).

8. Imunolipossomas

Os anticorpos aqui revelados podem também ser formulados na forma de imunolipossomas. Os lipossomas contendo o anticorpo são preparados por métodos conhecidos na especialidade, tal como descrito em Epstein et al., Proc. Natl. Acad. Sci. USA, 82:3688 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA, 77:4030 (1980); e Patentes U.S. 4 485 045 e 4 544 545. Revelam-se lipossomas com tempo de circulação aumentado na Patente U.S. 5 013 556.

Podem gerar-se lipossomas particularmente úteis pelo método da evaporação em fase inversa com uma composição lipídica compreendendo fosfatidilcolina, colesterol e fosfatidiletanolamina derivatizada com PEG (PEG-PE). Extrudem-se os lipossomas através de filtros de tamanho de poros definido para obter lipossomas com o diâmetro pretendido. Podem conjugar-se com os lipossomas fragmentos Fab' do anticorpo da presente invenção tal como descrito em Martin et al., J. Biol. Chem., 257:286-288 (1982) através de uma reacção de permuta de dissulfureto. O lipossoma contém no seu interior opcionalmente um agente quimioterapêutico (tal como doxorubicina). Veja-se, Gabizon et al., J. National Cancer Inst., 81(19):1484 (1989).

F. Identificação de Proteínas Capazes de Inibir o Crescimento ou a Proliferação de Células Neoplásicas

As proteínas reveladas no presente pedido de patente foram ensaiadas num painel de 60 linhas de células tumorais correntemente utilizadas no rastreio investigacional, orientado para doenças, *in vitro*, de descoberta de fármacos, do National Cancer Institute (NCI). A finalidade deste rastreio é identificar moléculas que possuem actividade

citotóxica e/ou citostática contra diferentes tipos de tumores. O NCI rastreia mais de 10 000 novas moléculas por ano (Monks et al., J. Natl. Cancer Inst., 83:757-766 (1991); Boyd, Cancer: Princ. Pract. Oncol. Update, 3(10):1-12 ([1989])). As linhas de células tumorais empregues neste estudo foram descritas em Monks et al., *supra*. As linhas celulares cujo crescimento foi significativamente inibido pelas proteínas do presente pedido de patente estão especificadas nos Exemplos.

Os resultados mostraram que as proteínas testadas apresentam actividades citostáticas e, em alguns casos e em algumas concentrações, citotóxicas, numa variedade de linhas celulares de cancro, e portanto são candidatos úteis para terapia tumoral.

Outros ensaios à base de células e modelos animais para tumores (e.g., cancros) podem também ser utilizados para confirmar as identificações do rastreio do cancro do NCI, e para adicionalmente entender o relacionamento entre a proteína aqui identificada e o desenvolvimento e a patogénese do crescimento de células neoplásicas. Por exemplo, culturas primárias derivada de tumores em animais transgénicos (como descrito adiante) podem ser utilizadas nos ensaios à base de células, embora se prefiram linhas celulares estáveis. As técnicas para derivar linhas celulares contínuas a partir de animais transgénicos são bem conhecidas na especialidade (veja-se, e.g., Small et al., Mol. Cell. Biol., 5:642-648 [1985])).

G. Modelos Animais

Podem utilizar-se vários modelos animais bem conhecidos para adicionalmente entender o papel das moléculas aqui identificadas no desenvolvimento e na patogénese de tumores e para testar a eficácia de agentes terapêuticos candidatos, incluindo anticorpos e outros agonistas dos polipéptidos nativos, incluindo agonistas de molécula pequena. A natureza *in vivo* de tais modelos torna-os particularmente adequados para a previsão de respostas em pacientes humanos. Os modelos animais de tumores e cancros (e.g., cancro da mama, cancro do cólon, cancro da próstata, cancro do pulmão, etc.) incluem animais (transgénicos) tanto não recombinantes como

recombinantes. Os modelos animais não recombinantes incluem, por exemplo, roedores, e.g., modelos murinos. Tais modelos podem ser gerados introduzindo células tumorais em ratinhos singénicos utilizando técnicas *standard*, e.g., injeção subcutânea, injeção na veia da cauda, implante no baço, implante intraperitoneal, implante sob a cápsula renal ou implante "Orthopin", e.g., células de cancro do cólon implantadas em tecido de cólon. (Veja-se, e.g., a publicação PCT WO 97/33551, publicado em 18 de Setembro de 1997).

Provavelmente, a espécie animal mais frequentemente utilizada em estudos oncológicos é a de ratinhos imunodeficientes, em particular, ratinhos nus. A observação de que ratinhos nus com hipoplasia/aplasia podem funcionar com sucesso como hospedeiros de xenoenxertos de tumores humanos, levou à vulgarização da sua utilização para este fim. O gene *nu* autossómico recessivo foi introduzido num grande número de estirpes congénicas distintas de ratinhos nus, incluindo, por exemplo, ASW, A/He, AKR, BALB/c, B10.LP, C17, C3H, C57BL, C57, CBA, DBA, DDD, I/st, NC, NFR, NFS, NFS/N, NZB, NZC, NZW, P, RIII e SJL. Adicionalmente, têm-se criado e utilizado uma grande variedade de outros animais com deficiências imunológicas herdadas, diferentes dos ratinhos nus, como recebedores de xenoenxertos de tumores. Para detalhes adicionais veja-se, e.g., The Nude Mouse in Oncology Research, E. Boven e B. Winograd, ed., CRC Press, Inc., 1991.

As células introduzidas nestes animais podem ser derivadas de linhas celulares conhecidas de tumor/cancro, tais como qualquer das linhas celulares tumorais descritas acima e por exemplo, a linha celular B104-1-1 (linha celular NIH-3T3 estável transfectada com o proto-oncogene *neu*); células NIH-3T3 transfectadas com *ras*; Caco-2 (ATCC HTB-37); uma linha celular de adenocarcinoma de cólon humano moderadamente bem diferenciada de grau II, HT-29 (ATCC HTB-38) ou de tumores e cancros. Podem obter-se amostras de tumores ou células de cancro a partir de pacientes submetidos a cirurgia, utilizando condições *standard* envolvendo congelação e armazenagem em azoto líquido (Karmali et al., Br. J. Cancer, 48: 689-696 [1983]).

Podem introduzir-se células tumorais em animais, tais com ratinhos nus, através de vários procedimentos. O espaço subcutâneo (s.c.) do ratinho é muito adequado para implante de tumores. Os tumores podem ser transplantados s.c. na forma de blocos sólidos, de biópsias de agulha utilizando um trocarte, ou de suspensões celulares. Para o implante de blocos sólidos ou de trocarte, introduzem-se os fragmentos de tecido tumoral de tamanho adequado no espaço s.c.. As suspensões celulares são preparadas de fresco a partir de tumores primários ou linhas celulares tumorais estáveis e injectadas por via subcutânea. As células tumorais podem também ser injectadas na forma de implantes subdérmicos. Nesta localização, deposita-se o inóculo entre a parte inferior do tecido conjuntivo dérmico e o tecido s.c.. Boven e Winograd (1991), *supra*. Podem gerar-se modelos animais de cancro da mama, por exemplo, por implante de células de neuroblastoma de rato (das quais se isolou inicialmente o oncogene *neu*) ou de células NIH-3T3 transformadas com *neu* em ratinhos nus, essencialmente como descrito por Drebin et al., Proc. Natl. Acad. Sci. USA, 83:9129-9133 (1986).

De igual modo, podem gerar-se modelos animais de cancro do cólon por passagem de células de cancro do cólon em animais, e.g., ratinhos nus, levando ao aparecimento de tumores nesses animais. Descreveu-se um modelo de transplante ortotópico de cancro do cólon humano em ratinhos nus, por exemplo, em Wang et al., Cancer Research, 54:4726-4728 (1994) e Too et al., Cancer Research, 55:681-684 (1995). Este modelo baseia-se no denominado "METAMOUSE" vendido por AntiCancer, Inc., (San Diego, California).

Os tumores que surgem em animais podem ser removidos e cultivados *in vitro*. As células das culturas *in vitro* podem então ser passadas para animais. Estes tumores podem servir como alvos para outros testes ou rastreio de fármacos. Alternativamente, os tumores que resultam da passagem podem ser isolados e o ARN das células antes das passagens e de células isoladas após um ou mais ciclos de passagem pode ser analisado relativamente à expressão diferencial de genes de interesse. Tais técnicas de passagem podem ser realizadas com quaisquer linhas celulares conhecidas de tumor ou cancro.

Por exemplo, Meth A, CMS4, CMS5, CMS21 e WEHI-164 são fibrossarcomas induzidos quimicamente em ratinhos BALB/c fêmea (DeLeo et al., J. Exp. Med., 146:720 [1977]), que proporcionam um sistema modelo altamente controlável para o estudo de actividades antitumorais de vários agentes (Palladino et al., J. Immunol., 138:4023-4032 [1987]). Sucintamente, propagam-se células tumorais *in vitro* em cultura celular. Antes da injeção nos animais, lavam-se as linhas celulares e suspendem-se em tampão, a um densidade celular de cerca de 10×10^6 a 10×10^7 células/ml. Os animais são então infectados por via subcutânea com 10 a 100 μ l da suspensão celular, decorrendo uma a três semanas até ao aparecimento de um tumor.

Adicionalmente, o carcinoma de Lewis do pulmão (3LL) de ratinho, que é um dos tumores experimentais mais amplamente estudados, pode ser utilizado como modelo de tumor para a investigação. A eficácia neste modelo de tumor foi correlacionada com efeitos benéficos no tratamento de pacientes humanos diagnosticados com carcinoma de células pequenas do pulmão (SCCL). Este tumor pode ser introduzido em ratinhos normais por injeção de fragmentos de tumor de um ratinho afectado ou de células mantidas em cultura (Zupi et al., Br. J. Cancer, 41; supl. 4:309 [1980]) e as evidências indicam que podem ser iniciados tumores a partir da injeção de apenas uma célula e que uma grande proporção de células tumorais infectadas sobrevive. Para informação adicional sobre este modelo de tumor veja-se Zacharski, Haemostasis, 16:300-320 [1986]).

Um modo de avaliar a eficácia de um composto de teste num modelo animal ou num tumor implantado consiste em medir a dimensão do tumor antes e após o tratamento. Tradicionalmente, a dimensão dos tumores implantados tem sido medida com uma craveira em duas ou três dimensões. A medição limitada a duas dimensões não reflecte precisamente a dimensão do tumor pelo que é habitualmente convertida no volume correspondente utilizando uma fórmula matemática. Contudo, a medição das dimensões do tumor é muito pouco precisa. Os efeitos terapêuticos de um fármaco candidato podem ser mais bem descritos como atraso de crescimento e atraso de crescimento específico induzidos pelo tratamento. Outra variável importante na descrição do crescimento do tumor é o tempo de

duplicação do volume do tumor. Encontram-se igualmente disponíveis programas de computador para o cálculo e descrição do crescimento de tumores, tais como o programa relatado por Rygaard e Spang-Thomsen, Proc. 6th Int Workshop on Immune-Deficient Animals. Wu e Sheng ed., Basel, 1989, 301. Note-se contudo que necrose e respostas inflamatórias a seguir ao tratamento podem resultar de facto num aumento das dimensões do tumor, pelo menos inicialmente. Consequentemente, estas alterações têm que ser cuidadosamente monitoradas através de uma combinação de um método morfométrico e de análise de citometria de fluxo.

Podem manipular-se modelos animais recombinantes (transgénicos) por introdução da porção de codificação dos genes aqui identificados no genoma de animais de interesse utilizando técnicas *standard* para a produção de animais transgénicos. Os animais que podem servir como alvo para manipulação transgénica incluem, sem limitação, ratinhos, ratos, coelhos, porquinhos-da-índia, ovelhas, cabras, porcos e primatas não humanos, e.g., babuínos, chimpanzés e macacos. As técnicas conhecidas na especialidade para introduzir um transgene nestes animais incluem micro-injecção pró-nucleica (Hoppe e Wanger, patente U.S. 4 873 191); transferência génica para linhas germinais mediada por retrovírus (e.g., Van der Putten et al., Proc. Natl. Acad. Sci. USA, 82:6148-615 [1985]); direccionamento de genes em células estaminais embrionárias (Thompson et al., Cell, 56:313-321 [1989]); electroporação de embriões (Lo, Mol. Cell Biol., 3:1803-1814 [1983]); transferência génica mediada por esperma (Lavitrano et al., Cell, 57:717-73 [1989]). Para uma revisão, veja-se, por exemplo, a patente U.S. 4 736 866.

Para os fins da presente invenção, os animais transgénicos incluem os que são portadores do transgene apenas em parte das suas células ("animais mosaico"). O transgene pode ser integrado quer como um transgene único quer na forma de concatâmeros, e.g., estruturas em tandem cabeça-com-cabeça ou cabeça-com-cauda. A introdução selectiva de um transgene num tipo de célula particular é também possível seguindo, por exemplo, a técnica de Lasko et al., Proc. Natl. Acad. Sci. USA, 89:6232-636 (1992).

A expressão do transgene em animais transgênicos pode ser monitorada por técnicas *standard*. Por exemplo, pode utilizar-se análise *Southern blot* ou amplificação por PCR para verificar a integração do transgene. O nível de expressão de ARNm pode então ser analisado utilizando técnicas tais como hibridação *in situ*, análise *Northern blot*, PCR ou imunocitoquímica. Os animais são adicionalmente examinados quanto a sinais de desenvolvimento de tumores ou cancro.

A eficácia de anticorpos que se ligam especificamente aos polipéptidos aqui identificados e a outros fármacos candidatos pode ser também testada no tratamento de tumores animais espontâneos. Constitui um alvo adequado para estes estudos o carcinoma de células escamosas (SCC) oral felino. O SCC oral felino é um tumor maligno altamente invasivo que é a malignidade oral mais comum em gatos, sendo responsável por mais de 60% dos tumores orais relatados nesta espécie. Raramente forma metástases em locais distantes, apesar desta reduzida incidência de metástases poder ser meramente um reflexo dos curtos tempos de sobrevivência dos gatos com este tumor. Estes tumores não são habitualmente passíveis de cirurgia, principalmente devido à anatomia da cavidade oral felina. Presentemente não existe tratamento eficaz para este tumor. Antes da entrada no estudo, cada gato é submetido a um exame clínico completo, biópsia, e é submetido a tomografia computadorizada (TC). Os gatos diagnosticados com tumores de células escamosas orais sublinguais são excluídos do estudo. A língua pode paralisar em resultado deste tumor e mesmo que o tratamento mate o tumor os animais podem não ser capazes de se alimentar sozinhos. Cada gato é tratado repetidamente por um período de tempo mais longo. Obter-se-ão diariamente fotografias dos tumores durante o período de tratamento e a cada reverificação subsequente. Após o tratamento submete-se cada gato a uma nova TC. Cada TC e radiografia torácica são avaliadas de 8 em 8 semanas daí em diante. Os dados são avaliados relativamente a diferenças em sobrevivência, resposta e toxicidade comparativamente com grupos de controlo. Uma resposta positiva pode requerer evidência de regressão do tumor, preferivelmente com melhoria da qualidade de vida e/ou aumento da esperança de vida.

