

DESCRIÇÃO
DA
PATENTE DE INVENÇÃO

N.º 86 381

REQUERENTE: GIST-BROCADES N.V., holandesa, industrial
e comercial, com sede em 1, Wateringseweg
P O Box 1, 2600 MA Delft, Holanda

EPÍGRAFE: "PROCESSO DE CLONAGEM MOLECULAR E DE
EXPRESSION DE INTERLEUQUINA-3(IL-3)
HUMANA"

INVENTORES: Lambertus Christiaan Johannes
Dorssers, Gerard Wagemaker, Yvonne
Johanna Vos e Robert Willem Van
Leen

Reivindicação do direito de prioridade ao abrigo do artigo 4º da Convenção de Paris
de 20 de Março de 1883.

Eurpeia, em 16 de Dezembro de 1986 sob o No. 86202285.2
e em 13 de Julho de 1987 sob o No. 87201322.2



Memória descritiva referente à patente de invenção de GIST-BROCADES N.V., holandesa, industrial e comercial, com sede em 1, Wateringseweg, P.O. Box 1,2600 MA Delft, Holanda, (inventores: Dr. Ir. Lambertus Christiaan Johannes Dorssers, Dr. Gerard Wagemaker, Yvonne Johanna Vos e Robert Willem Van Leen, residentes na Holanda), para "PROCESSO DE CLONAGEM MOLECULAR E DE EXPRESSÃO DE INTERLEUQUINA-3 (IL-3) HUMANA".

MEMÓRIA DESCRITIVA

Domínio da Invenção

A presente invenção descreve um processo para a obtenção de ADNc que codifica para a interleuquina-3 humana (hIL-3) e a um processo que utiliza este ADN para, nomeadamente, clonagem e expressão em vários organismos, incluindo microorganismos, em particular em leveduras, bactérias e fungos, células de culturas de tecidos e animais ou plantas transgênicos.

Bases da Invenção

Designa-se por hemopoiese o processo activo de proliferação e de diferenciação de células progeni-



toras pluripotentes nos diversos tipos de células do sangue maduras e em células de alguns tecidos especializados. A produção de células do sangue funcionais é regulada por proteínas específicas, os factores de crescimento hemopoiéticos (hemopoietic growth factors = HGFs). Alguns dos HGFs regulam a maturação duma linhagem de maturação específica enquanto que outros estimulam a proliferação e a diferenciação dos progenitores ao longo de múltiplas vias. Grande parte do nosso conhecimento sobre o processo de diferenciação hemopoiético foi obtido a partir de estudos em ratos in vitro e in vivo usando factores de crescimento purificados. O factor murino de crescimento conhecido por interleuquina-3 murina (mIL-3), também designado por multi-CSF (cell stimulation factor: factor de estimulação celular), por factor de crescimento de células em bastonete, por factor de activação das células indiferenciadas ou por várias outras designações, estimula a proliferação de células multipotentes (CFU-S) de desenvolvimento precoce, efeito que é detectado por contagem de colónias de baço, tendo como resultado a produção de células progenitoras bem como de eritróides, megacariócitos, granulócitos/macrófagos, osteoblastos e várias outras linhagens. Além disso, a mIL-3 tem sido implicada na replicação de células indiferenciadas pluripotentes, provavelmente em sinergismo com outros HGFs.

Em anos recentes vários grupos tiveram êxito na clonagem de ADNc da mIL-3. Até agora não foi relatado qualquer resultado de identificação de sequências homólogas em ADN humano usando ADN da mIL-3 como sonda. É de presumir que o gene humano divergiu em alto grau do gene da mIL-3 ou que perdeu a sua função durante a evolução dos primatas. No entanto descobriu-se que leucócitos humanos produzem um HGF ou HGFs que podem substituir a mIL-3 como agente da proliferação de CFU-S no rato. Deste modo foi postulada a existência dum HGF humano que partilha as propriedades biológicas com a mIL-3 e que portanto poderia ser o respectivo homólogo humano. Yang, Y-C et al., Cell (1986) 47 : 3-10, com data de 10 de Outubro, descreve ADNc que codifica para uma proteína que possui actividade semelhante à IL-3 das células T do gibbon e o isolamento de ADN genómico que codifica para uma proteína análoga humana. É possível deduzir a se



quência dum ADNc que codifica para a IL-3 humana a partir da sequência do gene humano publicada por Yang et al.. No entanto o trabalho referido não descreve um ADNc que codifica para a IL-3 humana nem ensina um método para o isolamento deste ADNc e também não contém qualquer descrição da produção de hIL-3. A presente invenção descreve pela primeira vez o isolamento dum ADNc que contém a sequência integral que codifica para a IL-3 humana.

A proteína IL-3 humana nunca foi preparada sob a forma pura nem nunca foram descritas as respectivas características, para além da sua actividade em análises de proliferação in vitro e da estrutura primária deduzida. A presente invenção permite o isolamento de IL-3 humana pura e a identificação das respectivas características por meio de produção por técnicas recombinantes a partir dum clone de ADNc.

Resumo da Invenção

Como anteriormente afirmado, a presente invenção descreve pela primeira vez o isolamento dum ADNc que contém a sequência integral que codifica para a IL-3 humana. O baixo grau de homologia entre as sequências de ADN que codificam para a IL-3 murina e para a IL-3 humana não permite o isolamento do ADNc da hIL-3 por hibridação com a sequência que codifica para a mIL-3. Verificou-se surpreendentemente que é possível isolar o clone de ADNc da hIL-3 aproveitando o grau de homologia relativamente alto na parte 3' não codificante dos ADN's. Quando se dispõe do clone de ADNc torna-se possível a produção de hIL-3 por uma grande variedade de organismos hospedeiros. Depois de se proceder à produção da proteína em grande escala é possível proceder à sua purificação e à utilização terapêutica.

A presente invenção permite a produção da proteína IL-3 humana recombinante numa grande variedade de células hospedeiras por transcrição e tradução duma sequência de ADNc que codifica para a proteína IL-3 humana. A produção da proteína é ilustrada seguidamente em vários hospedeiros, incluindo E. coli, células COS, células Cl27, B. subtilis e B. licheniformis, S. cerevisiae, e K. lactis. Também se tornou possível a



produção em outros hospedeiros mediante o uso de sistemas de expressão apropriados por meio da utilização de ADNc sem intrões. Mais geralmente, o facto de se dispôr de anticorpos de IL-3 anti-humanos que permitem a identificação de colónias que levam a uma produção viável da proteína recombinante possibilita a produção de IL-3 humana a partir de qualquer sistema recombinante.

De acordo com um aspecto da presente invenção, por conseguinte, refere-se a preparação de ADN recombinante sem intrões que codifica para a proteína IL-3 humana.

De acordo com outro aspecto da presente invenção descreve-se a preparação de sistemas de expressão com capacidade para efectuar a expressão da sequência referida de ADN que codifica para a hIL-3 num hospedeiro apropriado.

De acordo com outros aspectos, a presente invenção diz respeito à preparação da proteína IL-3 humana recombinante sob a forma glicosilada ou não glicosilada, de IL-3 humana isenta das substâncias que normalmente acompanham a referida proteína e de anticorpos que reagem de forma específica com estas proteínas recombinantes ou purificadas.

Breve Descrição dos Desenhos

A figura 1 apresenta uma comparação das sequências de ADN e das proteínas de multi-CSF humana e da IL-3 de rato. A sequência de hmulti-CSF e a do ADN (Clone D11, linhas da parte superior) foram alinhadas com a sequência do ADN de mIL-3 (11, 35) e a da respectiva proteína (30). Os nucleótidos idênticos estão indicados por uma linha vertical e os ácidos aminados idênticos estão contidos em caixas. Os sinais de polia-denilação são indicados por pontos e as unidades de repetição ATTTA estão marcadas por traços horizontais.

A figura 2 apresenta a construção do plasmídeo pLB4 que contém o ADNc da IL-3 humana. E = EcoRI, Sm = SmaI, B = BamHI, S = SstI e K = KpnI.

A figura 3 apresenta a actividade biológica de COS/pLB4 CM em progenitores da medula óssea humana. Apresentam-se os números médios ($\pm$ DP) de colónias de eritróides



(BFU-E), de granulócitos-macrófagos (CFU-GM), de granulócitos (CFU-G), de eosinófilos (CFU-Eo), de macrófagos (CFU-M) e de culturas mistas (CFU-MIX) para culturas em duplicado estimuladas com volumes crescentes de COS/pLB4 CM.

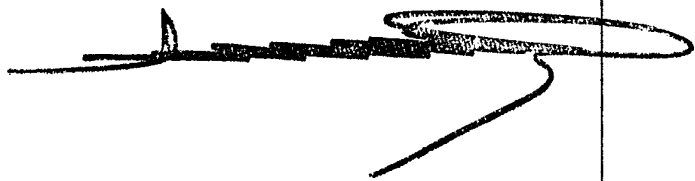
A figura 4 mostra a indução da proliferação de AML por COS/pLB4 CM medida por análise duma cultura de colônias (painel A) e por análise numa síntese de ADN (incorporação de $^3\text{H-TdR}$) (painel B).

A figura 5 mostra um diagrama de construção dos vectores de expressão em E. coli pGB/IL-301, GB/IL-302, pGB/IL-303, pGB/IL-304, pGB/IL-305 e pGB/IL-306. Nesta figura, X representa um local XhoI, E representa um local EcoRI, B representa um local BamHI e A representa um local AvaI.

A figura 6 apresenta a sequência do local de multiclonagem em pTZ18R (Pharmacia) e do seu derivado pT1.

A figura 7 mostra uma representação esquemática de clones de expressão de hmulti-CSF. Apenas se apresenta a inserção de ADNc de hmulti-CSF para os plasmídeos de expressão eucariótica pLB4 e pLH1. As regiões de codificação para o peptídeo guia (|||||) e para a proteína de hmulti-CSF madura (■) são apresentadas em caixas. Os clones de expressão bacteriana de hmulti-CSF (derivados de pLH1) contêm a região de codificação para lacZ e para a proteína multi-adaptador (|||||), a região 5' terminal não codificante de hmulti-CSF (□) e a região de codificação hmulti-CSF. A flecha indica o códon ATG de início utilizado no vector particular.

A figura 8 apresenta as sequências das regiões de fusão de ADN de lacZ/hmulti-CSF para vários vectores de expressão bacterianos. A sequência de clones é dada a partir do início da proteína lacZ tanto em pUC8 como em pTZ18R (letras minúsculas) e da sequência de ADN de hmulti-CSF até ao local ClaI na posição 158. As mutações na sequência de ADN de hmulti-CSF estão sublinhadas, dando como resultado: $\text{trp}^{13} \rightarrow \text{arg}^{13}$ (pGB/IL-302); $\text{leu}^9 \rightarrow \text{pro}^9$ e $\text{trp}^{13} \rightarrow \text{arg}^{13}$ (pGB/IL-303); $\text{met}^3 \rightarrow \text{thr}^3$ e uma alteração silenciosa (pGB/IL-304). Os números em expoente significam o número do resíduo de ácido aminado na proteína



madura.

A figura 9 apresenta uma electrofo-
rese em gel de poliacrilamida de hmulti-CSF bacteriano produzido
a partir de bactérias que contêm pGB/IL-301 e pGB/IL-302.

A figura 10 apresenta a titulação
de uma proteína de fusão hmulti-CSF em células blásticas AML.

A figura 11 mostra uma análise de
mancha de Western que demonstra a reacção específica de IL-3 de
anti-soro de coelho desenvolvido por reacção contra a proteína
de 21 kD isolada a partir do lisado de E. coli transformado com
pGB/IL-301.

A figura 12 mostra o efeito do anti-
soro da figura 11 na actividade de IL-3.

A figura 13 mostra uma representa-
ção esquemática do plasmídeo pGB/IL-307. A caixa (▨) indica a
sequência de codificação para a IL-3 humana. Os ácidos aminados
terminais da proteína de fusão estão representados sob o desenho.

A figura 14 mostra uma representa-
ção esquemática do plasmídeo pGB/IL-308. A sequência de nucleóti-
dos da região do promotor está representada sob o desenho.

A figura 15 apresenta a construção
do plasmídeo pGB/IL-309. A primeira caixa (▨) indica uma par-
te da sequência da IL-3 humana, nomeadamente a sequência de si-
nal mais 20 ácidos aminados da proteína madura. A outra caixa
(▨) indica parte da região 3' não codificante da sequência de
ADNc da IL-3.

A figura 16 é uma representação es-
quemática do plasmídeo pGB/IL-310.

A figura 17 mostra a sequência de
nucleótidos do plasmídeo pBHA1.

A figura 18 apresenta a construção
dos plasmídeos pGB/IL-311 e pGB/IL-312. A caixa (▨) indica a
região de codificação do precursor da IL-3 humana.

A figura 19 apresenta a construção
do plasmídeo pGB/IL-313. A sequência do lado 5' da sequência da
IL-3 está representada sob o desenho.

A figura 20 mostra uma representa-



ção esquemática do plasmídeo pGB/IL-317.

A figura 21 mostra uma representação esquemática do plasmídeo pGB/IL-316.

A figura 22 apresenta a sequência de nucleótidos do plasmídeo pGB/IL-316 entre o local SacII singular no promotor de lactase e o local HindIII depois do terminador (resíduos 4457 a 7204).

A figura 23 apresenta a sequência de nucleótidos do plasmídeo pGB/IL-318 entre o local SacII singular no promotor de lactase e o local HindIII depois do terminador (resíduos 4457 a 7190).

A figura 24 apresenta a sequência de nucleótidos do promotor de EF-1 α do adaptador SalI-BglII-XhoI e do terminador de actina tal como aparecem no plasmídeo pGB/TEFact.

Descrição pormenorizada da Invenção

A - Definições

Na presente descrição utilizam-se os termos "IL-3 humana", "hIL-3", "multi-CSF humano" e "hmulti-CSF" indiferentemente para designar uma proteína que apresenta as seguintes actividades:

1. A proteína estimula a formação de colónias por células progenitoras hemopoiéticas humanas e as colónias formadas incluem eritróides, granulócitos, granulócitos macrófagos e colónias mistas.

2. A proteína estimula a síntese de ADN por células blásticas de leucemia mielogénia aguda (acute myelogenous leukemia = AML), como se verifica, por exemplo, pela absorção de timidina marcada.

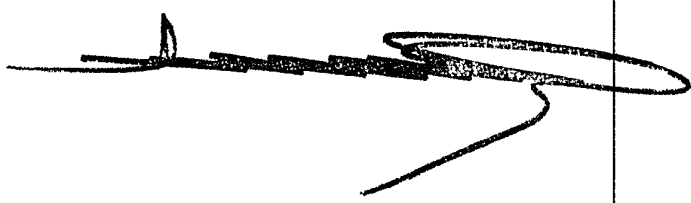
Para corresponder à definição de hmulti-CSF, a actividade nas análises descritas não deve ser inibida de modo substancial por anticorpos desenvolvidos em resposta a GM-CSF e imuno-específicos para este factor, a menos que estes anticorpos também inibam estas actividades pela forma ilus-



trativa de hmulti-CSF seguinte.

Na figura 1 apresenta-se uma forma ilustrativa de hmulti-CSF sob a forma de uma proteína madura de 133 ácidos aminados possuindo uma sequência de sinal de 19 ácidos aminados. A sequência de ácidos aminados da figura 1 é idêntica à descrita por Yang, Y-C., et al., Cell (1986) 47:3-10 (supra) excepto na posição 8 da proteína madura, sendo a Ser substituída na proteína de Yang por Pro nesta posição. Conforme demonstrado em seguida, esta sequência de ácidos aminados é efectiva na sua forma não glicosilada. No entanto esta sequência contém dois locais de glicosilação e a forma glicosilada está também incluída no âmbito da presente invenção. Mencione-se que a proteína pode existir sob a forma de um sal de adição de ácido, sob a forma de um sal básico ou pode ser neutra, dependendo do pH do meio. Por derivatização por meio de fosforilação, de acetilação ou por outros meios, desde que a actividade não seja destruída, é obtida uma proteína também incluída no âmbito da presente invenção,

Mencione-se também que pode acontecer que não seja necessária a sequência integral para que se manifeste a actividade. Conforme ilustrado na presente memória descritiva, a alanina na posição 1 pode ser suprimida, podendo também ser suprimidos os primeiros resíduos de ácidos aminados até um máximo de catorze se forem substituídos por uma sequência de resíduos duma sequência peptídica fundida. Além disso crê-se que a forma murina da proteína apenas requer os primeiros 79 resíduos para a actividade; esta situação corresponde aproximadamente aos 83 primeiros resíduos da forma humana. Deste modo, os fragmentos que contém apenas os 83 primeiros resíduos de ácidos aminados da proteína e as suas formas contendo substituições na extremidade N-terminal estão também incluídas no âmbito da presente invenção. Além disso deve ser considerado que a extremidade N-terminal da hIL-3 madura é formada pelos resíduos ala-pro-met etc. (ver figura 1). Sabe-se que a proteína, quando segregada por uma levedura hospedeira, pode, em alguns casos, ser encurtada em dois ácidos aminados (ala-pro), devido à interacção com uma dipeptidil-aminopeptidase (72). A hIL-3 sem a alanina e a prolina



N-terminais conserva ainda a actividade biológica. As estirpes de levedura que contêm uma mutação de anulação no gene de X-prolil-dipeptidilaminopeptidase produzirão hIL-3 completa (ácidos aminados 1 a 133). Deste modo consideram-se incluídos nos multi-CSFs da presente invenção tanto os que contêm como os que não contêm a alanina e a prolina N-terminais, produzidos por hospedeiros mutantes de X-prolil-dipeptidilaminopeptidase e por hospedeiros de tipo selvagem, respectivamente.

Quando a proteína é produzida sob a forma duma proteína madura num hospedeiro procariótico, a sequência de codificação para a proteína madura é prefaciada por um códon de início. A metionina N-terminal resultante pode em seguida ser eliminada, ou removida parcialmente, por processamento no interior da célula hospedeira, dependendo da natureza da sequência de ácidos aminados que se segue. Novamente ambas as formas de hIL-3 são biologicamente activas. Por conseguinte, consideram-se incluídos nos hmulti-CSFs da presente invenção tanto os que contêm como os que não contêm a metionina N-terminal.

Do que ficou dito é claro que podem ser introduzidas alterações de ácidos aminados da proteína IL-3 humana sem afectar a respectiva função biológica. Alterações menores nas sequências de ácidos aminados por modificação química do resíduo codificado, por substituição dum resíduo diferente ou por supressão ou por adição de um ou vários resíduos, de preferência apenas de um, têm como resultado proteínas que conservam a actividade. Deste modo, estas mutações não destrutivas estão também abrangidas na presente invenção, em particular as variações alélicas de ocorrência natural e outras mutações que não sejam letais para a actividade.

