

DESCRIÇÃO
DA
PATENTE DE INVENÇÃO

N.º 86 772

REQUERENTE: SCHERING BIOTECH CORPORATION, norte-americana,
com sede em 901 Califórnia Avenue, Palo Alto,
Califórnia 94304, Estados Unidos da América.

EPÍGRAFE: "PROCESSO PARA A PREPARAÇÃO DE INTERLEUQUINA-3
HUMANA E DE SUAS MUTEINAS".

INVENTORES: Takeshi Otsuka

Reivindicação do direito de prioridade ao abrigo do artigo 4.º da Convenção de Paris
de 20 de Março de 1883. Estados Unidos da América em 18 de
Fevereiro de 1987 sob o nº 016079

86-1112

MEMORIA DESCRITIVA

Resumo

Divulga-se um processo para a preparação de novas proteínas interleuquina-3 humanas contendo uma substituição não conservativa de aminoácidos (Ser → Pro na posição 7) relativamente a sequências já anteriormente divulgadas de IL-3 humana. Divulgam-se igualmente métodos para sintetizar muteínas neutras quanto à conformação e também antigenicamente. A proteína humana IL-3 favorece o crescimento e desenvolvimento de uma vasta gama de linhagens hematopoieticas e pode ser útil no tratamento de situações que

=====

SCHERING BIOTECH CORPORATION

"PROCESSO PARA A PREPARAÇÃO DE INTERLEUQUINA-3 HUMANA E DE SUAS MUTEINAS"

envolvam a regeneração de e/ ou populações de células sanguíneas enfraquecidas tais como as da hipoplasia mieloide, de infecções crônicas e de terapias utilizando transplantações de medula óssea.

Segundo o processo do invento é incorporado, nas referidas proteínas, um polipeptídeo glicosilado ou não glicosilado, 10 vezes substituído o qual é produzido pelo seguinte método: a) construção de um vector compreendendo a sequência ne nucleotídeos codificadora do referido polipeptídeo; b) incorporação do vector num hospedeiro; c) manutenção do hospedeiro numa meio de cultura em condições adequadas para expressar a referida sequência de nucleotídeos no referido polipeptídeo.

Processo para a preparação de interleuquinas-3 humanas e de suas muteínas

O invento diz respeito, de um modo geral, a mediadores proteicos de desenvolvimento do sistema hematopoiético e a ácidos nucleicos que codificam aqueles. Em particular, o invento diz respeito a uma variante da interleuquina-3 humana, a suas muteínas e aos ácidos nucleicos que as codificam.

As células sanguíneas em circulação estão a ser constantemente substituídas por novas células que se desenvolveram. As células sanguíneas de substituição formam-se num processo designado hematopoiesis que compreende a produção de, pelo menos, oito linhagens de células sanguíneas maturadas: células sanguíneas vermelhas (eritrócitos), macrófagos (monócitos), granulócitos eosinofílicos, megacariócitos (plaquetas), granulócitos neutrofílicos, granulócitos basofílicos (células másticas), T linfócitos e B linfócitos (Burgess and Nicola, Growth Factors and Stem Cells (Academic Press, New York, 1983)). Grande parte do controlo da formação de células sanguíneas é mediado por um grupo de glicoproteínas que interactuam, designados factores de estimulação de colónia ("colony stimulating factors", CSFs). Estas glicoproteínas são assim chamadas por causa dos ensaios in vivo e in vitro utilizados para detectar a sua presença. As técnicas de cultura clonal de células hematopoiéticas em meio de cultura semi-sólido foram especialmente importantes no desenvolvimento dos vários ensaios in vitro. Nessas culturas, células progenitor individuais (i.e. células destinadas, do ponto de vista do desenvolvimento, a uma linhagem particular, mas ainda assim capazes de proliferar) são capazes de proliferar a fim de formar uma colónia de descendentes em maturação, de um modo que se crê ser basicamente idêntico ao processo comparável que tem lugar in vivo. O papel dos CSFs na hematopoiesis foi objecto de vários trabalhos recentes, por ex. Metcalf,

The Hemopoietic Colony Stimulating Factors (Elsevier, New York, 1984); Metcalf, Science, Vol. 229, págs. 16-22 (1985); Nicola et al., Immunology Today, vol.5, págs. 76-80 (1984); e Whetton et al., TIBS, vol. 11, págs. 207-211 (1986).

A detecção, isolamento e purificação destes factores é extremamente difícil, uma vez que é frequentemente complicada pela natureza dos sobrenadantes em que geralmente se encontram, as divergências e cruzamentos de actividades dos vários componentes das misturas, a sensibilidade (ou ausência da mesma) dos ensaios utilizados para determinar as propriedades dos factores, a semelhança frequente de gamas de massas moleculares e de outras características dos factores, e a muito baixa concentração dos factores no seu ambiente natural.

A medida que se foram tornando disponíveis mais CSFs, sobretudo por intermédio de clonagem molecular, cresceu o interesse em encontrar para eles aplicações clínicas. As utilizações potenciais dos CSFs foram comparadas com as utilizações correntes de hormonas, devido às suas semelhanças fisiológicas (por ex. factores solúveis, mediadores do crescimento, acção por intermédio de receptores celulares), por ex. Dexter, Nature, Vol. 321, pág. 198 (1986). Foi sugerida a sua utilização em diversas situações em que é desejável a estimulação da produção de células sanguíneas, tal como em terapia reabilitadora após quimioterapia ou terapia por intermédio de radiações de tumores, tratamento de hipoplasias mielóides, tratamento de deficiência neutrofílica, tratamento para incrementar a regeneração hematopoiética após transplantação da medula óssea, e tratamento para aumentar a resistência do hospedeiro a infecções estabelecidas por ex. Dexter (atrás citado), e Metcalf, Science (atrás citado).

Ao longo dos últimos anos, foi identificada, purificada e clonada uma actividade de CSF murina, denominada multi-CSF ou interleuquina-3 (IL-3); ver

Yokota et al; Proc. Natl. Acad. Sci., Vol. 81 págs. 1070-1074 (1984); Fung et al., Nature, Vol. 307, págs. 233-237 (1984); Hapel et al., Blood, Vol. 65, págs. 1453-1459 (1985); Ihle et al., J. Immunol., Vol 129 págs 2431-2436 (1982); Ihle et al. J. Immunol., Vol 131, págs. 282-287 (1983); e Iscove et al. págs. 397-425 em Cellular and Molecular Biology of Lymphokines (Biologia Celular e Molecular de Linfoquinas) (Academic Press, Nova Yorque, 1985). A IL-3 murina exhibe um amplo espectro de actividades de CSF, sendo capaz de estimular a proliferação e o desenvolvimento das células progenitoras de todas as linhagens mieloides reconhecidas, por ex. Dexter, Nature, Vol 309, págs. 746-747 (1984); Ihle et al., Lymphokines, Vol. 9, págs 153-200 (1984). Devido aos amplos efeitos de estimulação do factor, tem sido grande o interesse em isolar os factores análogos noutros mamíferos, em particular no homem.

Recentemente, foram revelados homólogos de IL-3 de murganho relativos a rato, gibbon e ao homem por Cohem et al., Nucleic Acids Research, Vol. 14, págs. 3641-3658 (1986); e Yang et al., Cell, Vol. 47, págs. 3-10 (1986). As sequências de aminoácidos das moléculas referidas apresentam um grau de homologia em relação a IL-3 de murganho relativamente baixo: 54 por cento para o rato, 29 por cento para o gibbon e 29 por cento para o homem. Além disso, não se sabe ainda se o homólogo humano possui a mesma gama de actividades biológicas que o seu equivalente no murganho.

Tendo em conta que atrás ficou dito, a disponibilidade de variantes do homólogo humano de IL-3 de murganho com actividades biológicas confirmadas seria altamente vantajosa para efeitos de possíveis utilizações clínicas, em particular em situações em que a estimulação da geração de células sanguíneas poderia compensar hipoplasia mieloide induzida por doença, inverter os efeitos secundários perniciosos das terapias do cancro, incrementar a eficácia das transplantações de medula óssea ou outras situações aná-

logas.

O presente invento diz respeito a uma interleuquina-3 (IL-3) humana e as suas muteínas obtidas por manipulação genética. Inclui polipéptidos que compreendem uma sequência de aminoácidos seleccionada do conjunto adiante definido pela Fórmula I, e que possuem uma actividade de factor de estimulação de colónia (CSF) tal como é definida por ensaios convencionais. O invento inclui também ácidos nucleicos capazes de codificar os polipéptidos definidos pela Fórmula I, métodos de preparação desses polipéptidos utilizando os ácidos nucleicos do invento, e métodos de utilização dos polipéptidos do invento, quer isoladamente, quer em conjunto com outros CSFs a fim de tratar afecções envolvendo hipoplasia nucleíde ou para incrementar a regeneração do sistema hematopoiético após transplatação da medula óssea ou depois de certas terapias relacionadas com o cancro.

Um modelo de execução preferido do invento é o conjunto de polipéptidos glicosilados ou não glicosilados que compreende uma sequência de aminoácidos seleccionada do conjunto de sequências definido pela fórmula seguinte:

X(Pro) - X(Met) - X(Tre) - X(Gln) - X(Tre) - W(Tre) -
10
W(Pro) - W(Leu) - W(Lis) - X(Tre) - X(Ser) - X(Trp) -
X(Val) - X(Asn) - Cis - X(Ser) - W(Asn) - W(Met) -
20
W(Ile) - W(Asp) - W(Glu) - W(Ile) - W(Ile) - W(Tre) -
30
W(His) - W(Leu) - W(Lis) - W(Gln) - W(Pro) - W(Pro) -
X(Leu) - X(Pro) - X(Leu) - X(Leu) - X(Asp) - X(Fen) -
40
X(Asn) - X(Asn) - W(Leu) - W(Asn) - W(Gli) - W(Glu) -

W(Asp) - W(Gln) - W(Asp) - W(Ile) - W(Leu) - W(Met) -
 50
 W(Glu) - W(Asn) - W(Asn) - W(Leu) - W(Arg) - W(Arg) -
 60
 W(Pro) - W(Asn) - W(Leu) - W(Glu) - W(Ala) - W(Fen) -
 W(Asn) - W(Arg) - W(Ala) - W(Val) - W(L s) - W(Ser) -
 70
 W(Leu) - W(Gln) - W(Asn) - W(Ala) - W(Ser) - W(Ala) -
 W(Ile) - W(Glu) - W(Ser) - W(Ile) - W(Leu) - X(Lis) -
 80
 X(Asn) - X(Leu) - X(Leu) - X(Pro) - Cis - X(Leu) -
 90
 X(Pro) - X(Leu) - X(Ala) - X(Tre) - X(Ala) - W(Ala) -
 W(Pro) - W(Tre) - W(Arg) - W(His) - W(Pro) - W(Ile) -
 100
 W(His) - W(Ile) - W(Lis) - W(Asp) - W(Gli) - W(Asp) -
 W(Trp) - W(Asn) - W(Glu) - W(Fen) - W(Arg) - W(Arg) -
 110
 W(Lis) - W(Leu) - X(Tre) - X(Fen) - X(Tir) - X(Leu) -
 120
 X(Lis) - W(Tre) - W(Leu) - W(Glu) - W(Asn) - W(Ala) -
 W(Gln) - W(Ala) - W(Gln) - W(Gln) - X(Tre) - X(Tre) -
 130
 X(Leu) - X(Ser) - X(Leu) - X(Ala) - X(Ile) - X(Fen)

Fórmula I

na qual os termos W(Xaa) e X(Xaa) representam grupos de aminoácidos sinónimos do aminoácido Xaa localizados numa região hidrofílica ou hidrofóbica da proteína, respectivamente, tal como se determina pela análise de Hopp-Woods. Aminoácidos sinónimos no interior de um grupo possuem propriedades físico-químicas suficientemente semelhantes para que a substituição entre membros do grupo conserve a função biológica da molécula: Grantham, Science, vol.185, págs. 862-864 (1974); e Dayhoff et al., "A model of Avolutionary Change in Protein" (Um modelo de modificação evolutiva em proteínas), Atlas of Protein Sequence and Structure 1972 (Atlas de Sequências e Estruturas Proteicas 1972), vol. 5, págs. 89-99. É óbvio que é possível fazer inserções e deleções de aminoácidos na sequência atrás definida sem alterar a função biológica, em particular se as inserções e deleções envolverem apenas alguns aminoácidos, por ex. até dez, e não removerem ou deslocarem aminoácidos que são fundamentais para a conformação funcional, por ex. resíduos de cisteína: Anfinsen, "Principles That Govern The Folding of Protein Chains". (Princípios que governam a dobragem de cadeias proteicas), Science, Vol.181 págs. 223-230 (1973). As proteínas e muteínas produzidas por tais deleções e/ou inserções estão contidas no âmbito do presente invento. Onde quer que resíduos de aminoácidos da proteína de Fórmula I forem aqui referidos por meio de números, tais números são relativos à extremidade N da proteína.