Adicionalmente, podem também testar-se outros tumores animais espontâneos tais como fibrossarcoma, adenocarcinoma, linfoma, condroma, leiomiossarcoma de cães, gatos e babuínos. De entre estes, o adenocarcinoma da mama de cães e gatos é um modelo preferido dado que o seu aparecimento e comportamento são muito semelhantes aos dos seres humanos. Contudo, a utilização deste modelo está limitada à rara ocorrência deste tipo de tumor em animais.

H. Ensaaios de Rastreio para Candidatos a Fármacos

Os ensaios de rastreio para candidatos a fármacos são concebidos para identificar compostos que se ligam competitivamente, ou complexam, com o(s) receptor(es) dos polipéptidos aqui identificados ou que de outro modo assinalam através deste(s) receptor(es). Estes ensaios de rastreio incluirão ensaios passíveis de rastreio de alto rendimento de bibliotecas de compostos químicos, tornando-os particularmente adequados para identificar candidatos a fármacos de molécula pequena. As moléculas pequenas contempladas incluem compostos orgânicos ou inorgânicos sintéticos, incluindo péptidos, preferivelmente péptidos solúveis, fusões (poli)péptido-imunoglobulina e em particular anticorpos incluindo, sem limitação, anticorpos policlonais e monoclonais e fragmentos de anticorpos, anticorpos de cadeia simples, anticorpos anti-idiotípicos e versões quiméricas ou humanizadas desses anticorpos ou fragmentos, assim como anticorpos e fragmentos de anticorpos humanos. Os ensaios podem ser realizados em vários formatos incluindo ensaios de ligação proteína-proteína, ensaios de rastreio bioquímico, imunoensaaios e ensaios baseados em células, que se encontram bem caracterizados na especialidade.

Em ensaios de ligação, a interacção é a ligação e o complexo formado pode ser isolado ou detectado na mistura reaccional. Numa concretização particular, um receptor de um polipéptido codificado pelo gene aqui identificado ou o fármaco candidato é imobilizado numa fase sólida, e.g., numa placa de microtítulo, através de ligações covalentes ou não covalentes. A ligação não covalente é geralmente conseguida por revestimento da superfície sólida com uma solução do polipéptido e secagem. Alternativamente, pode utilizar-se um

anticorpo imobilizado, e.g., um anticorpo monoclonal, específico para o polipéptido a imobilizar para o ancorar a uma superfície sólida. O ensaio é realizado por adição do componente não imobilizado, que pode ser marcado através de um marcador detectável, ao componente imobilizado, e.g., a superfície revestida contendo o componente ancorado. Quando a reacção se completa, removem-se os componentes que não reagiram, e.g., por lavagem, e detectam-se os complexos ancorados à superfície sólida. Quando o componente originalmente não imobilizado possui um marcador detectável a detecção do marcador imobilizado na superfície indica que ocorreu complexação. Quando o componente originalmente não imobilizado não possui um marcador pode detectar-se a complexação, por exemplo, utilizando um anticorpo marcado que se liga especificamente ao complexo imobilizado.

Se o composto candidato interage, sem se ligar, com um receptor particular, a sua interacção com esse polipéptido pode ser ensaiada por métodos bem conhecidos de detecção de interacções proteína-proteína. Esses ensaios incluem abordagens tradicionais tais como reticulação, co-imunoprecipitação e co-purificação através de gradientes ou colunas cromatográficas. Adicionalmente, podem monitorar-se interacções proteína-proteína utilizando um sistema genético baseado em leveduras descrito por Fields e colaboradores [Fields e Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88: 9578-9582 (1991)] tal como revelado por Chevray e Nathans, Proc. Natl. Acad. Sci. USA, 89:5789-5793 (1991)]. Muitos activadores da transcrição, tais como GAL4 de levedura, consistem em dois domínios modulares fisicamente discretos actuando um deles como domínio de ligação ao ADN enquanto que o outro funciona como domínio de activação da transcrição. O sistema de expressão de levedura descrito nas publicações anteriores (referido genericamente como "sistema de dois híbridos") aproveita esta propriedade e utiliza duas proteínas híbridas, uma em que a proteína alvo se encontra fundida com o domínio de ligação ao ADN de GAL4 e outra na qual se fundem proteínas de activação candidatas com o domínio de activação. A expressão de um gene repórter GAL1-lacZ sob o controlo de um promotor activado por GAL4 depende da reconstituição da actividade de GAL4 através de interacção proteína-proteína. Detectam-se colónias contendo

polipéptidos em interacção com um substrato cromogénico para β -galactosidase. Encontra-se comercialmente disponível na Clontech um *kit* completo (MATCHMAKER™) para identificar interacções proteína-proteína entre duas proteínas específicas utilizando a técnica dos dois híbridos. Este sistema pode também ser estendido ao mapeamento de domínios da proteína envolvidos em interacções específicas da proteína assim como para identificar resíduos de aminoácido que são essenciais para essas interacções.

I. Composições Farmacêuticas

Os polipéptidos da presente invenção, anticorpos agonistas que se ligam especificamente a proteínas aqui identificadas, bem como outras moléculas identificadas pelos ensaios de rastreio aqui revelados, podem ser administrados para o tratamento de tumores, incluindo cancros, sob a forma de composições farmacêuticas.

Quando se utilizam fragmentos de anticorpo, é preferido o menor fragmento inibidor que se liga especificamente ao domínio de ligação da proteína alvo. Por exemplo, com base nas sequências da região variável de um anticorpo, podem conceber-se moléculas peptídicas que retêm a capacidade de ligar a sequência da proteína alvo. Tais péptidos podem sintetizar-se quimicamente e/ou produzir-se por tecnologia de ADN recombinante (veja-se, e.g., Marasco et al., Proc. Natl. Acad. Sci. USA, 90:7889-7893 [1993]).

A presente formulação pode também conter mais do que um composto activo, conforme necessário para a indicação particular a tratar, preferivelmente aqueles com actividades complementares que não se afectam adversamente uns aos outros. Alternativamente, ou em adição, a composição pode compreender um agente que melhore a sua função, tal como, por exemplo, um agente citotóxico, uma citóquina, um agente quimioterapêutico ou um agente inibidor do crescimento. Estas moléculas estão adequadamente presentes em combinação em quantidades que são eficazes para o fim pretendido.

Preparam-se formulações terapêuticas dos polipéptidos aqui identificados, ou de seus agonistas, para armazenagem,

por mistura do ingrediente activo com o grau de pureza desejado, com transportadores, excipientes ou estabilizantes farmacologicamente aceitáveis opcionais (Remington's Pharmaceutical Sciences, 16ª edição, Osol, A. ed. [1980]), sob a forma de formulações liofilizadas ou soluções aquosas. Os transportadores, excipientes ou estabilizantes aceitáveis não são tóxicos para os recebedores nas dosagens e concentrações empregues e incluem tampões tais como fosfato, citrato e outros ácidos orgânicos; antioxidantes incluindo ácido ascórbico e metionina; conservantes (tais como cloreto de octadecildimetilbenzilamónio; cloreto de hexametónio; cloreto de benzalcónio, cloreto de benzetónio; fenol, álcool butílico ou benzílico; alquilparabenos tais como metilparabeno ou propilparabeno; catecol; resorcinol; ciclo-hexanol; 3-pentanol; e *m*-cresol); polipéptidos de baixo peso molecular (menos de cerca de 10 resíduos); proteínas, tais como albumina sérica, gelatina ou imunoglobulinas; polímeros hidrófilos tais como polivinilpirrolidona; aminoácidos tais como glicina, glutamina, asparagina, histidina, arginina ou lisina; monossacáridos, dissacáridos e outros hidratos de carbono incluindo glucose, manose ou dextrinas; agentes quelantes tais como EDTA; açúcares tais como sacarose, manitol, trealose ou sorbitol; contra-íões formadores de sais tais como sódio; complexos metálicos (por exemplo, complexos Zn-proteína); e/ou tensioactivos não iónicos tais como TWEEN™, PLURONICS™ ou polietilenoglicol (PEG).

A presente formulação pode também conter mais do que um composto activo, conforme necessário para a indicação particular a tratar, preferivelmente os que apresentam actividades complementares que não se afectam adversamente uns aos outros. Alternativamente, ou em adição, a composição pode compreender um agente citotóxico, uma citoquina ou um agente inibidor do crescimento. Estas moléculas estão adequadamente presentes em combinação em quantidades que são eficazes para o fim pretendido.

Os ingredientes activos podem também ser encapsulados em microcápsulas preparadas, por exemplo, por técnicas de coacervação ou por polimerização interfacial, por exemplo, microcápsulas de hidroximetilcelulose ou gelatina e microcápsulas de poli(metilmetacrilato), respectivamente, em

sistemas de entrega de fármacos coloidais (por exemplo, lipossomas, microesferas de albumina, microemulsões, nanopartículas e nanocápsulas) ou em macroemulsões. Tais técnicas estão descritas em Remington's Pharmaceutical Sciences, 16ª edição, Osol, A. ed. (1980).

As formulações destinadas a administração *in vivo* têm que ser estéreis. Isto pode ser facilmente conseguido por filtração através de membranas de esterilização por filtração antes ou após a liofilização e a reconstituição.

As presentes composições terapêuticas geralmente são colocadas num recipiente possuindo uma porta de acesso estéril, por exemplo, um saco ou um frasco de solução intravenosa possuindo uma rolha perfurável com uma agulha de injeção hipodérmica.

Podem preparar-se preparações de libertação sustentada. Os exemplos adequados de preparações de libertação sustentada incluem matrizes semi-permeáveis de polímeros hidrófobos sólidos contendo o anticorpo, matrizes estas que se encontram sob a forma de artigos conformados, e.g., filmes ou microcápsulas. Os exemplos de matrizes de libertação sustentada incluem poliésteres, hidrogéis (por exemplo, poli(2-hidroxietilmetacrilato) ou poli(álcool vinílico)), polilactidos (Patente U.S. 3 773 919), copolímeros de ácido L-glutâmico e L-glutamato de γ -etilo, etileno-acetato de vinilo não degradável, copolímeros de ácido láctico-ácido glicólico degradáveis, tais como o LUPRON DEPOT™ (microesferas injectáveis compostas de copolímero de ácido láctico-ácido glicólico e acetato de leuprolida) e ácido poli-D-(-)-3-hidroxibutírico. Enquanto que os polímeros tais como etileno-acetato de vinilo e ácido láctico-ácido glicólico permitem a libertação de moléculas durante mais de 100 dias, determinados hidrogéis libertam proteínas durante períodos de tempo mais curtos. Quando os anticorpos encapsulados permanecem no corpo durante um longo período de tempo podem desnaturar ou agregar-se como resultado de exposição a humidade a 37°C, resultando perda de actividade biológica e possíveis alterações na imunogenicidade. Podem ser concebidas estratégias racionais para estabilização dependendo do mecanismo envolvido. Por exemplo, quando se verifica que o

mecanismo de agregação é a formação de ligações S-S intermoleculares através de permuta tio-dissulfureto, pode conseguir-se a estabilização por modificação de resíduos sulfidrido, liofilização a partir de soluções ácidas, controlo do teor de humidade, utilização de aditivos apropriados e desenvolvimento de composições de matrizes poliméricas específicas.

J. Métodos de Tratamento

Está contemplado que os polipéptidos da presente invenção e seus agonistas, incluindo anticorpos, péptidos e agonistas de molécula pequena, possam ser utilizados para tratar vários tumores, e.g., cancros. Os exemplos de condições e desordens a tratar incluem tumores benignos ou malignos (e.g., tumores renais, do fígado, do rim, da bexiga, da mama, gástricos, dos ovários, colorrectais, da próstata, pancreáticos, do pulmão, da vulva, da tiróide, carcinomas hepáticos; sarcomas; glioblastomas; e vários tumores da cabeça e pescoço); leucemias e malignidades linfóides; outras desordens, tais como desordens neuronais, gliais, astrocitais, hipotalâmicas e outras desordens glandulares, macrofágicas, epiteliais, do estroma e blastocélicas; e desordens inflamatórias, angiogénicas e imunológicas. Os agentes antitumorais da presente invenção (incluindo os polipéptidos aqui revelados e agonistas que imitam a sua actividade, e.g., anticorpos, péptidos e moléculas orgânicas pequenas), são administrados a um mamífero, preferivelmente um ser humano, de acordo com métodos conhecidos, tais como administração intravenosa em *bolus* ou por infusão contínua ao longo de um período de tempo, ou por via intramuscular, intraperitoneal, intracerebroespinal, intra-ocular, intra-arterial, intralesional, subcutânea, intra-articular, intra-sinovial, intratecal, oral, tópica ou por inalação.

Podem combinar-se outros regimes terapêuticos com a administração dos agentes anticancerosos da presente invenção. Por exemplo, o paciente a tratar com estes agentes anticancerosos pode também receber terapia de radiação. Alternativa ou adicionalmente, pode administrar-se um agente quimioterapêutico ao paciente. A preparação e o calendário de dosagem para tais agentes quimioterapêuticos podem ser

utilizados de acordo com as instruções do fabricante ou como determinado empiricamente pelo profissional especialista. A preparação e o calendário de dosagem para esta quimioterapia são também descritos em Chemotherapy Service Ed., M.C. Perry, Williams & Wilkins, Baltimore, MD (1992). O agente quimioterapêutico pode preceder ou seguir a administração do agente antitumoral da presente invenção, ou pode ser administrado simultaneamente com ele. Os agentes anticancerosos da presente invenção podem ser combinados com um composto anti-estrogénico, tal como tamoxifeno ou um composto anti-progesterona, tal como onapristona (veja-se, EP 616812) em dosagens conhecidas para estas moléculas.

Pode ser desejável administrar também anticorpos contra antigénios associados aos tumores, tais como anticorpos que se ligam a ErbB2, EGFR, ErbB3, ErbB4 ou factor endotelial vascular (VEGF). Alternativa ou adicionalmente, podem ser co-administrados ao paciente dois ou mais anticorpos que se ligam ao mesmo antigénio ou a dois ou mais antigénios diferentes associados ao cancro. Por vezes, pode ser benéfico administrar também uma ou mais citocinas ao paciente. Numa concretização preferida, os agentes anticancerosos aqui são co-administrados com um agente inibidor do crescimento. Por exemplo, o agente inibidor do crescimento pode ser administrado primeiro, seguido pela administração de um agente anticanceroso da presente invenção. Contudo, estão também contempladas a administração simultânea ou a administração do agente anticanceroso da presente invenção primeiro. As dosagens adequadas para o agente inibidor do crescimento são as utilizadas actualmente e podem ser diminuídas devido à acção combinada (sinergia) do agente inibidor do crescimento e do presente anticorpo.