Por outro lado, deve ter-se em conta que alterações de ácidos aminados na proteína IL-3 humana podem ser benéficas para a utilização terapêutica da proteína. Tal como reconhecido na presente especificação, a proteína madura contém quatro domínios conservados nos resíduos 15 a 36, 54 a 61, 74 a 91 e 107 a 118. São possíveis proteínas contendo alterações de ácidos aminados simples ou múltiplas nas regiões não conservadas, 1 a 14 (que podem ser, de qualquer modo, substituídas pelas



sequências de proteínas de fusão derivadas do hospedeiro), 37 a 53, 62 a 73, 92 a 106 e 119 a 133. No entanto crê-se que os resíduos de cisteína nas posições 16 e 84 podem ser necessários para formação de pontes de dissulfureto uma vez que estes resíduos são conservados de espécie para espécie. As alterações nos domínios conservados mencionadas anteriormente podem influenciar as propriedades biológicas da proteína, tais como ligação ao receptor e transdução do sinal. Considera-se que a hIL-3 com as propriedades alteradas terá utilidade terapêutica. Estes derivados de hIL-3, que podem ser preparados por intermédio de técnicas de manipulação de proteínas conhecidas, são considerados abrangidos pelo âmbito da presente invenção.

O preparado da proteína pode conter os peptídeos de hmulti-CSF sob uma forma monomérica ou agregada, desde que os agregados mantenham a actividade tal como acima definida.

Na presente memória descritiva o termo "sistema de expressão" diz respeito a uma sequência de ADN que contém tanto a sequência de codificação cuja expressão se pretende como as sequências de regulação apropriadas em ligação operacional com a sequência de codificação referida, de forma a possibilitar a respectiva expressão quando as sequências de regulação são compatíveis com o hospedeiro no qual o sistema de expressão está colocado. Em geral considera-se que o termo "sequências de regulação" diz respeito a segmentos de ADN necessários para regular a expressão da sequência de codificação com a qual esses segmentos se encontram ligados de modo operacional.

As sequências de regulação incluem, em todos os hospedeiros, os promotores, os quais podem ser ou não reguláveis por modificações no respectivo ambiente. Os promotores típicos adequados para procariontes incluem, por exemplo, o promotor de trp (inductível por privação de triptofano), o promotor de lac (inductível com o análogo de galactose IPTC), o promotor de beta-lactamase e o promotor P_L derivado de um fago (inductível por variação de temperatura). Adicionalmente, especialmente para a expressão em Bacillus, são também promotores úteis os promotores de alfa-amilase, de protease, de Spo2 e sequências



promotoras sintéticas. Os promotores adequados para expressão em leveduras incluem o promotor de 3-fosfoglicerato-quinase e os promotores para outras enzimas glicolíticas, bem como as regiões de promotor para desidrogenase de álcool e para fosfatase de levedura. Também são úteis o promotor do factor de elongação de transcrição (transcription elongation factor = TEF) e o promotor de lactase. A expressão em mamíferos geralmente utiliza promotores derivados de vírus tais como os sistemas de promotores de adenovírus e do promotor de SV40, mas também pode utilizar promotores reguláveis, tais como o promotor de metalotioneína, que é regulado pela concentração de metais pesados ou de glucocorticóides. Também se encontram disponíveis sistemas de expressão em células de insectos com base em vírus bem como sistemas de expressão baseados em promotores de células de plantas, como sejam os promotores de sintetase de nopalina.

Além da sequência de ADN do promotor, que é necessária para a transcrição do gene pela polimerase de ARN, são também úteis para a regulação da expressão uma variedade de sequências de regulação, incluindo as que comandam a terminação (por exemplo provocando sequências de poliadenilação em sistemas eucarióticos). Alguns sistemas contêm também elementos de reforço que são convenientes mas não necessariamente indispensáveis para a efectivação da expressão.

Os comandos de tradução incluem um local de ligação ribossômico (ribosome binding site = RBS) em sistemas procarióticos, enquanto que em sistemas eucarióticos a tradução pode ser comandada pela sequência de nucleótidos junto do codão AUG.

De acordo com as considerações anteriores, pode efectuar-se a produção de proteína recombinante numa larga variedade de hospedeiros, incluindo bactérias (predominantemente E. coli, Bacillus e Streptomyces), leveduras e fungos (tais como Saccharomyces, Kluyveromyces e Aspergillus) e culturas de células de mamífero ou outras, tais como células COS, células C127, células de ovário de hamster chinês, células de Spodoptera frugiperda (Sf9) e outras. A proteína pode ser produzida sob a forma duma proteína intracelular madura ou de fusão ou pode ser segregada quando o gene contém incluído o ADN que codifi-

ca para um sinal compatível apropriado.

A presente invenção possibilita pela primeira vez a produção em larga escala de IL-3 humana recombinante, podendo por conseguinte esta proteína ser agora usada - sob uma forma purificada - como agente terapêutico. Os métodos descritos seguidamente proporcionam meios para a produção de formas glicosilada e não glicosilada da proteína, as quais podem ser purificadas para se obter IL-3 humana sob a forma substancialmente pura. A expressão IL-3 humana "purificada" significa IL-3 humana tal como anteriormente definida que se encontra sob uma forma isenta de outras proteínas que normalmente a acompanham.

B. Isolamento de ADNc que codifica para IL-3 humana

A IL-3 humana foi isolada de acordo com a seguinte estratégia:

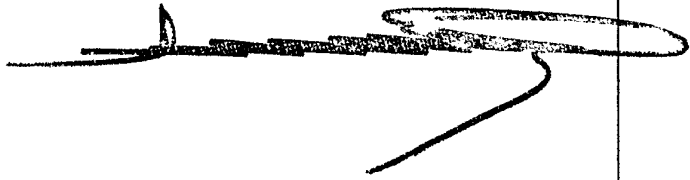
1. Foi desenvolvido um procedimento que possibilita a produção de modo reproductível de factores de crescimento hemopoiético (HGFs) por leucócitos humanos.

2. Preparou-se ARNm a partir das células produtoras referidas e transcreveu-se o mesmo em ADNc de dupla cadeia.

3. Seleccionou-se a partir do ADNc com um ADNc completo de mIL-3 contendo tanto a porção de codificação como a porção 3' não traduzida a juzante de forma a obter DII.

4. Inseriu-se o clone DII de ADNc que hibrida num vector de expressão pLO a fim de obter pLB4, o qual é expresso em células COS para confirmar a presença da sequência que codifica para a IL-3 humana. Meios condicionados obtidos a partir destas células apresentaram a actividade biológica esperada de hIL-3.

O ADNc humano foi isolado usando este procedimento apesar de uma considerável falta de homologia com a sequência de codificação murina porque se verificou existir um grau de homologia surpreendente na região 3' não traduzida. Os requerentes não têm conhecimento de qualquer descrição an



terior da utilização de uma homologia numa sequência 3' não traduzida para o isolamento dum gene duma espécie diferente.

Mais em pormenor, um meio condicionado de linfócitos na presença de 13-acetato de 12-o-tetradecanoilforbol (TPA) e de concanavalina A (Con A) constitui uma origem adequada de HGFs humanos, o que se pode verificar por análise do meio usando uma estimulação de CFU-S de rato em culturas em suspensão, proliferação de células DA-1 dependentes de mIL-3, análise de progenitores hemopoiéticos humanos por formação de colônias in vitro e estimulação in vitro de células blásticas de leucemia aguda. Construiu-se um banco de ADNc a partir de linfócitos humanos no fago gt-10 lambda (20) e seleccionou-se usando o fragmento HindIII-XbaI de ADNc de mIL-3 relativamente à ocorrência de sequências relacionadas com a mIL-3. Não foram identificados quaisquer clones que hibridassem.

No entanto, quando se usou para sonda o ADNc integral de IL-3 murina, foram identificados quatro clones. Por análise por enzimas de restrição do clone maior (D11) verificou-se a existência duma inserção de 910 pb contendo um local EcoRI interno (na posição 411 na figura 1).

(Investigou-se se este local EcoRI era originário de ligação de dois fragmentos independentes de ADNc ou se era um local de ocorrência natural. Por meio de análise de Southern de ADN humano digerido com enzimas de restrição usando fragmentos EcoRI 5' e 3' marcados do clone D11 como sonda, verificou-se a existência de fragmentos de ADN idênticos a seguir a digestão com HindIII (15 kb) e BamHI (4,6 kb). Para além disso, a sequência de ADN junto do local EcoRI não corresponde à sequência do adaptador (pCCGAATTCGG) usado para a inserção do ADNc em ADN fágico, indicando que estes fragmentos EcoRI são derivados dum único ARNm.)

A partir de experiências de hibridação e de sequenciação concluiu-se que os clones menores (II, IV e VI) são idênticos à sequência de nucleótidos 3' do clone D11 e derivam da mesma espécie de ARNm.

Por meio de alinhamento assistido por computador (figura 1) do ADNc de D11 e do ADNc de mIL-3 veri



ficou-se a existência de homologia de sequências na extremidade 5' terminal de 100 pb entre os nucleótidos 236 e 269 e entre os nucleótidos 598 e 803 na região 3' terminal (68%, 71% e 73% de homologia, respectivamente). Em particular, a região entre os nucleótidos 706 e 763 apresenta um alto grau de conservação (93% de homologia) e contém sequências repetitivas ricas em AT. A baixa homologia na extremidade 5' terminal de 600 pb do ADNc humano (52%) exclui a detecção por hibridação com o fragmento HindIII-XbaI da mIL-3.

Por análise do clone de ADNc humano quanto a uma proteína codificada verificou-se a existência dum quadro de leitura livre antes do codão de terminação TGA na posição 495 a 497 (figura 1). O primeiro tripleto ATG é provavelmente o verdadeiro codão de início do polipeptídeo codificado. A presumível proteína codificada consiste num peptídeo guia hidrofóbico de 19 ácidos aminados, o qual provavelmente é cindido entre os resíduos de glicina e de alanina (22 e 23).

O alinhamento dos resíduos previstos de ácidos aminados da IL-3 humana e do rato (figura 1) revela uma homologia de 50% para o peptídeo guia (resíduos -26 a +1) e de 28% para a proteína madura (resíduos 1 a 133). No interior do peptídeo guia existem duas regiões conservadas de quatro ácidos aminados (resíduos -13 a -10 e -3 a +1), dos quais o segundo contém o local de processamento. A proteína madura tem um comprimento de 133 ácidos aminados e tem um peso molecular de 15 kD. A proteína madura tem quatro domínios conservados (resíduos 15 a 36, 54 a 61, 74 a 91 e 107 a 118) e contém dois locais de potencial glicosilação (resíduos 15 a 17 e 70 a 72). Os dois resíduos de cisteína presentes na proteína humana (posições 16 e 84) são conservados e podem desempenhar uma função essencial na dobragem da proteína por formação de ponte de dissulfureto.

A fim de se verificar que este ADNc humano codifica para uma proteína funcional que é semelhante à mIL-3, inseriu-se o ADNc do D11 num vector de expressão eucariótico (pLO, contendo uma unidade de transcrição de SV40) a fim de obter o vector de expressão pLB4 e transfectou-se para células COS 1. Experimentou-se o meio condicionado por COS/pLB4 (CM) com



o fim de (1) verificar a respectiva capacidade para estimular a formação de colônias por células da medula óssea humana e (2) para estimular células blásticas de leucemia mielogênia aguda (AML) humana.

O crescimento in vitro de colônias de progenitores hemopoiéticos humanos desprovidos de células acessórias mielomonocíticas (positivas em relação a Vim-2) e linfocíticas T (positivas em relação a T-3) foi estimulado eficientemente por CM COS/pLB4. Os dados demonstram a existência de estimulação dos progenitores de várias linhagens de diferenciação hemopoiética e de uma subpopulação de BFU-E por CM COS/pLB4.

Numa experiência separada, enriqueceu-se em relação a células progenitoras medula óssea por centrifugação com gradiente de densidade, com sedimentação de rosácea E para remoção de linfócitos T e por aderência para remover os fagócitos mononucleares, e em seguida cultivou-se em meio enriquecido contendo soro fetal de vitela. Nestas condições, a maioria das colônias obtidas após estimulação com CM/COS/pLB4 continham duas ou mais linhagens de diferenciação hemopoiéticas: todas continham macrófagos, aproximadamente metade de células blásticas imaturas e/ou células eritróides imaturas e/ou granulócitos neutrofílicos e uma minoria, adicionalmente, de granulócitos basofílicos e eosinofílicos. Estes resultados demonstram as propriedades de estimulação multilinhagem da proteína codificada pelo clone D11 de ADNc humano e a respectiva acção nas células hemopoiéticas multipotentes de desenvolvimento precoce.

Com respeito à estimulação de AML, foram estimuladas células AML blásticas de cinco pacientes por meio de CM COS/pLB4 e analisou-se para verificar uma resposta por medida da incorporação de ^3H -TdR e de formação de colônias. Três das cinco amostras de células de leucemia responderam ao CM COS/pLB4 em ambas as análises; foram obtidas relações características dose-resposta para a formação de colônias e para a síntese de ADN de células blásticas AML de diferentes pacientes. As respostas a GM-CSF demonstraram adicionalmente diferenças fenotípicas entre as leucemias que responderam a CM COS/pLB4.

Estes dados demonstram que o clone



de ADNc D11 contém a informação genética completa para uma proteína biologicamente activa que é exportada para o meio de cultura nas células COS transformadas. Apesar da aparente falta de homologia com respeito à sequência proteica entre a proteína humana e a mIL-3 (apenas 30%), as proteínas são comparáveis com respeito à respectiva função biológica. Ambas as proteínas exercem o seu efeito em progenitores hemopoiéticos de várias linhagens de desenvolvimento precoce. A baixa homologia ao nível dos ácidos aminados é também reflectida por uma baixa homologia na sequência dos nucleótidos de codificação. No entanto, de forma inteiramente inesperada, verificou-se que existe um alto grau de homologia - suficiente para isolamento do clone de ADNc humano - na região 3' não traduzida.

Por análise de Southern do ADN humano verificou-se a existência de um único gene que hibrida, indicando que este ADNc não pertence a uma família de genes intimamente relacionados.

A partir dos resultados anteriores concluímos que a inserção de ADNc humano em D11 codifica para a proteína humana homóloga de mIL-3. Decidimos usar o termo operacional de hmulti-CSF para a proteína codificada pelo clone D11 de ADNc tendo em vista o seu principal efeito biológico e análise.

A identificação de clones de ADNc de hmulti-CSF em virtude de hibridação com a região terminal 3' do ADNc da mIL-3 era inesperada. Enquanto que as sequências de ADN homólogo são em geral encontradas predominantemente na região de codificação, a sequência de hmulti-CSF divergia extensivamente (45% de homologia) nesta parte do gene. Uma análise do domínio altamente conservado na região não codificante terminal 3' revela a ocorrência de 5 unidades repetidas ATTTA que se conservam todas no ADNc da mIL-3 (figura 1).

As proteínas hmulti-CSF e mIL-3 apresentam um grau de homologia consideravelmente inferior ao de outros factores de crescimento ou linfoquinas murinos e humanos, tais como GM-CSF (25), interleuquina-2 (25), interleuquina-1 (26) e interferões (27 a 29). A actividade biológica da mIL-3 madura



parece estar contida nos primeiros 79 ácidos aminados, incluindo uma exigência absoluta da presença dum resíduo de cisteína na posição 17 (30). Este resíduo de cisteína mantém-se conservado no hmulti-CSF (Fig. 1, pos. 16) e pode desempenhar um papel essencial na dobragem da proteína. A ocorrência dum local de potencial glicosilação na proximidade deste resíduo de cisteína pode interferir com a formação duma ponte de dissulfureto.

C. Produção e formulação de hmulti-CSF

Os requerentes conseguiram uma variedade representativa de sistemas de expressão capazes de produzir a proteína IL-3 humana sob uma variedade de formas - como proteínas de fusão, como proteínas intracelulares maduras e como proteínas segregadas. Os requerentes não têm conhecimento da existência anterior de acordo com o estado da técnica de formas recombinantes de IL-3 humana, ou mesmo de qualquer IL-3 humana num preparado isento das proteínas que normalmente acompanham a proteína pretendida. Por conseguinte, a presente invenção proporciona pela primeira vez a proteína IL-3 humana dum modo capaz de adaptação a utilizações terapêuticas e de diagnóstico.

A IL-3 humana pode ser produzida sob a forma duma proteína de fusão com sequências heterólogas em relação à sequência de ácidos aminados da IL-3 humana. Por "heteróloga" compreende-se uma sequência que não é encontrada na própria IL-3 humana mas é uma sequência não relacionada com aquela. Esta sequência heteróloga pode ser derivada duma proteína bacteriana, duma proteína duma levedura, duma proteína dum mamífero ou pode ser qualquer uma das variedades sequências fortuitamente codificadas tais como, por exemplo, as codificadas por poliadaptadores. É evidente, a partir dos resultados apresentados seguidamente, que pelo menos os primeiros 14 ácidos aminados da extremidade N-terminal da sequência da IL-3 humana podem ser substituídos por uma sequência heteróloga, pelo menos se a proteína de fusão se estende ainda para além da extremidade N-terminal.

A proteína pode também ser obtida sob a forma duma proteína madura intracelular por construção de sequências nas quais o codão de início ATG é colocado imediata-



mente a montante da extremidade N-terminal pretendida. Estas proteínas intracelulares, quer sejam proteínas maduras quer de fusão, podem ser recuperadas por lise das células e por purificação da IL-3 humana mediante o uso de técnicas normais de purificação de proteínas.

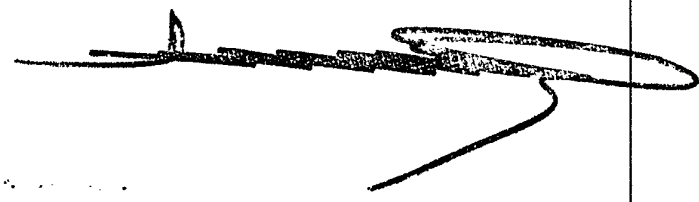
A purificação da proteína é mais simples se a IL-3 humana for segregada para o meio. Quando produzida em células de mamífero com as quais a sequência nativa de sinal é compatível, esta sequência nativa de sinal pode ser usada para provocar a secreção para o meio. Em sistemas bacterianos ou de leveduras, podem ser usadas sequências de sinal compatíveis com estes hospedeiros, tais como as sequências de penicilinase ou de alfa-amilase em bactérias ou a sequência de sinal do factor alfa em leveduras.

Quando produzida por técnicas recombinantes, a IL-3 humana apresenta-se isenta das proteínas que normalmente a acompanham e pode ser purificada das proteínas e de outros materiais indígenas do hospedeiro recombinante usando, por exemplo, métodos cromatográficos, filtração por gel, precipitação por sulfato de amónio e outras técnicas.

Conforme descrito seguidamente, a proteína é útil para fins terapêuticos e de diagnóstico. Para fins terapêuticos a proteína pode ser incorporada em formulações de acordo com as técnicas normalmente utilizadas para a preparação de composições farmacêuticas adequadas para a administração de proteínas. São excipientes adequados, por exemplo, solução salina fisiológica e solução de Ringer, entre outros. Também podem ser empregues formas alternativas de formulação, incluindo formulações sólidas (p. ex. liofilizadas).

D. Preparação de anticorpos

O facto de se poder dispor de proteína IL-3 recombinante ou de fracções desta proteína permite a produção de anticorpos dirigidos contra a proteína ou contra as suas fracções, conforme demonstrado seguidamente. Tais anticorpos são úteis, nomeadamente para a detecção in vitro de colónias produtoras de hIL-3, para uso terapêutico e para a purificação



de hIL-3 tanto de origem natural como recombinante.