Deve notar-se que o polipéptido IL-3 humano representado por aminoácidos Xaa na fórmula I (e aqui mais adiante indicado como fórmula II), contém uma substituição de aminoácido não conservativa de serina por prolina na posição 7 relativamente a sequências de IL-3 humana previamente referidas por Yang et al. em Cell, Vol. 47, págs. 3-10.

Os grupos aminoácidos sinônimos são, de preferência, aqueles que são definidos nos Quadros I e II. De modo mais preferido, os grupos de aminoácidos sinônimos são aqueles que são indicados antes do segundo traço em cada linha dos Quadros I e II; e, de modo preferido em primeiro lugar, os grupos de aminoácidos sinônimos são aqueles que são indicados antes do primeiro traço em cada linha nos Quadros I e II (cada grupo W(Xaa) preferido em primeiro lugar, consistindo apenas no próprio Xaa).

Quadro I. Grupos de aminoácidos sinónimos X(Xaa) preferidos, mais preferidos e preferidos em primeiro lugar

<u>Xaa</u>	<u>X(Xaa)</u>
Ser	Ser, // Tre, Gli, Asn, Cis
Arg	Arg, / Lis, His, / Gln, Glu
Leu	Leu, Ile, Met, / Fen/Val, Tir
Pro	Pro, / Ala, / Gli, Tre
Tre	Tre, // Pro, Ser, Ala, Gli, His, Gln
Ala	Ala, / Pro, / Gli, Tre
Val	Val, / Met, Ile, / Leu, Tir, Fen
Gli	Gli, // Ala, Tre, Pro, Ser
Ile	Ile, Met, Leu, / Fen, Val, / Tir
Fen	Fen, / Met, Tir, Ile, Leu, / Val, Trp
Tir	Tir, / Fen, / Trp, Met, Ile, Val, Leu
His	His, / Gln, Arg, / Glu, Lis, Tre
Gln	Gln, / Glu, His, Tre, Arg, Lis, Asn
Asn	Asn, / Asp, / Ser, Gln
Lis	Lis, / Arg, / Glu, Gln, His
Asp	Asp, / Asn, / Glu
Glu	Glu, / Gln, / His, Arg, Asp, Lis, Asn
Met	Met, Ile, Leu, / Fen, Val/
Trp	Trp//
Cis	Cis//

Quadro II Grupos de aminoácidos sinônimos W(Xaa) preferidos, mais preferidos e preferidos em primeiro lugar

<u>Xaa</u>	<u>W(Xaa)</u>
Ser	Ser, // Cis
Arg	Arg, // His, Lis
Leu	Leu, // Ile, Met, Fen
Pro	Pro, // Ala
Tre	Tre //
Ala	Ala, // Pro
Val	Val, // Met, Ile
Gli	Gli //
Ile	Ile, // Met, Leu, // Fen, Val
Fen	Fen, // Met, Tir, Ile, Leu
Tir	Tir, // Fen
His	His, // Gln, Arg
Gln	Gln, // Glu, His
Asn	Asn, // Asp
Lis	Lis, // Arg
Asp	Asp, // Asn
Glu	Glu, // Gln
Met	Met, // Ile, Leu, // Fen, Val
Trp	Trp //
Cis	Cis //

O invento inclui os polipéptidos de Fórmula I com substituições de aminoácidos (entre um aminoácido do polipéptido definido pela Fórmula II (a sequência nativa putativa) e um aminoácido sinônimo) numa única posição ou em várias posições. O termo "substituído N-vezes" é usado para descrever um subconjunto de polipéptidos definidos pela Fórmula I em que 0-N aminoácidos da sequência nativa foram substituídos por aminoácidos sinônimos. Assim, por exemplo, o grupo de polipéptidos de Fórmula I substituídos 1 vez consiste em 381 polipéptidos para os grupos de aminoácidos sinônimos preferidos, 245 para os grupos de aminoácidos sinônimos mais preferidos, e 38 para os grupos de aminoácidos sinônimos preferidos em primeiro lugar. Estes números são obtidos adicionando duas somas: (i) a soma do número de aminoácidos de cada espécie nas regiões hidrofóbicas da sequência nativa vezes a dimensão do grupo aminoácido sinônimo (grupo X) para esse aminoácido, e (ii) a soma do número de aminoácidos de cada espécie nas regiões hidrofílicas da sequência nativa vezes a dimensão do grupo aminoácido sinônimo (grupo W) para esse aminoácido.

De preferência, o grupo de polipéptidos 10 vezes substituídos de Fórmula I em que os aminoácidos das regiões definidas pelos resíduos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé, e 95-113 inclusivé não estão mais de 3 vezes substituídos. As sequências de aminoácidos nestas regiões estão mais altamente conservadas relativamente à restante molécula, quando comparadas com sequências de IL-3 no murganho, rato e gibão.

De modo mais preferido, o grupo de polipéptidos IL-3 humanos consiste nos polipéptidos 5 vezes substituídos de Fórmula I em que os aminoácidos nas regiões definidas pelos resíduos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusive e 95-113 inclusivé não estão substituídos mais de 1 vez. De modo ainda mais preferido, o grupo de polipéptidos IL-3 humanos consiste nos polipéptidos 2 ve-

zes substituídos de Formula I em que os aminoácidos nas regiões definidas pelos resíduos inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé não têm quaisquer substituições. De modo preferido em primeiro lugar, o polipéptido do invento compreende a sequência de aminoácidos nativa definida pela seguinte fórmula:

```

Pro - Met - Tre - Gln - Tre - Tre - Pro - Leu -
10
Lis - Tre - Ser - Trp - Val - Asn - Cis - Ser -
20
Asn - Met - Ile - Asp - Glu - Ile - Ile - Tre -
30
His - Leu - Lis - Gln - Pro - Pro - Leu - Pro -
40
Leu - Leu - Asp - Fen - Asn - Asn - Leu - Asn -
Glu - Glu - Asp - Gln - Asp - Ile - Leu - Met -
50
Glu - Asn - Asn - Leu - Arg - Arg - Pro - Asn -
60
Leu - Glu - Ala - Fen - Asn - Arg - Ala - Val -
70
Lis - Ser - Leu - Gln - Asn - Ala - Ser - Ala -
80
Ile - Glu - Ser - Ile - Leu - Lis - Asn - Leu -
Leu - Pro - Cis - Leu - Pro - Leu - Ala - Tre -
90
Ala - Ala - Pro - Tre - Arg - His - Pro - Ile -
100
His - Ile - Lis - Asp - Gli - Asp - Trp - Asn -
110
Glu - Fen - Arg - Arg - Lis - Leu - Tre - Fen -
120
Tir - Leu - Lis - Tre - Leu - Glu - Asn - Ala -
Gln - Ala - Gln - Gln - Tre - Tre - Leu - Ser -
130
Leu - Ala - Ile - Fen

```

Fórmula II

De modo análogo, o termo "inserido N-vezes" relativamente aos polipéptidos de Fórmula I é usado para descrever um conjunto de polipéptidos em que foram inseridos 1 a N aminoácidos na sequência definida pela fórmula I. De preferência, os aminoácidos inseridos são seleccionados dos grupos preferidos de aminoácidos sinónimos dos aminoácidos que flanqueiam a inserção de modo mais preferido, são seleccionados dos grupos mais preferidos de aminoácidos sinónimos dos aminoácidos que flanqueiam a inserção, e de modo preferido em primeiro lugar dos grupos preferidos em primeiro lugar de aminoácidos sinónimos dos aminoácidos que flanqueiam a inserção (ver Quadros I e II). Assim, por exemplo, um subgrupo do grupo de péptidos inseridos 1 vez compreende um aminoácido inserido entre o X(Pro) N-terminal e o X(Met) adjacente. As inserções que definem os membros deste subgrupo são seleccionadas, de preferência, entre Pro, Ala, Gli, Tre, Met, Fen, Ile, Val e Leu; com maior preferência são seleccionadas entre Ala, Pro, Met, Leu, Fen, Ile e Val, e preferido em primeiro lugar, não seleccionados entre Pro, Met, Ile e Leu. As inserções podem ser feitas entre quaisquer aminoácidos adjacentes de Fórmula I e também no início e no fim da sequência de aminoácidos. Uma vez que existem 133 possíveis locais de inserção, e uma vez que podem ser feitas no mesmo local inserções múltiplas, um péptido inserido 2 vezes de fórmula I dá origem a 17 689 subgrupos de péptidos, e a dimensão de cada subgrupo depende das dimensões dos grupos de aminoácidos sinónimos dos aminoácidos que flanqueiam as inserções.

O termo "eliminado N-vezes" relativamente aos polipéptidos de fórmula I é usado para descrever um conjunto de péptidos com 1 a N aminoácidos eliminados da sequência definida pela Fórmula I. Assim, o conjunto de polipéptidos de Fórmula I eliminados 1 vez consiste em 132 subgrupos de polipéptidos, cada qual com 131 aminoácidos (131

-ômeros). Cada um dos subgrupos consiste, por uma vez, em todos os 131-ômeros definidos pelos grupos de aminoácidos sinónimos preferidos, mais preferidos e preferidos em primeiro lugar, de acordo com a multiplicidade de substituições.

Quando quer que for definido um conjunto de polipéptidos em termos tanto de deleções relativamente à sequência de Fórmula I, e inserções relativamente à sequência de Fórmula I, forma-se primeiro o conjunto de polipéptidos definido pelas inserções e depois, para cada membro desse conjunto, forma-se um conjunto de polipéptidos definido por deleções. A soma dos membros de todos estes últimos conjuntos forma o conjunto de polipéptidos definidos por inserções e deleções.

Os polipéptidos de acordo com o invento podem fazer parte de um polipéptido mais longo que compreende um polipéptido adicional ligado à terminação N do polipéptido de acordo com o invento. O polipéptido adicional pode ser uma sequência sinal para dirigir a secreção da proteína do organismo hospedeiro, por ex. sequências sinal MF-alfa1 na levedura (Kurjan et al. Cell, Vol. 30, págs. 933-943 (1982)). Ou o péptido adicional pode ser usado para facilitar a purificação do polipéptido desejado, por ex. Sassenfeld et al. "A Polypeptide Fusion Designed for the Purification of Recombinant Proteins" (Uma fusão polipeptídica destinada à purificação de proteínas recombinantes), Biotechnology, págs 76-81 (Janeiro 1984). Neste caso, o péptido adicional contém um local de dissociação selectiva adjacente à terminação N do polipéptido desejado para separação ulterior da sequência de aminoácidos supérflua (relativamente à actividade biológica) por meios enzimáticos ou químicos.

O modelo de execução do invento preferido atrás mencionado inclui ainda sequências de nucleótidos capazes de codificarem os polipéptidos de Fórmula I dos grupos de aminoácidos sinónimos preferidos, mais preferidos

e preferidos em primeiro lugar. Em particular, o invento inclui ácidos nucleicos possuindo a sequência do fragmento de inserção de cDNA do clone pcD-huIL3-22-1 ilustrada na Figura 2. pcD-huIL3-22-1 foi depositado na American Type Culture Collection, Rockville, Maryland, sob o número de acesso 67318.

Em todo o texto são utilizadas abreviaturas usuais para designarem aminoácidos, nucleótidos, endonucleases de restrição, etc, por ex. Cohn, "Nomenclature and Symbolism of α -Amino Acids" (Nomenclatura e Simbologia de α -Aminoácidos), Methods in Enzymology (Métodos de Enzimologia), Vol 106, págs. 3-107 (1984); Wood et al. Biochemistry a Problems Approach (Bioquímica: uma abordagem problemática), 2ª ed. (Benjamin, Menlo Park, 1981); e Roberts, "Directory of Restriction Endonucleases" (Lista de endonucleases de restrição), Methods in Enzymology, Vol. 68, págs. 27-40 (1979).

O presente invento está indicado para problemas associados com a aplicação de agentes imunoreguladores para tratar afecções médicas e/ou veterinárias. Em particular, proporciona compostos capazes de estimular a regeneração de células sanguíneas e pode ser útil na terapia do cancro, no tratamento de certas doenças do sangue e no tratamento de infecções persistentes.

Para uma melhor compreensão do invento, serão agora descritos modelos de execução preferidos, que remetem para os diagramas anexos, nos quais:

a Figura 1 ilustra os locais de restrição e regiões de codificação mais importantes do plasmídeo pcDV1 usado na construção do vector de expressão e clonagem pcD;

a Figura 2 ilustra os locais de restrição e regiões de codificação mais importantes do plasmídeo pL1 usado na construção do vector de expressão e clonagem pcD;

a Figura 3 ilustra a sequência de nucleótidos e a sequência de aminoácidos prevista para um DNA de IL-3 humano de acordo com o invento; e

a Figura 4 ilustra o vector pCD modificado, pSP6, usado para gerar mRNA para ser injectado em oócitos.