Para a prevenção ou tratamento de uma doença, a dosagem apropriada de um agente antitumoral dependerá do tipo de doença a tratar, tal como definido acima, da gravidade e evolução da doença, se o agente é administrado com fins preventivos ou terapêuticos, de terapia prévia, da história clínica do paciente e da resposta ao agente e critério do médico assistente. O agente é administrado adequadamente ao paciente de uma vez ou ao longo de uma série de tratamentos. Experiências com animais proporcionam orientações fiáveis para

a determinação de doses eficazes para terapia humana. A passagem de escala inter-espécies das doses eficazes pode ser realizada seguindo os princípios estabelecidos por Mordenti, J. e Chappell, W. "The use of interspecies scaling in toxicokinetics" em Toxicokinetics and New Drug Development, Yacobi et al., eds., Pergamon Press, New York 1989, pp. 42-96.

Por exemplo, dependendo do tipo e gravidade da doença, cerca de 1 µg/kg a 15 mg/kg (e.g., 0,1-20 mg/kg) de um agente antitumoral constituem uma dosagem candidata inicial para administração ao paciente quer, por exemplo, numa só ou em mais administrações separadas quer por infusão contínua. Uma dosagem diária típica pode variar de cerca de 1 µg/kg a 100 mg/kg ou mais, dependendo dos factores supramencionados. Para administrações repetidas ao longo de vários dias ou mais longas, dependendo da condição, o tratamento é mantido até que ocorra uma supressão desejada dos sintomas da doença. Contudo, podem ser úteis outros regimes de dosagem. A progressão desta terapia é facilmente monitorada por técnicas e ensaios convencionais. São proporcionadas na literatura orientações quanto a dosagens e métodos de entrega particulares; veja-se, por exemplo, Pat. U.S. 4 657 760; 5 206 344; ou 5 225 212. Antecipa-se que diferentes formulações serão eficazes para diferentes compostos de tratamento e diferentes desordens, que a administração com alvo num órgão ou num tecido, por exemplo, pode necessitar de entrega de uma maneira diferente da que tem como alvo outro órgão ou outro tecido.

K. Artigos de Fabrico

Noutra concretização da invenção, proporciona-se um artigo de fabrico contendo materiais úteis para o diagnóstico ou tratamento das desordens descritas acima. O artigo de fabrico compreende um recipiente e um rótulo. Os recipientes adequados incluem, por exemplo, frascos, garrafas, seringas e tubos de ensaio. Os recipientes podem ser feitos de vários materiais, tais como vidro ou plástico. O recipiente contém uma composição que é eficaz para o diagnóstico ou tratamento da condição e pode ter uma porta de acesso estéril (por exemplo, o recipiente pode ser um saco ou um frasco de solução intravenosa com uma rolha perfurável com uma agulha de injeção hipodérmica). O agente activo na composição é um

agente antitumoral da presente invenção. O rótulo no recipiente ou a ele associado indica que a composição é utilizada para diagnóstico ou tratamento da doença em questão. O artigo de fabrico pode compreender adicionalmente um segundo recipiente compreendendo um tampão farmacologicamente aceitável, tal como solução salina tamponada com fosfato, solução de Ringer e solução de dextrose. Pode incluir adicionalmente outros materiais desejáveis do ponto de vista comercial e do utilizador, incluindo outros tampões, diluentes, filtros, agulhas, seringas e folhetos inclusos com instruções para utilização.

Os exemplos que se seguem são oferecidos com fins meramente ilustrativos e não se pretendem de modo algum limitativos do âmbito da presente invenção.

EXEMPLOS

Os reagentes disponíveis comercialmente referidos nos exemplos foram utilizados de acordo com as instruções do fabricante, salvo indicação em contrário. A fonte das células identificadas nos exemplos seguintes e ao longo de todo o fascículo pelos números de acesso na ATCC é a American Type Culture Collection, Manassas, VA.

EXEMPLO 1

Rastreamento de Homologia do Domínio Extracelular para Identificar Novos Polipéptidos e ADNs que os Codificam

As sequências do domínio extracelular (ECD) (incluindo a sequência de sinal de secreção, caso exista) de cerca de 950 proteínas excretadas conhecidas da base de dados pública Swiss-Prot foram utilizadas para uma busca em base de dados EST. As bases de dados EST incluíam bases de dados públicas (por exemplo, Dayhoff, GenBank) e bases de dados registadas (por exemplo LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA). A busca foi efectuada utilizando o programa de computador BLAST ou BLAST-2 (Altschul et al., Methods in Enzymology, 266:460-480 (1996)) como comparação das sequências de proteínas ECD com uma tradução de 6 quadros das sequências EST. As comparações com uma pontuação BLAST de 70 (ou em

alguns casos 90) ou superior que não codificam proteínas conhecidas foram agrupadas e montadas em sequências de ADN de consenso com o programa "phrap" (Phil Green, University of Washington, Seattle, WA).

Utilizando este rastreio de homologia do domínio extracelular, as sequências de ADN de consenso foram montadas relativamente às outras sequências EST identificadas utilizando o phrap. Adicionalmente, as sequências de ADN de consenso obtidas foram frequentemente (mas nem sempre) alongadas utilizando ciclos repetidos de BLAST ou BLAST-2 e phrap para alongar a sequência de consenso o mais possível utilizando as fontes de sequências EST discutidas acima.

Com base nas sequências de consenso obtidas como descrito acima, sintetizaram-se então oligonucleótidos que se utilizaram para identificar por PCR uma biblioteca de ADNc que continha a sequência de interesse e para utilização como sondas para isolar um clone da sequência de codificação de comprimento completo para um polipéptido PRO. Os iniciadores de PCR directos e inversos têm, em geral, de 20 a 30 nucleótidos e são frequentemente concebidos para originarem um produto de PCR de cerca de 100-1000 pb de comprimento. As sequências sonda têm tipicamente 40-55 pb de comprimento. Em alguns casos, sintetizam-se oligonucleótidos adicionais quando a sequência de consenso é maior do que cerca de 1-1,5 kpb. De modo a rastrear várias bibliotecas quanto a um clone de comprimento completo, rastreou-se o ADN das bibliotecas por amplificação por PCR, tal como em Ausubel et al., Current Protocols in Molecular Biology, com o par de iniciadores de PCR. Utilizou-se então uma biblioteca positiva para isolar clones que codificam o gene de interesse, utilizando o oligonucleótido sonda e um dos pares de iniciadores.

As bibliotecas de ADNc utilizadas para isolar os clones de ADNc foram construídas por métodos *standard* utilizando reagentes disponíveis comercialmente, tais como os de Invitrogen, San Diego, CA. O ADNc foi iniciado com oligo dT contendo um local *NotI*, ligado com extremidades lisas a adaptadores tratados com hemiquinase *SalI*, clivados com *NotI*, ajustou-se o tamanho para ficar adequado para electroforese em gel e clonou-se numa orientação definida num vector de

clonagem adequado (tal como pRKB ou pRKD; pRK5B é um precursor de pRK5D que não contém o local *SfiI*; veja-se, Holmes et al., Science, 253:1278-1280 (1991)) nos locais *XhoI* e *NotI* únicos.

EXEMPLO 2

Isolamento de Clones de ADNc por Rastreo com Amilase

1. Preparação da biblioteca de ADNc iniciada com oligo dT

Isolou-se ARNm a partir de um tecido humano de interesse utilizando reagentes e protocolos da Invitrogen, San Diego, CA (Fast Track 2). Utilizou-se este ARN para gerar uma biblioteca de ADNc iniciada com oligo dT no vector pRK5D utilizando reagentes e protocolos da Life Technologies, Gaithersburg, MD (Super Script Plasmid System). Neste procedimento, o ADNc de cadeia dupla foi dimensionado para superior a 1000 pb e o ADNc ligado *SalI/NotI* foi clonado no vector clivado com *XhoI/NotI*. O pRK5D é um vector de clonagem que possui um local de iniciação da transcrição de sp6 seguido por um local para a enzima de restrição *SfiI* precedendo os locais de clonagem de ADNc *XhoI/NotI*.

2. Preparação da biblioteca de ADNc iniciada aleatoriamente

Gerou-se uma biblioteca secundária de ADNc de modo a representar preferencialmente as extremidades 5' dos clones de ADNc primários. Gerou-se ARN de Sp6 a partir da biblioteca primária (descrita acima), e utilizou-se este ARN para gerar uma biblioteca de ADNc iniciada aleatoriamente no vector pSST-AMY.0 utilizando reagentes e protocolos da Life Technologies (Super Script Plasmid System, acima referenciado). Neste procedimento, o ADNc de cadeia dupla foi dimensionado para 500-1000 pb, ligado com extremidades lisas a adaptadores *NotI*, clivado com *SfiI*, e clonado no vector clivado com *SfiI/NotI*. O pSST-AMY.0 é um vector de clonagem que possui um promotor de álcool-desidrogenase de levedura precedendo os locais de clonagem do ADNc e a sequência de amilase de ratinho (a sequência madura sem o sinal de secreção) seguida pelo terminador da álcool-desidrogenase de levedura, depois dos locais de clonagem. Assim, os ADNc clonados neste vector que são fundidos em enquadramento com a

sequência da amilase conduzirão à secreção de amilase por colónias de leveduras apropriadamente transfectadas.

3. Transformação e Detecção

ADN da biblioteca descrita no parágrafo 2 anterior foi arrefecido em gelo ao qual se adicionaram bactérias DH10B electrocompetentes (Life Technologies, 20 ml). A mistura de bactérias e vector foi então electroporada como recomendado pelo fabricante. Subsequentemente, adicionou-se meio SOC (Life Technologies, 1 ml) e incubou-se a mistura a 37°C durante 30 minutos. Plaquearam-se então os transformantes sobre 20 placas *standard* de LB de 150 mm contendo ampicilina e incubou-se durante 16 horas (37°C). As colónias positivas foram raspadas das placas e o ADN foi isolado a partir da pelete bacteriana utilizando protocolos *standard*, e.g., gradiente de CsCl. O ADN purificado foi então levado aos protocolos com leveduras que se seguem.

Os métodos com leveduras foram divididos em três categorias: (1) Transformação de levedura com o vector combinado plasmídeo/ADNc; (2) Detecção e isolamento de clones de levedura que excretam amilase; e (3) amplificação por PCR da inserção directamente a partir da colónia de leveduras e purificação do ADN para sequenciação e análise posterior.

A estirpe de levedura utilizada foi HD56-5A (ATCC-90785). Esta estirpe possui o seguinte genotipo: MAT alfa, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL⁺, SUC⁺, GAL⁺. Preferivelmente, podem ser empregues mutantes de levedura que possuem vias pós-tradução deficientes. Estes mutantes podem ter alelos deficientes em translocação em *sec71*, *sec72*, *sec62*, sendo mais preferido o *sec71* truncado. Alternativamente, antagonistas (incluindo nucleótidos anti-sentido e/ou ligandos) que interferem com a operação normal destes genes, outras proteínas implicadas nesta via pós-tradução (e.g., SEC61p, SEC72p, SEC62p, SEC63p, TDJ1p ou SSA1p-4p) ou a formação de complexos destas proteínas, podem também ser preferivelmente empregues em combinação com a levedura que expressa amilase.

A transformação foi realizada com base no protocolo delineado por Gietz et al., Nucl. Acid. Res., 20: 1425 (1992).

As células transformada foram então inoculadas a partir de ágar em caldo de meio complexo YEPD (100 ml) e crescidas durante a noite a 30°C. O caldo YEPD foi preparado como descrito em Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). A cultura de uma noite foi então diluída para cerca de 2×10^6 células/ml (aprox. $DO_{600} = 0,1$) em caldo YEPD fresco (500 ml) e novamente crescida até 1×10^7 células/ml (aprox. $DO_{600} = 0,4-0,5$).

As células foram então colhidas e preparadas para transformação por transferência para frascos de rotor GS3 num rotor Sorval GS3 a 5000 rpm durante 5 minutos, o sobrenadante foi rejeitado e depois ressuspendeu-se em água estéril, e centrifugou-se novamente em tubos Falcon de 50 ml a 3500 rpm numa centrífuga Beckman GS-6KR. O sobrenadante foi rejeitado e as células foram subsequentemente lavadas com LiAc/TE (10 ml, Tris-HCl 10 mM, EDTA 1 mM, pH 7,5, Li_2OOCCH_3 100 mM), e ressuspensas em LiAc/TE (2,5 ml).

A transformação ocorreu misturando as células preparadas (100 μ l) com ADN de testículos de salmão de cadeia simples recentemente desnaturado (Lofstrand Labs, Gaithersburg, MD) e transformando o ADN (1 μ g, vol. <10 μ l) em tubos de centrífuga. A mistura foi brevemente agitada num vórtex, e depois adicionaram-se PEG a 40%/TE (600 μ l, polietilenoglicol-4000 a 40%, Tris-HCl 10 mM, EDTA 1 mM, Li_2OOCCH_3 100 mM, pH 7,5). Esta mistura foi suavemente agitada e incubada a 30°C enquanto agitando durante 30 minutos. As células foram então submetidas a choque térmico a 42°C durante 15 minutos, e o vaso reaccional foi centrifugado numa microcentrífuga a 12 000 rpm durante 5-10 segundos, decantou-se e ressuspendeu-se em TE (500 μ l, Tris-HCl 10 mM, EDTA 1 mM, pH 7,5) seguindo-se nova centrifugação. As células foram então diluídas em TE (1 ml) e espalharam-se alíquotas (200 μ l) sobre os meios selectivos previamente preparados em placas de crescimento de 150 mm (VWR).

Alternativamente, em vez de múltiplas pequenas reacções, a transformação foi realizada utilizando uma única reacção, em grande escala, em que as quantidades de reagentes foram aumentadas de escala proporcionalmente.

O meio selectivo utilizado foi um ágar de dextrose completo sintético sem uracilo (SCD-Ura) preparado como descrito em Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Os transformantes foram crescidos a 30°C durante 2-3 dias.

A detecção de colónias que excretam amilase foi realizada incluindo amido vermelho no meio de crescimento selectivo. Acoplou-se amido ao corante vermelho (Reactive Red-120, Sigma) conforme o procedimento descrito por Biely et al., Anal. Biochem., 172:176-179 (1988). Incorporou-se o amido acoplado nas placas de ágar SCD-Ura numa concentração final de 0,15% (p/v), e tamponou-se com fosfato de potássio a um pH de 7,0 (concentração final de 50-100 mM).

Repicaram-se as colónias positivas e riscaram-se em meio selectivo fresco (em placas de 150 mm) de modo a obter colónias individuais bem isoladas e identificáveis. As colónias individuais bem isoladas positivas para secreção de amilase foram detectadas por incorporação directa de amido vermelho no ágar SCD-Ura tamponado. As colónias positivas foram determinadas pela sua capacidade de clivar o amido resultando um halo límpido em redor da colónia positiva visualizado directamente.