Declaração de Utilidade

A sequência de nucleótidos do ADNC da IL-3 humana integral ou de frações deste ADNC ou sequências intimamente relacionadas com este ADN possibilitam com vantagem a detecção de anormalidades genéticas, incluindo rearranjos genômicos, polimorfismos de dimensão de fragmentos de restrição, mutações e expressão de genes alterados mediante o uso de técnicas como a análise de ADN cromossomal usando enzimas de restrição, de técnicas de mancha de ADN e de ARN bem como de técnicas de hibridação (Maniatis et al. 1982) e de electroforese em gel bidimensional (Fisher e Lerman, 1983).

O hmulti-CSF recombinante proporcionado pela presente invenção facilitará uma análise em pormenor da respectiva função na hemopoiese humana, em particular o possível sinergismo do hmulti-CSF e de vários outros HGFs. Para além disso, o hmulti-CSF apresenta um interesse considerável devido à sua aplicabilidade para o diagnóstico in vitro de doenças humanas nas quais as células progenitoras hemopoiéticas se encontram envolvidas, doenças essas que incluem a leucemia, bem como em aplicações terapêuticas potenciais no sentido duma expansão da hemopoiese in vivo. O efeito do hmulti-CSF sobre várias afecções malignas hemopoiéticas com respeito à diferenciação terminal das células leucêmicas também necessita de ser explorado. Além disso o hmulti-CSF pode ser necessário para o estabelecimento dum estado de proliferação de células indiferenciadas humanas em protocolos de terapia genética, uma vez que se verificou que a estimulação com IL-3 é necessária para se conseguir resultados positivos na infecção de células indiferenciadas de rato com retrovírus recombinantes deficientes com respeito à replicação.

A proteína IL-3 também pode ser usada com vantagem para a detecção de células precursoras hemopoiéticas precoces em culturas in vitro (Wagemaker e Visser, 1980; Metcalf et al., 1982; Merchav e Wagemaker, 1984; Metcalf, 1986).

A proteína IL-3 e as respectivas va



riantes podem adicionalmente ser usadas para a multiplicação de células indiferenciadas hemopoiéticas in vitro, possivelmente em conjunção com outros factores de crescimento, para a transplantação de medula óssea e para a manipulação genética de células indiferenciadas (Lowenberg e Dicke, 1977; Wagemaker e Petem, 1978; Lemischka et al., 1986).

A proteína IL-3 pode ser usada para a determinação do padrão de resposta de células hemopoiéticas malignas em experiências in vitro (Touw e Lowenberg, 1985; Griffin et al., 1986; Griffin e Lowenberg, 1986).

A proteína IL-3 pode ainda ser usada para a detecção de células leucémicas remanescentes por métodos in vitro (Touw e Lowenberg, 1986; Griffin et al., 1986; Griffin e Lowenberg, 1986).

Para além disso, a proteína IL-3 pode ser usada in vivo para o tratamento e para a prevenção de afecções malignas e não malignas, quer por si própria quer em combinação, podendo uma resposta específica obtida pelo sistema hemopoiético ter como resultado um benefício clínico.

Estas aplicações incluem:

- citopenias e/ou imunossupressão devida a infecções, como por exemplo no caso de SIDA
- citopenias devidas a quimioterapia e/ou a irradiação
- afecções ósseas tais como fracturas ósseas e osteoporose
- imunodeficiências devidas a procedimentos de anestesia geral
- recuperação a seguir a transplantação de medula óssea
- adjuvante de vacinas e terapia adjuvante para infecções.

A sequência de ADN de IL-3 humana clonada ou de um ADN intimamente relacionado pode ser usada para terapia genética em desvios genéticos do gene normal da IL-3.

A fim de facilitar a análise acima descrita torna-se necessário dispor duma grande quantidade de IL-3 humana. O meio mais fácil de obtenção de quantidades suficientes desta proteína consiste na sua produção por intermédio de microorganismos, em particular de leveduras, de bactérias ou de fungos, por exemplo, espécies de Saccharomyces, Kluyveromyces,



Aspergillus, Streptomyces, Bacillus e E. coli. Também é possível a produção em células de mamíferos ou em outros sistemas eucarióticos, tais como células C127, células de Spodoptera e animais ou plantas transgênicos, por peritos na matéria mediante os ensinamentos da presente invenção.

Com um fim ilustrativo de como é possível obter células humanas que produzem a proteína IL-3 humana por expressão do ADNc da hIL-3, construíram-se vários plasmídeos e transferiram-se estes plasmídeos para E. coli, B. subtilis, B. licheniformis, S. cerevisiae, K. lactis e células C127. Mediante o uso destas estirpes hospedeiras conseguiu-se a produção de IL-3 humana recombinante. Os produtos obtidos foram experimentados para verificação da respectiva capacidade de estimulação de células blásticas de AML humanas, tal como descrito anteriormente para o meio condicionado de COS/pLB4. A partir destas experiências verificou-se que as proteínas preparadas eram biologicamente activas. Os exemplos que se seguem têm por finalidade ilustrar a presente invenção sem que de algum modo restrinjam o seu âmbito.

EXEMPLO 1

Isolamento do ADNc que codifica para o multi-CSF humano (hmulti-CSF)

Leucócitos humanos estimulados por TPA (5 ng/ml) e ConA (10 ug/ml) produziram quantidades consideráveis de HGFs de acordo com determinações efectuadas por meio da análise de proliferação de células indiferenciadas e por várias outras análises de colónias. As células foram separadas 24 h depois da estimulação, uma vez que a produção de ARNm muitas vezes tem um regime transiente a seguir a estimulação com ésteres de forbol e lecitinas. Após 24 h os HGFs já eram facilmente detectáveis no CM.

Preparação de ARNm

Separaram-se as células, lavaram-se com PBS e homogeneizaram-se em solução de isotiocianato de guaní-



dínio (36). O ARN foi isolado através dum leite de cloreto de cé-
sio. Para a selecção dos ARN's usou-se cromatografia em oligo
(dT)-celulose (36).

Síntese de ADNc

Sintetizou-se o ADNc essencialmente de acordo com a técnica de Gubler e Hoffman (37) usando oligo (dT) como iniciador e transcriptase inversa de AMV. A segunda cadeia foi sintetizada com RNaseH e polimerase I de ADN de E. coli. As interrupções foram fechadas com ligase de ADN T4 e as extremidades foram alisadas com polimerase de ADN T4. A fim de proteger os locais de restrição EcoRI internos, metilou-se o ADNc com metilase de EcoRI. Em seguida ligou-se o ADNc a adaptadores de EcoRI fosforilados com ligase de ADN T4. Após digestão com EcoRI o excesso de adaptadores foi eliminado por cromatografia em Sepharose CL-4B. O material isolado no volume vazio da coluna tinha um comprimento superior a 250 pb e foi usado para a construção dos bancos.

Construção do banco de ADNc fágico

Ligou-se o ADNc a braços do fago gt10 lambda (20) e acondicionou-se com extractos de acondicionamento comerciais (Gigapack, Vector Cloning Systems). Os fagos recombinantes foram propagados em E. coli C600 hfl.

Selacção do banco de fagos

De cada cultura contendo 1 a 5000 placas foram preparadas duas réplicas em filtro de nitrocelulose de acordo com os procedimentos normais. Os filtros foram em seguida hibridados com uma sonda de mIL-3 marcada com radioisótopos do fragmento HindIII-XbaI do ADNc da mIL-3 ou com o clone de ADNc de mIL-3 integral marcado com isótopos radioactivos com iniciadores aleatórios. O clone de ADNc da mIL-3 (pL101) foi isolado a partir dum banco de ADNc de WEHI-3B. O ARNm de WEHI-3B foi isolado mediante o uso do método de isotiocianato de guanidínio e CsCl, sendo em seguida fraccionado em função da dimensão em



gradiente de sucrose e injectado em oócitos de Xenopus laevis. As fracções de ARN que induziram a produção de um factor com capacidade para manter a proliferação de células indiferenciadas murinas nos oócitos foram usadas para a síntese de ADNc como anteriormente descrito; acrescentaram-se ao ADNc extremidades com resíduos de dC e inseriu-se no local PstI do plasmídeo pUC9. Identificaram-se os clones de mIL-3 mediante o uso de oligonucleótidos sintéticos (a partir da sequência de mIL-3 publicada, 11). A inserção de pL101 foi purificada em gel de poliacrilamida e foi utilizada para efectuar a selecção do banco de ADNc humano. A sonda de ADN foi marcada utilizando o método do iniciador aleatório (38). As placas potencialmente positivas foram novamente seleccionadas e purificadas. Deste modo foram identificados quatro clones, incluindo o fago D11.

Sequenciação de clones de ADNc

Cultivaram-se fagos recombinantes em grande escala e purificaram-se, separaram-se as inserções de ADNc dos braços dos fagos por digestão com EcoRI e purificaram-se em gel de poliacrilamida. Os fragmentos purificados foram ligados em ADN de M13mp18 e pTZ18R digerido com EcoRI e foram usados para transformação de E. coli JM109. Preparou-se ADN de cadeia simples e sequenciou-se de acordo com os procedimentos habituais (39). Os resultados da sequenciação foram analisados por meio de vários programas de computador (40-43).

A sequência obtida para a inserção no fago D11 é apresentada na figura 1. Esta sequência de 910 pb contém a região de codificação integral para o hmulti-CSF e a respectiva sequência de sinal e apresenta um alto grau de homologia com o clone murino pL101 na região não traduzida 3'. A homologia a montante na sequência de codificação é relativamente mais limitada. Tal como descrito anteriormente, a proteína tem uma presumível sequência de sinal de 19 ácidos aminados seguida pela proteína madura com 133 ácidos aminados contendo dois locais de glicosilação (15 a 17 e 70 a 72) e dois resíduos de cisteína nas posições 16 e 84.

A sequência de ácidos aminados dedu



zida é idêntica à sequência codificada pelo ADN genômico descrito por Yang, Y-C, et al. (ver acima), excepto no que se refere a um ácido aminado - o ácido aminado na posição 8 da presumível proteína madura; o ADN de Yang codifica para Ser enquanto que o ADNc da presente invenção codifica para Pro.

A sequência isenta de intrões obtida no fago D11 pode ser usada para expressão procariótica bem como para expressão em sistemas eucarióticos, de acordo com a ilustração apresentada seguidamente.

EXEMPLO 2

Expressão em células de mamífero

A. Construção do vector de expressão eucariótico pLB4

Digeriu-se o fago D11 (contendo a inserção maior de ADNc) com HindIII e BglII e subclonou-se no plasmídeo pT1 (um derivado do plasmídeo pTZ18R contendo vários locais de restrição adicionais no multi-adaptador, ver Exemplo 3A). Identificaram-se por análise de restrição clones contendo o fragmento fágico que continha a inserção de ADNc. A inserção de ADNc foi retirada deste plasmídeo por digestão parcial com EcoRI e foi purificada por electroforese em gel de poliacrilamida. O fragmento apropriado foi inserido num vector de expressão eucariótico (pLO) numa unidade de transcrição de SV40.

O vector pLO é constituído por: EcoRI (preenchido) - PstI de pBR322 (1-755), PstI-AvaI de pBR329 (756-1849), adaptador AvaI-PvuII (1850-1868), PvuII-HindIII (preenchido) de SV40 (promotor) (1869-2211), adaptador PvuII-BamHI contendo um sítio singular EcoRI (2211-2251), fragmento de encaixe MboI de SV40 (2252-2861), "fragmento poli A" BclI-BamHI (preenchido) de SV40 (2862-3098), fragmento promotor PvuI-HindIII de SV40 (3099-3440), gene Eco gpt HindIII-BamHI (3441-4501), fragmento de encaixe MboI de SV40 (4502-5111) e o "fragmento poli A" BclI-BamHI (preenchido) de SV40 (5112-5348).

A unidade de transcrição Eco gpt



não tem qualquer importância na expressão transiente de proteínas em células COS 1. O plasmídeo de expressão resultante para hmulti-CSF foi denominado pLB4 e foi purificado em CsCl. Este plasmídeo foi depositado em E. coli no Centraal Bureau von Schimmelcultures (CBS), Baarn, Países Baixos, de acordo com as disposições do Tratado de Budapeste em 12 de Dezembro de 1986 sob o número 568.86. A construção é apresentada na figura 2.

B. Expressão de hmulti-CSF em células COS 1 e análises biológicas

Transfectou-se o ADN do vector pLB4 em células COS 1 usando o método de co-precipitação com fosfato de cálcio (45). As células foram cultivadas durante 48 a 72 horas em meio alfa contendo 10% de soro fetal de vitela. O meio de cultura foi separado e filtrado, sendo usado em análises para determinação da respectiva actividade biológica. Levaram-se a efeito análises por meio de colónias de células progenitoras de medula óssea e de células blásticas mielóides agudas e análises de proliferação do modo seguinte. A medula óssea foi obtida a partir de voluntários adultos hematologicamente normais por punção ilíaca posterior após consentimento consciente. As células mononucleadas foram separadas por centrifugação com gradiente de densidade obtido com ficol (Nijegaard and Co., Oslo, Noruega), sendo em seguida lavadas e suspensas novamente em solução salina equilibrada de Hanks (Hanks balanced salt solution = HBSS). Nesta fase as células mielóides e os linfócitos T foram separados. Para este fim as células da medula foram lisadas após incubação com anticorpos monoclonais OKT-3 (CD3; Ortho, Ravitan, N.Y.) e Vim 2 (células mielo-monocíticas, 46) a concentrações de saturação na presença de complemento de coelho (40%; 30 minutos, 25°C) de acordo com procedimentos normais (47). As células foram lavadas duas vezes em HBSS, suspensas novamente em meio de Dulbecco modificado de acordo com Iscove (Iscove's modified Dulbecco's medium = IMDM) e cultivadas na presença de plasma autólogo de acordo com Fauser e Messner (16), tal como descrito anteriormente (48), até uma concentração de 1,5 a 3×10^4 /ml. Adicionaram-se 1 U/ml de eritropoietina (carneiro, fase III, Connaught, Willowdale, Canadá) e CM COS/pLB4 para estimulação do crescimento. Também são a-



presentados resultados de culturas normalizadas com leucócitos CM estimulados por fito-hemaglutinina (PH-LCM) em comparação directa com CM COS/pLB4. Colheram-se sessenta por cento das colónias e identificaram-se por análise ao microscópio. O CM das células COS transfectado com o vector sem inserção (pLO) não exerceu por si só qualquer acção de estímulo da formação de colónias.

Os resultados são apresentados na figura 3. Na figura apresentam-se os números médios de colónias de eritróides (BFU-E), de granulócitos-macrófagos (CFU-GM), de granulócitos (CFU-G), de eosinófilos (CFU-Eo), de macrófagos (CFU-M) e de colónias mistas (CFU-MIX) ($\pm$ DP) em culturas realizadas em duplicado estimuladas com volumes crescentes de CM COS/pLB4.

Indução da proliferação de AML (ver figura 4)

Purificaram-se células blásticas AML por gradiente de densidade usando uma albumina de bovino (BSA). Os linfócitos T residuais foram removidos das amostras de AML por sedimentação de rosácea E (17, 49, 50). Determinou-se a formação de colónias de AML (paciente 1) não apenas no sistema de alimentação de leucócitos PHA estabelecido (PHA 1.f) mas igualmente numa versão modificada da técnica na qual os leucócitos foram substituídos por CM COS/pLB4, permitindo a avaliação da respectiva actividade de estimulação das colónias (17, 18, 49, 50), tal como se apresenta na figura 4A. Todas as experiências foram levadas a efeito em triplicado. A síntese de ADN de AML blásticos (paciente 2) foi determinada por meio da integração de timidina, como descrito (51) com resultados apresentados na figura 4B. Ambas as determinações mostraram uma relação de dependência da dose para o CM COS/pLB4 adicionado. A adição de meio COS de comparação não teve qualquer efeito sobre a proliferação de AML em qualquer das análises.

C. Construção do vector de expressão eucariótico pLB4/BPV

A fim de estabelecer linhas de células estáveis que efectuassem a expressão de IL-3 humana, trans-



fectaram-se células C127 (ATCC CRL 1616) com um derivado de pLB4. Este derivado foi construído por inserção do genoma BPV-1 integral (69) no plasmídeo pLB4 de acordo com a seguinte estratégia. O fragmento BamHI de BPV-1 foi retirado do vector p δ BPV-MMTneo (342-12) (70). As extremidades coesivas BamHI foram preenchidas mediante o uso de polimerase de Klenow. Em seguida o vector pLB4 foi cortado no local EcoRV singular dentro do gene Eco gpt. Em seguida o fragmento BPV-1 de extremidades alinhadas foi clonado no plasmídeo pLB4 cortado com EcoRV, obtendo-se como resultado o vector pLB4/BPV que tem capacidade de se replicar em células C127. Transfectou-se o vector pLB4/BPV para células C127 usando o método de precipitação de fosfato de cálcio (45). As células transfectadas foram cultivadas durante 16 dias e em seguida os focos foram colhidos das placas de cultura. Foram estabelecidas várias linhas de células independentes. Verifica-se que o vector pLB4/BPV é mantido de forma estável no interior das células, de acordo com análise de mancha de Southern de extractos de Hirt (71) de várias linhas de células. Experimentou-se a actividade de IL-3 do meio de cultura condicionado por meio da análise de proliferação de AML. As linhas de células estáveis produzem IL-3 humana activa.

EXEMPLO 3

Construção de vectores de expressão de E. coli

A. Construção de pGB/IL-301 (ver figuras 5, 6, 7 e 8)

Para a construção de vectores de expressão de E. coli introduziram-se as seguintes modificações nos procedimentos normais (36).

1. Suprimiram-se as sequências terminais 3' não codificantes entre o local AvaI (posição 541) e o local XhoI (posição 856) no plasmídeo pLB4 por fusão dos fragmentos de ADN a seguir ao preenchimento das extremidades coesivas com enzima de Klenow (figura 5).

2. Para introdução da inserção de hmulti-CSF num vector de expressão bacteriano, levaram-se a efeito



to as seguintes fases. Digeriu-se o vector pLH1 com AvaI e preencheram-se as extremidades reentrantes com polimerase de Klenow. Após ligação dum adaptador de BglII (CAGATCTG), digeriu-se o ADN com BglII e BamHI. Purificou-se o fragmento BglII-BamHI de hmulti-CSF em gel de poliacrilamida e subclonou-se no local BglII do plasmídeo pT1, um derivado do plasmídeo pTZ18R (Pharmacia) modificado no local de clonagem múltipla (ver figura 6). Foram obtidos dois clones contendo a inserção em orientação oposta em relação ao promotor de lacZ (ver figura 5). Isolaram-se inserções destes dois clones em gel de poliacrilamida após digestão com BglII e EcoRV e subclonaram-se no plasmídeo pT1 digerido com BglII e HindII. A junção do adaptador de BglII e do ADN de hmulti-CSF foi verificada por análise de sequência verificando-se a existência duma junção do adaptador ao local AvaII localizada no nt 1 do clone de ADNc (este local AvaII resultou da ligação do adaptador de EcoRI com a molécula de ADNc). Uma vez que esta construção (pGB/IL-300) não estava em fase com a proteína lacZ, subclonou-se a inserção BglII-EcoRV no plasmídeo pUC8 (52) digerido com BamHI e HindII. A construção resultante (pGB/IL-301, ver figuras 5, 7 e 8) foi experimentada para verificação da produção da proteína de fusão lacZ/hmulti-CSF.