O presente invento inclui polipéptidos glicosilados ou não glicosilados que apresentam actividade de IL-3 humana, e que podem ser derivados dos polipéptidos IL-3 aqui revelados usando técnicas de engenharia proteica usuais. O invento inclui também ácidos nucleicos possuindo sequências capazes de codificar os polipéptidos do invento. Finalmente, o invento inclui métodos de preparação dos polipéptidos glicosilados ou não glicosilados do invento que utilizam as sequências de nucleótidos aqui revelados, e métodos de utilização dos polipéptidos do invento.

As técnicas de preparação, utilização e identificação dos polipéptidos e ácidos nucleicos do invento são a seguir discutidos em termos genéricos. Em seguida, são apresentados vários exemplos em que as técnicas genéricas são empregadas utilizando-se tipos de células, vectores reagentes, etc. específicos.

I. Preparação de cDNAs de IL-3 humanas a partir de fontes naturais

cDNAs de acordo com o presente invento podem ser obtidos construindo uma sonda de cDNA a partir da sequência de cDNA revelada na Figura 2. A sonda é depois, usada para identificar um clone de cDNA complementar

numa biblioteca de cDNA. De preferência, a biblioteca de cDNA é constituída a partir de uma população de clones de células T induzidos por mitogénio ou linfócitos sanguíneos periféricos ("peripheral blood lymphocytes", PBLs). A construção de bibliotecas de cDNA a partir de células induzidas a crescer e/ou a sintetizar certos produtos celulares é uma técnica bem conhecida em biologia molecular: por ex., recentes artigos que passam em revista o assunto são os de Doherty, "Cloning and Expression of cDNA" (clonagem e expressão de cDNA), Capítulo 10 de Gottesman, ed. Molecular Cell Genetics (Genética Celular Molecular) (John Wiley & Sons, Nova Iorque, 1985); e Brandis et al. "Preparation of cDNA Libraries and the Detection of Specific Gene Sequences" (Preparação de bibliotecas de cDNA e a detecção de sequências de genes específicas) em Setlow et al., eds. Genetic Engineering (Engenharia Genética), vol. 8, págs. 299-316 (Plenum Press, Nova Iorque, 1986).

A extracção da totalidade do mRNA, a sua conversão em cDNA de cadeia dupla, a clonagem do cDNA e a transformação de hospedeiros adequados são levadas a cabo do modo indicado no parágrafo que liga a página 30 à página 31 da Application PCT/US 86/02464, publicada como W087/02990.

A extracção da totalidade do mRNA, a sua conversão em cDNA de cadeia dupla, a clonagem do cDNA e a transformação de hospedeiros adequados são levados a cabo do modo indicado no parágrafo que liga a página 30 à página 31 da Application PCT/US 86/02464, publicada como W087/02990.

Uma fonte preferida de mRNA que codifica os polipéptidos desejados são células cujos sobrenadantes contêm uma actividade associada com os polipéptidos do presente invento. De um modo geral, células T adequadas podem ser obtidas a partir de uma multiplicidade de fontes, tal como baço, amígdalas e sangue periférico humanos. Clones de células T, tais como aqueles que são isolados a partir de T-lin

fócitos de sangue periférico, podem também ser usados (ver Research Monographs in Immunology, eds. von Doehmer, H. e Haaf, V.; Secção D: "Human T-Cell Clones" (Clones de células T humanas), vol. 8, págs. 243-333; Elsevier Science Publishers, Nova Iorque /1985/).

Sondas para a biblioteca de cDNA são preparados utilizando técnicas convencionais, por ex. tradução de incisão ("nick-translation") da totalidade ou da quase totalidade do comprimento de clones de cDNA do presente invento, ou por intermédio de oligonucleótidos sintéticos marcados. Maniatis et al. (atrás referidos) forneceu instruções pormenorizadas para a técnica de "nick-translation", e os oligonucleótidos sintéticos marcados com qualquer sequência de bases podem ser construídos usando agentes sintetizadores de DNA disponíveis no mercado, por ex. o modelo 381A de Applied Biosystems, Inc. (Foster City, CA) ou outro análogo.

A biblioteca de cDNA é depois percorrida pela sonda utilizando técnicas convencionais, por ex. Beltz et al., "Isolation of Multigene Families and Determination of Homologies by Filter Hybridization Methods" (Isolamento de famílias multigênicas e determinação de homologias por métodos de hibridização de filtro), Methods in Enzymology, vol. 100, págs. 266-285 (1983); Callahan et al. "Detection and Cloning of Human DNA Sequences Related to the Mouse Mammary Tumor Virus Genome", (Detecção e clonagem de sequências de DNA humano apresentadas do genoma do vírus do tumor mamário do murganho), Proc. Natl. Acad. Sci., Vol. 79, págs 5503-5507 (1982), Grunstein et al., Proc. Natl. Acad. Sci., Vol. 72, págs. 3961-3965 (1975); ou Benton et al. Science, Vol. 196, págs. 180-183 (1977).

II. Preparação de muteínas de IL-3 humanas por meio de engenharia proteica

Técnicas genéricas para mutar DNA e produzir muteínas são descritas na application PCT/US 86/02464 (atrás referida) na página 33, linha 4 a partir do fim até à pág. 3, linha 11; estas podem ser também aplicadas no presente invento.

Muteínas do polipéptido natural podem ser desejáveis numa actividade de variedade de circunstâncias. Por exemplo, efeitos secundários não desejados podem ser reduzidos por certas muteínas, em particular se a actividade de efeito secundário estiver associada com uma parte do polipéptido diferente da da actividade de sejada. Além disso, nalguns sistemas de expressão, o polipéptido nativo é susceptível de degradação por meio de proteasos; nesse caso, substituições e/ou deleções seleccionadas de aminoácidos que modifiquem as sequências susceptíveis podem aumentar significativamente o rendimento: por ex. Pedido de Patente Britânica 2173-804-A em que Arg na posição 275 do activador de plasminogénio de tecido humano é substituído por Gli ou Glu. As muteínas podem também aumentar o rendimento em processos de purificação e/ou aumentar o tempo de vida das proteínas "na prateleira" mediante a eliminação de aminoácidos susceptíveis à oxidação, acilação, alquilação ou a outras modificações químicas. Por exemplo, a metionina sofre facilmente oxidação dando origem a um sulfóxido, que em muitas proteínas provoca uma perda de actividade biológica: por ex. Brot e Weissbach, Arch. Biochem. Biophys., Vol. 223, pág. 271 (1983). As metioninas podem frequentemente ser substituídas por aminoácidos mais inertes com pequena ou nenhuma perda de actividade biológica; por ex. Pedido de Patente Australiana AU-A-52451/86. Em sistemas de expressão bacteriana, os sedimentos podem por vezes ser aumentados mediante eliminação ou substituição de resíduos de cisteína não essenciais do ponto de vista conformacional: por ex. Mark et al., Patente dos E.U.A. 4 518 584.

Nas secções que se seguem, a notação usada por Estell et al. (atrás referido) para identificação de muteínas é seguida e generalizada. Por exemplo, "muteína IL-3 Leu¹³ Humana" (ou simplesmente "Leu¹³" se do contexto se depreender a proteína nativa) indica um polipéptido cuja sequência de aminoácidos é idêntica à da proteína nativa à excepção da posição 13 relativamente à terminação N. Val foi substituída nessa posição por Leu. Pode-se indicar de modo análogo mais de uma substituição; por ex., uma muteína em que Leu substitui Val na posição 13 e Feu substitui Met na posição 18 será referida como muteína IL-3 (Leu¹³, Fen¹⁸) humana. As deleções são indicadas por " Δ 's". Por exemplo, uma muteína em que falta Gen na posição 4 é referida como muteína IL-3 ⁴ humana. Uma inserção é indicada por "IN (Xaa)". Por exemplo, uma muteína com Leu inserida a seguir a Gen na posição 4 é preferida como muteína IL-3 IN⁴ (Leu) humana. Assim, a muteína IL-3 (Ser¹⁰, Δ ⁴, IN⁶ (Gli)) humana representa a sequência de IL-3 humana nativa que foi modificada substituindo Tre por Ser na posição 10, eliminando Gln na posição 4 e inserindo Gli imediatamente a seguir a Tre na posição 6. A inserção múltipla de aminoácidos no mesmo local é indicada por INⁱ(Xaa₁-Xaa₂-Xaa₃-...), em que Xaa₁-Xaa₂-Xaa₃... é a sequência inserida a seguir à posição i. Adições na terminação N são indicadas pelo índice superior "0", por ex. IN⁰(Xaa); e uma sequência de deleções dos aminoácidos 6-10, por exemplo, é referida como ⁶⁻¹⁰ ou como (Δ ⁶, Δ ⁷, Δ ⁸, Δ ⁹, Δ ¹⁰).

Prefere-se em primeiro lugar a utilização de mutagénese de cassette para produzir muteínas IL-3 humanas. Tal como adiante se descreve mais pormenorizada-mente, um gene de IL-3-humana sintético é construído com uma sequência de locais de restrição de endonuclease cínicos espaçados de modo aproximadamente uniforme ao longo do gene. A singularidade dos locais de restrição deverá manter-se quando o gene for inserido num vector apropriado, de modo que o gene possa ser convenientemente excisado e substituído por

oligonucleotidos sintéticos (i.e. "cassettes") que codificam mutações desejadas.

A determinação do número e distribuição de locais de restrição únicos implica tomar em linha de conta vários factores incluindo (1) locais de restrição pré-existent no vector a utilizar na expressão, (2) se é ou não desejada a utilização de códons específicos de espécies ou géneros, (3) o número de diferentes endonucleases disponíveis que não contam o vector (e as suas multiplicidades no interior do gene sintético), e (4) a conveniência e fiabilidade de sintetizar e/ou determinar a sequência dos segmentos entre os locais de restrição únicos.

De um modo geral, a Fórmula I define conjuntos de polipéptidos com substituições, inserções e/ou deleções de aminoácidos conservativas relativamente à sequência de IL-3 humana nativa putativa. "Conservativo", na acepção em que aqui é utilizado, significa (i) que as alterações são do ponto de vista da conformação, tão neutras quanto possível, isto é, concebidas de modo a produzir alterações mínimas na estrutura terciária dos polipéptidos mutantes quando comparados com a IL-3 humana nativa, e (ii) que as alterações são, do ponto de vista antigénico, tão neutras quanto possível, isto é, concebidas de modo a produzir alterações mínimas nos determinantes antigénicos dos polipéptidos mutantes quando comparados com a IL-3 humana nativa. A neutralidade conformacional é desejável para conservar a actividade biológica, e a neutralidade antigénica é desejável para utilizar o desencadear de reacções imunogénicas em pacientes ou animais tratados com os compostos do invento. Embora seja dificil seleccionar com certeza absoluta quais das alternativas serão neutras do ponto de vista conformacional e antigénico, existem regras que podem ser seguidas pelos especialistas na técnica para fazer alterações que tenham uma elevada probabilidade de serem neutras do ponto de vista conformacioo.

nal e antigénico, por ex. Anfinsen (atrás referido) e Berzopsly, Science, Vol. 229, págs. 932-940 (1985). Algumas das vezes mais importantes são (1) substituições de resíduos hidrofóbicos têm menor probabilidade de produzir alterações na antigenicidade uma vez que se situam provavelmente no interior da proteína, por ex., Berzofsky (atrás referido); (2) substituições de resíduos físico-quimicamente análogos, i.e. sinónimos têm menor probabilidade de produzir alterações na conformação uma vez que os aminoácidos de substituição podem desempenhar o mesmo papel estrutural que os aminoácidos substituídos; e (3) alterações de sequências conservadas pela evolução têm maior probabilidade de produzir efeitos conformacionais perniciosos uma vez que o facto de terem sido conservadas pela evolução sugere que essas sequências podem ser funcionalmente importantes.

Para além dessas regras básicas para seleccionar mutefinas, existem ensaios para confirmar a actividade biológica e conformação das maléculas criadas. Os ensaios biológicos para os polipéptidos do invento são descritos mais pormenorizadamente adiante. As alterações de conformação podem ser testadas pelo menos por intermédio de dois ensaios bem conhecidos: o método de fixação de microcomplemento, por ex. Wasserman et al., J. Immunol., Vol. 87, págs. 290-295 (1961), ou Levine et al. Methods in Enzymology, Vol. 11, págs. 928-936 (1967), amplamente utilizado em estudos evolutivos da estrutura terciária das proteínas; e afinidades relativamente a conjuntos de anticorpos monoclonais específicos relativamente a certas conformações, por ex. Lewis et al. Biochemistry Vol. 22, págs. 948-954 (1983).