4. Isolamento de ADN por Amplificação por PCR

Quando foi isolada uma colónia positiva, uma sua porção foi repicada com um palito e diluída em água estéril (30 µl) numa placa de 96 poços. Neste momento, as colónias positivas ou foram congeladas e armazenadas para análise subsequente ou foram imediatamente amplificadas. Utilizou-se uma alíquota de células (5 µl) como molde para a reacção PCR num volume de 25 µl contendo: 0,5 µl de Klentaq (Clontech, Palo Alto, CA); 4,0 µl de dNTP's 10 mM (Perkin Elmer-Cetus); 2,5 µl de tampão de Kentaq (Clontech); 0,25 µl de oligo 1 directo; 0,25 µl de oligo 2 inverso; 12,5 µl de água destilada. A sequência do oligonucleótido 1 directo era:

5'-TGTAACGACGGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:57)

A sequência do oligonucleótido 2 inverso era:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:58)

A PCR foi então realizada como se segue:

- a. Desnaturação 92°C, 5 minutos
- b. 3 ciclos de:
 - Desnaturação 92°C, 30 segundos
 - Hibridação 59°C, 30 segundos
 - Alongamento 72°C, 60 segundos
- c. 3 ciclos de:
 - Desnaturação 92°C, 30 segundos
 - Hibridação 57°C, 30 segundos
 - Alongamento 72°C, 60 segundos
- d. 25 ciclos de:
 - Desnaturação 92°C, 30 segundos
 - Hibridação 55°C, 30 segundos
 - Alongamento 72°C, 60 segundos
- e. Manutenção 4°C

As regiões sublinhadas dos oligonucleótidos hibridaram com a região promotora de ADH e a região de amilase, respectivamente, e amplificaram uma região de 307 pb do vector pSST-AMY.0 quando não estava presente a inserção. Tipicamente, os primeiros 18 nucleótidos da extremidade 5' destes oligonucleótidos continham locais de hibridação para os iniciadores de sequenciação. Assim, o produto total da reacção PCR a partir de um vector vazio tinha 343 pb. Contudo, o ADNc fundido com a sequência de sinal resultou em sequências nucleotídicas consideravelmente mais longas.

Após a PCR, examinou-se uma alíquota da mistura reaccional (5 µl) por electroforese em gel de agarose num gel de agarose a 1% utilizando um sistema de tamponamento Tris-Borato-EDTA (TBE) como descrito por Sambrook *et al.*, *supra*. Os clones que resultaram num único produto de PCR forte maior que 400 pb foram adicionalmente analisados por sequenciação do ADN após purificação com uma coluna de limpeza de PCR 96 Qiaquick (Qiagen Inc., Chatsworth, CA).

EXEMPLO 3

Isolamento de Clones de ADNc Utilizando Análise de Algoritmo de Sinal

Identificaram-se várias sequências de ácido nucleico que codificam polipéptidos, por aplicação de um algoritmo de busca de sequências de sinal registado desenvolvido por Genentech, Inc., (South San Francisco, CA) nas EST, bem como fragmentos EST agrupados e montados de bases de dados públicas (e.g., GenBank) e/ou privadas (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA). O algoritmo de sequências de sinal calcula uma pontuação de sinal de secreção com base no carácter dos nucleótidos do ADN que circundam o primeiro e, opcionalmente, o segundo codão(ões) de metionina (ATG) na extremidade 5' da sequência ou do fragmento de sequência em questão. Os nucleótidos a seguir ao primeiro ATG tem que codificar pelo menos 35 aminoácidos não ambíguos sem quaisquer codões de paragem. Se o primeiro ATG tiver os aminoácidos necessários, o segundo não é examinado. Se nenhum deles cumprir os requisitos, a sequência candidata não é pontuada. De modo a determinar se a sequência EST contém uma sequência de sinal autêntica, os ADN e as sequências de aminoácidos correspondentes que circundam o codão ATG são pontuados utilizando um conjunto de sete sensores (parâmetros de avaliação) que se sabe estarem associados a sinais de secreção. A utilização deste algoritmo resultou na identificação de numerosas sequências de ácido nucleico que codificam polipéptidos.

EXEMPLO 4

Isolamento de Clones de ADNc que Codificam PRO440 Humano

Montou-se uma sequência de ADN de consenso relativamente a outras sequências EST utilizando o phrap conforme descrito no Exemplo 1 anterior. As bases de dados EST incluíam bases de dados EST públicas (e.g., GenBank) e uma base de dados EST privada (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) e EST registadas da Genentech. Esta sequência de consenso é aqui designada por DNA77634. Com base na sequência de consenso de DNA77634, sintetizaram-se oligonucleótidos: 1) para identificar por PCR uma biblioteca de ADNc que continha a

sequência de interesse, e 2) para utilizar como sondas para isolar um clone da sequência de codificação de comprimento completo para PRO4400.

Sintetizaram-se um par de iniciadores de PCR (directo e inverso):

iniciador de PCR directo:

5'-GCTGCTGCCGTCCATGCTGATG-3' (SEQ ID NO:3)

iniciador de PCR inverso:

5'-CTCGGGGAATGTGACATCGTCGC-3' (SEQ ID NO:4)

Adicionalmente, construiu-se uma sonda de hibridação oligonucleotídica sintética a partir da sequência de consenso de DNA77634 que tinha a seguinte sequência nucleotídica:

sonda de hibridação:

5'-GCTGCCGTCCA TGCTGA TGTITGCGGTGATCGTGG-3' (SEQ ID NO: 5)

O ARN para a construção das bibliotecas de ADNc foi isolado a partir de uma linha celular de adenocarcinoma humano. As bibliotecas de ADNc utilizadas para isolar os clones de ADNc foram construídas por métodos *standard* utilizando reagentes comercialmente disponíveis tais como os da Invitrogen, San Diego, CA. O ADNc foi iniciado com oligo dT contendo um local *NotI*, ligado com extremidades lisas a adaptadores *Sall* tratados com hemi-quinase, clivado com *NotI*, apropriadamente dimensionado por electroforese em gel e clonado num orientação definida num vector de clonagem adequado (tal como pRKB ou pRKD: pRK5B é um precursor de pRK5D que não contém o local *SfiI*; veja-se, Holmes et al., Science, 253:1278-1280 (1991)) nos locais únicos *XhoI* e *NotI*.

A sequenciação do ADN dos clones isolados como acima descrito originou a sequência de ADN de comprimento completo para o polipéptido PRO4400 (aqui designada por DNA87974-2609 [Figura 1, SEQ ID NO:1]) e a sequência de proteína derivada para esse polipéptido PRO4400.

A totalidade da sequência de codificação de DNA87974-2609 está incluída na Figura 1 (SEQ ID NO:1, o Clone DNA87974-2609 contém um único enquadramento de leitura aberto com um local de iniciação da tradução aparente nas posições nucleotídicas 27-29, e um codão de paragem aparente nas posições nucleotídicas 1026-1028. O precursor de polipéptido previsto tem 333 aminoácidos de comprimento, e tem um peso molecular estimado de cerca de 38 618 daltons e um pI de cerca de 9,27. A análise da sequência de PRO4400 de comprimento completo mostrada na Figura 2 (SEQ ID NO:2) evidencia a presença de uma variedade de importantes domínios polipeptídicos, em que as localizações dadas para esses importantes domínios polipeptídicos são aproximadas como acima descrito. A análise do polipéptido PRO4400 de comprimento completo mostrada na Figura 2 evidencia a presença do seguinte: um péptido de sinal desde cerca do aminoácido 1 a cerca do aminoácido 23; locais de N-glicosilação desde cerca do aminoácido 67 a cerca do aminoácido 71 e desde cerca do aminoácido 325 a cerca do aminoácido 329; locais de fosforilação por tirosina-quinase desde cerca do aminoácido 152 a cerca do aminoácido 159 e a cerca do aminoácido 183; e locais de N-miristoilação desde cerca do aminoácido 89 a cerca do aminoácido 95, de desde cerca do aminoácido 128 a cerca do aminoácido 134. O clone DNA87974-2609 foi depositado na ATCC em 27 de Abril, 1999, e foi-lhe atribuído o número de depósito na ATCC 203963.

Uma análise da base de dados Dayhoff (versão 35.45 SwissProt 35), utilizando uma análise de alinhamento de sequências WU-BLAST2 da sequência de comprimento completo mostrada na Figura 2 (SEQ ID NO:2, evidenciou significativa homologia entre a sequência de aminoácidos de PRO4400 e as seguintes sequências Dayhoff: AF033827_1, AF070594_1, AF022729_1, CEC34F6_4, SYFB_THETH, G70405, SD_DROME, S64023, ALK1_YEAST e VG04_HSVII.

EXEMPLO 5

Hibridação *in situ*

A hibridação *in situ* é uma técnica poderosa e versátil para a detecção e localização de sequências de ácido nucleico

em preparações de células ou tecidos. Pode ser útil, por exemplo, para identificar locais de expressão génica, analisar a distribuição da transcrição nos tecidos, identificar e localizar uma infecção viral, seguir alterações na síntese de ARNm específico e auxiliar no mapeamento de cromossomas.

A hibridação *in situ* é realizada seguindo uma versão optimizada do protocolo de Lu e Gillett, Cell Vision, 1:169-176(1994), utilizando ribossondas marcadas com ^{33}P geradas por PCR. Sucintamente, seccionaram-se tecidos humanos fixados com formalina e embebidos em parafina, desparafinizaram-se, desproteínaram-se com proteinase K (20 g/ml) durante 15 minutos a 37°C e processaram-se adicionalmente para hibridação *in situ* tal como descrito por Lu e Gillett, *supra*. Gera-se uma ribossonda anti-sentido marcada com (^{33}P)UTP a partir de um produto de PCR e hibrida-se a 55°C durante a noite. Mergulham-se as lâminas em emulsão de traçador nuclear Kodak NTB2™ e expõem-se durante 4 semanas.

Síntese de ^{33}P -Ribossonda

Secaram-se num Speed-Vac 6,0 μl (125 mCi) de ^{33}P -UTP (Amersham BF 1002, SA <2000 Ci/mmol). A cada tubo contendo o ^{33}P -UTP seco adicionam-se os seguintes ingredientes:

20 μl de tampão de transcrição 5x
1,0 μl de DTT (100 mM)
2,0 μl de mistura de NTP (2,5 mM: 10 μl de cada de GTP, CTP & ATP, 10 mM, + 10 μl de H_2O)
1,0 μl de UTP (50 μM)
1,0 μl de RNAsin
1,0 μl de molde de ADN (μg)
1,0 μl de H_2O
1,0 μl de ARN-polimerase (para produtos de PCR T3= AS, T7= S, usualmente)

Incubaram-se os tubos a 37°C durante uma hora. Adicionou-se um total de 1,0 μl de ADNase RQ1, seguido por incubação a 37°C durante 15 minutos. Adicionou-se um total de 90 μl de TE (Tris 10 mM, pH 7,6/EDTA 1 mM, pH 8,0) e pipetou-se a mistura para papel DE81. A solução restante é carregada numa unidade de ultrafiltração MICROCON-50™ e

centrifugada utilizando o programa 10 (6 minutos). Inverte-se a unidade de filtração para um segundo tubo e centrifuga-se utilizando o programa 2 (3 minutos). Após a centrifugação de recuperação final, adiciona-se um total de 100 µl de TE, seguidamente pipeta-se 1 µl do produto final em papel DE81 e conta-se em 6 ml de BIOFLUOR II™.

Corre-se a sonda num gel de TBE/ureia. Adiciona-se um total de 1-3 µl da sonda ou 5 µl de ARN Mrk III a 3 µl de tampão de carga. Após aquecimento a 95°C num bloco de aquecimento durante três minutos, coloca-se o gel imediatamente em gelo. Lavam-se os poços do gel e carrega-se a amostra e corre-se a 180-250 volts durante 45 minutos. Embrulha-se o gel em película plástica (marca SARAN™) e expõe-se a película XAR com um ecrã intensificador num congelador a -70°C durante, de uma hora a durante a noite.

Hibridação com ³³P

A. Pré-tratamento de secções congeladas

Removem-se as lâminas do congelador, colocam-se em tabuleiros de alumínio e descongelam-se à temperatura ambiente durante 5 minutos. Colocam-se os tabuleiros num incubador a 55°C durante cinco minutos para reduzir a condensação. Fixam-se as lâminas durante 10 minutos em paraformaldeído a 4% em gelo numa *hotte* de extracção e lava-se com SSC 0,5x durante 5 minutos, à temperatura ambiente (25 ml de SSC 20x + 975 ml de H₂O SQ). Após desproteinação com proteinase K a 0,5 µg/ml durante 10 minutos a 37°C (12,5 µl de solução-mãe a 10 mg/ml em 250 ml de tampão de ARNase isento de ARNase pré-aquecido), lavaram-se as secções com SSC 0,5x durante 10 minutos à temperatura ambiente. Desidratam-se as secções em etanol a 70%, 95% e 100%, 2 minutos cada.

B. Pré-tratamento de secções embebidas em parafina

Desparafinizaram-se as lâminas, colocaram-se em H₂O SQ e enxaguaram-se duas vezes com SSC 2x à temperatura ambiente, durante 5 minutos de cada vez. Desproteínaram-se as secções com proteinase K a 20 µg/ml (500 µl de 10 mg/ml em 250 ml de tampão de ARNase isento de ARNase; 37°C, 15 minutos) para

tecido de embrião humano ou 8 x proteinase K (100 µl em 250 ml de tampão de ARNase, 37°C, 30 minutos) para tecidos em formalina. Subsequentemente enxagua-se com SSC 0,5x e desidrata-se como descrito acima.

C. *Pré-hibridação*

Dispõem-se as lâminas numa caixa de plástico forrada com papel de filtro saturado com tampão Box (SSC 4x, formamida a 50%). Cobre-se o tecido com 50 µl de tampão de hibridação (3,75 g de sulfato de dextrano + 6 ml de H₂O SQ), submete-se a vórtex e aquece-se no microondas durante 2 minutos com a tampa solta. Após arrefecer em gelo, adicionam-se 18,75 ml de formamida, 3,75 ml de SSC 20x e 9 ml de H₂O SQ e submete-se bem o tecido a vórtex e incuba-se a 42°C durante 1-4 horas.

D. *Hibridação*

Aquece-se a 95°C durante 3 minutos, 1,0 x 10⁶ cpm de sonda e 1,0 µl de ARNt (solução-mãe a 50 mg/ml) por lâmina. Arrefecem-se as lâminas em gelo e adicionam-se 48 µl de tampão de hibridação por lâmina. Após submissão a vórtex, adicionam-se 50 µl de mistura ³³P a 50 µl de pré-hibridação na lâmina. Incubam-se as lâminas durante a noite a 55°C.

E. *Lavagens*

A lavagem é efectuada 2x 10 minutos com SSC 2x, EDTA à temperatura ambiente (400 ml de SSC 20x + 16 ml de EDTA 0,25 M, V₁=4L), seguida por tratamento com ARNase a 37°C durante 30 minutos (500 µl de 10 mg/ml em 250 ml de tampão de ARNase = 20 µg/ml). Lavam-se as lâminas 2 x 10 minutos com SSC 2x, EDTA à temperatura ambiente. As condições de rigor da lavagem são as seguintes: 2 horas a 55°C, SSC 0,1x, EDTA (20 ml SSC 20x + 16 ml EDTA, V₁=4L).