B. Construção dos vectores pGB/IL-302, pGB/IL-303, pGB/IL-304 e pGB/IL-305 (figuras 5, 7 e 8)

Introduziram-se várias alterações de bases na sequência de codificação para a parte N-terminal das proteínas de fusão por introdução de oligo-nucleótidos sintéticos no vector pGB/IL-300. Os novos vectores de expressão denominados pGB/IL-303 e pGB/IL-304 foram construídos do modo seguinte: Isolou-se o fragmento Hind II - Hind III do vector pGB/IL-300 em gel de agarose e ligou-se este fragmento num oligonucleótido sintético contendo os nucleótidos 99 a 137 do hmulti-CSF e uma sequência de reconhecimento Sal I terminal 5, inserindo-se no vector pTZ18R digerido com Sal I e Hind III. Analisaram-se as sequências de vários clones. Constatou-se a ocorrência de várias alterações de bases tendo como resultado modificações na proteína hmulti-CSF. Transferiram-se para o vector pUC8 inserções de vá-



rios clones para a expressão da proteína de fusão lacZ (pGB/IL-302, pGB/IL-303). Fez-se com que o clone pGB/IL-304 ficasse em fase com a sequência lacZ por ligação do local Sal I após preenchimento das extremidades reentrantes. A construção foi verificada por digestão com Pvu I. Vários clones tinham um oligonucleótido sintético a menos e verificou-se que estavam fundidos em fase com a proteína lacZ. Um destes clones foi denominado pGB/IL-305.

C. Construção de pGB/IL-306 (ver figuras 5, 7 e 8)

Preparou-se um vector de expressão que codifica para uma proteína sem os ácidos aminados na extremidade N-terminal de lacZ a partir de pGB/IL-300 por meio dum ciclo de supressão tal como descrito em (53). O oligonucleótido sintético continha 22 nucleótidos a montante do gene lacZ de PTZ incluindo o codão de início ATG e os primeiros 24 nucleótidos que codificam para a IL-3 madura. Este plasmídeo foi designado por pGB/IL-306 (figuras 5, 7 e 8).

Foram depositadas no CBS em 13 de Julho de 1987 estirpes de E. coli contendo os plasmídeos pGB/IL-300, pGB/IL-301 e pGB/IL-302 sob os números CBS 377.87, CBS 379.87 e CBS 378.87, respectivamente.

A figura 8 mostra a sequência de regiões de fusão para os diversos plasmídeos construídos. A sequência dos clones é dada a partir do início da região de codificação da proteína lacZ tanto no vector pUC8 como no pTZ18R (letras minúsculas) e na região de codificação do hmulti-CSF (letras maiúsculas) até ao local Cla I na posição 158. As mutações na sequência do ADN do hmulti-CSF estão sublinhadas; estas mutações têm como resultado as alterações $trp^{13} \rightarrow arg^{13}$ (pGB/IL-302); $leu^9 \rightarrow pro^9$ e $trp^{13} \rightarrow arg^{13}$ (pGB/IL-303); $net^3 \rightarrow thr^3$ e uma modificação silenciosa (pGB/IL-304).

No pedido de patente da prioridade EP 87201322.2, apresentado em 13 de Julho de 1987, foram usadas outras designações para estes plasmídeos, do modo seguinte:

pGB/IL-300 = pT-hIL3;
pGB/IL-301 = pUC/ hmulti;



pGB/IL-302	=	pUC/ hmulti Δ 1A;
pGB/IL-303	=	pUC/ hmulti Δ 1B;
pGB/IL-30	=	pUC/ hmulti Δ 1C;
pGB/IL-30	=	pUC/ hmulti Δ 2
pGB/IL-30	=	pTZ/ hmulti;

D. Expressão das proteínas de fusão lacZ/ hmulti-CSF e hmulti-CSF madura em E. Coli

Cultivaram-se estirpes de E. coli (JM 109) contendo 50 g/ml de ampicilina a 37° C até se atingir uma densidade óptica de 0,5 a 550 nm. Em seguida adicionou-se IPTG (isopropil-beta-tiogalactosido, Pharmacia) à cultura até uma concentração final de 1 mM e continuou-se a incubação durante 3 a 4 horas.

Os plasmídeos pGB/IL-306 e pGB/IL-302 foram também utilizados para transformação em E. Coli DH1 (operação lacZ tipo selvagem). As estirpes foram cultivadas em meio LB ou em meio TY 2 x contendo 50 µg/ml de ampicilina a 37° C durante 16 horas.

As bactérias foram separadas por centrifugação e foram desintegradas por ultra-sons em tampão contendo Tris/HCl 0,1 M a pH 8,0; EDTA 5 M, Nonidet p40 (NP-40) a 0,2% e fluoreto de fenil-metilsulfonilo (PMSF) 1 M. Centrifugou-se durante 30 min. a 20 000 x g. Por electroforese em gel de poliacrilamida das fracções do sólido e do sobrenadante verificou-se que a maior parte das proteínas hmulti-CSF se encontra na bactéria sob uma forma insolúvel.

Extraiu-se novamente a fracção sólida com tampão com 0,5% de NP-40 e em seguida solubilizou-se com ureia 8 M, Tris/HCl 0,1 M a pH 8,0 e ditiotreitol 5 M. Deste modo conseguiu-se uma purificação considerável das proteínas de fusão (figura 9).

Tal como se pode verificar pela figura isolaram-se corpos de inclusão das bactérias (E. Coli) contendo os plasmídeos pGB/IL-301 e pGB/IL-302 tal como se descreve. As pistas referem-se ao sobrenadante de 0,2% de NP-40 (a amostra corresponde a 0,1 ml da cultura de bactérias original). As pis-



tas 2 referem-se ao sobrenadante de 0,5% de NP-40 (0,2 ml) e as pistas 3 ao sólido da centrifugação solubilizado em tampão de ureia 8 M (A: 0,05 ml; B 0,2 ml). As proteínas foram separadas num gel de poliacrilamida com 13,5% de SDS e foram coradas com Azul Brilhante de Coomassie. À direita representam-se os pesos moleculares (em kd) das proteínas de marcação (pista M). As proteínas de fusão de multi-CSF humano são indicadas por flechas. A proteína de fusão codificada pelo plasmídeo pGB/IL-301 tem um peso molecular, tal como esperado, de cerca de 20 kd ; a proteína de fusão produzida por intermédio do plasmídeo pGB/IL-302 tem um p.m. de cerca de 16 kd.

E. Determinação da actividade biológica de preparados de hmulti-CSF bacteriano

Diluíram-se preparados da proteína bacteriana em meio alfa contendo 1% de albumina de soro de bovino, esterilizaram-se por filtração e analisaram-se por meio da análise de proliferação de AML blásticos. Adicionaram-se amostras diluídas a células blásticas AML e cultivaram-se durante quatro dias. A síntese de ADN foi medida usando ^3H - timidina, como descrito (51). Uma unidade por ml é definida como a quantidade de hmulti-CSF necessária para a proliferação de metade do máximo de células blásticas AML. A figura 10 apresenta esta análise. Analisaram-se várias diluições da preparação de proteína extraída por ureia de bactérias contendo o plasmídeo pGB/IL-302 para verificação da estimulação da proliferação de células blásticas AML usando ^3H -timidina. A concentração de proteína de fusão deste preparado proteico era de 33 $\mu\text{g/ml}$. Com base na curva de titulação apresentada calculou-se a actividade deste preparado em 16 000 unidades/ml.

A quantidade de proteína de fusão bacteriana nos preparados foi estimada a partir de electroforese em gel de poliacrilamida e usou-se o resultado para cálculo das actividades específicas.

Os resultados são apresentados no quadro seguinte:

.
. .
.

QUADRO I

Actividade Biológica de Preparados de hmulti-SCF bacteriano

	P.m. ($\times 10^{-3}$) de lacZ/ hmulti (1)	$\mu\text{g/de}$ proteína por ml (2)	Unidades por ml (3)	Actividade es- pecífica em unidades pr mg de IL-3
pGB/IL-301	20	20	45	4,500
pGB/IL-302/303	16	5	2400	480,000
pGB/IL-304	18	ND (4)	18	-
pGB/IL-305	16	1	300	300,000
pGB/IL-306	15	ND	70	ND

1. Os pesos moleculares aproximados foram estimados a partir da sequência de ADN correspondente à proteína de fusão (figura 8).
2. As concentrações de IL-3 foram estimadas em gel de SDS-polia-
crilamida e foram calculadas por ml da cultura de origem.
3. A actividade da proteína solubilizada por ureia foi determina-
da na análise de proliferação de AML e está expressa por ml
da cultura de origem.
4. Não determinado.

A partir dos resultados anteriores concluiu-se que o multi-CSF humano expresso sob a forma duma pro-
teína de fusão em E. Coli foi obtido sob uma forma biologicamen-
te activa. Os resultados mostram que as alterações introduzidas
na extremidade N-terminal das proteínas de fusão podem influenci-
ar a actividade específica destas proteínas.

EXEMPLO 4

Preparação de preparados de anticorpos com capacidade de reacção
imuno específica com a proteína IL-3 humana

A. Anti-soro IL-3 anti-humano de coelho policlonal

Fez-se um gel preparativo a partir



lisado de E. Coli contendo o plasmídeo pGB/IL-301. A banda a 20 kd com a proteína de fusão IL-3 foi destacada, moída num almofariz com solução salina e emulsificada numa proporção de 1:1 com adjuvante de Freund completo contendo 1 mg de Mycobacterium tuberculosis H37RA por ml.

Imunizaram-se coelhos brancos da Nova Zelândia (spf) com 1 ml da emulsão (com $\pm$ 100 μ g da proteína de fusão de IL-3) dividida por 5 locais de injeção (2 x i.m. nas coxas, 3 x s.c. na região lombar). Administraram-se injeções de reforço do mesmo antígeno em adjuvante de Freund incompleto nas semanas 2, 4 e 6. Colheu-se soro na semana 8 por punção da veia da orelha.

Um volume de soro foi absorvido com 9 volumes de E. Coli contendo pUC8 tratado por ultra-sons (dum dia para o outro a 4°C) a fim de remover os anticorpos não específicos. Por análise de mancha imunológica de todas as construções de IL-3 feitas em E. Coli, B. licheniformis, B. subtilis, S. cerevisiae e K. lactis verificou-se a ocorrência de reação imuno-específica com os soros absorvidos numa diluição de 1 em 6500.

Alguns destes resultados são apresentados na figura 11. As proteínas foram isoladas a partir dos hospedeiros recombinantes tal como descrito anteriormente e foram separados num gel de poliacrilamida a 13,5% e depositadas sobre uma membrana de nitrocelulose. Pista 1: E. Coli contendo pTZ18R (comparação); pista 2: pGB/IL-301; pista 3: pGB/IL-301; pista 4: pGB/IL-302; pista 5: pUC19 (comparação); pista 6: pGB/IL-301; pista 7: pGB/IL-302. As pistas 6 e 7 mostram as proteínas presentes no sólido de centrifugação após tratamento por ultra-sons das bactérias. As pistas 3, 4 e 5 mostram as proteínas presentes no sólido de centrifugação após a primeira fase de lavagem. As pistas 1 e 2 mostram as frações proteicas finais solubilizadas por ureia.

As flechas indicam as proteínas de fusão (da dimensão esperada) expressas pelos plasmídeos pGB/IL-301 e pGB/IL-302.

A figura 12A apresenta a inibição

da proliferação de células blásticas AML dependente de IL-3 por anti-soro anti-IL-3.

A figura 12B mostra que o soro pré-imune não afecta a acção da IL-3 sobre a proliferação de células blásticas AML. Em ambos os painéis utilizou-se a simbologia: ▲ = IL-3 a 10 U/ml; ■ = IL-3 a 1 U/ml; ● = comparação sem qualquer adição.

A figura 12A mostra que o crescimento dependente de IL-3 na análise de proliferação de células blásticas AML (51) foi inibido pelo soro de um modo dependente da dose; a figura 12B mostra que o soro pré-imune não tem este efeito. Como comparação, o crescimento dependente de GM-CSF não foi afectado por estes soros do mesmo modo (figura 12A, em que ◆ = GM-CSF a 100 U/ml).

B. Anticorpos monoclonais de rato anti-IL-3 humana

Imunizaram-se ratos Balb/C com 3 x 0,1 ml (s.c.) da mesma emulsão utilizada para os coelhos. Deu-se um reforço (0,1 ml i.p.) do antigene em adjuvante de Freund incompleto na 2ª semana e três dias mais tarde fundiram-se linfócitos de baço com células de mieloma SP2/0 de acordo com as técnicas normais (65). Seleccionaram-se os hibridomas resultantes por análise de imunosorvente ligado a enzima (ELISA) usando um lisado de E. Coli pGB/IL-302 (contendo o produto de fusão da IL-3 de 17 kd) para referência de comparação positiva e um lisado de E. coli pUC8 para referência de comparação negativa. No total seleccionaram-se 29 culturas de hibridoma de IL-3 que segregavam anticorpos específicos para IL-3, tendo-se estabilizado estas culturas.

EXEMPLO 5

Construção de vectores de expressão de Bacillus

Utilizaram-se técnicas gerais de clonagem (36).

A. Construção de pGB/IL-307 (figura 13)



Para a construção do vector pGB/IL-307 ligou-se o fragmento Sma I menor do plasmídeo pLB4 contendo o gene de hmulti-CSF ao plasmídeo pUB110 (54) digerido com PvuII. Após transformação de células competentes (56) de DB 105 (uma estirpe spo^- derivada da estirpe DB104 deficiente em relação a protease (55)), foram obtidos dois clones, como se esperava: o fragmento tinha sido clonado nas duas orientações possíveis. O plasmídeo que continha o fragmento na orientação correcta com respeito ao designado "promotor de Hpa II" (57) foi denominado pGB/IL-307. Neste caso resulta uma proteína de fusão (ver figura 13).

B. Construção de pGB/IL-310

Preparou-se um plasmídeo de expressão hmulti-CSF como descrito seguidamente.

1. Clonagem do promotor (figura 14)

Para expressão em *Bacillus* usou-se um promotor sintético σ^{43} como descrito (58) (o promotor é habitualmente designado por σ^{55}).

Digeriram-se os plasmídeos pPROM55s (58), o plasmídeo contendo o promotor, e pGPA14 (59) com EcoRI e XbaI. O fragmento do promotor foi ligado no fragmento do vector que tinha sido purificado num gel de agarose. Após transformação de *E. Coli* (JM 101) obteve-se o plasmídeo correcto, que foi designado pGB/IL-308 (fig. 14).

2. Introdução dum oligonucleótido sintético em pGB/IL-308 (figura 15)

Ligou-se no plasmídeo pGB/IL-308 digerido com SalI-Xma III um oligonucleótido sintético contendo os nucleótidos 39 a 158 e 484 a 546 do hmulti-CSF, uma sequência de reconhecimento Sal I terminal 5' e um local Xma III terminal 3'. A mistura de ligação foi introduzida na estirpe JM 101. Após análise de várias transformantes, encontrou-se o plasmídeo correcto a que se deu o nome de pGB/IL-309.

3. Introdução de hIL-3 (figura 16)

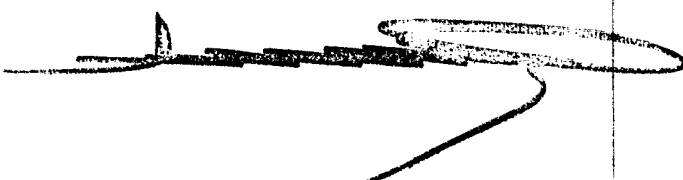


Após transformação e isolamento a partir de B. subtilis DB 105, digeriu-se o plasmídeo pGB/IL-309 com Xma III. As extremidades reentrantes foram preenchidas com polimerase de Klenow e o plasmídeo foi cortado com ClaI. Digeriu-se o plasmídeo pGB/IL-307 com AvaI, preencheram-se as extremidades com Klenow e em seguida digeriu-se com ClaI. Em seguida ligou-se o fragmento contendo o hmulti-CSF no fragmento de pGB/IL-309 e transformou-se em JM101. O plasmídeo resultante foi denominado pGB/IL-310 (figura 16). Este plasmídeo continha o gene de hIL-3 com a sua própria sequência de sinal. Após isolamento do plasmídeo correcto, este foi também introduzido em B. subtilis DB 105.

C. Construção de pGB/IL-311 e pGB/IL-312 (figuras 17 e 18)

Digeriu-se parcialmente o plasmídeo pGB/IL-310 com Hind III e totalmente com Pvu II. Os dois fragmentos Pvu-II - Hind III contendo limulti-CSF foram digeridos com Hind III e Sma I.

A figura 17 apresenta a sequência de nucleótidos do plasmídeo pBHA1. O plasmídeo é constituído por: posições 11 a 105 e 121 a 215; terminador de bacteriófago FD (duplo): posições 221 a 307; uma parte do plasmídeo pBR322 (com as posições 2069 a 2153): posições 313 a 768; origem de replicação do bacteriófago F1 (com as posições 5482 a 5943): posições 772 a 2571; parte do plasmídeo pBR322 com a origem de replicação e o gene de beta-lactamase: posições 2572 a 2772; terminador de triptofano (duplo): posições 2773 a 3729; transposição Tn9, o gene de acetiltransferase de cloranfenicol. Os nucleótidos nas posições 3005 (A), 3038 (C), 3302 (A) e 3409 (A) diferem da sequência de codificação do tipo selvagem. Estas mutações foram introduzidas a fim de eliminar os locais NcoI, BalI EcoRI e Pvu II: posições 3730 a 3804; local de clonagem múltipla: posições 3807 a 7264; parte do plasmídeo pUB110 com a função de replicação e o gene de resistência à canamicina (fragmento EcoRI-Pvu II) (66,67): posições 7267 a 7331; local de clonagem múltipla. Os fragmentos foram reunidos por técnicas de clonagem conhecidas, por exemplo por preenchimento das extremidades coesivas com Klenow, por clonagem



de adaptadores, etc.. Todos os dados foram recebidos de Genbank^R, National Nucleic Acid Sequence Data Bank, NH, EUA.

Após transformação em JM101 e análise de várias colônias resistentes à ampicilina, encontraram-se dois plasmídeos diferentes: pJB/IL-312 que contém o gene completo com sequências de regulação completas e pGB/IL-311 que contém o gene completo e o promotor sem a região -35 em orientação contrária (ver figura 18).

O plasmídeo pGB/IL-311 foi utilizado para transformação de B. subtilis DB105 e da estirpe B. licheniformis T399 (Δ amy, spo⁻, exo⁻ protease-negativo, rif^r, ver ref. 68 em que esta estirpe é denominada T9).

D. Construção de pGB/IL-313 (figura 19)

A fim de obter um plasmídeo menor com o gene de hmulti-CSF depois do "promotor de HpaII", digeriu-se o plasmídeo pGB/IL-312 com BamHI e ligou-se novamente. A mistura da ligação foi utilizada para transformar células competentes DB105. Analisaram-se várias colônias resistentes à neomicina e obteve-se o plasmídeo correcto. Este plasmídeo foi denominado pGB/IL-313.