III. Ensaio de actividade de CSF

A fim de determinar a actividade de CSF, células hemopoiéticas, por ex. células de medula óssea ou células sanguíneas de cordão umbilical são transformadas numa suspensão de célula única. As células individuais são depois "imobilizadas" num meio semi-sólido (agar) ou viscoso (metilcelulose) contendo substâncias nutritivas e, geralmente, soro fetal de vitelo. Na presença de um factor de estimulação apropriado, células individuais irão proliferar e diferenciar-se. Uma vez que as células iniciais estão imobilizadas, as colónias desenvolvem-se à medida que as células proliferam e maturam. Estas colónias podem ser analisadas passados 7-14 dias. São dadas instruções pormenorizadas para a realização de ensaios de CSF por Burgess, A. Growth Factors and Stem Cells (Factores de Crescimento e Células), págs. 52-55 (Academic Press, Nova Iorque 1984) e Metcalf, The Hemopoietic Colony Stimulating Factors (Os Factores de Estimulação de Colónia Hemopoiéticos), págs. 103-125 (Elsevier, Nova Iorque 1984). Caso seja desejado, colónias individuais podem ser extraídas, colocadas em lâminas de microscópio, fixadas e coradas com Wright/Geimsa (Todd-Samford, Clinical Diagnosis by Laboratory Methods (Diagnóstico Clínico por intermédio de Métodos Laboratoriais), 15ª edição, eds. Davidson e Henry (1974). Pode-se então determinar a análise morfológica de tipos de célula presentes por colónia individual.

Células de medula óssea recolhidas de pacientes sofrendo de doença não-hematológica são colocadas em camada sobre Ficoll (tipo 400, Sigma Chemical Co. St. Louis, Mo) e centrifugadas ($600 \times g$, 20 min), sendo removidas as células na interface. As células são lavadas por duas vezes em Iscove's Modified Dumbecco's Medium contendo 10% do soro fetal de vitelo (FCS) e suspensas de novo para

no mesmo meio, e as células aderentes são removidas por aderência para caixas de Petri plásticas (as células aderentes são frequentemente células produtoras de GM-CSF (Metcalf, atrás referido)). As células são aderentes são adicionadas a 10^5 células/ml a Iscove's Medium contendo 20% de FCS, 50 uM de 2-mercaptoetanol, 0,9% de metilcelulose e concentrações variadas ou de sobrenadantes que se sabem possuírem actividade de estimulação de colônia ou sobrenadantes de ensaio. Partes alíquotas de um ml são semeadas em caixas de Petri de 35 mm e cultivadas a 37°C numa atmosfera completamente humidificada de CO₂ a 6% em ar. Três dias após o início da cultura, adiciona-se 1 unidade de eritropoietina a cada placa. As colônias de granulociternacrofagos e as acumulações eritroides são contadas passadas 10-14 dias usando um microscópio invertido.

Células sanguíneas de cordão recolhidas em hepasina são resolvidas a 600×g durante 6 min. Os glóbulos brancos na interface entre o plasma e o pico de glóbulos vermelhos são transferidos para um tubo contendo cloreto de amônio 0,17 N e FCS a 6%. Após 5 min. em gelo, a suspensão é tratada com 4 ml de FCS e centrifugada durante 6 minutos a 600 × g. O sedimento celular é lavado com soro fisiológico tamponado com fosfato Dulbecco e submetido às etapas de Ficoll e aderência plástica atrás descritas para células de medula óssea. As células não-aderentes de baixa densidade são recolhidas e colocadas a 10^5 células/cultura no meio de cultura semi-sólido tal como atrás se descreveu.

No final dos ensaios, a composição celular é determinada após aplicação das colônias individuais a lamínas de vidro e coloração com Wright-Geimsa. Eosinófilos são determinadas corando com Luxol Fast Blue, por ex. Johnson, G. e Metcalf, D, Exp. Hematol, Vol. 8, págs. 549-561, (1980).

IV. Purificação e Composições Farmacêuticas

Métodos gerais de purificação dos polipéptidos do presente invento, e de sua utilização em laboratórios ou como ingredientes activos em composições farmacêuticas são reveladas na secção 5 de Application PCT/US 86/02464, especialmente da página 43, linha 20, à página 44, linha 44, linha 15, e da página 44, linha 27 à página 45, linha 27. A quantidade de polipéptido activo numa dose unitária de acordo com o presente invento pode ir de 1µg a 100 mg, dependendo da potência e aplicação.

Além disso, populações de células hematopoiéticas de uma pessoa podem ser mantidas e/ou expandidas ou feitas diferenciar-se ex vivo para uma eventual reintrodução na mesma ou noutra pessoa a fim de terem um efeito benéfico. As composições podem estimular selectivamente vários componentes do sistema imunitário, quer isoladamente quer com outros agentes bem conhecidos dos especialistas na técnica. Em particular, as composições podem ser incluídos outros agentes imunorreactivos, tal como linfoquinas (por ex. IL-1, IL-2 IL-4, GM-CSF, etc.)

V. Sistemas de expressão

Pode ser escolhido qualquer sistema de expressão adequado; a sua natureza não constitui um aspecto critico do presente invento, desde que aquele seja apropriado. Discutem-se sistemas de expressão adequados na Secção VI, Sistemas de Expressão, páginas 45-50 da Application PCT/

US 86-02464.

EXEMPLOS

Os exemplos que se seguem servem para ilustrar o presente invento. A selecção de bibliotecas de cDNA, vectores e hospedeiros, bem como a concentração de reagentes, temperaturas e os valores de outras variáveis destinam-se apenas a exemplificar a aplicação do presente invento, não devendo ser considerados como constituindo uma sua limitação.

EXEMPLO I

Preparação de cDNA de IL-3 humana mediante rastreio de uma biblioteca de pcD cDNA com uma sonda sintética e expressão em células de macaco COS 7 e oócitos de Xenopus.

Uma biblioteca de pcD DNA foi construída de mRNA isolado de um clone de T-células humanas (6D11) induzido a sintetizar mRNA mediante exposição a concanvalina A (conA) (2-10 micrograma/ml durante 4-12 horas) usando as técnicas reveladas por Okayama e Berg, Mol. Cell. Biol., Vol. 2, págs. 161-170 (1982) e vol 3, págs. 280-289 (1983). Outros compostos de indução, tal como fitohemaglutinina (PHA), "pokeweed mitogen", ionóforo de cálcio, antigénio anti-T3, etc. podem ser mais eficazes para alguns tipos de células e sob

algumas condições. Plasmídeos precursores de pcD, pcDV1 e pL1, podem ser obtidos convencionalmente de Pharmacia, Inc. (Piscataway, NJ).

A construção da biblioteca pcD cDNA de Okayama e Berg é descrita de modo sumário em Application PC/US 86/02464 no parágrafo que liga as páginas 31 e 32.

Uma vez terminada a biblioteca de cDNA nos vectores plasmídeos Okayama/Berg, ilustrada sob a forma de diagrama na Figura 1, os clones de cDNA são recolhidos, e amostras aleatórias analisadas relativamente à presença dos cDNAs desejados por métodos standardizados, por ex. selecção híbrida, detecção de determinantes antigénicos em produtos exprimidos, e/ou ensaios funcionais.

A. Isolamento de mRNA

RNA celular total foi isolado de células 6D11 usando o procedimento de isotiocianato de guanidina de Chirgwin et al. Biochemistry, Vol. 18, págs. 5294-5299 (1979). O procedimento é também revelado por Maniates et al (atrás referido). Células 6D11 estimuladas com conA lavadas e sedimentadas foram suspensas em solução de lise de isotiocianato de guanidina (50 g de tiocianato de guanidínio com 0,5 g de N-lauroil sarcosina de sódio (concentração final 0,5%), 2,5 ml de citrato de sódio 1M, pH 7,0 (25 mM), 0,7 ml de 2-mercaptoetanol (0,1M) e 0,33 ml de Sigma 305 Antifoam A (0,1%); juntou-se água desionizada até o volume perfazer 100 ml à temperatura ambiente e o volume final foi filtrado e o pH ajustado para o valor 7 com NaOH 1N). Vinte ml de solução de lise foram usados para $1,5 \times 10^8$ células.

Os sedimentos resultantes foram depois processados de acordo com o procedimento da Application PCT/US 86/02464, Secção B, página 55, linha 22, até à página 56, linha 24.

B. Construção da biblioteca de pcD

O procedimento de Okayama e Berg (Mol. & Cell Biol. Vol. 2, págs. 161-170 [1982]) foi usado com apenas pequenas alterações. Os plasmídeos pcDV1 e pL1 são descritos por Okayama e Berg (Mol. & Cell Biol. 3:380-389 [1983]) e podem ser obtidos de Pharmacia (Piscataway, N.J.). Especificamente, utilizou-se um plasmídeo pcDV1 modificado que continha um local NsiI no sítio onde previamente havia um local KpnI e utilizou-se um pL1 modificado que continha uma inserção no local XhoI consistindo numa porção do longo segmento terminal repetido (LTR) de um retrovirus HTLV(I). A presença do segmento LTR provou de um modo geral, aumentar a expressão em hospedeiros de células de mamíferos por um factor 20-50. As construções modificadas são ilustradas nas Figuras 1 e 2.

LTRs retrovirais revelaram incrementar a expressão numa variedade de sistemas; ver por ex. Chen et al Proc. Natl. Acad. Sci., Vol. 82, págs. 7285-7288 (1985); Gorman et al Proc. Natl. Acad. Sci. Vol. 79 págs. 6777-6781 (1982); Temin, Cell, Vol. 33, págs. 1-3 (1981); e Luciw et al. Cell, Vol. 33, págs. 705-716 (1983). A porção do LTR de HTLV(I) inserida no local XhoI de pL1 consiste num segmento de 267 bases incluindo a totalidade da região R e parte da região U5 (designada U5' na Figura 2) que se prolonga da fronteira R/U5 até ao primeiro local TaqI a jusante

(39 bases na totalidade). A sequência do LTR de HTLV (I) é revelada por Yoshida et al. em Current Topics in Microbiology, Vol. 115, págs. 157-175 (1985).

O segmento LTR de HTLV(I) é inserido em pL1 em quatro passos utilizando procedimentos revelados no Exemplo II. Primeiro, pL1 é digerido com BglI e XhoI, e o fragmento grande é isolado. O fragmento grande é depois misturado com os sintéticos 1A/B e 2A/B (adiante definidos) numa mistura de ligação comum. Os sintéticos 1A/B e 2A/B consistem, em conjunto, na sequência 5'→3", na sequência do pequeno fragmento BglI/XhoI, no fragmento 5' de 26 bases da região LTR R, e num polilinker que consiste em locais SmaI, SmaI, SacI, NaeI e XhoI na ordem indicada.

(BglI)

TCGGCCTCTGAGCTATTCCAG-
CGGAGCCGGAGACTCGATAAGGTC-

AAGTAGTGAGGAGGCTTTTTTGG-
TTCATCACTCCTCCGAAAAACCT-

GGCCTAGGCTTTT
CCGG

Sintéticos 1A/B

SV40 ori -->| R---->
GCAAAAAGCTCGGCTCG-
ATCCGAAAACGTTTTTCGAGCCGAGC-

SmaI
R --->| SacI
CATCTCTCCTTCACGCGCCCGGGGAG-
GTAGAGAGGAAGTGCGCGGGCCCCTC-

NaeI XhoI
CTCGCCGGCC
GAGCGCCGGAGCT

Sintéticos 2A/B

Após ligação, o plasmídeo resultante do passo 1 e propagado, isolado e digerido com SmaI e SacI e o fragmento grande é isolado. O fragmento grande é depois misturado com sintéticos 3A/B e 4A/B (definidos adiante) numa

mistura de ligação comum.

GCCGCCCTACCTGAGGCCGCCATC-
CGGCCGGGATGGACTCCGGCGGTAG-

CACGCCGGTTGAGTCGCGTTCT
GTGCGGCCAACTC

Sintéticos 3A/B

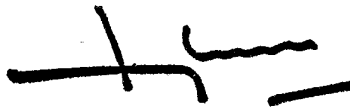
GCCGCCTCCCGCCTG-
AGCGCAAGACGGCGGAGGGCGGAC-

TGGTGCTCCTGAAC TGCGTCCGC-
ACCACGGAGGACTTGACGCAGGCG-

SacI
CGTCTAGGTAAGTTTAGAGCT
GCAGATCCATTCAAATC

Sintéticos 4A/B

Após ligação o plasmídeo resultante do passo 2 é propagado, isolado e digerido com SacI e NacI e o fragmento é isolado. O fragmento grande é depois misturado com sintéticos 5A/B (a seguir definidos) numa mistura de ligação comum.