EXEMPLOS 6

Utilização de PRO como Sonda de Hibridação

O seguinte método descreve a utilização de uma sequência nucleotídica que codifica PRO como sonda de hibridação. O ADN

compreendendo a sequência de codificação de PRO de comprimento completo ou maduro, tal como aqui revelado e/ou seus fragmentos, são empregues como sonda para rastrear ADN homólogo (tal como os que codificam variantes de ocorrência natural de PRO) em bibliotecas de ADNc de tecido humano ou bibliotecas genómicas de tecido humano.

A hibridação e a lavagem dos filtros contendo qualquer das bibliotecas de ADN são realizadas sob as seguintes condições de elevado rigor. A hibridação da sonda marcada radioactivamente derivada do gene que codifica um polipéptido PRO, com os filtros, é efectuada numa solução de formamida a 50%, SSC 5x, SDS a 0,1%, pirofosfato de sódio a 0,1%, fosfato de sódio 50 mM, pH 6,8, solução de Denhardt 2x e sulfato de dextrano a 10% a 42°C durante 20 horas. A lavagem dos filtros é efectuada com uma solução aquosa de SSC 0,1x e SDS a 0,1% a 42°C.

Podem então identificar-se os ADN com a identidade de sequência desejada em relação ao ADN que codifica o PRO de sequência nativa de comprimento completo, utilizando técnicas *standard* conhecidas na especialidade.

EXEMPLO 7

Expressão de PRO em *E. coli*

Este exemplo ilustra a preparação de uma forma não glicosilada de PRO por expressão recombinante em *E. coli*.

A sequência de ADN que codifica PRO é inicialmente amplificada utilizando iniciadores de PCR seleccionados. Os iniciadores devem conter locais para enzimas de restrição que correspondem aos locais para enzimas de restrição do vector de expressão seleccionado. Podem ser empregues vários vectores de expressão. Constitui um exemplo de um vector adequado o pBR322 (derivado de *E. coli*; veja-se Bolivar et al., *Gene*, 2:95 (1977)) que contém genes para resistência a ampicilina e tetraciclina. O vector é digerido com enzima de restrição e desfosforilado. Ligam-se então as sequências amplificadas por PCR ao vector. O vector incluirá, preferivelmente, sequências que codificam um gene de resistência a antibiótico, um

promotor *trp*, um comando poli-His (incluindo os primeiros seis codões STII, sequência poli-His e o local de clivagem para enteroquinase), a região de codificação de PRO, o terminador de transcrição lambda e um gene *argU*.

A mistura de ligação é então utilizada para transformar uma estirpe de *E. coli* seleccionada utilizando os métodos descritos em Sambrook et al., supra. Identificam-se os transformantes pela sua capacidade de crescer em placas LB e seleccionam-se as colónias resistentes ao antibiótico. O ADN plasmídico pode ser isolado e confirmado por análise de restrição e sequenciação de ADN.

Podem crescer-se os clones seleccionados durante a noite em meio de cultura líquido, tal como caldo LB suplementado com antibióticos. A cultura durante a noite pode ser subsequentemente utilizada para inocular uma cultura em maior escala. Crescem-se então as células até uma densidade óptica desejada, durante o que a expressão do promotor é iniciada.

Após cultura das células durante mais algumas horas, as células podem ser recolhidas por centrifugação. A pelele celular obtida na centrifugação pode ser solubilizada utilizando vários agentes conhecidos na especialidade e a proteína PRO solubilizada pode ser então purificada utilizando uma coluna quelante de metais, sob condições que permitem a ligação forte da proteína.

Os PRO podem ser expressos em *E. coli* numa forma marcada com poli-His, utilizando o seguinte procedimento. O ADN que codifica PRO é inicialmente amplificado utilizando iniciadores de PCR seleccionados. Os iniciadores conterão locais para enzimas de restrição que correspondem aos locais para enzimas de restrição no vector de expressão seleccionado, e outras sequências úteis que proporcionam um início de tradução fiável e eficiente, purificação rápida numa coluna quelante de metais e remoção proteolítica com enteroquinase. Ligam-se então as sequências amplificadas por PCR marcadas com poli-His num vector de expressão, que é utilizado para transformar um hospedeiro *E. coli* baseado na estirpe 52 (W3110 fuhA(tonA) *lon* gale rpoHts(htpRts) clpP(lacIq)). Crescem-se primeiro os transformantes em LB contendo carbenicilina a 50 mg/ml a 30°C,

com agitação, até se atingir uma D.O.₆₀₀ de 3-5. Diluem-se então as culturas 50-100 vezes em meio CRAP (preparado por mistura de 3,57 g de $(\text{NH}_4)_2\text{SO}_4$, 0,71 g de citrato de sódio·2H₂O, 1,07 g de KCl, 5,36 g de extracto de levedura Difco, 5,36 g de Hycase SF de Sheffield em 500 ml de água, bem como MPOS 110 mM, pH 7,3, glucose a 0,55% (p/v) e MgSO_4 7 mM) e crescem-se durante aproximadamente 20-30 horas a 30°C com agitação. Removem-se as amostras para verificar a expressão por análise SDS-PAGE e centrifuga-se a cultura em bruto para sedimentar as células. Congelam-se as peletes de células até à purificação e redobragem.

Ressuspende-se a pasta de *E. coli* de fermentações de 0,5 a 1 L (peletes de 6-10 g) em 10 volumes (p/v) em tampão de guanidina 7 M, Tris 20 mM, pH 8. Adiciona-se sulfito de sódio sólido e tetrationato de sódio para obter concentrações finais de 0,1 M e 0,02 M, respectivamente, e agita-se a solução durante a noite a 4°C. Este passo resulta numa proteína desnaturada com todos os resíduos de cisteína bloqueados por sulfitação. Centrifuga-se a solução a 40 000 rpm numa ultracentrífuga Beckman durante 30 min. Dilui-se o sobrenadante com 3-5 volumes de tampão de coluna quelante de metais (guanidina 6 M, Tris 20 mM, pH 7,4) e filtra-se através de filtros de 0,22 microns para clarificar. O extracto clarificado é carregado numa coluna quelante de metais Qiagen Ni^{2+} -NTA de 5 ml equilibrada no tampão de coluna quelante de metais. Lava-se a coluna com tampão adicional contendo imidazole 50 mM (Calbiochem, grau Utrol), pH 7,4. Elui-se a proteína com tampão contendo imidazole 250 mM. Reúnem-se as fracções que contêm a proteína desejada e armazenam-se a 4°C. Estima-se a concentração de proteína através da sua absorvância a 280 nm, utilizando o coeficiente de extinção calculado com base na sua sequência de aminoácidos.

As proteínas são redobradas por diluição lenta da amostra em tampão de redobragem recentemente preparado que consiste em: Tris 20 mM, pH 8,6, NaCl 0,3 M, ureia 2,5 M, cisteína 5 mM, glicina 20 mM e EDTA 1 mM. Os volumes de redobragem são seleccionados de modo a que a concentração final de proteína esteja entre 50 a 100 microgramas/ml. Agita-se suavemente a solução de redobragem a 4°C durante 12-36 horas. Para-se a reacção de redobragem por adição de TFA até uma concentração

final de 0,4% (pH aproximadamente de 3). Antes da purificação adicional da proteína, filtra-se a solução através de um filtro de 0,22 microns e adiciona-se acetonitrilo até uma concentração final de 2-10%. Cromatografa-se a proteína redobrada numa coluna de fase inversa Poros R1/H utilizando um tampão móvel de TFA a 0,1% com eluição com um gradiente de acetonitrilo de 10 a 80%. Analisam-se alíquotas das fracções com absorvância A_{280} em géis de SDS-poliacrilamida e reúnem-se as fracções contendo proteína redobrada homogénea. Em geral, as espécies correctamente redobradas da maioria das proteínas são eluídas com as concentrações mais baixas de acetonitrilo, dado que essas espécies são as mais compactas com os seus interiores hidrófobos protegidos da interacção com a resina de fase inversa. As espécies agregadas são usualmente eluídas com concentrações de acetonitrilo mais elevadas. Além de resolver as formas com dobragem incorrecta das proteínas relativamente à forma desejada, o passo de fase inversa também remove endotoxina das amostras.

As fracções contendo o polipéptido PRO dobrado desejado são reunidas e remove-se o acetonitrilo utilizando uma corrente suave de azoto dirigida para a solução. As proteínas são formuladas em Hepes 20 mM, pH 6,8 com cloreto de sódio 0,14 M e manitol a 4% por diálise ou por filtração em gel utilizando resinas G25 Superfine (Pharmacia) equilibradas em tampão de formulação, e esterilizadas por filtração.

Muitos dos polipéptidos PRO aqui revelados foram expressos com sucesso como acima descrito.

EXEMPLO 8

Expressão de PRO em Células de Mamífero

Este exemplo ilustra a preparação de uma forma potencialmente glicosilada de PRO por expressão recombinante em células de mamífero.

O vector, pRK5 (veja-se EP 307 247, publicado em 15 de Março de 1989), é utilizado como vector de expressão. Opcionalmente, liga-se o ADN de PRO a pRK5 com enzimas de restrição seleccionadas para permitir a inserção do ADN de PRO

utilizando métodos de ligação, tal como descrito em Sambrook *et al.*, *supra*. O vector resultante é denominado pRK5-PRO.

Numa concretização, as células hospedeiras seleccionadas podem ser células 293. Cultivam-se células 293 humanas (ATCC CCL 1573) até à confluência em placas de cultura de tecidos em meio tal como DMEM suplementado com soro fetal de vitelo e opcionalmente, componentes nutrientes e/ou antibióticos. Misturam-se cerca de 10 µg de ADN de pRK5-PRO com cerca de 1 µg de ADN que codifica o gene VA RNA [Thimmappaya *et al.*, Cell, 31:543 (1982)] e dissolve-se em 500 µl de Tris-HCl 1 mM, EDTA 0,1 mM, CaCl₂ 0,227 M. A esta mistura adicionam-se, gota a gota, 500 µl de HEPES 50 mM (pH 7,35), NaCl 280 mM, NaPO₄ 1,5 mM, e deixa-se formar um precipitado durante 10 minutos a 25°C. Suspende-se o precipitado e adiciona-se a células 293 e deixa-se assentar durante cerca de quatro horas a 37°C. Aspira-se o meio de cultura e adicionam-se 2 ml de glicerol a 20% em PBS durante 30 segundos. Lavam-se então as células 293 com meio isento de soro, adiciona-se meio fresco e incubam-se as células durante cerca de 5 dias.

Aproximadamente 24 horas após as transfecções, remove-se o meio de cultura e substitui-se por meio de cultura (sozinho) ou meio de cultura contendo ³⁵S-cisteína a 200 µCi/ml e ³⁵S-metionina a 200 µCi/ml. Após uma incubação de 12 horas, recolhe-se o meio condicionado, concentra-se num filtro rotativo e carrega-se num gel de SDS a 15%. O gel processado pode ser seco e exposto a película durante um período de tempo seleccionado para revelar a presença do polipéptido PRO. As culturas contendo células transfectadas podem ser submetidas a incubação adicional (em meio isento de soro) e o meio é testado em bioensaio seleccionados.

Numa técnica alternativa, pode introduzir-se o PRO em células 293 transientemente utilizando o método do sulfato de dextrano descrito por Sompariyac *et al.*, Proc. Natl. Acad. Sci., 12:1575 (1981). Cultivam-se as células 293 até à densidade máxima num balão rotativo e adicionam-se 700 µg de ADN de pRK5-PRO. As células são primeiro concentradas no balão rotativo por centrifugação e lavam-se com PBS. Incuba-se o precipitado de ADN-dextrano na pelete celular durante quatro horas. Tratam-se as células com glicerol a 20% durante

90 segundos, lavam-se com meio de cultura de tecidos e reintroduzem-se num balão rotativo contendo meio de cultura de tecidos, insulina bovina a 5 µg/ml e transferrina bovina a 0,1 µg/ml. Após cerca de quatro dias, o meio condicionado é centrifugado e filtrado para remover células e resíduos. A amostra contendo PRO expresso pode então ser concentrada e purificada por qualquer método seleccionado, tal como diálise e/ou cromatografia em coluna.

Noutra concretização, o PRO pode ser expresso em células CHO. O vector pRK5-PRO pode ser transfectado em células CHO utilizando reagentes conhecidos, tais como CaPO_4 ou DEAE-dextrano. Tal como descrito acima, as culturas celulares podem ser incubadas e o meio substituído por meio de cultura (sozinho) ou meio contendo um marcador radioactivo, tal como ^{35}S -metionina. Após determinar a presença de um polipéptido PRO, o meio de cultura pode ser substituído por meio isento de soro. Preferivelmente, as culturas são incubadas durante cerca de 6 dias e seguidamente recolhe-se o meio condicionado. O meio contendo o polipéptido PRO expresso pode então ser concentrado e purificado por qualquer método seleccionado.

Pode também expressar-se o PRO marcado com epítopo em células hospedeiras CHO. O PRO pode ser subclonado fora do vector pRK5. A inserção do subclone pode ser sujeita a PCR para se fundir em enquadramento com um marcador epitópico seleccionado, tal como um marcador poli-His, num vector de expressão de baculovírus. A inserção de PRO marcada com poli-His pode então ser subclonada num vector dirigido por SV40 contendo um marcador de selecção, tal como DHFR, para a selecção de clones estáveis. Finalmente, as células CHO podem ser transfectadas (tal como descrito acima) com o vector dirigido por SV40. Pode-se efectuar a marcação tal como descrito acima, para verificar a expressão. O meio de cultura contendo o PRO marcado com poli-His expresso pode então ser concentrado e purificado por qualquer método seleccionado, tal como por cromatografia de afinidade em Ni^{2+} -quelato.

O PRO pode também ser expresso em células CHO e/ou COS através de um procedimento de expressão transiente ou em

células CHO através de outro procedimento de expressão estável.

A expressão estável em células CHO é realizada utilizando o seguinte procedimento. As proteínas são expressas na forma de uma construção de IgG (imunoadesina), em que as sequências de codificação para as formas solúveis (e.g., domínios extracelulares) das proteínas respectivas, são fundidas a uma região constante de IgG1 contendo os domínios de charneira, CH2 e CH2 e/ou numa forma marcada com poli-His.

Após amplificação por PCR, os ADN respectivos são subclonados num vector de expressão de CHO utilizando técnicas *standard*, tal como descrito em Ausubel et al., Current Protocols of Molecular Biology, Unidade 3.16, John Wiley and Sons (1997). Os vectores de expressão de CHO são construídos de modo a terem locais de restrição compatíveis a 5' e 3' do ADN de interesse, para permitir o vaivém conveniente dos ADNc. O vector utilizado para a expressão em células CHO é tal como descrito em Lucas et al., Nucl. Acids Res., 24:9 (1774-1779) (1996) e utiliza o promotor precoce/potenciador de SV40 para conduzir a expressão do ADNc de interesse e di-hidrofolato-redutase (DHFR). A expressão de DHFR permite a selecção relativamente à manutenção estável do plasmídeo após transfecção.