E. Construção de pGB/IL-317 (figura 20)

A fim de clonar o gene de hmulti-CSF depois da região de início de transcrição e de tradução de alfa-amilase de B. licheniformis e da sequência de sinal, utilizou-se um dos vectores pOL5-delta já descritos (68) nomeadamente o vector pOL5-2-delta. Além da sequência de sinal de alfa-amilase (com um comprimento de 29 ácidos aminados) este plasmídeo contém um ácido aminado da sequência da alfa-amilase madura (uma Ala) seguido dum local de clonagem múltipla: EcoRI - XmaIII - XmaI - Sal I - Hind III (68).

O fragmento Sal I - Pvu II do plasmídeo pGB/IL-310 contendo o gene de hmulti-CSF foi ligado no vector pOL5-2-delta digerido com Sal I - Pvu II e utilizado para transformar a estirpe DB105. O plasmídeo resultante foi denominado pGB/IL-317 (figura 20). O gene de hIL-3 ainda contém a sua



própria sequência de sinal neste plasmídeo. O plasmídeo foi também introduzido na estirpe B. licheniformis T 399.

F. Expressão de cinco plasmídeos de expressão em estirpes de Bacillus

Cultivaram-se estirpes de B. subtilis e de B. licheniformis contendo os plasmídeos de expressão abaixo mencionados em meio TSB contendo 20 µg/ml de neomicina ou 10 µg/ml de eritromicina a 37° C (durante 16 a 24 horas); centrifugaram-se 300 µg/ml da cultura. O sólido foi novamente suspenso em tampão de amostras e foi analisado por meio de electroforese em gel de poliacrilamida seguida e análise de mancha de Western. Precipitou-se o sobrenadante com TCA e suspendeu-se de novo o sólido em tampão de amostras. Tanto o sobrenadante como o sólido foram analisados para detectar a proteína IL-3 (ver Quadro 2).

A fim de determinar a actividade biológica das proteínas produzidas, levaram-se a efeito as seguintes fases operacionais: As células foram novamente suspensas num tampão contendo Tris/HCl 0,1 M a pH 8,0 e Mg Cl₂ 10 mM. Adicionou-se lisozima até uma concentração final de 1mg/ml e PMSF até uma concentração final de 1 mM. Incubou-se a solução durante 30 min. a 37° C. Em seguida adicionou-se DNase (concentração final 20 µg/ml) e incubou-se a solução durante 15 min. a 20° C. Por fim determinou-se a actividade biológica desta preparação bem como do sobrenadante da cultura celular, como descrito. Os resultados são apresentados no Quadro 2.

QUADRO 2

Expressão de vectores de Bacillus

Plasmídeo	Estirpe	P.M. da IL-3		Actividade Bioló- gica	
		Sólido (kd)	Sobrenadante (kd)	Sólido	Sobrenadante
pGB/IL-307	DB105	21	-	+	-
pGB/IL-310	DB105	15;17	15;17	-	-
pGB/IL-311	DB105	12.5;15	-	+	-
	T399	-	-	+	-
pGB/IL-313	DB105	15;17	12.5;15	+	-
	T399	-	-	+	-
pGB/IL-317	DB105	12.5;15	12.5;15	+	+
		17;20	17		
	T399	12.5;15	12.5;15	+	+
		17;20	17		

Pode concluir-se que em B. subtilis usando o plasmídeo pGB/IL-307 se produz uma proteína de fusão com actividade de IL-3. Quando o gene de IL-3 humana contém apenas a sua própria sequência de sinal não se obtém qualquer secreção significativa de IL-3 humana. Toda a actividade de IL-3 é encontrada intracelular. Nestes casos parece que além de precursor de IL-3 se formou IL-3 madura (15 kd) na célula. Deste modo pode ter tido lugar algum transporte através da membrana mas a proteína não é transportada através da parede celular. No entanto, usando os sinais de regulação e de secreção de alfa-amilase (pGB/IL-317) verifica-se que a maior parte da actividade de IL-3 é segregada para o meio de cultura. Para além dum produto de degradação, detectam-se duas proteínas no sobrenadante, uma de cerca de 15 kd e uma cerca de 17 kd, com toda a probabilidade IL-3 madura e IL-3 precursora, respectivamente. Estes resultados indicam que são usados os dois locais de processamento, a saber os locais de processamento de alfa-amilase e de hmulti-CSF. Na célula, o produto mais abundante é a IL-3 precursora, contendo a sequência de

— sinal de alfa-amilase (a proteína de 20 kd), tal como se constata por análise de mancha de Western. Por vezes detecta-se um produto de degradação.

EXEMPLO 6

Construção de vectores de expressão de Kluyveromyces lactis

A. Construção de pGB/IL-316

Inseriu-se no local Sma I do vector pUC19 (63) um fragmento de ADN contendo o gene Tn5 (61), que confere resistência à gentamicina G 418, sob a direcção do promotor de álcool-deidrogenase I (ADHI) de S. cerevisiae, semelhante ao descrito por Bennetzen e Hall (62). Uma estirpe de E. Coli contendo o plasmídeo obtido, pUC - G418, foi depositada no CBS em 4 de Dezembro de 1987 sob o nº. CBS 872.87.

Inseriu-se no vector pUC-418 cortado com XbaI - Hind III do plasmídeo pGB903 (64) contendo o promotor de K. lactis e ADN de proquimosina de vitela, dando origem ao plasmídeo pGB/IL-314.

O fragmento SalI-Hind III deste plasmídeo foi substituído por um fragmento de ADN sintético contendo um pequeno local de clonagem múltipla e o terminador de lactase (ver figura 21 e 22). O plasmídeo resultante foi denominado pGB/IL-315.

No vector pGB/IL-315 cortado por Sac II - Xho I ligaram-se os seguintes fragmentos:

1. O fragmento Sac II - Xba I de pKs 105 (pedido de patente U.S. Ser. 078 539 64) contendo a fracção 3' do promotor de lactase e a fracção 5' da sequência de sinal do factor alfa de S. cerevisiae.

2. Um oligonucleótido sintético contendo a fracção 3' da sequência de sinal do factor alfa com início no local XbaI e a fracção 5' da sequência de ADNc da hIL-e madura até à metade 5' do local Hpa I (resíduo de aa nº. 14).

3. O fragmento HpaI - Xho I contendo a maior parte da sequência do ADNc da hIL-3 (resíduos 15 a 133



mais a região 3' não codificante).

O plasmídeo resultante, denominado pGB/IL-316, é apresentado esquematicamente na figura 21. Na figura 22 apresenta-se a sequência completa do vector desde o local Sac II na sequência do promotor de lactase até ao local Hind III na extremidade do terminador sintético.

A figura 22 apresenta a sequência de nucleótidos do plasmídeo pGB/IL-316 entre o local singular Sac II no promotor de lactase e o local Hind III depois do terminador (resíduos 4457 e 7204). Os resíduos 4457 a 6100 contêm a sequência do promotor de lactase. Os resíduos 6101 a 6355 contêm a sequência de sinal do factor alfa. Os resíduos 6356 a 7115 contêm a sequência para a IL-3 humana madura mais a sequência do ADNc 3' não codificante. Finalmente os resíduos 7116 a 7204 contêm a sequência do terminador sintético.

B. Construção de pGB/IL-318

Construiu-se um vector de expressão similar a pGB/IL-316 mas substituindo a informação de codificação para a sequência de sinal do factor alfa de S. cerevisiae pela sequência de sinal do factor alfa de K. lactis (64). O resto do plasmídeo é idêntico ao plasmídeo pGB/IL-316. A sequência do plasmídeo pGB/IL-318 entre o local Sac II no promotor de lactase e o local Hind III depois do terminador (resíduos 4457 a 7190) é apresentada na figura 23.

Os resíduos 4457 a 6087 contêm a sequência do promotor de lactase e uma pequena sequência dum adaptador. Os resíduos 6088 a 6342 contêm a sequência de sinal do factor alfa de K. lactis. Os resíduos 6343 a 7102 contêm a sequência para a IL-3 humana madura mais a sequência do ADNc 3' não codificante. Finalmente os resíduos 7103 a 7190 contêm a sequência do terminador sintético.

C. Transformação de Kluyveromyces Lactis e análise da hIL-3 segregada

Digeriram-se os plasmídeos pGB/IL-316 e pGB/IL-318 no local singular Sac II na região do promotor



de lactase e foram utilizados para transformar a estirpe K. lac-
tis CBS 2360 (ver 64). Deste modo a integração dos plasmídeos é
dirigida para a região do promotor do gene de lactase cromosso-
mal. Os transformantes resistentes G418 resultantes foram culti-
vados até saturação em meio líquido YEPD e os sobrenadantes das
culturas e os lisados celulares foram analisados para se verifi-
car a existência de actividade de IL-3 por meio da análise de
síntese de ADN das células AML.

Verificou-se que virtualmente toda
a IL-3 era segregada para o meio de cultura e era activa. As pro-
teínas do sobrenadante da cultura foram precipitadas por etanol
e foram analisados usando electroforese em gel de poliacrilamida
desnaturante seguida de análise de mancha de Western. O produto
predominante tinha um peso molecular aparente de cerca de 21 kd,
observando-se também uma banda distinta a cerca de 15 kd. Este
último produto corresponde com toda a probabilidade à IL-3 madu-
ra não glicosilada, enquanto que o produto de 21 kd é o produto
que contém uma glicosilação do núcleo nos dois locais de potenci-
al glicosilação. Por incubação com endoglicosidase H obtém-se uma
proteína que migra na região de 15 kd sugerindo que toda a IL-3
é correctamente processada durante o processo de secreção e que
a maior parte da proteína se apresenta glicosilada.

EXEMPLO 7

Construção dum vector de expressão de *Saccharomyces cerevisiae*

A. Construção de pGB/IL-319

Em primeiro lugar construiu-se um
vector de expressão que se denominou pGB/TEFact. Neste plasmídeo
derivado do plasmídeo pTZ18R (Pharmacia), a sequência do promo-
tor do factor de alongação (EF-1 alfa) de tradução de S. cerevi-
siae, clonado e sequenciado conforme já descrito (73, 74), foi
acoplado por meio dum pequeno adaptador Sal I - Bgl II - XhoI à
sequência do terminador de transcrição de actina de S. cerevisi-
ae (75), a qual foi sintetizada por meio dum sintetizador de ADN
da firma Applied Biosystems. A sequência do bloco integrado de

expressão é dada na figura 24. Os resíduos 1 a 949 contêm o promotor de EF-1 alfa. Os resíduos 950 a 967 contêm a sequência do adaptador de Sali-BglIII-XhoI. Finalmente os resíduos 968 a 1113 contêm a sequência do terminador de actina.

O sítio SmaI singular no plasmídeo pGB/TEFact foi utilizado para introduzir o bloco integrado de resistência ao antibiótico G418 descrito no Exemplo 6. O plasmídeo resultante denominado pGB/TEFact G418.

Por fim, construiu-se o vector pGB/IL-318 de expressão de hIL-3 por introdução das seguintes sequências de ADN no plasmídeo pGB/TEFact G418 cortado por Sali e XhoI:

- O fragmento Sali-NruI do plasmídeo pGB/IL-316 contendo a sequência de sinal do factor alfa de S. cerevisiae e a sequência que codifica para a hIL-3 até ao local NruI.

- Um fragmento de ADN sintético NruI-XhoI contendo os restantes nucleótidos que codificam para a hIL-3 e a sequência de reconhecimento de XhoI imediatamente a seguir ao codão de paragem TGA.

B. Transformação de Saccharomyces cerevisiae e análise da hIL-3 segregada

Cortou-se o plasmídeo pGB/IL-319 no local EcoRI singular no promotor de EF-12. A integração do plasmídeo é dirigida deste modo para a região cromossomal de EF-1. Transformou-se a estirpe (alfa; ATCC 25 657) tal como se descreveu para K. lactis (64). As colónias resistentes a G418 foram isoladas e os transformantes foram cultivados em meio YEPD líquido. O sobrenadante da cultura foi analisado por meio de análise de AML a fim de se detectar a actividade de hIL-3. Verificou-se que a proteína produzida por S. cerevisiae era biologicamente activa.

As proteínas do sobrenadante foram precipitadas por meio de etanol e em seguida foram analisadas por electroforese em gel de poliacrilamida seguida de análise de mancha de Western. Foi possível distinguir dois produtos principais na análise de mancha de Western, um produto glicosilado de 21 kD e um produto não glicosilado de 15 kD.



R E I V I N D I C A Ç Õ E S

- 1ª -

Processo para a produção de interleuquina-3 (IL-3) humana por uma célula hospedeira caracterizado por compreender:

a introdução na referida célula hospedeira duma construção de ADN contendo um bloco integrado de expressão (cassette), o qual contém, na direcção da transcrição, uma região de regulação de início de transcrição funcional na referida célula hospedeira: uma sequência de ADN que codifica para a IL-3 humana: e uma região de regulação de terminação de transcrição funcional na referida célula hospedeira;

o crescimento da referida célula hospedeira, contendo a referida construção de ADN, num meio nutriente sob condições de cultura adequadas, produzindo-se IL-3 humana; e a recuperação da IL-3 humana como produto final.

- 2ª -

Processo de acordo com a reivindicação 1 caracterizado por a célula hospedeira ser uma célula hospedeira transformada contendo material genético derivado de material de ADN recombinante e que codifica para a IL-3.

- 3ª -

Processo de acordo com a reivindicação 2 caracterizado por a célula hospedeira ser uma célula hospedeira escolhida de entre o grupo constituído por leveduras, bac-



térias, fungos e células de cultura de tecidos.

- 4ª -

Processo de acordo com a reivindicação 3 caracterizado por a célula hospedeira ser um fungo escolhido de entre o grupo constituído por Saccharomyces e Kluyveromyces.

- 5ª -

Processo de acordo com a reivindicação 3 caracterizado por a célula hospedeira ser uma bactéria escolhida de entre o grupo constituído por E. coli e Bacillus.

- 6ª -

Processo de acordo com a reivindicação 3 caracterizado por a célula hospedeira ser uma célula duma cultura de tecidos escolhidos de entre o grupo constituído por células COS, C127 e células de insectos.

- 7ª -

Processo para a preparação dum sistema de expressão operacional num hospedeiro recombinante caracterizado por se inserir num sistema de expressão adequado uma sequência de ADN que codifique para a IL-3 humana ligado de modo operacional a sequências de regulação efectivas no referido hos-



pedeiro.

- 8ª -

Processo de acordo com a reivindica-
ção 7 caracterizado por o ADN que codifica para a IL-3-humana
não conter intrões.

- 9ª -

Processo de acordo com qualquer das
reivindicações 7 ou 8 caracterizado por as sequências de regula-
ção conterem um promotor seleccionado de entre o grupo constituí-
do pelo promotor lac, pelo promotor HpaII, pelo promotor 43, pe-
lo promotor de alfa-amilase, pelo promotor EF-1 alfa e pelo pro-
motor SV40.

- 10ª -

Processo para a preparação duma cé-
lula hospedeira recombinante caracterizado por se transformar
uma célula hospedeira com um sistema de expressão preparado de
acordo com qualquer das reivindicações 7 a 9.

- 11ª -

Processo para a preparação duma se-
quência de ADN que não contem intrões e que codifica para a IL-
-3 humana caracterizado por compreender as fases de:

- 
- a) produção de factores de crescimento hemopoiéticos por leucócitos humanos,
 - b) preparação de ARN a partir das células referidas e transcrição do mesmo em ADNc de cadeia dupla.
 - c) selecção do ADNc com um ADNc murino completo de IL-3 de rato (mIL-3) contendo tanto a porção de codificação como a porção 3' não traduzida a juzante de forma a obter DII.
 - d) inserção do clone DII de ADNc que hibrida num vector de expressão pLO.

- 12^a -

Processo de acordo com a reivindicação 11 caracterizado por o ADN recombinante obtido conter a sequência de nucleótidos apresentada como codificando para os ácidos aminados 1 a 133 na sequência "H" da Figura 1.

- 13^a -

Processo de acordo com qualquer das reivindicações 1 a 6 caracterizado por a célula hospedeira ser uma célula hospedeira recombinante obtida de acordo com a reivindicação 10.

- 14^a -

Processo para a preparação duma proteína dotada de actividade de IL-3 humana caracterizado por se obter a expressão da referida proteína numa célula hospedeira transformada obtida de acordo com a reivindicação 10.



- 15^a -

Processo de acordo com qualquer das reivindicações 1 a 6 ou 13 ou 14 caracterizado por a proteína ser obtida sob uma forma substancialmente isenta de outras substâncias que normalmente acompanham a referida proteína.

- 16^a -

Processo de acordo com a reivindicação 15 caracterizado por a proteína ser obtida por meio da expressão recombinante da sequência de ADN apresentada como codificando para os ácidos aminados 1 a 133 na sequência "H" da Figura 1 ou dum mutante desta sequência de ocorrência natural.

- 17^a -

Processo de acordo com qualquer das reivindicações 15 ou 16 caracterizado por a proteína ser obtida sob forma glicosilada ou não glicosilada.

- 18^a -

Processo para a preparação de um preparado de anticorpos dotado de capacidade de reacção imuno-específica com IL-3 humana caracterizado por compreender uma fase de injeção dum hospedeiro vertebrado com a IL-3 humana purificada obtida de acordo com qualquer das reivindicações 14 a 17.

•
•
•

A requerente declara que os primeiros pedidos desta patente foram apresentados como pedidos de patente europeus em 16 de Dezembro de 1986 e em 13 de Julho de 1987, sob os nº.s 86202285.2 e 87201322.2, respectivamente.

Lisboa, 16 de Dezembro de 1987.

A handwritten signature in black ink, consisting of a series of loops and a long horizontal stroke.

RESUMO

"PROCESSO DE CLONAGEM MOLECULAR E DE EXPRESSÃO DE
INTERLEUQUINA-3 (IL-3) HUMANA"

A invenção refere-se a um processo para a produção de interleuquina-3 (IL-3) humana (hIL-3) que compreende as fases de cultivar células transformadas por um vector de expressão apropriado em condições adequadas para o respectivo crescimento e para a expressão do ADN que codifica para a hIL-3 isolar a hIL-3 produzida.