CAGGTCGAGACCGGGCCTTG -
TCGAGTCCAGCTCTGGCCCGGAAAC -

TCCGGCGCTCCCTTGGAGCCTACCT -
AGGCCGCGAGGGAACCTCGGATGGA -

NaeI
AGACTCAGCC
TCTGAGTCGG

Sintéticos 5A/B

Após ligação, o plasmídeo resultante do passo 3 é propagado, isolado e digerido com NaeI e XhoI, e o fragmento grande é isolado. O fragmento grande é depois misturado com sintéticos 6A/B e 7A/B (adiante definidos) numa mistura de ligação comum a fim de dar origem ao fragmento de plasmídeo pL1 modificado, designado pL1'.

NaeI

GGCTCTCCACGCTTTGCCTGACCC-
CCGAGAGGTGCGAAACGGACTGGG-

TGCTTGCTCAACTCTACG
ACGAACGAG

Sintéticos 6A/B

TCTTTGTTTCGTTTT-
TTGAGATGCAGAAACAAAGCAAAA-

XhoI

CTGTTCTGCGCCGTTACAGATC
GACAAGACGCGCAATGTCTAGAGCT

Sintéticos 7A/B

Uma amostra de 80 ug de pcDV1 DNA foi digerida a 30°C com 20 U de endonuclease KpnI numa mistura de reacção de 450 µl contendo Tris.HCl (pH7,5) 6mM, MgCl₂ 6 mM, NaCl 6 mM, 2-Me 6mM e 0,1 mg de albumina de soro bovino (BSA) por ml. Após 16 horas a digestão foi terminada com 40 µl de EDTA (pH8,0) 0,25 M e 20 µl de dodecilsulfato de sódio (SDS) a 10%; o DNA foi recuperado após extracção com fenol-CHCl₃ 1:1 saturado de água (daqui para a frente referido como fenol CHCl₃) e precipitado com etanol. Adicionaram-se extremidades homopoliméricas com uma média de 60 mas não de 80, resíduos de desoxitimidilato (dT) por extremidade às extremidades produzidas pela endonuclease NsiI com transferase

terminal de timo de vitelo, como se segue: A mistura de reacção (38 µl) continha cacodilato de sódio 30 mM Tris HCl pH 6,8 como tampão, com CoCl_2 1mM, ditiotretol 0,1 mM, dTTP 0,25 mM, o DNA digerido com endonuclease NsiI, e 68 U da transferase de desoxinucleotidilo terminal (P-L Biochemicals, Inc., Milwaukee, WI). Após 30 min a 37°C a reacção foi gerada com 20 µl de EDTA 0,25 M (pH 8,0) e 10 µl de SDS a 10% e após várias extracções com fenol- CHCl_3 o DNA foi recuperado por precipitação com etanol. O DNA foi então digerido com 15 U de endonuclease EcoRI em 50 µl contendo Tris. HCl pH 7,4 10 mM, MgCl_2 10 mM, ditiotretol 1 mM e 0,1 mg de BSA por ml durante 5 horas a 37°C. O fragmento grande contendo o local de poliadenilação de SV 40 e o gene jBR322 de origem de replicação e resistência à ampicilina foi purificado por meio de electroforese de gel de agarose (1%) e recuperado do gel por intermédio de uma modificação do método de pó de vidro (Vogelstein, B. & Gillespie, D., Proc. Natl. Acad. Sci. 76: 615-619 [1979]). O DNA com extremidade de dT foi ainda purificado por meio de absorção e eluição a partir de uma coluna de oligo (dA)-celulose como se segue: O DNA foi dissolvido em 1 ml tampão de Tris.HCl pH 7,3 10 mM contendo EDTA 1mM e NaCl 1M, arrefecido para 0°C e aplicado a uma coluna de oligo (dA)-celulose (0,6 por 2,5 cm) e equilibrada com o mesmo tampão a 0°C e eluída com água à temperatura ambiente. O DNA eluído foi precipitado com etanol e dissolvido em Tris.HCl pH 7,3 10 mM com EDTA 1mM.

O DNA ligado com extremidade de oligo (dG) foi preparado digerindo 75 µg de pL1' DNA com 20 U de endonuclease PstI em 450 µl contendo Tris.HCl pH 7,4 6 mM, MgCl_2 6 mM, 2-ME 6 mM, NaCl 50 mM, e 0,01 mg de BSA por ml. Passadas 16 horas a 30°C a mistura de reacção foi extraída com fenol- CHCl_3 e o DNA foi precipitado com álcool. Extremidades com 10 a 15 resíduos desoxiguanilato (dG) foram depois adicionadas por extremidade com 46 U de transferase terminal de desoxinucleotidilo na mesma mistura de reacção (38 µl) que atrás foi descrita, com a excepção de dGTP 0,1 mM substituiu

dTTP. Passados 20 min. a 37°C a mistura foi extraída com fenol-CHCl₃ e o DNA foi precipitado com etanol e depois digerido com 35 U de endonuclease HindIII em 50 µl contendo Tris. HCl pH7,4 20 mM, MgCl₂ 7 mM, NaCl 60 mM e 0,1 mg de BSA a 37°C durante 4 horas. O pequeno DNA linker com extremidade oligo (dG) foi purificado por meio de electroforese de gel de agarose (1,8%) e recuperado tal como atrás se descreveu.

Construção da biblioteca: Foi feita de acordo com a Secção (2), Preparação da biblioteca de cDNA, da página 48, linha 9 a contar do fim, à página 61, linha 10 da Application PCT/US 86/02464. Depois, amostras da cultura foram congeladas em DMSO para serem armazenadas. Separadamente antes da adição de DMSO, vinte caixas de cultura de 15 cm foram inoculadas cada qual com cerca de 3000 clones.

Sondas marcadas radioactivamente para as colónias atrás referidas foram construídas como se segue: (1) utilizando os dados sobre a sequência da IL-3 humana revelados por Yang et al. Cell, Vol. 47, págs. 3-10 (1986), sintetizou-se um oligonucleótido sintético compreendendo os primeiros 55 nucleótidos da cadeia de codificação da referida sequência da IL-3 humana (começando com ATG); (2) usando os mesmos dados sobre a sequência sintetizou-se uma cadeia de oligonucleótido anti-sense correspondente que compreendia as bases 45-100 da referida sequência; (3) as duas cadeias sintéticas foram aneladas e usadas como fragmentos de iniciação ("primers") para fragmento Klenow de DNA polimerase I na presença de ATP, GTP, TTP e CTP marcado radioactivamente; (4) finalmente a reacção foi "cold chased" com precursores não marcadas para garantir extensão máxima de cadeia. ("Cold chased" significa que o percurso radioactivo CTP é substituído pela sua forma não radioactiva).

As 20 caixas atrás preparadas foram submetidas a rastreio relativamente a clones utilizando

as sondas sintéticas por um protocolo de hibridização de colônia padrão, por ex Maniatis et al. (atrás referido). Foram detectadas duas colônias positivas. Um dos clones, designado pcD-huIL3-22-1 foi seleccionado, amplificado e reclonado no plasmídeo M13 para efeitos de determinação de sequência pelo procedimento de terminação de cadeia didesoxi, por ex. Sanger et al., Proc. Natl. Acad. Sci., Vol. 74, págs. 5363-5367 (1977). A sequência de cDNA é ilustrada na Figura 3 juntamente com a sequência de aminoácidos prevista do contexto de leitura ("reading frame") aberto maior. A análise dos aminoácidos da terminação N de acordo com as regras empíricas publicadas por Perlman et al., J. Mol. Biol., Vol. 167, págs. 391-409 (1983) indica que a terceira prolina a partir do início do contexto de leitura ("reading frame") aberto e a terminação N mais provável da proteína finalizada. Todavia, apesar destas regras, a Ala adjacente pode constituir a terminação N correcta da proteína nativa tendo em conta a descoberta de que a terminação N da IL-3 de murganho é Ala. De qualquer modo, a sequência sinal não é uma porção crítica da molécula activa, e pode compreender uma grande variedade de sequência sem quaisquer efeitos negativos sobre a secreção ou actividade, por ex. Kaiser et al. Science, Vol. 235, págs. 312-317 (1987). Os locais de N-glicosilação potenciais estão assinalados por caixas na Figura 3.

Surpreendentemente, descobriu-se que o cDNA de IL-3 humanas isolado continha uma diferença de aminoácidos significativa da sequência referida por Yang et al. (atrás referidos): o cDNA isolado codifica uma prolina na posição 7 (relativamente à terminação N da Fórmula II) em vez de uma serina tal como referido por Yang et al. (atrás citados). Uma tal modificação deverá corresponder a diferenças de conformação significativas nos dois polipéptidos dadas as dissemelhanças entre prolina e serina.

Ao mesmo tempo, o mesmo clone foi amplificado separadamente; os vectores de clonagem pcD foram

isolados usando técnicas correntes, por ex. Maniatis et al. (atrás referidos); e o cDNA foi expresso em dois sistemas diferentes: células de macaco COS7 e oócitos de Xenopus. Os vectores pcD foram usados para transfectar células COS7 como se segue. Um dia antes da transfecção, aproximadamente 10^6 células de macaco COS 7 foram semeadas em caixas individuais de 100 mm em Dulbecco's modified Eagle's medium (DME) contendo 10% de soro fetal de vitelo e glutamina 2 mM. A fim de se realizar a transfecção, o meio foi aspirado de cada caixa e substituído por 4 ml de DME contendo Tris.HCl pH 7,4 50 mM, 400 µg/ml DEAE-dextrano e 50 µg dos DNAs plasmídeo a serem testados. As caixas foram incubadas durante quatro horas a 37°C, depois o meio contendo DNA foi removido e as caixas foram lavadas por duas vezes com 5 ml de DME isento de soro. DME contendo Cloroquina 150 µM foi adicionado às caixas que foram depois incubadas durante mais 3 horas a 37°C. As caixas foram lavadas uma vez com DME, e depois adicionaram-se DME contendo 4% de soro fetal de vitelo, glutamina 2 mM, penicilina e estreptomicina. As células foram depois incubadas durante 72 horas a 37°C. O meio de cultura foi recolhido e analisado relativamente a actividade de CSF usando o ensaio de sangue de cordão umbilical atrás descrito. Observou-se uma variedade de tipos de colónia.

Sobrenadantes de oócitos de Xenopus injectados com RNA transcrito do fragmento de inserção de cDNA de pcD-huIL3-22-1 foram analisados relativamente a actividade biológica. O protocolo seguido foi essencialmente aquele que é revelado por Melton et al. Nucleic Acids Research, Vol. 12, págs. 7035-7056 (1984) e por Krieg et al. Nucleic Acids Research Vol. 12 págs. 7057-7060 (1984). Primeiro, o fragmento de inserção de cDNA de pcD-huIL3-22-1 foi reclonado no vector pcD modificado, pSP6. pSP6 foi construído substituindo a origem de replicação de SV40 em pcD com o promotor de SP6, tal como se indica na Figura 3. O vector pcD modificado continha também um local PstI desactivado no gene de resistência à ampicilina (indicando por Δ PstI na Figura 3) e uma re-

gião polilinker inserida no local XhoI da região poliA do pcD. A região polilinker continha locais EcoRI, SacI, SmaI, BamHI, XbaI, SalI, SstI e HindIII. O local PstI desactivado permite que a região de origem de SV40 seja extraída por meio de digestão (antes da inserção do polilinker) com HindIII e PstI. O fragmento de inserção promotor de SP6 adequado (com cerca de 60 bases de comprimento) pode ser sintetizado usando técnicas usuais a partir da sequência revelada por Melton et al. (atrás referidos). Após inserção do promotor de SP6, o plasmídeo resultante foi amplificado, isolado e digerido com XhoI. O polilinker sintetizado usando técnicas correntes é inserido no local XhoI. Alternativamente, plasmídeos à venda no mercado contendo o promotor de SP6 e regiões polilinker podem ser usados, por ex. pSP62-PL que pode ser obtido de NEN Research Products (Boston, MA) ou psP18 ou psP19 que podem ser obtidos de BRL Life Technologies (Gaithersburg, MD).

cDNA de IL-3 humana foi extraída de pcD-hmIL-3-22-1 digerindo-o com PstI e BamHI, isolado e ligado com SP6 digerido com PstI/BamHI para formar pSP6-huIL3. pSP6-huIL3 foi amplificado em E. coli MC1061, isolado e linearizado com XbaI. RNA para injeção em oócitos foi preparado transcrevendo o psP6-hmIL3 linearizado com SP6 RNA polimerase que pode ser obtida de BRL Life Technologies) seguindo o protocolo de Kreig et al. (atrás referidos) e Melton et al. (atrás referidos). A SP6 RNA polimerase foi tratada com DNase (por ex. DNase livre de RNAsina e RNase a 1 unidade/microlitro e 20 microgramas (ml, respectivamente) e injectada no citoplasma dos oócitos de Xenopus e o sobrenadante das culturas de Xenopus foi analisado relativamente a actividade de CSF. Observou-se uma variedade de tipos de colônia.