Introduzem-se doze miligramas do ADN do plasmídeo desejado em aproximadamente 10 milhões de células CHO utilizando reagentes de transfecção Superfect® (Qiagen), Dosper® ou Fugene® (Boehringer Mannheim) comercialmente disponíveis. Crescem-se as células como descrito em Lucas et al., *supra*. Congelam-se aproximadamente 3×10^7 células numa ampola para crescimento e produção adicionais, tal como se descreve adiante.

Descongela-se as ampolas contendo o ADN plasmídico colocando-as num banho de água e mistura-se com vórtex. Pipeta-se o conteúdo para um tubo de centrífuga contendo 10 ml de meio e centrifuga-se a 1000 rpm durante 5 minutos. Aspira-se o sobrenadante e ressuspendem-se as células em 10 ml de meio selectivo (PS20 filtrado por 0,2 µm com soro bovino fetal diafiltrado por 0,2 µm a 5%). As células são então

divididas em alíquotas para um balão rotativo de 100 ml contendo 90 ml de meio selectivo. Após 1-2 dias, transferem-se as células para um balão rotativo de 250 ml contendo 150 ml de meio de crescimento selectivo e incuba-se a 37°C. Após mais 2-3 dias, inoculam-se balões rotativos de 250 ml, 500 ml e 2000 ml com 3×10^5 células/ml. O meio celular é trocado por meio fresco por centrifugação e ressuspensão em meio de produção. Embora possa ser empregue qualquer meio adequado para CHO, utiliza-se de facto um meio de produção descrito na Patente US 5 122 469, concedida em 16 de Junho de 1992. Inocula-se um balão rotativo de produção de 3 L a $1,2 \times 10^6$ células/ml. No dia 0, determina-se o número de células e o pH. No dia 1, amostra-se o balão rotativo e inicia-se a aspersão com ar filtrado. No dia 2, amostra-se o balão rotativo, muda-se a temperatura para 33°C e adicionam-se 30 ml de glucose a 500 g/L e 0,6 ml de anti-espuma a 10% (e.g., emulsão de polidimetilsiloxano a 35%, Dow Corning 365 emulsão de qualidade medicinal). Ao longo da produção, ajusta-se o pH conforme necessário para o manter a cerca de 7,2. Após 10 dias, ou quando a viabilidade cair abaixo de 70%, recolhe-se o meio de cultura celular por centrifugação e filtra-se através de um filtro de 0,22 μm . O filtrado é armazenado a 4°C ou carregado imediatamente em colunas para purificação.

Para as construções marcadas com poli-His, as proteínas são purificadas utilizando uma coluna Ni^{2+} -NTA (Qiagen). Antes da purificação, adiciona-se imidazole ao meio condicionado numa concentração de 5 mM. Bombeia-se o meio condicionado para uma coluna de Ni^{2+} -NTA de 6 ml equilibrada em tampão Hepes 20 mM, pH 7,4, contendo NaCl 0,3 M e imidazole 5 mM, a um caudal de 4-5 ml/min. a 4°C. Após a carga, lava-se a coluna com tampão de equilíbrio adicional e elui-se a proteína com tampão de equilíbrio contendo imidazole 0,25 M. A proteína altamente purificada é subsequentemente dessalinizada num tampão de armazenagem contendo Hepes 10 mM, NaCl 0,14 M e manitol a 4%, pH 6,8, com uma coluna de 25 ml de G25 Superfine (Pharmacia) e armazenada a -80°C.

As construções de imunoadesina (contendo Fc) são purificadas a partir do meio condicionado da seguinte forma. Bombeia-se o meio condicionado para uma coluna de 5 ml de Proteína A (Pharmacia) que foi equilibrada em tampão de

fosfato de Na 20 mM, pH 6,8. Após a carga, lava-se a coluna extensivamente com tampão de equilíbrio antes de eluição com ácido cítrico 100 mM, pH 3,5. Neutraliza-se a proteína eluída imediatamente por recolha de fracções de 1 ml para tubos contendo 275 µl de tampão Tris 1 M, pH 9. A proteína altamente purificada é subsequentemente dessalinizada em tampão de armazenamento tal como descrito acima para as proteínas marcadas com poli-His. Avalia-se a homogeneidade em géis de SDS-poliacrilamida e por sequenciação dos aminoácidos N-terminais por degradação de Edman.

Muitos dos polipéptidos PRO aqui descritos foram expressos com sucesso como acima descrito.

EXEMPLO 9

Expressão de PRO em levedura

O método seguinte descreve a expressão recombinante de PRO em levedura.

Primeiro, constroem-se vectores de expressão de levedura para a produção intracelular ou secreção de PRO a partir do promotor ADH2/GAPDH. O ADN que codifica PRO e o promotor são inseridos em locais para enzimas de restrição adequados no plasmídeo seleccionado para dirigir a expressão intracelular do PRO. Para secreção, pode clonar-se ADN que codifica PRO no plasmídeo seleccionado, conjuntamente com ADN que codifica o promotor ADH2/GAPDH, um péptido de sinal de PRO nativo ou outro péptido de sinal de mamífero ou, por exemplo, uma sequência de sinal/comando de secreção do factor alfa ou de invertase de levedura, e sequências ligantes (caso necessário) para a expressão de PRO.

As células de levedura, tal como a levedura da estirpe AB110, podem ser então transformadas com os plasmídeos de expressão descritos acima e cultivadas em meios de fermentação seleccionados. Os sobrenadantes da levedura transformada podem ser analisados por precipitação com ácido tricloroacético a 10% e separação por SDS-PAGE, seguidas de coloração dos géis com corante azul de Coomassie.

O PRO recombinante pode subsequentemente ser isolado e purificado por remoção das células de levedura do meio de fermentação por centrifugação e depois concentração do meio utilizando filtros de cartucho seleccionados. O concentrado contendo PRO pode ser purificado adicionalmente utilizando resinas de cromatografia em coluna seleccionadas.

Muitos dos polipéptidos PRO aqui descritos foram expressos com sucesso como acima descrito.

EXEMPLO 10

Expressão de PRO em Células de Insecto Infectadas Com Baculovírus

O método seguinte descreve a expressão recombinante em células de insecto infectadas com baculovírus.

A sequência de codificação para PRO é fundida a montante de um marcador epitópico contido num vector de expressão de baculovírus. Estes marcadores epitópicos incluem marcadores poli-His e marcadores de imunoglobulina (tal como regiões Fc de IgG). Podem ser utilizados vários plasmídeos, incluindo plasmídeos derivados de plasmídeos disponíveis comercialmente, tal como o pVL1393 (Novagen). Sucintamente, a sequência de codificação de PRO ou a porção desejada da sequência de codificação de PRO [tal como a sequência que codifica o domínio extracelular de uma proteína transmembranar ou a sequência que codifica a proteína madura caso a proteína seja extracelular] é amplificada por PCR com iniciadores complementares às regiões 5' e 3'. O iniciador 5' pode incorporar locais para enzimas de restrição flanqueadores (seleccionados). O produto é então digerido com essas enzimas de restrição seleccionadas e subclonado no vector de expressão.

Gera-se o baculovírus recombinante por co-transfecção do plasmídeo anterior e ADN de vírus BaculoGold™ (Pharmingen) em células de *Spodoptera frugiperda* ("Sf9") (ATCC CRL 1711) utilizando Lipofectine (disponível comercialmente de GIBCO-BRL). Após 4-5 dias de incubação a 28°C, recolhem-se os vírus libertados e utilizam-se para amplificações adicionais. A

infecção viral e a expressão da proteína são realizadas tal como descrito por O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

O PRO marcado com poli-His expresso pode então ser purificado, por exemplo, por cromatografia de afinidade em Ni^{2+} -quelato, tal como se segue. Preparam-se os extractos a partir de células Sf9 infectadas com vírus recombinante, tal como descrito por Rupert et al., Nature, 362:175-179 (1993). Sucintamente, lavam-se as células Sf9, ressuspendem-se em tampão de ultra-sons (25 ml de Hepes, pH 7,9; MgCl_2 12,5 mM; EDTA 0,1 mM; glicerol a 10%; NP-40 a 0,1%; KCl 0,4 M) e tratam-se duas vezes com ultra-sons durante 20 segundos em gelo. Clarificam-se os resultantes por centrifugação e dilui-se o sobrenadante 50 vezes em tampão de carga (fosfato 50 mM, NaCl 300 mM, glicerol a 10%, pH 7,8) e filtra-se através de um filtro de 0,45 μm . Prepara-se uma coluna de Ni^{2+} -NTA agarose (disponível comercialmente da Qiagen) com um volume de leito de 5 ml, lava-se com 25 ml de água e equilibra-se com 25 ml de tampão de carga. Carrega-se o extracto celular filtrado na coluna a 0,5 ml por minuto. Lava-se a coluna até uma linha de base de A_{280} com tampão de carga, ponto em que se inicia a recolha de fracções. Seguidamente, lava-se a coluna com um tampão de lavagem secundário (fosfato 50 mM; NaCl 300 mM, glicerol a 10%, pH 6,0), que elui a proteína ligada não especificamente. Após atingir novamente a linha de base de A_{280} , desenvolve-se a coluna com um gradiente de imidazole de 0 a 500 mM no tampão de lavagem secundário. Recolhem-se fracções de um ml e analisam-se por SDS-PAGE e coloração com prata ou *Western blot* com Ni^{2+} -NTA conjugado com fosfatase alcalina (Qiagen). Reúnem-se as fracções contendo o PRO marcado com His10 eluído e dialisam-se contra tampão de carga.

Alternativamente, a purificação do PRO marcado com IgG (ou marcado com Fc) pode ser realizada utilizando técnicas de cromatografia conhecidas, incluindo por exemplo, cromatografia em coluna de Proteína A ou proteína G.

Muitos dos polipéptidos PRO aqui descritos foram expressos com sucesso como acima descrito.

EXEMPLOS 11

Preparação de Anticorpos Que Ligam PRO

Este exemplo ilustra a preparação de anticorpos monoclonais que podem ligar especificamente PRO.

São conhecidas na especialidade técnicas para a produção de anticorpos monoclonais e são descritas, por exemplo, em Goding, *supra*. Os imunogénios que podem ser empregues incluem as proteínas de fusão de PRO purificadas contendo PRO e células que expressam PRO recombinante na superfície celular. A selecção do imunogénio pode ser realizada pelo perito na especialidade sem experiências desnecessárias.

Imunizam-se ratinhos, tais como Balb/c, com o imunogénio de PRO emulsionado em adjuvante completo de Freund que se injecta subcutânea ou intraperitonealmente numa quantidade de 1-100 microgramas. Alternativamente, emulsiona-se o imunogénio em adjuvante MPL-TDM (Ribi Immunochemical Research, Hamilton, MT) e injecta-se nas almofadas da pata posterior do animal. Os ratinhos imunizados recebem então um reforço 10 a 12 dias mais tarde com imunogénio adicional emulsionado no adjuvante seleccionado. Seguidamente, durante várias semanas, os ratinhos podem também receber reforços com injeções de imunização adicionais. Podem obter-se amostras de soro dos ratinhos periodicamente por sangramento retro-orbital para testar em ensaios ELISA para detectar anticorpos anti-PRO.

Depois de se detectar um título de anticorpo adequado, os animais "positivos" para anticorpos podem ser injectados com uma injeção intravenosa final de PRO. Três a quatro dias mais tarde, sacrificam-se os ratinhos e recolhem-se células do baço. As células do baço são então fundidas (utilizando polietilenoglicol a 35%) com uma linha celular de mieloma murino seleccionada tal como P3X63AgU.1, disponível na ATCC, com o número CRL 1597. As fusões geram células de hibridoma que podem ser então plaqueadas em placas de cultura de tecidos de 96 poços contendo meio HAT (hipoxantina, aminopterina e timidina) para inibir a proliferação de células não fundidas, híbridos de mieloma e híbridos de células de baço.

As células de hibridoma serão rastreadas num ELISA quanto a reactividade contra o PRO. A determinação de células de hibridoma "positivas" que excretam os anticorpos monoclonais contra PRO desejados faz parte da perícia na especialidade.

As células de hibridoma positivas podem ser injectadas intraperitonealmente em ratinhos Balb/c singénicos para produzir ascites contendo os anticorpos monoclonais anti-PRO. Alternativamente, as células de hibridoma podem ser crescidas em balões de cultura de tecidos ou frascos rotativos. A purificação dos anticorpos monoclonais produzidos nas ascites pode ser efectuada utilizando precipitação com sulfato de amónio, seguida de cromatografia de exclusão em gel. Alternativamente, pode ser empregue cromatografia de afinidade baseada na ligação do anticorpo a proteína A ou proteína G.

EXEMPLO 12

Purificação de Polipéptidos PRO Utilizando Anticorpos Específicos

Os polipéptidos PRO nativos ou recombinantes podem ser purificados através de uma variedade de técnicas *standard* na especialidade da purificação de proteínas. Por exemplo, pro-polipéptido PRO, polipéptido PRO maduro ou pre-polipéptido PRO, são purificados por cromatografia de imunoafinidade utilizando anticorpos específicos para o polipéptido PRO de interesse. Em geral, uma coluna de imunoafinidade é construída por acoplamento covalente do anticorpo anti-polipéptido PRO com uma resina cromatográfica activada.

As imunoglobulinas policlonais são preparadas a partir de soros imunes por precipitação com sulfato de amónio ou por purificação sobre Proteína A imobilizada (Pharmacia LKB Biotechnology, Piscataway, N.J.). Do mesmo modo, os anticorpos monoclonais são preparados a partir de fluido ascítico de ratinhos por precipitação com sulfato de amónio ou cromatografia sobre Proteína A imobilizada. A imunoglobulina parcialmente purificada é ligada covalentemente a uma resina cromatográfica tal como SEPHAROSE[™] activada por CnBr (Pharmacia LKB Biotechnology). O anticorpo é acoplado à

resina, a resina é bloqueada e a resina derivatizada é lavada de acordo com as instruções do fabricante.

Esta coluna de imunoafinidade é utilizada na purificação do polipéptido PRO preparando uma fracção a partir de células contendo o polipéptido PRO numa forma solúvel. Esta preparação é derivada por solubilização das células completas ou de uma sua fracção subcelular obtida por centrifugação diferencial pela adição de detergente ou por outros métodos bem conhecidos na especialidade. Alternativamente, o polipéptido PRO solúvel contendo uma sequência de sinal pode ser excretado numa quantidade útil para o meio em que as células foram crescidas.

Uma preparação contendo polipéptido PRO solúvel é passada através da coluna de imunoafinidade e a coluna é lavada sob condições que permitem a absorvência preferencial do polipéptido PRO (e.g., tampões de elevada força iónica na presença de detergente). Depois, a coluna é eluída sob condições que rompem a ligação anticorpo/polipéptido PRO (e.g., um tampão de baixo pH tal como aproximadamente pH 2-3, ou uma elevada concentração de um caotrope tal como ureia ou ião tiocianato), e o polipéptido PRO é recolhido.