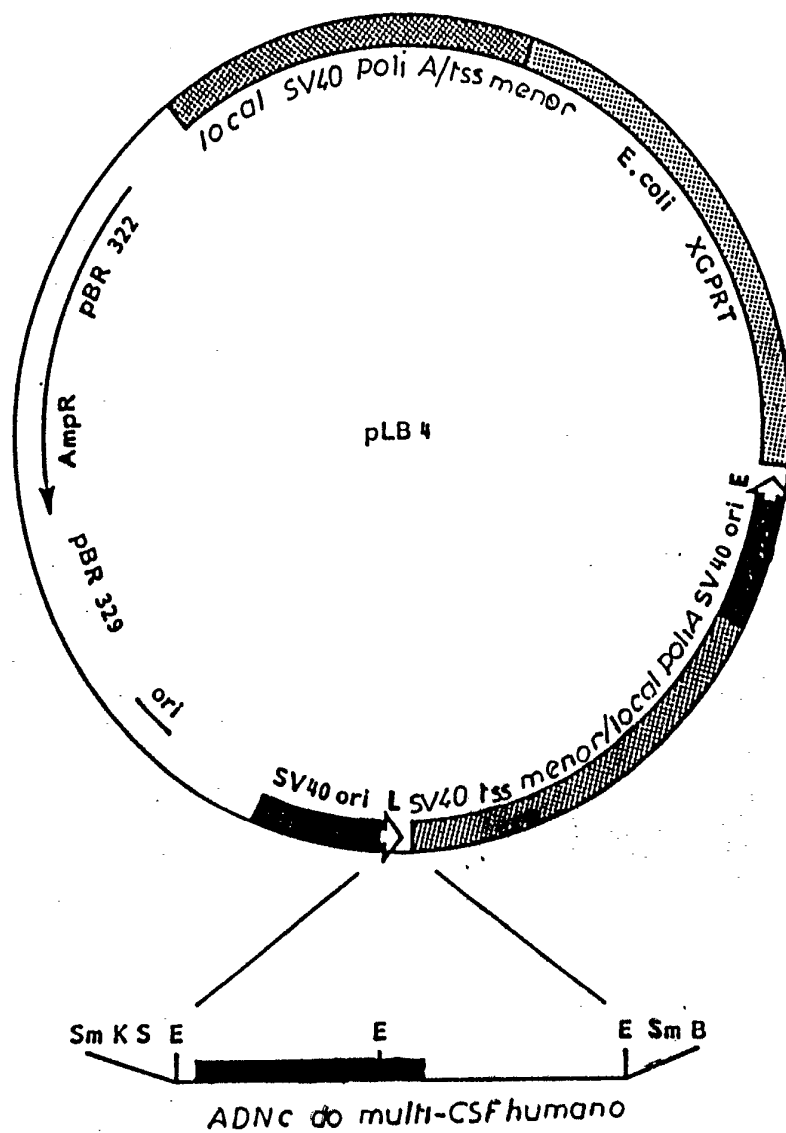


FIGURA 2

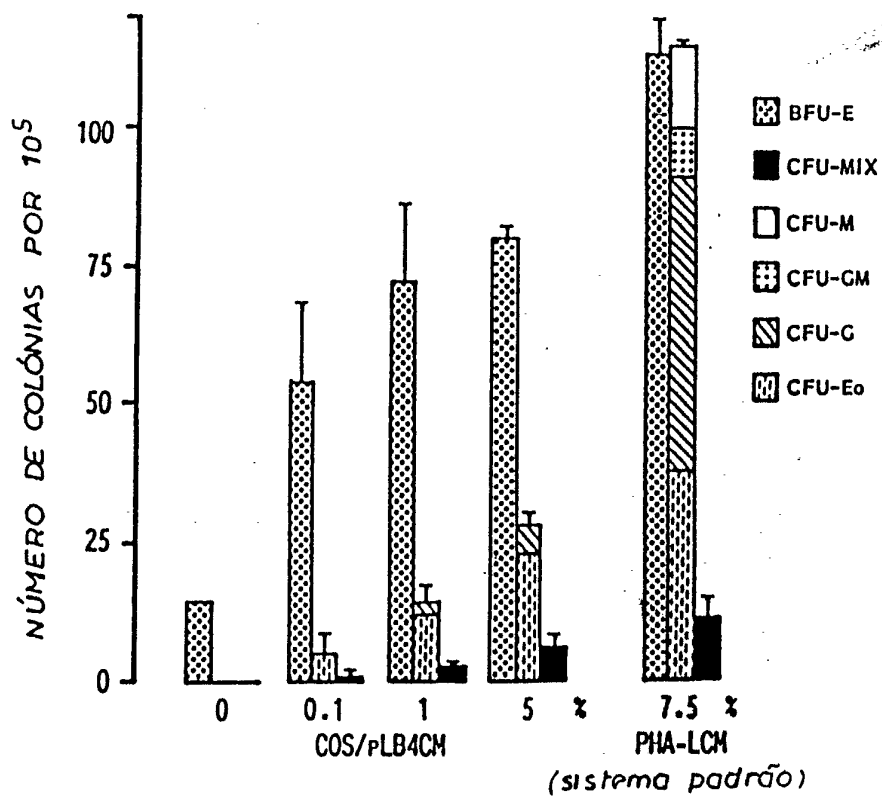


FIGURA 3

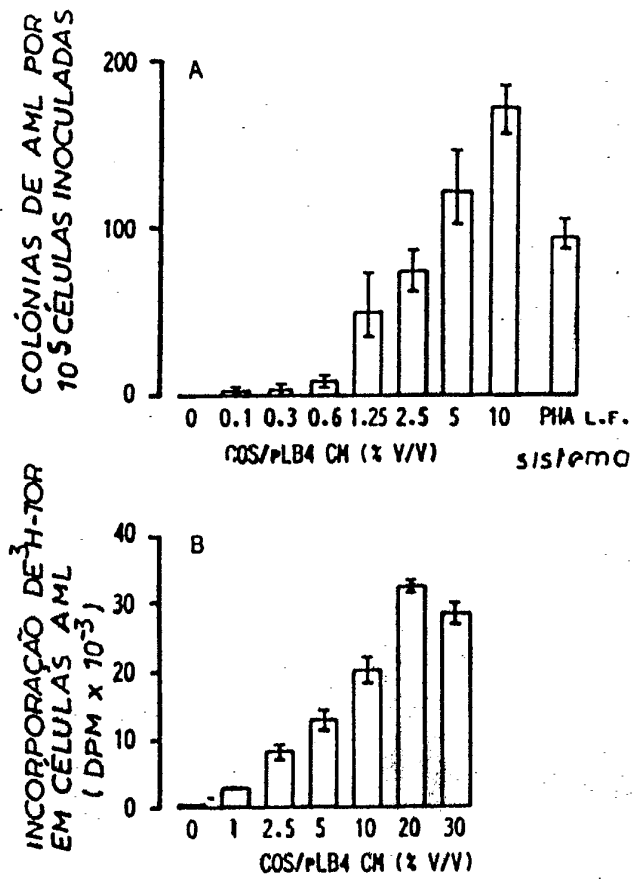
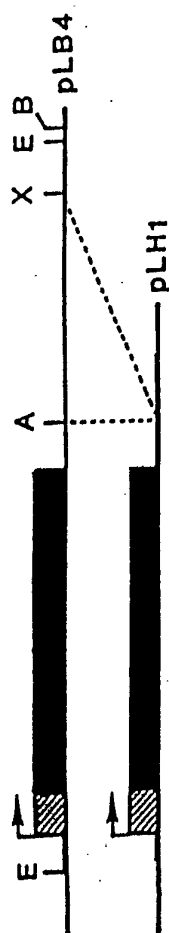
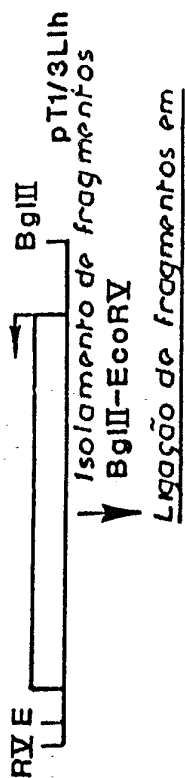


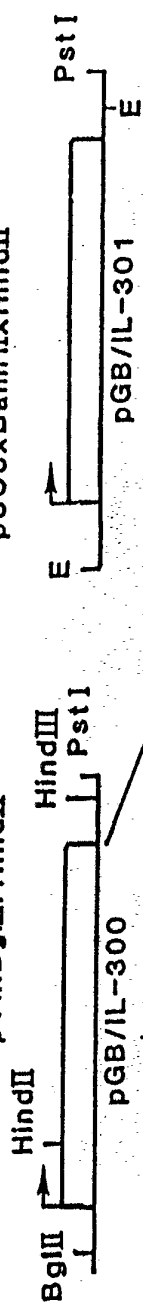
FIGURA 4



X *Ava*II
Klenow
ligação do adaptador *Bgl*III
X *Bgl*III
X *Bam*HI
isolamento de fragmentos
ligação de pT1 x *Bgl*III



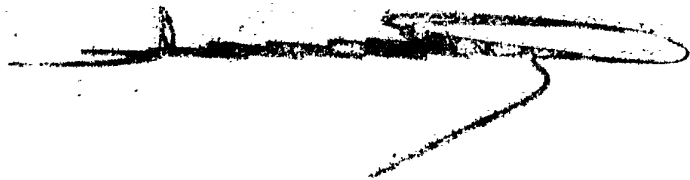
pT1x*Bgl*III/*Hind*II pUC8xBamHIx*Hind*II



Fragmento *Hind*II-*Hind*III
oligonucleotido

pGB/IL-302
pGB/IL-303
pGB/IL-304
pGB/IL-305

FIGURA 5



EcoRI SacI KpnI SmaI BamHI XbaI SalI PstI SphI HindIII
pTZ18R gggaattcgagctcggtaccggggatcctctagagtcgacctgcaggcatgcaagcttg

EcoRI SacI EcoRV HindIII BglII BamHI XbaI SalI PstI
pT1 gggaattcgagctcgatatcaagcttagatctcgaggggatoctctagagtcgacctgcag

SphI NdeI
gcattgcaagctgcattatgcagcttg

FIGURA 6

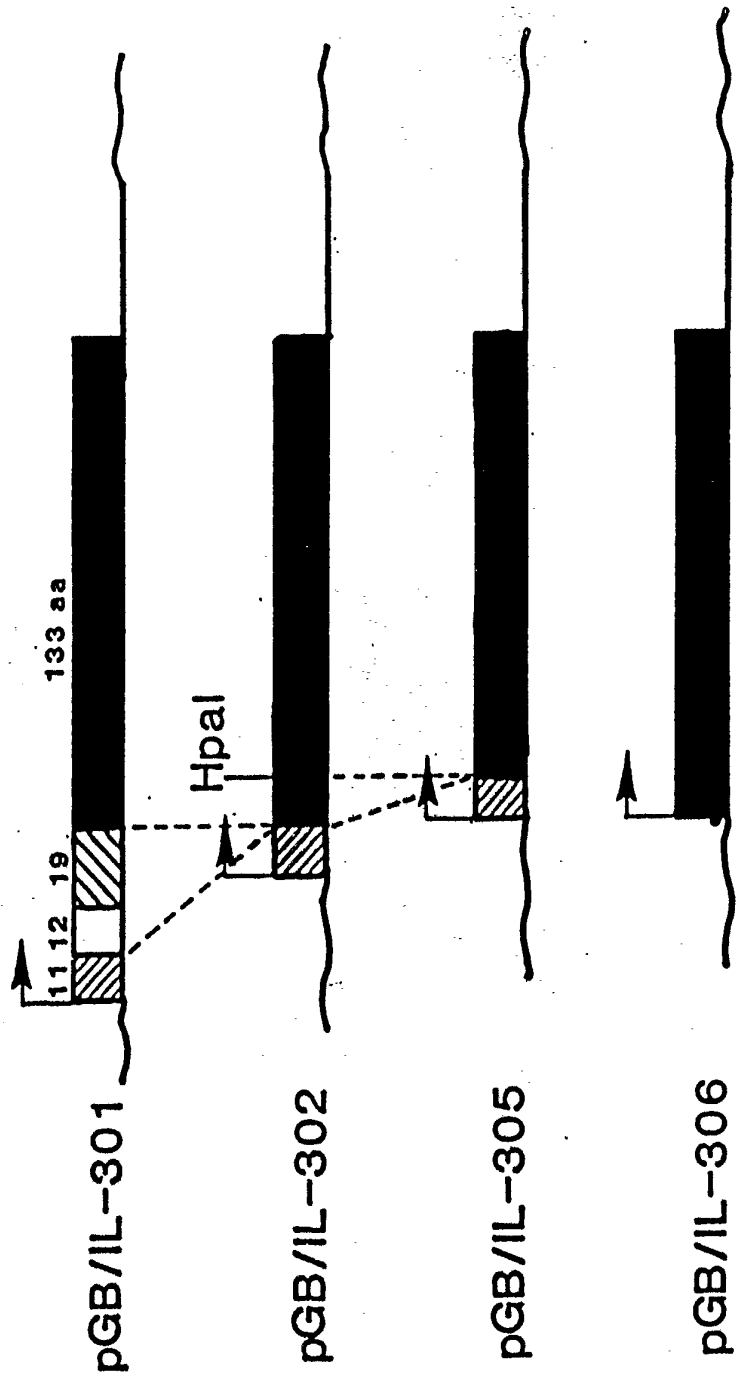
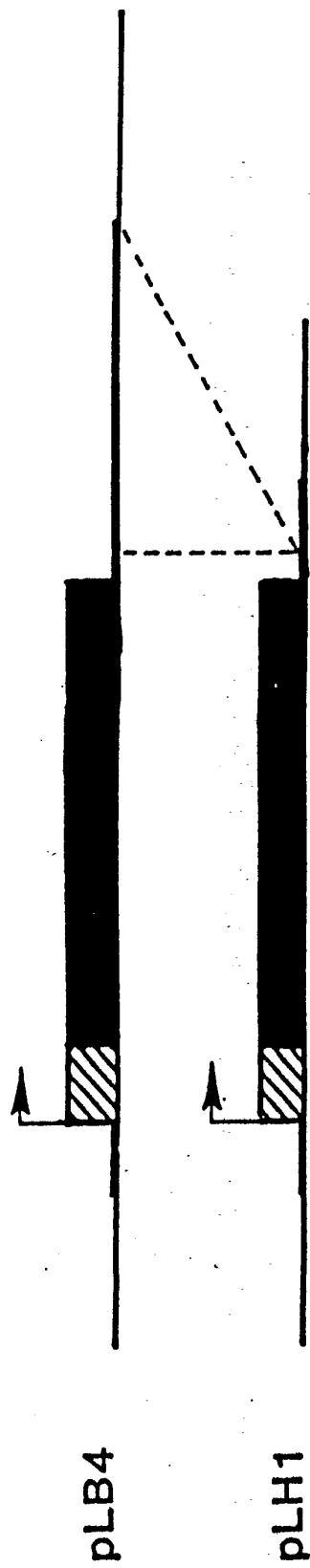


FIGURA 7

pGB/IL-301

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Gly Pro
atg acc atg att acg aat tcc cgg gga tct gGA CCA

Glu Gln Asp Arg Val Pro Pro Ala Asp Pro Asn Met
GAA CAA GAC AGA GTG CCT CCT GCC GAT CCA AAC ATG

Ser Arg Leu Pro Val Leu Leu Leu Leu Gln Leu Leu
AGC CGC CTG CCC GTC CTG CTC CTG CTC CAA CTC CTG

Val Arg Pro Gly Leu Gln Ala¹Pro Met Thr Gln Thr
GTC CGC CCC GGA CTC CAA GCT CCC ATG ACC CAG ACA

Thr Pro Leu Lys Thr Ser Trp Val Asn Cys Ser Asn
ACG CCC TTG AAG ACA AGC TGG GTT AAC TGC TCT AAC

Met Ile Asp
ATG ATC GAT

pGB/IL-302

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ser Arg
atg acc atg att acg aat tcc cgg gga tcc tct aga

Val Asp Pro²Met Thr Gln Thr Thr Pro Leu Lys Thr
gtc gac CCC ATG ACC CAG ACA ACG CCC TTG AAG ACA

Ser Arg Val Asn Cys Ser Asn Met Ile Asp
AGC CGG GTT AAC TGC TCT AAC ATG ATC GAT

pGB/IL-303

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ser Arg
atg acc atg att acg aat tcc cgg gga tcc tct aga

Val Asp Pro²Met Thr Gln Thr Thr Pro Pro Lys Thr
gtc gac CCC ATG ACC CAG ACA ACG CCC CCG AAG ACA

Ser Arg Val Asn Cys Ser Asn Met Ile Asp
AGC CGG GTT AAC TGC TCT AAC ATG ATC GAT

pGB/IL-304

Met Thr Met Ile Thr Asn Leu Ile Arg Leu Thr Ile
atg acc atg att acg aat tta ata cga ctc act ata

Gly Asn Ser Ser Ser Val Pro Gly Asp Pro Leu Glu
ggg aat tcg agc tcg gta ccc ggg gat cct cta gag

Ser Ile Asp Pro²Thr Thr Glu Thr Thr Pro Leu Lys
tcg atc gac CCC ACG ACC CAG ACA ACG CCC CTG AAG

Thr Ser Trp Val Asn Cys Ser Asn Met Ile Asp
ACA AGC TGG GTT AAC TGC TCT AAC ATG ATC GAT

pGB/IL-305

Met Thr Met Ile Thr Asn Leu Ile Arg Leu Thr Ile
atg acc atg att acg aat tta ata cga ctc act ata

Gly Asn Ser Ser Ser Val Pro Gly Asp Pro Leu Glu
ggg aat tcg agc tcg gta ccc ggg gat cct cta gag

Asn¹⁵Cys Ser Asn Met Ile Asp
AAC TGC TCT AAC ATG ATC GAT

pGB/IL-306

Met Ala¹Pro Met Thr Gln Thr Thr Pro Leu Lys Thr
atg GCT CCC ATG ACC CAG ACA ACG CCC TTG AAG ACA