EXEMPLO II

Construção de um gene de IL-3 humana sintético para mutagênese de cassete em pcD e expressão em células de macaco COS 7

Técnicas de construção e expressão do gene sintético deste exemplo são correntes na técnica da biologia molecular, por ex. especialmente Sproat e Gait, Nucleic Acids Research Vol. 13, págs. 2959-2977 (1985), e Perretti et al., Proc. Natl. Acad. Sci., Vol. 83, págs. 599-603 (1986); e também Mullenbach et al. J. Biol Chem, Vol. 261, págs. 719-722 (1986); Wells et al. Gene, Vol. 34 págs. 315-323 (1985); e Estell et al. Science, Vol 233, págs. 659-663 (1986). Sucintamente, o gene de IL-3 humana sintético é montado a partir de uma variedade de fragmentos de DNA de cadeia dupla sintetizadas quimicamente. São seleccionadas sequências de bases do gene sintético de modo que o gene sintético contenha uma série de locais de restrição únicos relativamente ao vector que transporta o gene.

A série de locais de restrição únicos define uma série de segmentos que podem ser facilmente extraídos e substituídos por segmentos com sequências de bases alteradas. Os fragmentos sintéticos são inseridos quer directamente quer após ligação com outros fragmentos num vector adequado, tal como um plasmídeo pcD ou outro análogo. Os segmentos atrás referidos correspondem às "cassettes" de Wells et al. (atrás citados). Os fragmentos sintéticos são sintetizados usando técnicas correntes, por ex. Gait, Oligonucleotide Synthesis: A Practical Approach (Síntese de Oligonucleótidos: Uma Abordagem Prática) (IRL Press, Oxford, Reino Unido, 1984). Utiliza-se de preferência um sintetizador automatizado, Tal como um modelo 380A de Applied Biosystems Inc (Foster City, CA). Plasmídeos pcD e outros vectores adequados estão disponíveis comercialmente, por ex. Pharmacia. Para pcD, a clonagem pode ser a cabo um E.coli MC1061 ou HB101 e a expres-

são pode ser levada a cabo em células de macaco COS ou células de ovários de hamsters chineses.

Neste exemplo um gene de IL-3 humana sintético é construído para inserção num plasmídeo pcD. Brevemente, o vector pcD é construído ligando primeiro quatro fragmentos uns aos outros com T4 DNA ligase: o fragmento grande HindIII/BamHI de pcDV1 contendo pBR322 ori e Ampr, o fragmento HindIII/PstI de pL1 contendo o promotor de região primeira SV40 e intrões de região posterior, e os sintéticos 11A/B e 12A/B, adiante descritos. A seguir, numa série de três passos adicionais de amplificação, purificação e inserção, o resto dos sintéticos 13A/B a 18A/B são adicionados até estar presente no vector pcD um gene de IL-3 sintético completo.

As digestões de endonucleases de restrição e as reacções de ligases são levados a cabo utilizando protocolos convencionais, por ex. Maniatis et al. Molecular Cloning: A Laboratory Manual (Clonagem Molecular: Um Manual de Laboratório)(Cold Spring Harbor Laboratory, Nova Iorque, 1982).

O método alcalino é usado para preparações de plasmídeos em pequena escala. Para preparações em grande escala emprega-se uma modificação do método alcalino em que se usa um igual volume de isopropanol para precipitar ácidos nucleicos a partir do lisato clarificado. A precipitação acetato de amónio 2,5M frio é usada para retirar RNA antes da centrifugação de densidade em equilíbrio com cloreto de cério e da detecção com brometo de etídio.

Cadeias complementares de DNAs sintéticos a serem clonadas (cerca de 400 mg cada qual) são misturadas e fosforiladas com polinucleótido quinase num volume de reacção de 50 µl. Este DNA é ligado com 1 µg de DNA vector digerido com enzimas de restrição apropriadas, e as ligações

estão num volume de 50 μ l à temperatura ambiente durante 4 a 12 horas. As condições de fosforilação, digestão por enzimas de restrição, reacções de polimevases, ligações e transfecções bacterianas foram já descritas (Maniatis et al., atrás referidos).

No passo 1, pcDV1 é amplificado, isolado e digerido com Hind III e BamHI. O fragmento contendo as regiões pBR322 ori e Amp^r é isolado por meio de electroforese de gel usando técnicas correntes, por ex. Maniatis et al. (atrás referido). pL1 é amplificado, isolado e digerido com Hind III e PstI. O fragmento contendo o promotor de região primeira de SV40 é isolado por meio de electroforese de gel. Os dois fragmentos isolados de pcDV1 e pL1 são depois misturados com sintéticos 11A/B e 12A/B (a seguir ilustrados) numa solução de T4 SNA ligase convencional. Os sintéticos 11A/B e 12A/B anulam-se para formar um fragmento de inserção no vector pcD agora criado, que é amplificado em E. Coli de depois isolado.

No passo 2, o vector pcD isolado do passo 1 é digerido com SpeI e SauI. O fragmento grande é separado e misturado com os sintéticos 13A/B e 14A/B numa solução de T4 DNA ligase convencional para formar o vector pcD do passo 2, que é amplificado em E. coli MC1061 e isolado.

No passo 3, o vector pcD isolado do passo 2 é digerido com MluI e SauI. O fragmento grande é separado e misturado com os sintéticos 15A/B e 16A/B numa solução de T4 DNA ligase convencional para formar o vector pcD do passo 3, que é amplificado em E. coli MC1061 e isolado.

No passo 4, o vector pcD isolado do passo 3 é digerido com EcoRI e SauI. O fragmento grande é separado e misturado com os sintéticos 17A/B e 18A/B numa solução de T4 DNA ligase convencional para formar o gene de

IL-3 sintético completo em pcD. O vector pcD do passo 4 é amplificado em E. coli MC1061, isolado e transfectado para células COS 7 tal como se descreve no Exemplo I. Os sobrenadantes da cultura são recolhidos e analisados relativamente a actividade de CSF tal como atrás se descreveu.

As sequências dos sintéticos 11A/B a 18A/B são a seguir indicadas. As localizações e os tipos dos locais de restrição únicos são indicados por cima das sequências.

(PstI)BclI
GTG ATCAATG AG CCG CCTG CCCG TCCTG CTC-
ACG TCACTAG TTACTCGG CGG ACGGG CAGG ACG AG -

CTG CTCCA ACTCCTGG TC
G ACG AGG TT

Sintéticos 11A/B

NarI
CGCCCCGG ACTCCAGG CGCCCATG ACC-
G AGG ACCAGG CGGGG CCTG AGG TCCG CGGG TACTGG -

SpeI SauI (BamHI)
CAG ACAACGCCCTTG AAG ACTAG TCCTTAGGG
G TCTG TTG CCGG AACTTCTG ATCAGGAATCCCCTAG

Sintéticos 12A/B

SpeI
CTAG TTGGG TTAAC TG CTCTAACATG ATCG ATCAA-
AACCCAATTG ACG AG ATTG TACTAGCTACTT-

XbaI
ATTATAACACACTTAAAG CAG CCACCTTTG CCTCTT-
TAATATTG TG TG AATTTCG TCGG TGG AAACGG AG AA-

CTAG ACTTCAACAACCTCAAT
GATCTG AAG TTG

Sintéticos 13A/B

GGGG AAG ACCAAG ACATTCTG ATGG AA-
TTGG AG TTACCCCTTCTGG TTCTG TAAG ACTACCTT-

XmaIII
AATAACCTTCG ACGG CCG AACCTGG AGG CATTCAAC-
TTATTGG AAGCTG CCGGCTTGG ACCTCCG TAAG TTG -

MluI (SauI)
AGGGCTG TCAAG AG TTTACAG AACGCG TCC
TCCCG ACAG TTCTCAAATG TCTTG CGCAGG AAT

Sintéticos 14A/B

MluI
CGCGTCAGCAATTGAGAGCATTCTTAAAAATCTC-
AGTCGTTAACTCTCGTAAGAAATTTTAGAG-

CTGCCATGTCGCCC
GACGGT

Sintéticos 15A/B

SacII
CTGGCCACCGCGGCACCCACGCGACAT-
ACAGACGGGGACCGGTGGCGCGTGGGTGCGCTGTA-

CCAATCCATATCAAGGACGGTGACTGGATT-
GGTTAGGTATAGTTCCTGCCACTGACCTAA-

EcoRI (SauI)
GAATTCCC
CTTAAGGGAAT

Sintéticos 16A/B

EcoRI
TTCCGGAGGAACTGACGTTCTATCTG-
AAGGCCTCCTTTGACTGCAAGATAGAC-

AAAACCCTTGAGAAAT
TTTTGG

Sintéticos 17A/B

GCGCAGGCTCAACAGACGACT-
GAACTCTTACGCGTCCGAGTTG TCTGCTGA-

NheI (SauI)
TTG TCGCTAGCGATCTTTTAGCC
AACAGCGATCGCTAGAAAATCGGATT

Sintéticos 18A/B

Exemplo III. Construção e expressão de muteína Leu¹³¹ de IL-3 humana

Ile na posição 131 é alterada para Leu a fim de formar muteína Leu¹³¹ de IL-3 humana. O plasmídeo pcD do Exemplo II contendo o gene completo de IL-3 humana é digerido com SauI e NheI. O fragmento grande é isolado e combinado com o sintético seguinte numa solução de T4 DNA ligase convencional.

O vector pcD resultante é amplificado em E. coli MC1061, isolado e transfectado para células de macaco COS 7 de acordo com os protocolos atrás descritos. Após incubação durante 3-4 dias os sobrenadantes da cultura de COS 7 são recolhidos e analisados relativamente a actividade de CSF.

Exemplo IV Construção e expressão de muteína Ile² de IL-3 humana

Met na posição 2 é alterada para Ile para formar a muteína Ile² da IL-3 humana. O plasmídeo pcD do Exemplo II contendo o gene de IL-3 humana completo é digerido com NarI e SpeI. O fragmento grande é isolado e combinado com o sintético seguinte numa solução de T4 DNA ligase convencional:

CGCCCATAACCCAGACAACGCCC-
GGG TATTGGG TCTGTTGCGGG-

TTGAAGA
AACTTCTGATC

O vector pcD resultante é amplificado em E. coli MC1061, isolado, e transfectado para células de macaco COS 7 de acordo com os protocolos atrás descritos. Após incubação durante 3-4 dias os sobrenadantes da cultura de COS 7 são recolhidos e analisados relativamente a actividade de CSF.

As descrições dos modelos de execução do invento precedentes foram apresentadas apenas a título de ilustração e descrição. Não são exaustivas e não restringem o invento às formas particulares aqui reveladas.

Aqueles que pediram esta patente depositaram pcD em-IL3-22-1 junto da American Type Culture Collection, Rockville, MD, EUA (ATCC) sob o número de acesso 67318.

REIVINDICAÇÕES:

1ª. - Processo para a preparação de uma proteína revelando uma actividade de factor de estimulação de colónias, caracterizado por se incorporar na referida proteína um polipeptídeo glicosilado ou não glicosilado 10 vezes substituído, defendido pela fórmula

X(Pro) - X(Met) - X(Tre) - X(Gln) - X(Tre) - W(Tre) -
 10
 W(Pro) - W(Leu) - W(Lis) - X(Tre) - X(Ser) - X(Trp) -
 X(Val) - X(Asn) - Cis - X(Ser) - W(Asn) - W(Met) -
 20
 W(Ile) - W(Asp) - W(Glu) - W(Ile) - W(Ile) - W(Tre) -
 30
 W(His) - W(Leu) - W(Lis) - W(Gln) - W(Pro) - W(Pro) -
 X(Leu) - X(Pro) - X(Leu) - X(Leu) - X(Asp) - X(Fen) -
 40
 X(Asn) - X(Asn) - W(Leu) - W(Asn) - W(Gl) - W(Glu) -
 W(Asp) - W(Gln) - W(Asp) - W(Ile) - W(Leu) - W(Met) -
 50
 W(Glu) - W(Asn) - W(Asn) - W(Leu) - W(Arg) - W(Arg) -
 60
 W(Pro) - W(Asn) - W(Leu) - W(Glu) - W(Ala) - W(Fen) -
 W(Asn) - W(Arg) - W(Ala) - W(Val) - W(Lis) - W(Ser) -
 70
 W(Leu) - W(Gln) - W(Asn) - W(Ala) - W(Ser) - W(Ala) -
 W(Ile) - W(Glu) - W(Ser) - W(Ile) - W(Leu) - X(Ls) -
 80
 X(Asn) - X(Leu) - X(Leu) - X(Pro) - Cis - X(Leu) -
 90
 X(Pro) - X(Leu) - X(Ala) - X(Tre) - X(Ala) - W(Ala) -
 W(Pro) - W(Tre) - W(Arg) - W(His) - W(Pro) - W(Ile) -
 100
 W(His) - W(Ile) - W(Lis) - W(Asp) - W(Gli) - W(Asp) -
 W(Trp) - W(Asn) - W(Glu) - W(Fen) - W(Arg) - W(Arg) -
 110
 W(Lis) - W(Leu) - X(Tre) - X(Fen) - X(Tir) - X(Leu) -
 120
 X(Lis) - W(Thr) - W(Leu) - W(Glu) - W(Asn) - W(Ala) -
 W(Gln) - W(Ala) - W(Gln) - W(Gln) - X(Tre) - X(Tre) -
 130
 X(Leu) - X(Ser) - X(Leu) - X(Ala) - X(Ile) - X(Fen)

em que:

X (Ser) representa Ser, Cis, Tre, Gli ou Asn;

X (Arg) representa Arg, His, Gln, Lis ou Glu;

X (Leu) representa Leu, Ile, Fen, Tir, Met ou Val;

X (Pro) representa Pro ou Ala;

X (Tre) representa Tre, Pro, Ala, Gli, His ou Gln;

X (Ala) representa Ala, Gli, Tre ou Pro;

X (Val) representa Val, Met, Tir, Fen, Ile ou Leu;

X (Gli) representa Gli, Ala, Tre, Pro ou Ser;

X (Ile) representa Ile, Met, Tir, Fen, Val ou Leu;

X (Fen) representa Fen, Trp, Met, Tir, Ile, Val ou Leu;

X (Tir) representa Tir, Trp, Met, Fen, Ile, Val ou Leu;

X (His) representa His, Glu, Lis, Gln, Tre ou Arg;

X (Gln) representa Gln, Glu, Lis, Asn, His, Tre ou Arg;

X (Asn) representa Asn, Glu, Asp, Gln ou Ser;

X (Lis) representa Lis, Glu, Gln, His ou Arg;

X (Asp) representa Asp, Glu ou Asn;

X (Glu) representa Glu, Asp, Lis, Asn, Gln, His ou Arg;

X (Met) representa Met, Fen, Ile, Val, Leu ou Tir;

X (Trp) é Trp;

W (Ser) é Ser;

W (Arg) representa Arg, His ou Lis;

W (Leu) representa Leu, Ile, Fen ou Met;

W (Pro) representa Pro ou Ala;

W (Tre) é Tre;

W (Ala) representa Ala ou Pro;

W (Val) representa Val, Met ou Ile;

W (Gli) é Gli;

W (Ile) representa Ile, Met, Fen, Val ou Leu;

W (Fen) representa Fen, Met, Tir, Ile ou Leu;

W (Tir) representa Tir ou Fen;

W (His) representa His, Gln ou Arg;

W (Gln) representa Gln; Glu ou His;

W (Asn) representa Asn ou Asp;

W (Lis) representa Lis ou Arg;

W (Asp) representa Asp ou Asn;

W (Glu) representa Glu ou Gln;

W (Met) representa Met, Fen, Ile, Val ou Leu;

W (Trp) é Trp;

e em que as regiões do referido polipeptídeo definidas pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé não são em conjunto, mais de 3 vezes substituídas.

2ª. - Processo para a preparação da proteína de acordo com a reivindicação 1, caracterizado por o referido polipeptídeo ser 5 vezes substituído de tal modo que as referidas regiões definidas pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé e 95-113 inclusivé não são, em conjunto, mais de 1 vez substituídas.

3ª. - Processo para a preparação da proteína de acordo com a reivindicação 2, caracterizado por o referido polipeptídeo ser 2 vezes substituído de tal modo que as referidas regiões definidas pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé não sofrem substituições.

4ª. - Processo para a preparação da proteína de acordo com a reivindicação 1, caracterizado por:

X (Ser) ser Ser;

X (Arg) representar Arg, His ou Lis;

X (Ile) representar Leu, Ile, Fen ou Met;

X (Pro) representar Pro ou Ala;

X (Tre) ser Tre;

X (Ala) representar Ala ou Pro;

X (Val) representar Val, Met ou Ile;

X (Gli) ser Gli;

X (Ile) representar Ile, Met, Fen, Val ou Leu;

X (Fen) representar Fen, Met, Tir, Ile ou Leu;

X (Tir) representar Tir ou Fen;

X (His) representar His, Gln ou Arg;

X (Gln) representar Gln, Glu ou His;

X (Asn) representar Asn ou Asp;

X (Lis) representar Lis ou Arg;

X (Asp) representar Asp ou Asn;

X (Glu) representar Glu ou Gln;

X (Met) representar Met, Fen, Ile, Val ou Leu;

W (Arg) ser Arg;

W (Leu) representar Leu, Ile ou Met;

W (Pro) ser Pro;

W (Ala) ser Ala;

W (Val) ser Val;

W (Ile) representar Ile, Met ou Leu;

W (Fen) ser Fen;

W (Tir) ser Tir;

W (His) ser His;

W (Gln) ser Gln;

W (Asn) ser Asn;

W (Lis) ser Lis;

W (Asp) ser Asp;

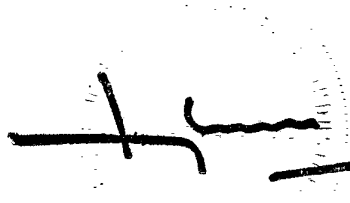
W (Glu) ser Glu e

W (Met) representar Met, Ile ou Leu.

5ª. - Processo para a preparação da proteína de acordo com a reivindicação 4, caracterizado por o referido polipeptídeo ser 5 vezes substituído de tal modo que as regiões pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé não são, em conjunto, mais de 1 vez substituídas.

6ª. - Processo para a preparação da proteína de acordo com a reivindicação 5, caracterizado por o referido polipeptídeo ser 2 vezes substituído de tal modo que as referidas regiões definidas pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé não sofrem substituições.

7ª. - Processo para a preparação da proteína de acordo com a reivindicação 4, caracterizado



por:

X (Arg) ser Arg;

X (Leu) representar Leu, Ile ou Met;

X (Pro) ser Pro;

X (Ala) ser Ala;

X (Val) ser Val;

X (Ile) representar Ile, Met ou Leu;

X (Fen) ser Fen;

X (Tir) ser Tir;

X (His) ser His;

X (Gln) ser Gln;

X (Asn) ser Asn;

X (Lis) ser Lis;

X (Asp) ser Asp;

X (Glu) ser Glu;

X (Met) representar Met, Ile ou Leu;

W (Leu) ser Leu;

W (Ile) ser Ile e

W (Met) ser Met.

8ª. - Processo para a preparação da proteína de acordo com a reivindicação 7, caracterizado por o referido polipeptídeo ser 5 vezes substituído de tal modo que as referidas regiões definidas pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé não são, em conjunto, mais de 1 vez substituídas.

9ª. - Processo para a preparação da proteína de acordo com a reivindicação 8, caracterizado por o referido polipeptídeo glicosilado ou não-glicosilado ser 2 vezes substituído de tal modo que as referidas regiões definidas pelos aminoácidos 10-32 inclusivé, 71-86 inclusivé e 95-113 inclusivé não sofrem substituições.

10ª. - Processo para a preparação da proteína de acordo com a reivindicação 9, caracterizado por o referido polipeptídeo glicosilado ou não-glicosilado compreender uma sequência de aminoácidos definida pela fórmula

Pro	-	Met	-	Tre	-	Gln	-	Tre	-	Tre	-	Pro	-	Leu	-
		10													
Lis	-	Tre	-	Ser	-	Trp	-	Val	-	Asn	-	Cis	-	Ser	-
				20											
Asn	-	Met	-	Ile	-	Asp	-	Glu	-	Ile	-	Ile	-	Tre	-
										30					
His	-	Leu	-	Lis	-	Gln	-	Pro	-	Pro	-	Leu	-	Pro	-
												40			
Leu	-	Leu	-	Asp	-	Fen	-	Asn	-	Asn	-	Leu	-	Asn	-
Gli	-	Glu	-	Asp	-	Gln	-	Asp	-	Ile	-	Leu	-	Met	-
		50													
Glu	-	Asn	-	Asn	-	Leu	-	Arg	-	Arg	-	Pro	-	Asn	-
				60											
Leu	-	Glu	-	Ala	-	Fen	-	Asn	-	Arg	-	Ala	-	Val	-
										70					
Lis	-	Ser	-	Leu	-	Gln	-	Asn	-	Ala	-	Ser	-	Ala	-
												80			
Ile	-	Glu	-	Ser	-	Ile	-	Leu	-	Lis	-	Asn	-	Leu	-

Leu - Pro - C_{is} - Leu - Pro - Leu - Ala - Tre -
 90
 Ala - Ala - Pro - Tre - Arg - His - Pro - Ile -
 100
 His - Ile - L_{is} - Asp - Gli - Asp - Trp - Asn -
 110
 Glu - Fen - Arg - Arg - L_{is} - Leu - Tre - Fen -
 120
 Tir - Leu - Lis - Tre - Leu - Glu - Asn - Ala -
 Gln - Ala - Gln - Gln - Tre - Tre - Leu - Ser -
 130
 Leu - Ala - Ile - Fen.

11^a. - Processo para a preparação de uma proteína revelando uma actividade de factor de estimulação de colónias, caracterizado por se incorporar na referida proteína um polipeptídeo glicosilado ou não glicosilado com 5 inserções, 5 deleções e 5 substituições, apresentando uma sequência de aminoácidos definidos pela fórmula da reivindicação 1, e em que as regiões do referido polipeptídeo definidas pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé não são, em conjunto, mais de 3 vezes substituídas.

12^a. - Processo para a preparação da proteína de acordo com a reivindicação 11, caracterizado por o referido polipeptídeo ter 2 inserções, 2 deleções e 2 substituições, de tal modo que a referida região definida pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé, não sofre substituições, deleções e inserções.

13^a. - Processo para a preparação da proteína de acordo com a reivindicação 12, caracterizado por o polipeptídeo ser definido pela fórmula:

Ala - Pro - Met - Tre - Gln - Tre - Tre - Pro -
10
Leu - Lis - Tre - Ser - Trp - Val - Asn - Cis -
20
Ser - Asn - Met - Ile - Asp - Glu - Ile - Ile -
30
Tre - His - Leu - Lis - Gln - Pro - Pro - Leu -
40
Pro - Leu - Leu - Asp - Fen - Asn - Asn - Leu -
Asn - Gli - Glu - Asp - Gln - Asp - Ile - Leu -
50
Met - Glu - Asn - Asn - Leu - Arg - Arg - Pro -
60
Asn - Leu - Glu - Ala - Fen - Asn - Arg - Ala -
70
Val - Lis - Ser - Leu - Gln - Asn - Ala - Ser -
80
Ala - Ile - Glu - Ser - Ile - Leu - Lis - Asn -
Leu - Leu - Pro - Cis - Leu - Pro - Leu - Ala -
90
Tre - Ala - Ala - Pro - Tre - Arg - His - Pro -
100
Ile - His - Ile - Lis - Asp - Gli - Asp - Trp -
110
Asn - Glu - Fen - Arg - Arg - Lis - Leu - Tre -
120
Fen - Tir - Leu - Lis - Tre - Leu - Glu - Asn -
Ala - Gln - Ala - Gln - Gln - Tre - Tre - Leu -
130
Ser - Leu - Ala - Ile - Fen.

14^a. - Processo para a preparação de um ácido nucleico, caracterizado por se incorporar no referido ácido nucleico uma sequência de nucleotídeos seleccionada entre sequências de nucleotídeos capazes de codificar polipeptídeos 10 vezes substituídos definidos pela fórmula da

reivindicação 1, em que as regiões do referido polipeptídeo definidas pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé não são, em conjunto, mais de 3 vezes substituídas.

15^a. - Processo para a preparação do ácido nucleico de acordo com a reivindicação 14, caracterizado por a referida sequência de nucleotídeos ser seleccionada entre sequências de nucleotídeos capazes de codificar um polipeptídeo definido pela fórmula da reivindicação 10.

16^a. - Método para a produção de um polipeptídeo revelando actividade de factor de estimulação de colónias, caracterizado por compreender os seguintes passos:

(a) construção de um vector compreendendo uma sequência de nucleotídeos que codificam o referido polipeptídeo, em que a sequência de nucleotídeo é capaz de ser expressa por um hospedeiro contendo o vector e em que a sequência de nucleotídeos é seleccionada entre sequências de nucleotídeos capazes de codificar um polipeptídeo 10 vezes substituído definido pela fórmula da reivindicação 1; e em que as regiões do referido polipeptídeo definidas pelos aminoácidos 10-32 inclusivé, 47-60 inclusivé, 71-86 inclusivé e 95-113 inclusivé não são, em conjunto mais de 3 vezes substituídas;

(b) incorporação do vector no hospedeiro; e

(c) manutenção do hospedeiro num meio de cultura em condições adequadas para expressar a referida sequência de nucleotídeos no referido polipeptídeo.