EXEMPLO 13

Rastreio de Fármacos

A presente invenção é particularmente útil para o rastreio de compostos utilizando polipéptidos PRO ou um seu fragmento de ligação em qualquer de uma variedade de técnicas de rastreio de fármacos. O polipéptido PRO ou o fragmento empregues neste teste podem estar livres em solução; estar fixados a um suporte sólido, estar suportados por uma superfície celular ou estar localizados intracelularmente. Um método de rastreio de fármacos utiliza células hospedeiras eucariotas ou procariotas que são estavelmente transformadas com ácidos nucleicos recombinantes que expressam o polipéptido PRO ou o fragmento. Os fármacos são rastreados contra estas células transformadas em ensaios de ligação competitiva. Estas células, quer na forma viável quer na forma fixada, podem ser utilizadas para ensaios de ligação *standard*. Pode medir-se, por exemplo, a formação de complexos entre um polipéptido PRO ou um fragmento e o agente a testar.

Alternativamente, pode examinar-se a diminuição da formação de complexo entre o polipéptido PRO e a sua célula alvo ou receptores alvo, causada pelo agente a testar.

Assim, a presente invenção proporciona métodos de rastreio de fármacos ou quaisquer outros agentes que possam afectar uma doença ou desordem associadas a um polipéptido PRO. Estes métodos compreendem o contacto de um destes agentes com um polipéptido PRO ou um seu fragmento e o ensaio (i) quanto à presença de um complexo entre o agente e o polipéptido PRO ou o fragmento, ou (ii) quanto à presença de um complexo entre o polipéptido PRO ou o fragmento e a célula, através de métodos bem conhecidos na especialidade. Nestes ensaios de ligação competitiva, o polipéptido PRO ou o fragmento são, tipicamente, marcados. Após incubação adequada, o polipéptido PRO ou fragmento livres são separados dos que estão presentes na forma ligada, e a quantidade de marcador livre ou não complexado constitui uma medida da capacidade do agente particular se ligar ao polipéptido PRO ou interferir com o complexo polipéptido PRO/célula.

Outra técnica para o rastreio de fármacos proporciona um rastreio de elevado rendimento para compostos possuindo afinidade de ligação adequada para com um polipéptido e está descrita com detalhes em WO 84/03564, publicado em 13 de Setembro, 1984. Sucintamente, grandes números de diferentes compostos de teste peptídicos pequenos são sintetizados sobre um substrato sólido, tal como pinos de plástico ou uma outra superfície. Conforme aplicado a um polipéptido PRO, os compostos de teste peptídicos são feitos reagir com o polipéptido PRO e lavados. O polipéptido PRO ligado é detectado por métodos bem conhecidos na especialidade. O polipéptido PRO purificado pode também revestir directamente placas para utilização no ensaio de rastreio de fármacos supramencionado. Em adição, anticorpos não neutralizantes podem ser utilizados para capturar o péptido e imobilizá-lo sobre o suporte sólido.

A presente invenção contempla também a utilização de ensaios de rastreio de fármacos competitivos em que anticorpos neutralizantes capazes de ligar um polipéptido PRO, competem especificamente com um composto de teste pela ligação ao

polipéptido PRO ou seus fragmentos. Desta maneira, os anticorpos podem ser utilizados para detectar a presença de qualquer péptido que partilhe um ou mais determinantes antigénicos com um polipéptido PRO.

EXEMPLO 14

Desenho Racional de Fármacos

O objectivo do desenho racional de fármacos é produzir análogos estruturais de um polipéptido biologicamente activo de interesse (*i.e.*, um polipéptido PRO) ou de moléculas pequenas com as quais aqueles interagem, *e.g.*, agonistas, antagonistas ou inibidores. Quaisquer destes exemplos podem ser utilizados para desenhar fármacos que são formas mais activas ou estáveis do polipéptido PRO ou que aumentam ou interferem com a função do polipéptido PRO *in vivo* (*c.f.*, Hodgson, Bio/Technology 9 : 19-21 (1991)).

Numa abordagem, a estrutura tridimensional do polipéptido PRO ou de um complexo polipéptido PRO-inibidor, é determinada por cristalografia de raios-x, por modelação em computador ou, mais tipicamente, através de uma combinação das duas abordagens. Tanto a forma como as cargas do polipéptido PRO têm que ser determinadas para elucidar a estrutura e para determinar local(ais) activos(s) da molécula. Menos frequentemente, a informação útil relativamente à estrutura do polipéptido PRO pode ser obtida por modelação com base na estrutura de proteínas homólogas. Em ambos os casos, a informação estrutural relevante é utilizada para desenhar moléculas semelhantes a polipéptidos PRO análogas ou para identificar inibidores eficientes. Os exemplos úteis de desenho racional de fármacos podem incluir moléculas que possuem actividade ou estabilidade melhoradas como mostrado por Braxton and Wells, Biochemistry, 31-7796-7801 (1992) ou que actuam como agonistas inibidores, ou antagonistas de péptidos nativos como mostrado por Athauda *et al.*, J. Biochem., 113: 742-746 (1993).

É igualmente possível isolar um anticorpo específico do alvo, seleccionado por ensaio funcional, como descrito acima, e depois resolver a sua estrutura cristalina. Esta abordagem, em princípio, origina um produto nuclear com potencial

farmacêutico (*pharmacore*) no qual pode ser baseado o subsequente desenho de fármacos. É possível ultrapassar a cristalografia da proteína gerando anticorpos anti-idiotípicos (anti-id) contra um anticorpo farmacologicamente activo, funcional. Como uma imagem no espelho de uma imagem no espelho, será de esperar que o local de ligação dos anti-id seja um análogo do receptor original. O anti-id pode então ser utilizado para identificar e isolar péptidos de bancos de péptidos produzidos quimicamente ou biologicamente, e os péptidos isolados actuarão então como *pharmacore*.

Em virtude da presente invenção, podem ser disponibilizadas quantidades suficientes do polipéptido PRO para realizar estudos analíticos como cristalografia de raios-x. Em adição, o conhecimento da sequência de aminoácidos do polipéptido PRO aqui proporcionado será uma orientação para quem empregue técnicas de modelação em computador em vez de, ou em adição a, cristalografia de raios-x.

EXEMPLO 15

Ensaio Antitumoral *In Vitro*

A actividade antiproliferativa do polipéptido foi determinada no ensaio investigacional, orientado para doenças, de descoberta, *in vitro*, de descoberta de fármacos anticancerosos do National Cancer Institute (NCI), utilizando um ensaio de ligação ao corante sulforrodamina B (SRB) essencialmente como descrito por Skehan et al., J. Natl. Cancer Inst., 82 1107-1112 (1990). As 60 linhas de células tumorais empregues neste estudo ("o painel NCI"), bem como as condições para a sua manutenção e cultura *in vitro* foram descritos por Monks et al., J. Natl. Cancer Inst., 83:757-766 (1991). A finalidade deste rastreio consiste em inicialmente avaliar a actividade citotóxica e/ou citostática dos compostos de teste contra diferentes tipos de tumores (Monks et al., *supra*: Boyd, Cancer: Prine. Pract. Oncol. Update, 3(10):1-12 [1989]).

Células de aproximadamente 60 linhas de células tumorais humanas foram colhidas com tripsina/EDTA (Gibco), lavadas uma vez, ressuspensas em IMEM e a sua viabilidade foi determinada. As suspensões celulares foram adicionadas através de uma

pipeta (volume de 100 µl) a placas de microtítulo de 96 poços separadas. A densidade celular da incubação de 6 dias foi inferior à da incubação de 2 dias para impedir o sobrecrecimento. Permitiu-se aos inoculados um período de pré-incubação de 24 horas a 37°C para estabilização. Adicionaram-se diluições de duas vezes a concentração de teste pretendida no tempo zero em alíquotas de 100 µl, aos poços das placas de microtítulo (diluições de 1:2). Avaliaram-se os compostos de teste em diluições de cinco semi-log (de 1000 para 100 000 vezes). As incubações ocorreram durante dois dias e seis dias num atmosfera de CO₂ a 5% e 100% de humidade.

Após incubação, removeu-se o meio e fixaram-se as células em 0,1 ml de ácido tricloroacético a 10% a 40°C. Enxaguaram-se as placas cinco vezes com água desionizada, secaram-se e coraram-se durante 30 minutos com 0,1 ml de corante sulforrodamina B a 0,4% (Sigma) dissolvido em ácido acético a 1%, enxaguou-se quatro vezes com ácido acético a 1% para remover corante não ligado, secou-se, e extractou-se o corante durante cinco minutos com 0,1 ml de base Tris 10 mM [tris(hidroximetil)aminometano], pH 10,5. A absorvância (DO) da sulforrodamina B a 492 nm foi medida utilizando um leitor de placas de microtítulo de 96 poços em interface com um computador.

Uma amostra de teste é considerada positiva se apresentar pelo menos 40% de efeito inibidor do crescimento numa ou mais concentrações. Os resultados estão apresentados na Tabela 7 que se segue, onde as abreviaturas do tipo de célula tumoral são as seguintes: NSCL= carcinoma do pulmão de células não pequenas; SNC= sistema nervoso central

Tabela 7

Composto de teste	Dias	Tipo de Linha Celular Tumoral	Designação da Linha Celular
PRO4400	N/A	NSCL	HOP-92; NC1-H226
PRO4400	N/A	NSCL	A549/ATCC; EKVX; HOP-62
PRO4400	N/A	NSCL	NCI-H23; NCI-H322
PRO4400	N/A	NSCL	NCI-H522
PRO4400	N/A	Cancro do Cólon	HCC-2998; HCT-15; HT29
PRO4400	N/A	Cancro do Cólon	KM12; SW620

Composto de teste	Dias	Tipo de Linha Celular Tumoral	Designação da Linha Celular
PRO4400	N/A	Cancro da Mama	HS 578T; BT-549; MCF7
PRO4400	N/A	Cancro da Mama	MDA-MB-231/ATCC
PRO4400	N/A	Cancro da Mama	NCI/ADR-RES
PRO4400	N/A	Cancro Ovariano	IGROVI*; OVCAR-3
PRO4400	N/A	Cancro Ovariano	OVCAR-4; OVCAR-5
PRO4400	N/A	Cancro Ovariano	OVCAR-8
PRO4400	N/A	Leucemia	CCRF-CEM; HL-60(TB); SR
PRO4400	N/A	Leucemia	RPMI-8226; K-562; MOLT-4
PRO4400	N/A	Cancro da Próstata	PC-3
PRO4400	N/A	Melanoma	MALME-3M; M 14; UACC-257
PRO4400	N/A	Cancro Renal	A498; ACHN; CAKI-I
PRO4400	N/A	Cancro Renal	RXF 393; SN 12C; TK-10
PRO4400	N/A	Cancro Renal	786-0
PRO4400	N/A	Cancro do SNC	SF-268; SNB-19; U251

* = CITOTÓXICA

Depósito de material

Os seguintes materiais foram depositados na American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, EUA (ATCC):

Material	N.º de depósito ATCC	Data de depósito
DNA87974-2609	203963	27 de Abril, 1999

Estes depósitos foram efectuados segundo os termos do "Tratado de Budapeste sobre o Reconhecimento Internacional do Depósito de Microorganismos para Efeitos do Procedimento em Matéria de Patentes" e seus Regulamentos (Tratado de Budapeste). Tal assegura a manutenção de uma cultura viável do depósito durante 30 anos a partir da data de depósito. O depósito será disponibilizado pela ATCC segundo os termos do Tratado de Budapeste e sujeito a um acordo entre a Genentech, Inc. e a ATCC, que assegura a disponibilidade permanente e não restrita da progénie da cultura do depósito

ao público após publicação da patente norte-americana pertinente ou após disponíveis ao público quaisquer pedidos de patente norte-americana ou estrangeira, o que ocorrer primeiro, e assegura a disponibilidade da progénie a quem determinado pelo U.S. Commissioner of Patents and Trademarks com esse direito de acordo com 35 U.S.C. §122 e as regras do Commissioner daí decorrentes (incluindo 37 C.F.R. §1.14 com particular referência a 886 OG 638).

O cessionário do presente pedido de patente concordou que, caso uma cultura dos materiais depositados morra ou se perca ou seja destruída quando cultivada em condições adequadas, os materiais serão prontamente substituídos após notificação, por outros equivalentes. A disponibilidade do material depositado não deve ser entendida como uma permissão para a prática da invenção em contravenção dos direitos concedidos pela autoridade de qualquer governo de acordo com as suas leis relativas a patentes.

O fascículo aqui escrito é considerado suficiente para permitir que um perito na especialidade ponha em prática a invenção. A presente invenção não está limitada no seu âmbito pela construção depositada, dado que a concretização depositada se destina unicamente a ilustração de certos aspectos da invenção e quaisquer construções que sejam funcionalmente equivalentes estão no âmbito da presente invenção. O depósito de material aqui não constitui uma admissão de que a descrição escrita aqui contida é inadequada para permitir a prática de qualquer aspecto da invenção, incluindo o seu melhor modo, nem deve ser interpretada como limitante do âmbito das reivindicações às ilustrações específicas que representa. De facto, várias modificações da invenção, para além das que são aqui apresentadas e descritas, serão evidentes para um perito na especialidade a partir da descrição anterior e caem no âmbito das reivindicações anexas.

Listagem de Sequências

<110> Genentech, Inc.