Ser Trp Val Asn Cys Ser Asn Met Ile Asp
AGC TGG GTT AAC TGC TCT AAC ATG ATC GAT

FIGURA 8 (continuação)

pGB/IL-301 pGB/IL-302

1 2 A₃ B M A₃ B 2 1

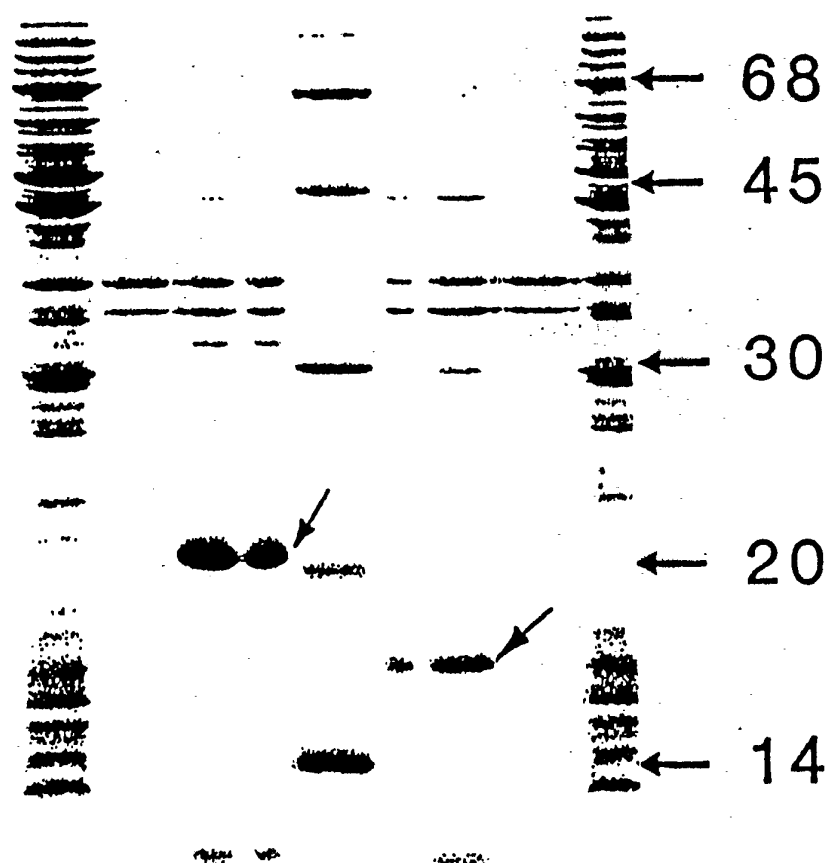


FIGURA 9

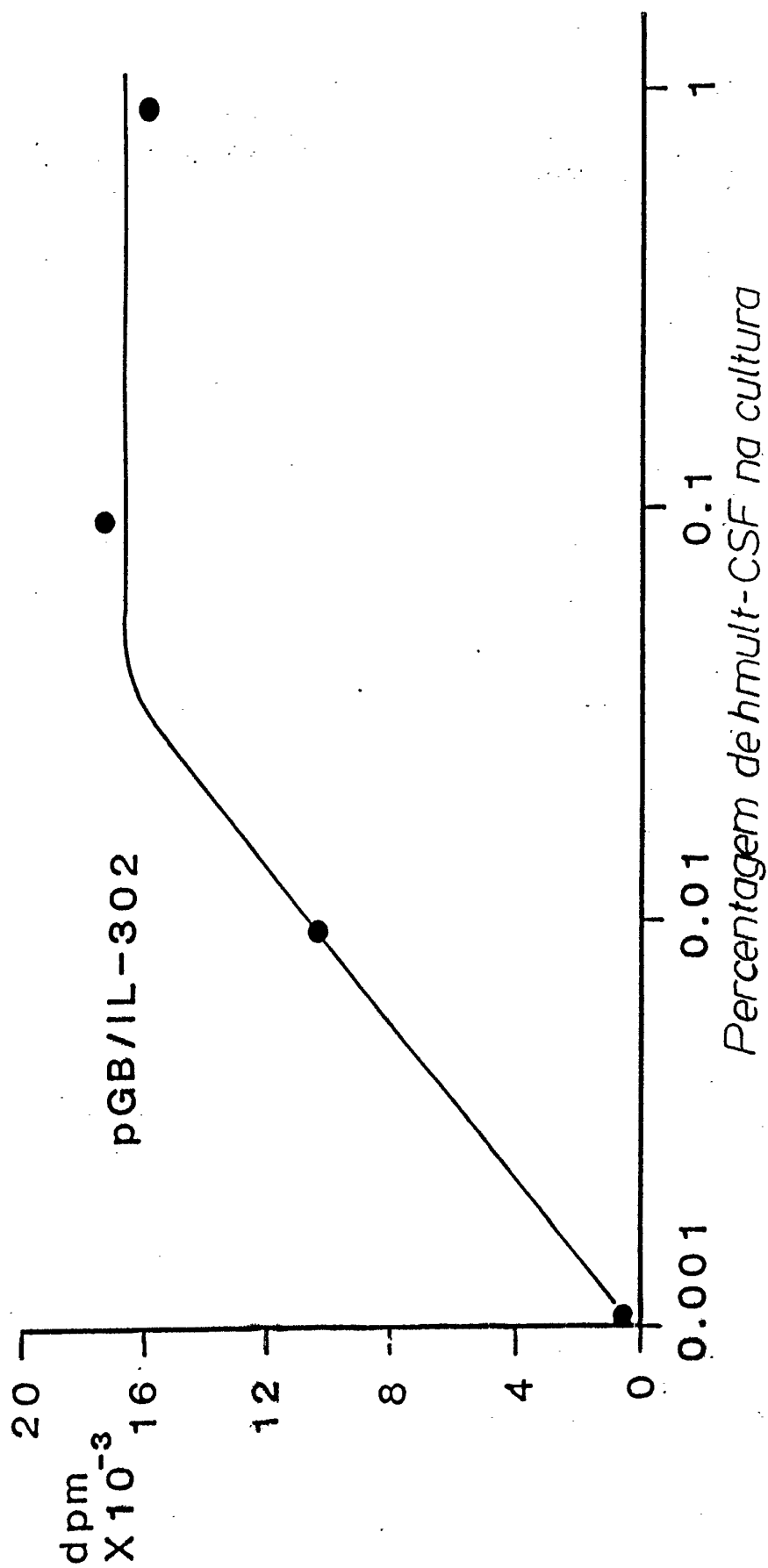


FIGURA 10

1 2 3 4 5 6 7

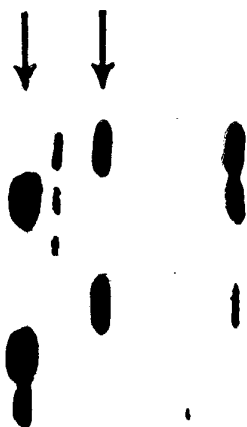
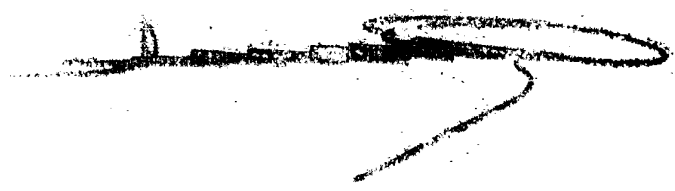


FIGURA 11



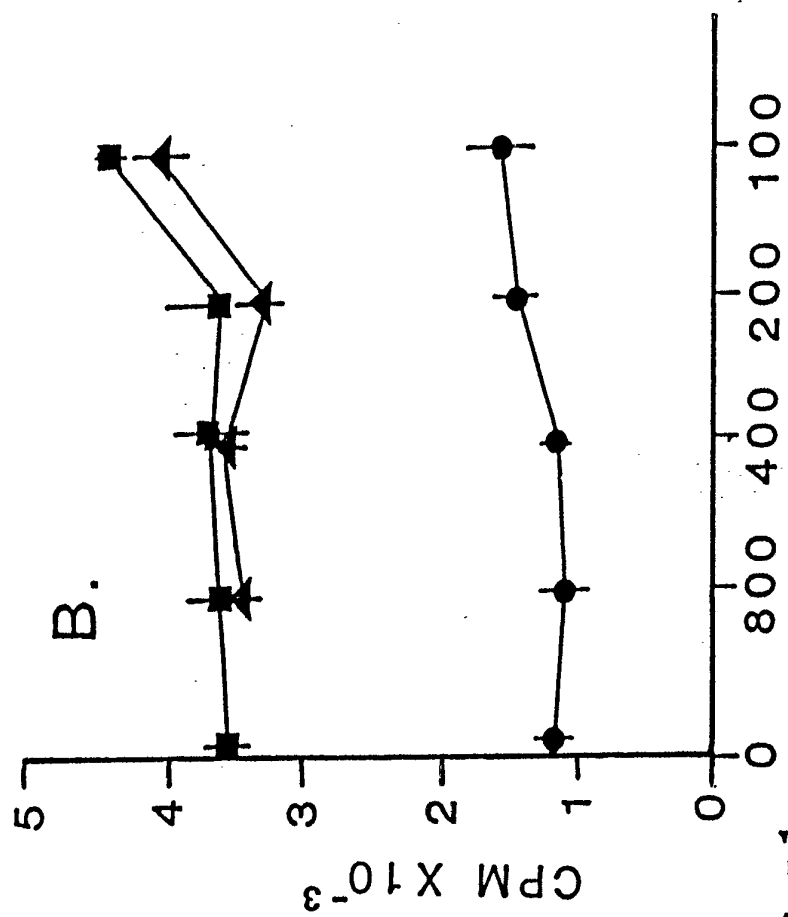
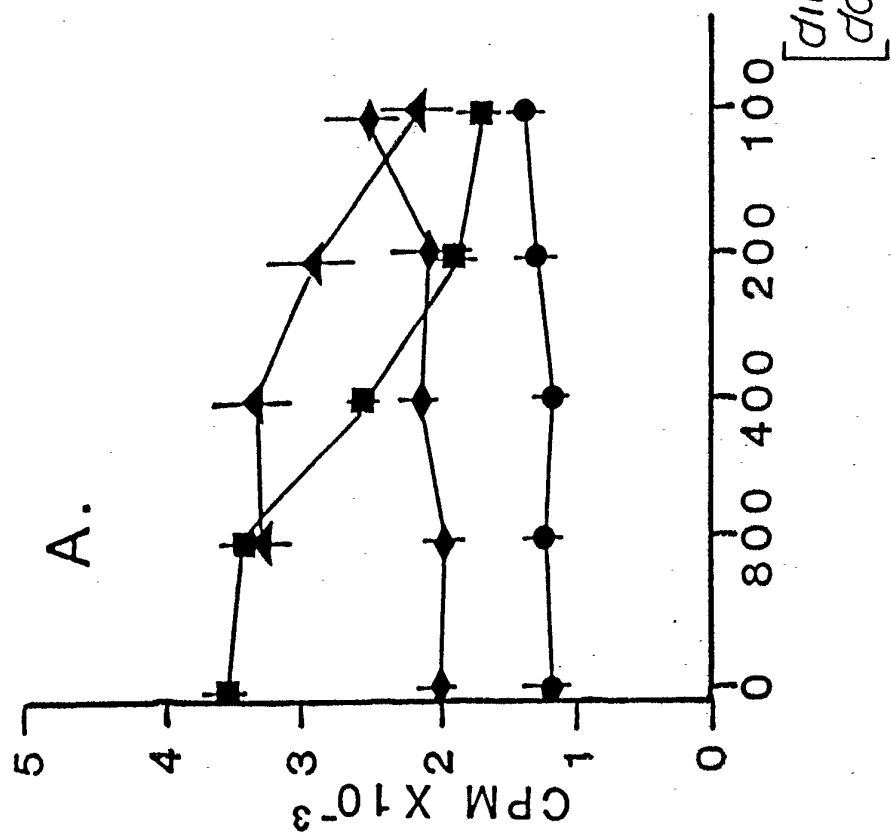
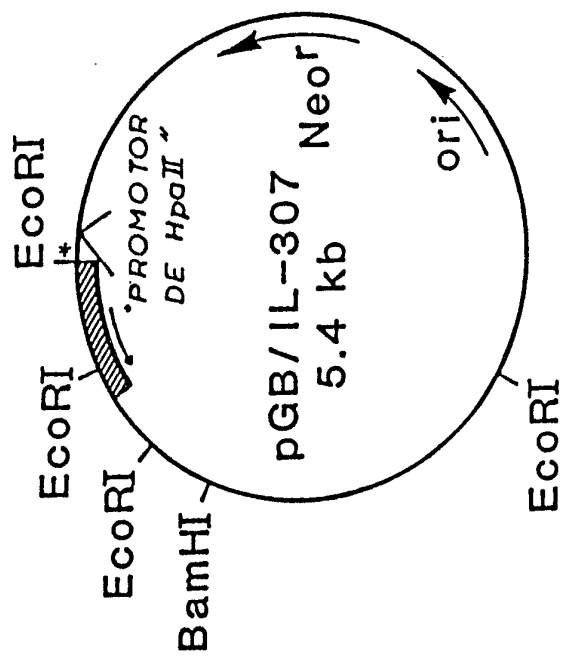


FIGURA 12



Sequência da extremidade N-terminal da proteína de fusão

Met Ser Tyr Ala Val Cys Arg Met Glu Lys

Val Lys Ser Gly Val Pro Ser Ser Asn Ser Gly

Pro Glu Gln Asp Arg Val Pro Pro Ala Asp Pro

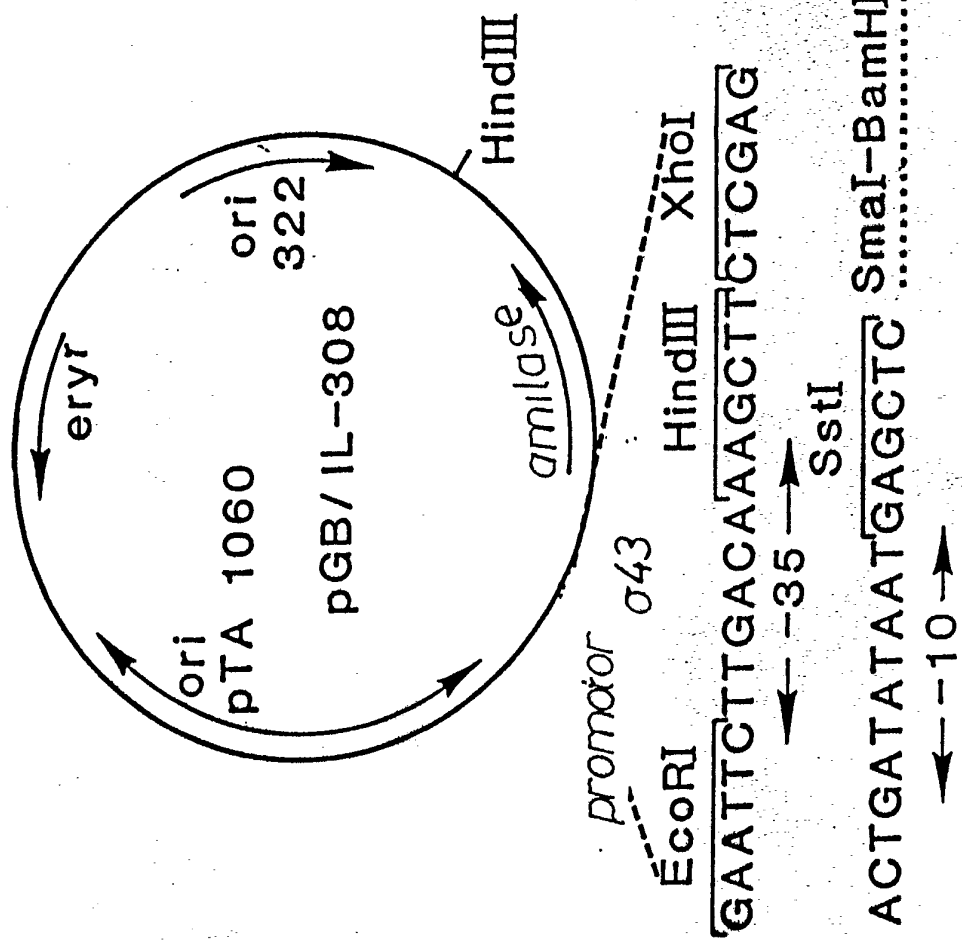
Asn Met Ser Arg Leu Ala Pro

-19 →

+1 →

sequência de sinal da hIL-3 sequência da hIL-3 madura

FIGURA 13



Shine Dalgarno...XbaI..Sall

FIGURA 14

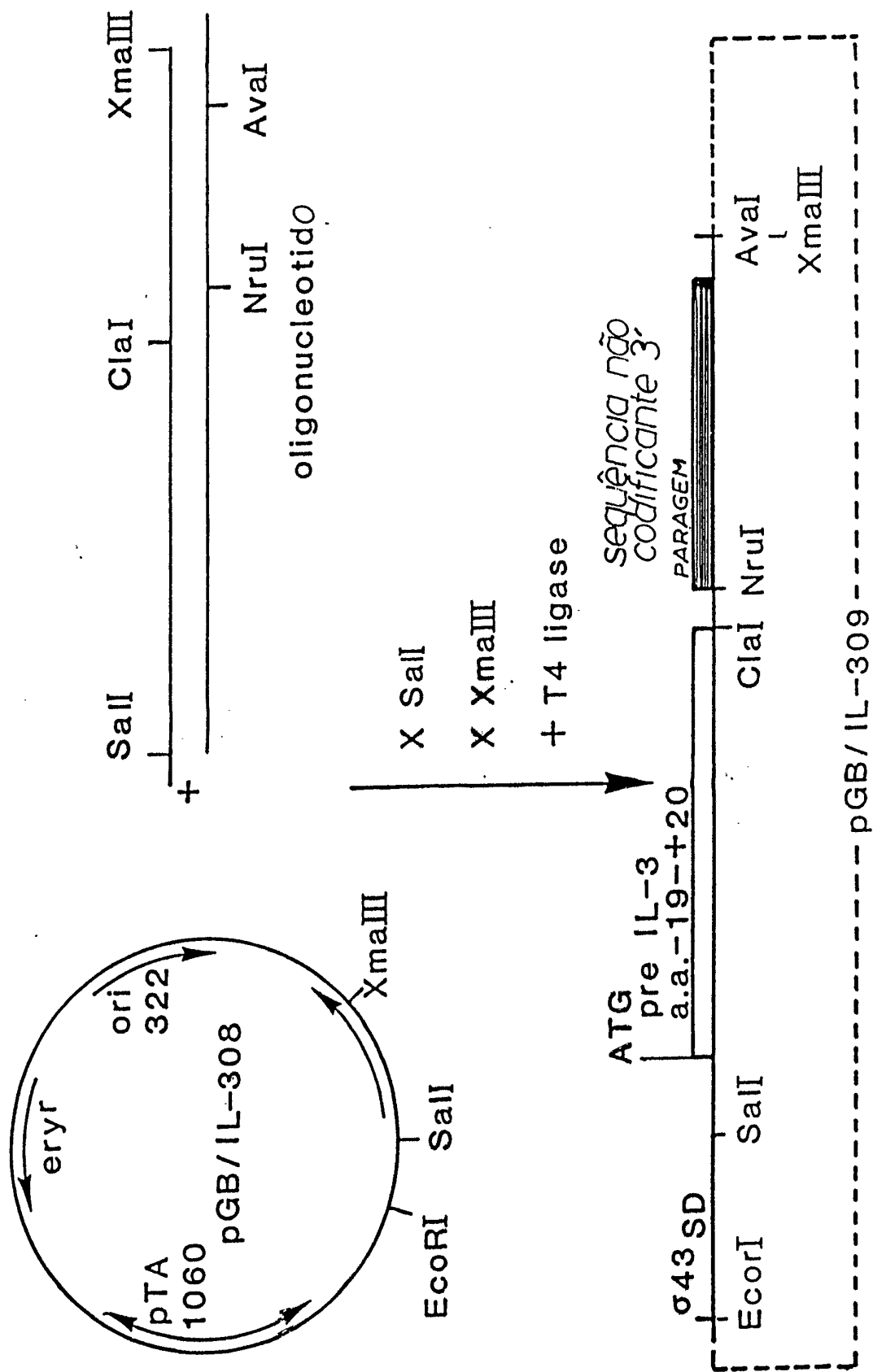
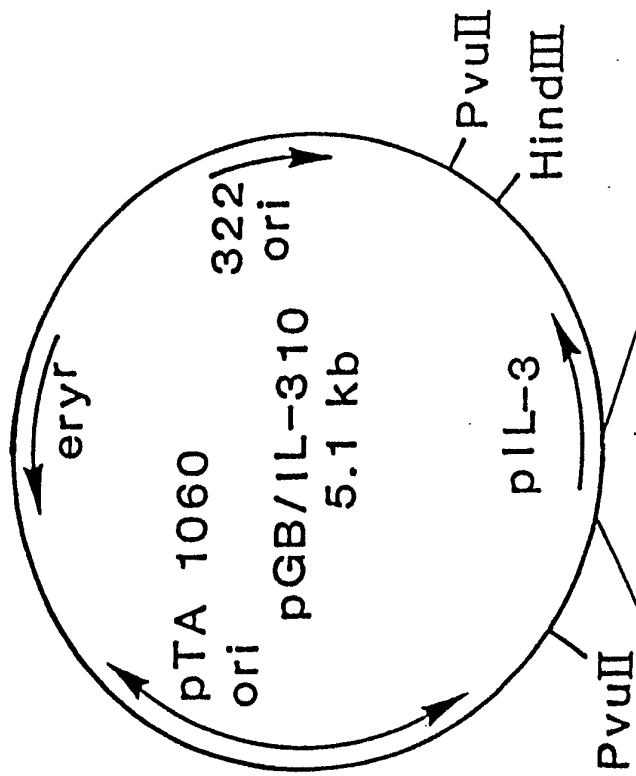


FIGURA 15



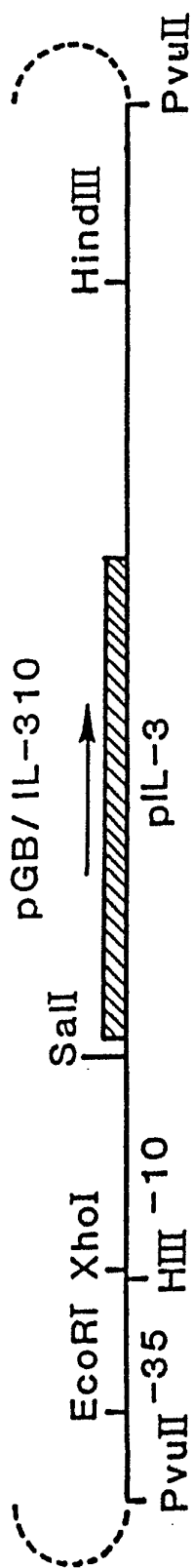
EcoRI-35 HindIII XhoI-10 SstI SmaI BamHI SD XbaI SalI

σ 43

pIL-3: gene do precursor da IL-3 humana

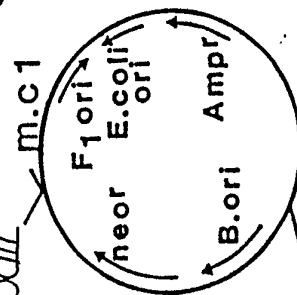
FIGURA 17 (cont. n.)