17^a. - Método de acordo com a reivindicação 16, caracterizado por a referida sequência de nucleotídeos ser seleccionada entre sequências de nucleotídeos capazes de codificar um polipeptídeo definido pela fórmula da reivindicação 10.

18^a. - Método de acordo com a reivindicação 17, caracterizado por compreender ainda a fase de separação do referido polipeptídeo do referido meio de cultura e do referido hospedeiro.

19^a. - Processo de preparação de uma composição farmaceutica para a estimulação in vitro ou in vivo do crescimento e desenvolvimento de células sanguíneas, caracterizado por se incluir na referida composição um veículo terapeuticamente compatível e uma quantidade eficaz de uma proteína obtida de acordo com a reivindicação 1.

20^a. - Processo para a preparação de uma composição farmaceutica de acordo com a reivindicação 19 caracterizado por a proteína compreender uma sequência de aminoácidos definidos pela fórmula da reivindicação 10 ou p pela referida fórmula precedida por Ala.

Lisboa, 17 de Fevereiro de 1988

J. PEREIRA DA CRUZ
Agente Oficial da Propriedade Industrial
RUA VICTOR CORDON, 10-A, 1.^o
1200 LISBOA

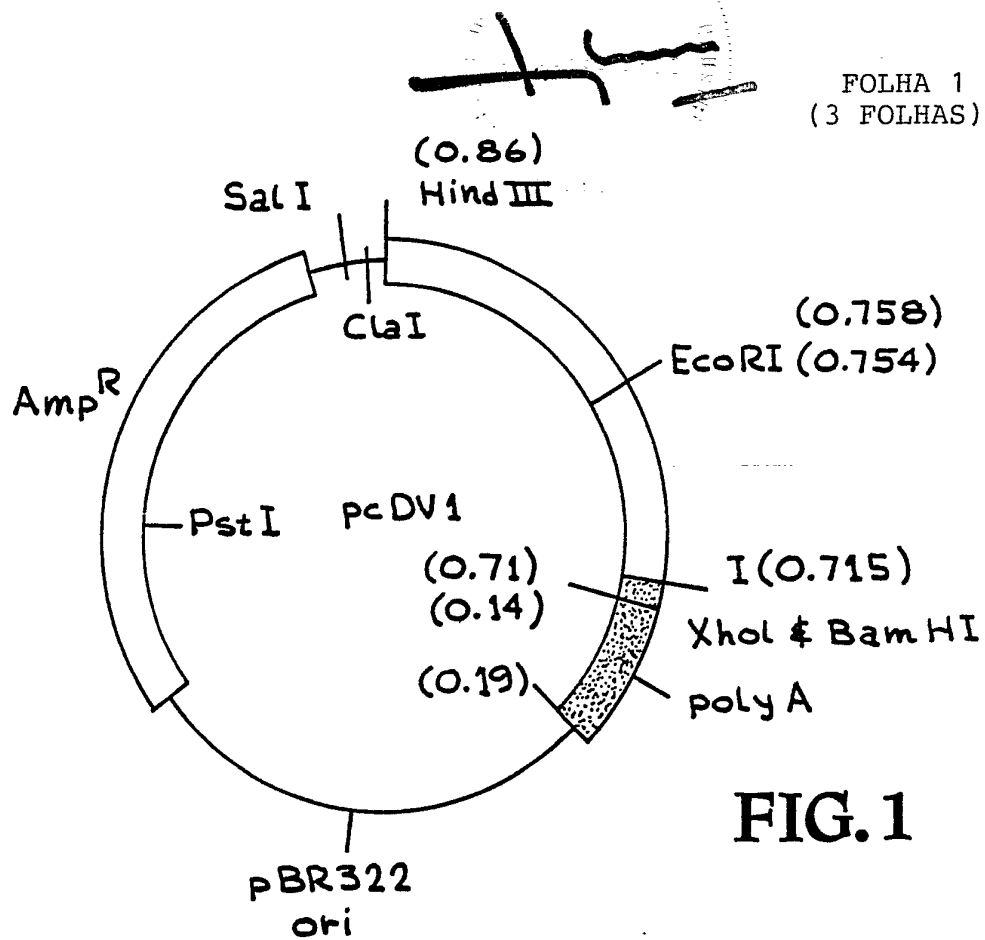


FIG. 1

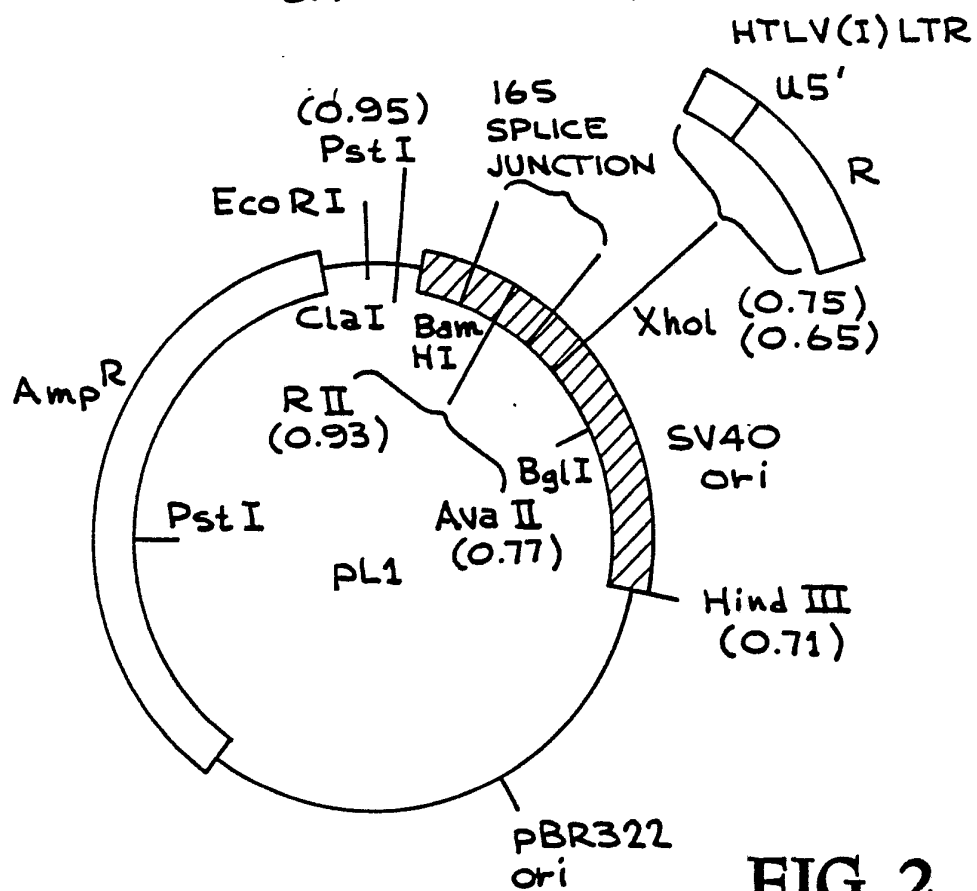


FIG. 2

10 20 30 40 50
 CAGAGCCCCA CGAAGGACCA GAACAAGACA GAGTGCCTCC TGCCGATCCA AAC ATG AGC CGC
 MET Ser Arg
 68 83 98 113
 CTG CCC GTC CTC CTG CTC CAA CTC GTC GTC CGC CCC GGA CTC CAA GCT CCC
 Leu Pro Val Leu Leu Leu Gln Gln Leu Leu Val Val Arg Pro Gly Leu Ala Pro
 128 143 158
 ATG ACC CAG ACA ACG CCC TTG AAG ACA AGC TGG GTT AAC TGC TCT AAC ATG ATC
 MET Thr Gln Thr Thr Pro Pro Leu Lys Thr Ser Trp Val Asn Cys Ser Asn MET Ile
 173 188 203 218
 GAT GAA ATT ATA ACA CAC TTA AAG CAG CCA CCT TTG CCT TTG CTG GAC TTC AAC
 Asp Ile Ile Thr His Leu Lys Gln Pro Pro Leu Pro Leu Leu Asp Phe Asn
 233 248 263 278
 AAC CTC AAT GGG GAA GAC CAA GAC ATT CTG ATG GAA AAT AAC CTT CGA AGG CCA
 Asn Leu Asn Gly Glu Asp Gln Asp Ile Leu MET Glu Asn Asn Leu Arg Arg Pro
 293 308 323 338
 AAC CTG GAG GCA TTC AAC AGG GCT GTC AAG AGT TTA CAG AAC GCA TCA GCA ATT
 Asn Leu Glu Ala Phe Phe Asn Arg Ala Val Lys Ser Leu Leu Gln Asn Ala Ser Ala Ile
 353 368 383
 GAG AGC ATT CTT AAA AAT CTC CTG CCA TGT CTG CCC CTG GCC ACG GCC GCA CCC
 Glu Ser Ile Leu Lys Asn Leu Leu Pro Cys Leu Pro Leu Ala Thr Ala Ala Pro
 398 413 428 443
 ACG CGA CAT CCA ATC CAT ATC AAG GAC GGT GAC TGG AAT GAA TTC CGG AGG AAA
 Thr Arg His Pro Ile His Ile Lys Asp Gly Asp Trp Asn Glu Phe Arg Arg Lys
 458 473 488
 CTG ACG TTC TAT CTG AAA ACC CTT GAG AAT GCG CAG GCT CAA CAG ACG ACT TTG
 Leu Thr Phe Tyr Leu Lys Thr Leu Glu Asn Ala Gln Ala Gln Thr Thr Leu
 503 522 532 542 552
 AGC CTC GCG ATC TTT TAG TCCAACGTCC AGCTCGTTCT CTGGGCTTC TCACCACAGA
 Ser Leu Ala Ile Phe

FIG. 3

[Handwritten signature]

FOLHA 3^a
(3 FOLHAS)

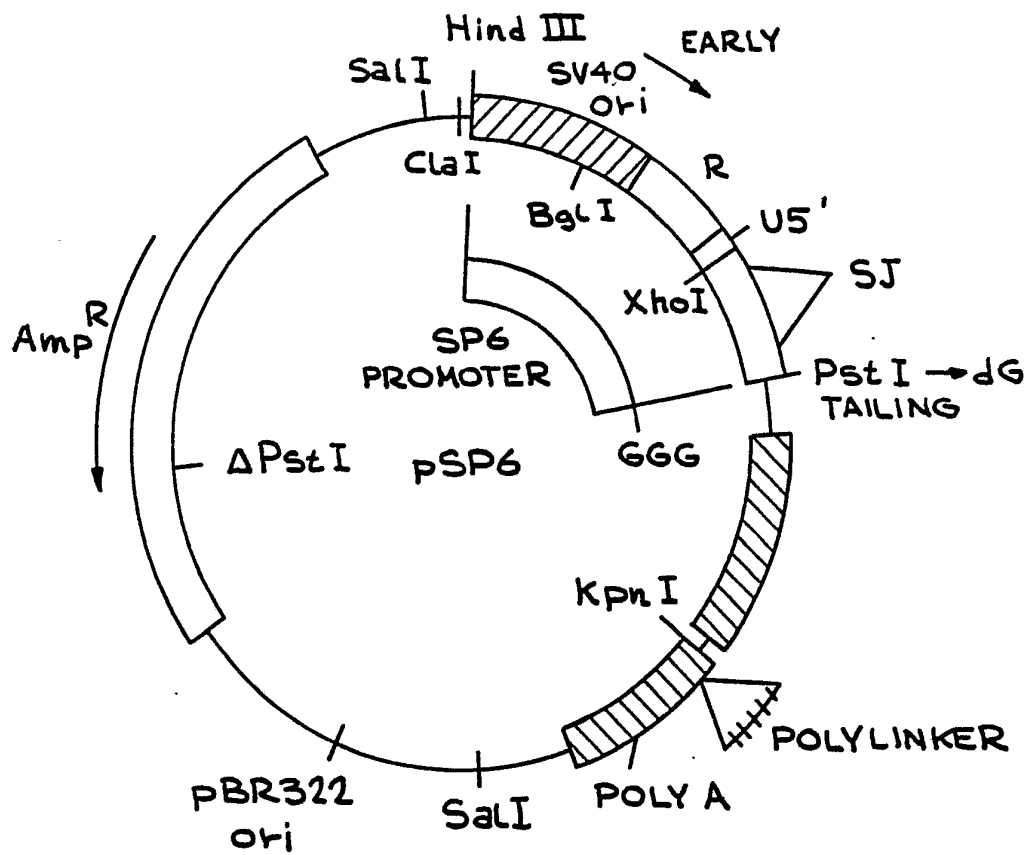


FIG. 4