<120> MÉTODOS E COMPOSIÇÕES PARA INIBIÇÃO DO CRESCIMENTO DE CÉLULAS
NEOPLÁSICAS

<130> CMD/FP5967989

<140> EP 00941164.6

<141> 2000-05-30

<150> PCT/US99/12252

<151> 1999-06-02

<150> US 60/140,650

<151> 1999-06-22

<150> US 60/141,037

<151> 1999-06-23

<150> US 60/144,758

<151> 1999-07-20

<150> PCT/US99/20111

<151> 1999-09-01

<150> PCT/US99/20594

<151> 1999-09-08

<150> US 60/162,506

<151> 1999-10-29

<150> PCT/US99/28313

<151> 1999-11-30

<150> PCT/US99/28634

<151> 1999-12-01

<150> PCT/US99/28551

<151> 1999-12-02

<150> US 60/170,262

<151> 1999-12-09

<150> PCT/US99/30095

<151> 1999-12-16

<150> PCT/US99/30999

<151> 1999-12-20

<150> PCT/US00/00376

<151> 2000-01-06

<150> PCT/US00/03565

<151> 2000-02-11

<150> PCT/US00/04341

<151> 2000-02-18

<150> PCT/US00/04342

<151> 2000-02-18

<150> PCT/US00/05841

<151> 2000-03-02

<150> US 60/187,202

<151> 2000-03-03

<150> PCT/US00/06319

<151> 2000-03-10

<150> PCT/US00/06884

<151> 2000-03-15

<150> PCT/US00/08439

<151> 2000-03-30

<150> PCT/US00/13705

<151> 2000-05-17

<160> 5

<210> 1

<211> 1840

<212> ADN

<213> *Homo sapiens*

<400> 1

```
gtgggccc cctgctgctg cgtccatgc tgatgtttgc ggtgatcgtg 50
gcctccagcg ggctgctgct catgatcgag cggggcatcc tggccgagat 100
gaagcccctg cccctgcacc cggccggccg cgagggcaca gcctggcgcg 150
ggaaagcccc caagcctggg ggctgtccc tcagggctgg ggacgcggac 200
ttgcaagtgc ggcaggacgt ccggaacagg accctgcggg cgggtgtcgg 250
acagccaggc atgccccggg acccctggga cttgccggtg gggcagcggc 300
gcaccctgct gcgccacatc ctgtaagtg accgttaccg cttcctctac 350
tgctacgtcc ccaagggtgg ctgctctaac tggaagcggg tgatgaaggt 400
gctggcaggc gtcctggaca gcgtggacgt ccgcctcaag atggaccacc 450
gcagtgaacct ggtgttcctg gccgacctgc ggcctgagga gattcgctac 500
cgctgcagc actactttaa gttcctgttt gtgcgggagc ccttggaacg 550
cctcctctct gcctaccgca acaagtttgg cgagatccga gagtaccagc 600
aacgctatgg ggctgagata gtgaggcggg acagggctgg agcggggccc 650
agccctgcag gcgacgatgt cacattcccc gagttcctga gatacctggt 700
ggatgaggac cctgagcgca tgaatgagca ttggatgccc gtgtaccacc 750
```

tgtgccagcc ttgtgccgtg cactatgact ttgtgggctc ctatgagagg 800
 ctggaggctg atgcaaatca ggtgctggag tgggtacggg caccacctca 850
 cgtccgattt ccagctcgcc aggcctggta ccggccagcc agccccgaaa 900
 gctgcatta ccacttgtgc agtcccccc gggccctgct gcaggatgtg 950
 ctgcctaagt atatcctgga cttctccctc tttgctacc cactgcctaa 1000
 tgtcaccaag gaggcggtgc agcagtgacc atgggtgtgg ggccagcagc 1050
 tgggtggggac tggtttcaac gccagcttctc tgtgcttctg cctgtcattc 1100
 ggagaaactc tggctctggg gcttggggct tctcaggatc ctggatggca 1150
 gagactgccc tcagaagttc cttgtccagg gtgggcaccc acagtgactc 1200
 agaggacagg gctaggcagg agacctgctg ctccctcattg gggggatctc 1250
 ttggggggca gacaccagtt tgccaatgaa gcaacacatc tgatctaaag 1300
 actggctcca gaccccgggc tgccaggatt atgcagtcca cttggtctac 1350
 cttaatthaa cctgtggcca aactcagaga tgggtaccagc caggggcaag 1400
 catgaccaga gccagggacc ctgtggctct gatcccccatt tatccaccc 1450
 catgtgcctc aggactagag tgagcaatca taccttataa atgacttttg 1500
 tgcctttctg ctccagtctc aaaatttctc acacctgcca gttctttaca 1550
 tttttccaag gaaaggaaaa cggaagcagg gttcttgctt ggtagctcca 1600
 ggaccagct ctgcaggcac ccaaagacc tctgtgcca gcctcttct 1650
 tgagtctctg gaacctctc cctaattctc ccttctctcc ccacaaggcc 1700
 tttgaggttg tgactgtggc tggatatctt ggctgccatt tttctgatgc 1750
 attttatthaa aatttgact ttttgataga acccttgtaa gggctttgtt 1800
 ttcctaatag ctgacttttt aataaagcag ttttatatat 1840

<210> 2

<211> 333

<212> PRT

<213> *Homo sapiens*

<400> 2

Met	Leu	Met	Phe	Ala	Val	Ile	Val	Ala	Ser	Ser	Gly	Leu	Leu	Leu
1				5					10				15	
Met	Ile	Glu	Arg	Gly	Ile	Leu	Ala	Glu	Met	Lys	Pro	Leu	Pro	Leu
				20					25				30	
His	Pro	Pro	Gly	Arg	Glu	Gly	Thr	Ala	Trp	Arg	Gly	Lys	Ala	Pro
				35					40				45	
Lys	Pro	Gly	Gly	Leu	Ser	Leu	Arg	Ala	Gly	Asp	Ala	Asp	Leu	Gln
				50					55				60	

Val Arg Gln Asp	Val Arg Asn Arg Thr	Leu Arg Ala Val Cys Gly
65	70	75
Gln Pro Gly Met	Pro Arg Asp Pro Trp	Asp Leu Pro Val Gly Gln
80	85	90
Arg Arg Thr Leu	Leu Arg His Ile Leu	Val Ser Asp Arg Tyr Arg
95	100	105
Phe Leu Tyr Cys	Tyr Val Pro Lys Val	Ala Cys Ser Asn Trp Lys
110	115	120
Arg Val Met Lys	Val Leu Ala Gly Val	Leu Asp Ser Val Asp Val
125	130	135
Arg Leu Lys Met	Asp His Arg Ser Asp	Leu Val Phe Leu Ala Asp
140	145	150
Leu Arg Pro Glu	Glu Ile Arg Tyr Arg	Leu Gln His Tyr Phe Lys
155	160	165
Phe Leu Phe Val	Arg Glu Pro Leu Glu	Arg Leu Leu Ser Ala Tyr
170	175	180
Arg Asn Lys Phe	Gly Glu Ile Arg Glu	Tyr Gln Gln Arg Tyr Gly
185	190	195
Ala Glu Ile Val	Arg Arg Tyr Arg Ala	Gly Ala Gly Pro Ser Pro
200	205	210
Ala Gly Asp Asp	Val Thr Phe Pro Glu	Phe Leu Arg Tyr Leu Val
215	220	225
Asp Glu Asp Pro	Glu Arg Met Asn Glu	His Trp Met Pro Val Tyr
230	235	240
His Leu Cys Gln	Pro Cys Ala Val His	Tyr Asp Phe Val Gly Ser
245	250	255
Tyr Glu Arg Leu	Glu Ala Asp Ala Asn	Gln Val Leu Glu Trp Val
260	265	270
Arg Ala Pro Pro	His Val Arg Phe Pro	Ala Arg Gln Ala Trp Tyr
275	280	285
Arg Pro Ala Ser	Pro Glu Ser Leu His	Tyr His Leu Cys Ser Ala
290	295	300
Pro Arg Ala Leu	Leu Gln Asp Val Leu	Pro Lys Tyr Ile Leu Asp
305	310	315
Phe Ser Leu Phe	Ala Tyr Pro Leu Pro	Asn Val Thr Lys Glu Ala
320	325	330

Cys Gln Gln

<210> 3

<211> 22

<212> ADN

<213> Sequência Artificial

<220>

<223> Sonda oligonucleotídica sintética.

<400> 3

gctgctgccg tccatgctga tg 22

<210> 4
<211> 23
<212> ADN
<213> Sequência Artificial

<220>
<223> Sonda oligonucleotídica sintética.

<400> 4
ctcggggaat gtgacatcgt cgc 23

<210> 5
<211> 35
<212> ADN
<213> Sequência Artificial

<220>
<223> Sonda oligonucleotídica sintética.

<400> 5
gctgccgtcc atgctgatgt ttgcggtgat cgtgg 35

Lisboa,

REIVINDICAÇÕES

1. Polipéptido isolado para utilização num método de tratamento médico, o polipéptido possuindo:

- (i) a sequência de aminoácidos apresentada na Fig. 2 (SEQ ID NO:2);
- (ii) a sequência de aminoácidos codificada pelo ADNc de comprimento completo depositado na ATCC com o número de acesso 203963;
- (iii) pelo menos 80% de identidade de sequência de aminoácidos com o polipéptido de (i) ou (ii), em que a identidade de sequências é determinada utilizando o programa de computador ALIGN-2, e em que o polipéptido é capaz de inibir a proliferação de células tumorais.

2. Polipéptido de acordo com a reivindicação 1, para utilização na inibição de crescimento de células neoplásicas.

3. Polipéptido de acordo com a reivindicação 2, para utilização no tratamento de tumores.

4. Polipéptido de acordo com a reivindicação 3, em que o tumor é um cancro.

5. Polipéptido de acordo com a reivindicação 4, em que o cancro é cancro da mama, cancro ovariano, cancro renal, cancro colorrectal, cancro uterino, cancro da próstata, cancro do pulmão, cancro da bexiga, cancro do sistema nervoso central, melanoma ou leucemia.

6. Ácido nucleico isolado que codifica um polipéptido tal como definido na reivindicação 1, para utilização num método de tratamento médico.

7. Ácido nucleico de acordo com a reivindicação 6, em que a utilização é como definida em qualquer uma das reivindicações 2 a 4.

8. Composição farmacêutica compreendendo um polipéptido de acordo com a reivindicação 1, ou ácido nucleico tal como

definido na reivindicação 6, em mistura com um transportador farmacêuticamente aceitável.

9. Composição de acordo com a reivindicação 8, compreendendo adicionalmente um outro agente inibidor do crescimento, um agente citotóxico ou um agente quimioterapêutico.

10. Composição de acordo com a reivindicação 8 ou a reivindicação 9, para utilização num método de tratamento.

11. Composição de acordo com a reivindicação 10, para utilização como definido em qualquer uma das reivindicações 2 a 5.

12. Utilização de um polipéptido tal como definido na reivindicação 1, ou de um ácido nucleico tal como definido na reivindicação 7, no fabrico de um medicamento para o tratamento de um tumor.

13. Utilização de acordo com a reivindicação 12, em que o tumor é como definido na reivindicação 4 ou na reivindicação 5.

14. Artigo de fabrico, compreendendo:

- (1) a composição farmacêutica de acordo com a reivindicação 8 ou a reivindicação 9;
- (2) um recipiente contendo a referida composição; e
- (3) um rótulo afixado no referido recipiente, ou um folheto incluído no referido recipiente, indicando a utilização da referida composição na inibição do crescimento de células neoplásicas.

Lisboa,

GTGGGCGGCCCCCTGCTGCTGCGCTCCATGCTGATGTTTGCGGTGATCGTGGCCCTCCAGCGGGCTGCTGCTC
ATGATCGAGCGGGGATCTCTGGCCGAGATGAAGCCCTGCCCCCTGCACCCCGCCGCGGAGGGGACAGC
CTGGCGCGGAAAGCCCCAAGCTTGGGGCGCTGTCTCTCAGGGCTGGGGACCGGACCTTCCAAGTCCGCGAC
GACGTCCGGAAACAGGACCTCGCGGGGGGTGTCGGACAGCCAGGCGATGCCCCGGACCCTTGGGAGTTGCC
GGTGGGCGACGGCGCACCTCTGCTCGGCCACATCTCTGTAAGTAGACCTTACCCTCTCTCTACTGCTTAC
TCCCACAAGGTGGCCTGCTCTAACTGGAACGGGTGATGAAGGTGCTGGCAGGCGCTCTGGACAGCGCTGGAC
TCCCGCTCAAGATGGACACCCGACCTGGTGTCTTCTGGCGACCTCGCGCTAGGAGATTCGCTA
CGGCTCGACCACTACTTAAGTTCTGTGTTGTGCGGGAGCCCTTGGAAACGCTCTCTCTGCTACCGCA
ACAAGTTTGGCGAGATCCAGAGATCCAGCAACGCTATGGGCTGAGTAGTGAGGCGGTACAGGCTTGA
CGGGGCGCAGCCTCGACGGCGACGATGTACATCCCCGAGTTCTGAGATACCTTGTGGATGAGGACCC
TGAGCGCATGAATGACATTGATGCCCGGTGACCACCTGTGCCAGCCTTGTGCCGTGCACTATGACTTTG
TGCGCTCTCATGAGAGCTGGAGGCTGATGCAAACTCAGGTGCTGGAGTGGGTACGGGACCACTCAGCTC
CGATTTTCAGCTCGCGAGGCTGTGATCGGGCAGCCGCCGAAGCCTGATACCACCTGTGCGATG
CCCCCGGCCCTGCTGCAGGATGTGCTGCCTAAGTATATCTTGGACTTCTCCCTCTTTGCCCTACCCACTGC
CTAATGTCAACAAGGAGCGGTGTCAGCATGACATGGGTGTGGGGCGACGCTGTTGGGACTGTTTTC
AACGCGAGCTTTCTGTCTTCTGCTCTGTATTCGGAGAACTCTGGCTCTGGGCTTCTCAGGA
TCTCTGATGCGCAGAGACTGCCCTCAGAAGTTCTTGTGTCAGGTTGGGCGACCCACAGTACCTCAGAGGACAG
GGCTCAGCAGGAGACTCTGCTCTCTCATTTGGGGGATCTCTTGGGGGGCAGACACAGTTTGCCAATGAA
CGAACACATCTGATCTAAGAACTGGCTCCAGACCCCGGCTGCCAGGATATGACGCTCACTTGGCTACC
TTAATTTAACTTGTGGCCAACTCAGAGATGGTACAGCCAGGGGCAAGCATGACAGAGCCAGGGACCTGT
GTGGCTCTGATCCCCCACTTCACTACCCCATGTGCCTCAGGACTAGATGAGCAATCATATCTTTATAAATG
ACTTTTGTGCTTTCTGCTCCAGTCACTCAAAATTTCTACACCTGCCAGTTCTTTACATTTTCCAAGGAA
GGAARACGGAAGCAGGGTTCTTGTCTGTTAGCTTCAGGACCGAGCTCTGACGAGCCAAAGACCTCTGT
GCCACGACTTCTCTTGAGTTTCTCGAACTCTCTCTTAATTTCTCCCTTCTTCCCAACAGGCTTTGAG
TTGTGCTGCTGGCTGGTATATCGGCTGCCATTTTCTGATGCACTTTATTTAAATTTGTGACTTTTGTAT
AGAACCCTTGAAGGCTTTGTTTCTTAATAGTGACTTTTAATAAAGCAGTTTATATAT

FIGURA 2

MLMFAVIVASSGLLLMIERGILAEKPLPLHPPGREGTAWRGKAPKPGGLSLRAGDADLQVRQDVNRNLT
AVCCQPGMPRDPWDLFVQQRRTLLRHILVSDRYRFLYCYVPKVACSNWKRVMKVLGVLDSVDVRLKMDHR
SDLVFLADLRPEEIRYRLQHYFKFLFVREPLERLLSAYRNKFGEIREYQORYGAEIVRRYRAGAGPSPAGD
DVTFFPEFLRYLVDEDPERMNEHWMPVYHLCQPCAVHYDFVGSYERLEADANQVLEWVRAPPHVRFPARQAW
YRPASPESLHYHLCAPRALLDVLPKYILDFSLFAYPLPNVTKEACQQ

Péptido de sinal:	Aminoácidos 1-23
Locais de N-glicosilação:	Aminoácidos 67-71;325-329
Locais de fosforilação de tirosina-quinase:	Aminoácidos 152-159;183-183
Locais de N-miristoilação:	Aminoácidos 89-95;128-134