FIGURA 17 (contin.)



HindIII parcial } dois *PvuII*-*HindIII*
 x *PvuII*
 fragmentos
 promotor de *HpaII*

X *HindIII*
 X *SmaI*



+

m.c2 *HindIII* *XbaI* *KpnI* *SphI* *EcoRV* *SstI* *SmaI* *EcoRI*

m.c 1

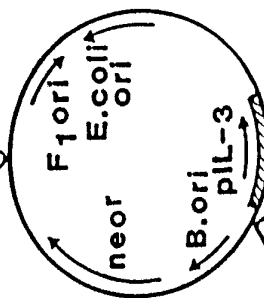
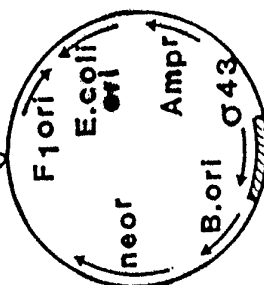
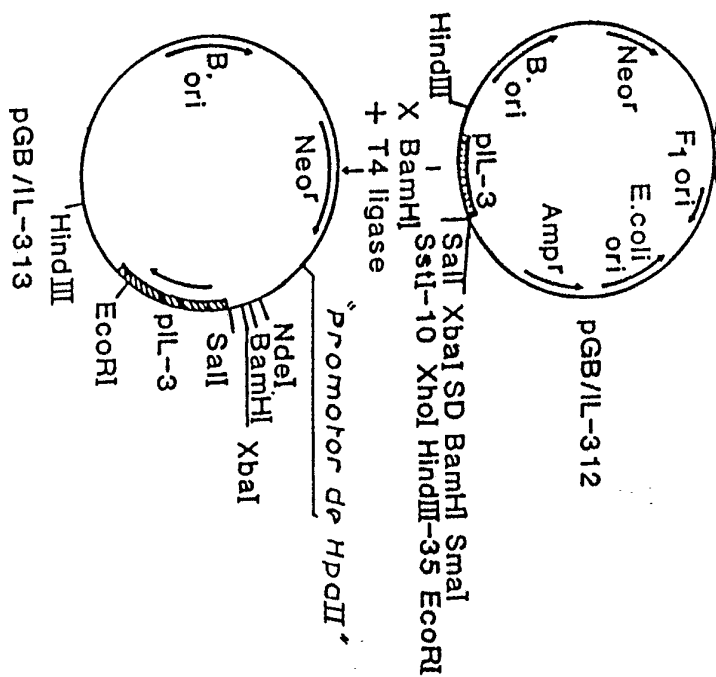


FIGURA 18

m.c.1 NdeI-BamHI-XbaI-SstI-BglII-MluI-XbaI



Sequência;

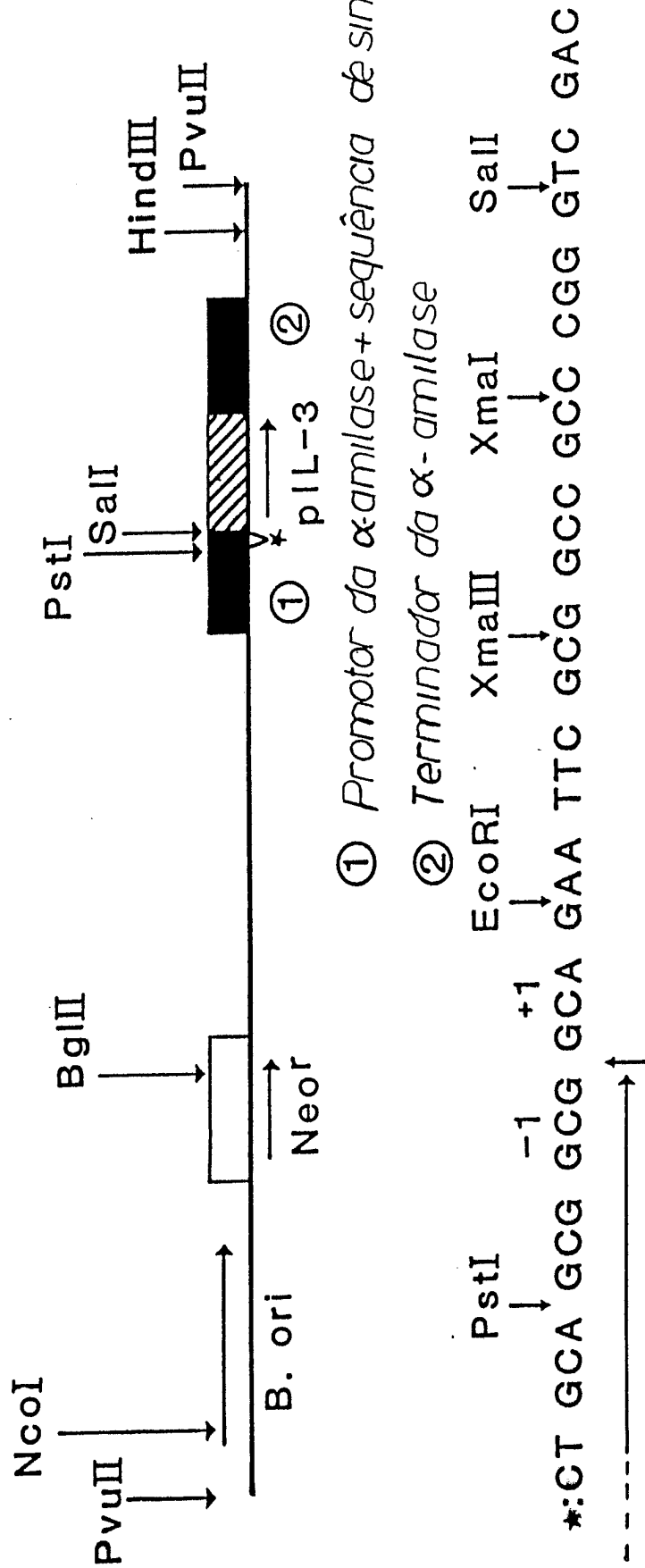
Promotor de HPcII -AAAGGAGCGATTACAT ATG AGT

TAT GCA GTT TGT AGA ATG CAA AAA GTG
tyr ala val cys arg met gln lys val met ser

AAA TCA GGG GGA TCC AAG GAG GTG ATC
lys ser gly ser lys glu val ile
TAG AGT CGA CATG...pil-3
porogem met

FIGURA 19

pGB/IL-317 : 4.8 kb



sequência de sinal da
α-amilase

local de maturação
da α-amilase

- ① Promotor da α-amilase + sequência de sinal
- ② Terminador da α-amilase

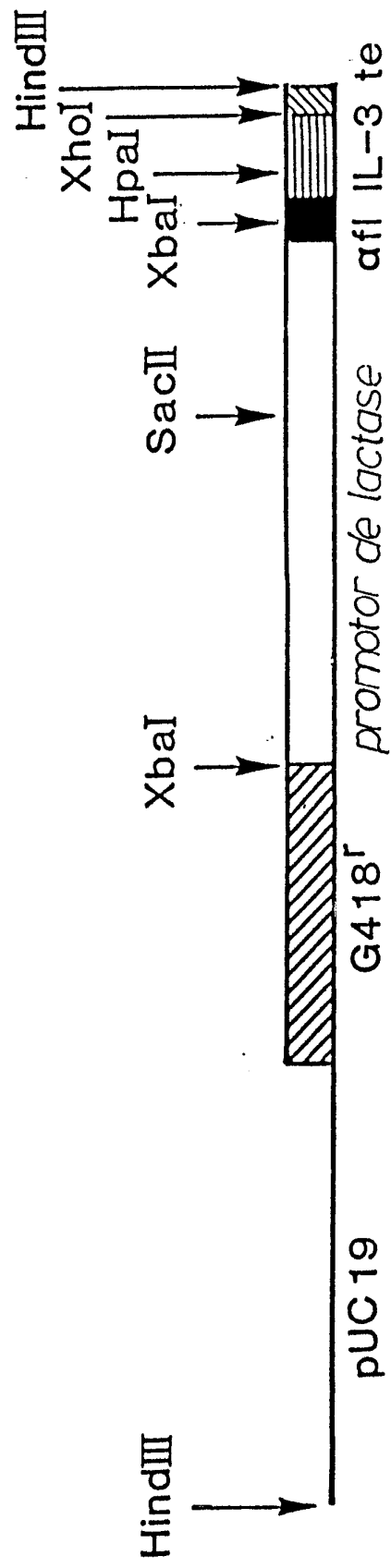


FIGURA 21

4406 4476 4486 4496 4506 4516 4526 4536 4546 4556
 CCGCGGCGAT CAGCTATATA AATAGTAAGC TTCTAATGCG TATCTATTGA CTACCAACCA TTAGTGTGGT TGCAGAAAGC GGAATTCTCC CTCTCTGGAA
 4506 4576 4586 4596 4606 4616 4626 4636 4646 4656
 TTCAGCTTGC TTTTCATTTT TATTTTCCAT TTTTCAGTTT TTGTTTGTGT CGAATTTAGC CAGTTGCTTC TCCAAGATGA AAAAAATCTT GCGCATTTTC
 4666 4676 4686 4696 4706 4716 4726 4736 4746 4756
 TGTGTGGCAA GATCTAATC GACTTTTCCA CCGGACACAA AAGTAAATGT TTCTTTGTTA CATTGCGGTG GGTAGCTAGC TCCCGAATC TCAAAAGACT
 4766 4776 4786 4796 4806 4816 4826 4836 4846 4856
 TAGGACTGCT ACTACATGAG AGTGTGTGCA CCTGGTTTGC TGGCTGTTT GAAAGAAAAG AGCGGGAAGT CCGCGGTTC CCGGGAATAA TCATGATATA
 4866 4876 4886 4896 4906 4916 4926 4936 4946 4956
 CTCCTTTGCG CTTCAGATTC GCATGTAGAG TAGACAACAG ACAGGAGGGG CAGGAAGGAT CTTTCACTGA GATCCTGTAT CTGTTTGGGT AAGTGGATG
 4966 4976 4986 4996 5006 5016 5026 5036 5046 5056
 AAAGGGGAAT GGTATGAGAT TGGAGAGGAT GCGGAAGAGG TAACGCTTTT TGTTAATCTG TTTAATTATT ATGGGGCAGG CGAGAGGGGG AGGAATTTAT
 5066 5076 5086 5096 5106 5116 5126 5136 5146 5156
 GTGTGTGAGG CCGGAGAGAG GAGGCACTCC AGGCCAGGTA GAAATAGAGA AAGCCGAATG TTAGACAATA TGGCAGCGTA GTAGATAGG TAGGTAGGTA
 5166 5176 5186 5196 5206 5216 5226 5236 5246 5256
 AGTACTGCTA GCAAAAGAGA GAAGGTAAG CTCAGCTTTC GCATTGCGCA CCGTTAGTGT GTCAGTTTAA ACAAAAAAAC AACTACTATA GCAATTAGTA
 5266 5276 5286 5296 5306 5316 5326 5336 5346 5356
 GACTGTGAACT TGACTTTTGG AAGGGTTTTT CCGACTGGCA TTATTGCTGA GGAATCAAGG TAGGAATTTG GTCATATTTA CCGAGAGAGG TGGGTGATTT
 5366 5376 5386 5396 5406 5416 5426 5436 5446 5456
 GCATATGAGT TAGGAAAGAG AGATCATGAT ATCTCATGAT ATGTTGCGGA ATTCTGTTC CCGCAAGATT CAGCGTGCTC TGGTGGTTTT GGTGTGGTTT
 5466 5476 5486 5496 5506 5516 5526 5536 5546 5556
 TTGCTTTGCT TGTGCTTTCT GTTGCTTTCT AATAATAGCC TAGGCTGTGA GCGGAAACTT AGGTAGGCTT TANTGTGGA ACCTATATAT GTATCATGTT
 5566 5576 5586 5596 5606 5616 5626 5636 5646 5656
 GACTGTGTTT **AACCAAGGCA** CCGGTAGGCT AGGCTAGGCT ACACAGTTT TTTGTATTCT TCAGTATAGT TGTGAAAGAT GTAGGGAATA TATGTGTTCT
 5666 5676 5686 5696 5706 5716 5726 5736 5746 5756
 GAGCAAGAGT **GTCTTTTCT** AGTATGAGG TGGTTACTTT GGTGACATT GTATTTTGA CTTTGTGCT ACAGGATTCA GTATTTGAGG TGGAGTTGTA
 5766 5776 5786 5796 5806 5816 5826 5836 5846 5856
 AGGTATGAT TTCTAGTCTT GAACAGTTT AGTTAGGTAA TGTTTTCTTT GCTGTTTTAC TTGAGATTTT GATTGAGAAA AAGTATTTA ATAGTTGAAA
 5866 5876 5886 5896 5906 5916 5926 5936 5946 5956
 TCAATGTGTT ATCATTTCTA AGATGTTCTT CCGTAAGTCG AAAGGTATAT GAGGCTTGTG TTTCTTAGGA GAATTATTAT TCTTTTGTTA TGTGCTTTCT
 5966 5976 5986 5996 6006 6016 6026 6036 6046 6056
 GTAGTTGGAA AAGGTCAAGA GACAAAAGCT TAACACTTGA AATTTAGGAA AGAGCAGAAT TTGGCAAAAA AAATAAAAAA AAAATAAACA CCGTCACTTG
 6066 6076 6086 6096 6106 6116 6126 6136 6146 6156
 TGAGCGGATA ACAATGATA CATACTCATC GAGAACTGAA AGATATGAGA TTTCCATCGA TTTTACTGCG AGTTTATTTC GCAGCATGCT CCGGATTAGC
 6166 6176 6186 6196 6206 6216 6226 6236 6246 6256
 TGCTCCAGTC AACACTACAA CAGAAGATGA AAGGACACAA ATTCCGCGTG AAGCTGTCAT CCGTTACTTA GATTTAGAAG GGGATTTTGA TGTGCTTTCT
 6266 6276 6286 6296 6306 6316 6326 6336 6346 6356
 TTGCCATTTT CCAACAGGAG AAATAACGCG TTATTGTTTA TAAATACTAC TATTGCCAGC ATTGCTGCTA AAGAAAGAGG GGTATTTCTA GATAAAGAGC
 6366 6376 6386 6396 6406 6416 6426 6436 6446 6456
 CTCCTCATGAC CCAAGAGAGG CCGTTGAAGA CAAGCTGGGT TAAGTCTCTT AACATGATCG ATGAAATTAT AACACACTTA AAGGAGGAGC CTTTGGTTTT
 6466 6476 6486 6496 6506 6516 6526 6536 6546 6556
 GCTGGATTTT **AACACCTCA** ATGAGGAGAA CCAAGACATT CTGATGAAA ATAGCTTCTT AAGGCAAAAC CTGAGGATAT TCAAGGAGG TGTGAGGAT
 6566 6576 6586 6596 6606 6616 6626 6636 6646 6656
 TTACAGAAGC CATCAAGCAAT TGAGAGCATT CTAAAAAATC TCCTGCCATG TCTGCCCTCG GCCACGGCGG CACCCAGGCG ACATGCAATC CATATCAAGG
 6666 6676 6686 6696 6706 6716 6726 6736 6746 6756
 ACGGTGACTG GAATGAATTC CGAAGGAAAC TGACCTTCTA TCTGAAAAAC CTTGAGAAAT CCGAGGCTCA ACAGAGGACT TTGAGGCTCG GATCTTTTTC
 6766 6776 6786 6796 6806 6816 6826 6836 6846 6856
 AGTCCAACGT CCAAGCTGTT CTCTGGGCTT TCTCACCACA GAGGCTCGGC ACATCAAAAA CAGCAGAACT TCTGAAACCT CTGGGTCATC TCTCAGACAT
 6866 6876 6886 6896 6906 6916 6926 6936 6946 6956
 TCCAGGACCA GAAGCATTTT ACCTTTTCTT CCGGCTCAG ATGAATTGTT AATTATCTAA TTTCTGAAAT GTGAGCTCC CATTGGGCTT TGTGCGGTTG
 6966 6976 6986 6996 7006 7016 7026 7036 7046 7056
 TGTCTCATTT TTTATCCGAT TGAGACTATT TATTTATGTA TGTATGTATT TATTTATTTA TTGCTGGAG TGTGAAGCT ATTTATTTTA GCAGAGGAGC
 7066 7076 7086 7096 7106 7116 7126 7136 7146 7156
 CATGCTCTGC TGCTCTGCA AAAAAGCTAG AGTGGGCTGG GAGGCTGTT CATTGTACC TCGAGAATTT ATACTTAGAT AAGTATGATC TTACAGGTAT
 7166 7176 7186 7196
 ATTTCTATGA GATACGATG TATACATGCA TGATAATATT TAAAGCTT

FIGURA 22

4468 4476 4488 4496 4506 4516 4526 4536 4546 4556
 CCGCGGGGAT CGACTCATAA AATAAGTAACC TTCTAATGGG TATCTATTGA CTACCAACCA TTAGTGTGGT TGCAGAAAGC GGAATTCTGC GTTCTTTGAA
 4568 4576 4586 4596 4606 4616 4626 4636 4646 4656
 TTCAGCTTGC TTTTCATTTT TATTTTCCAT TTTTCAGTTT TTGTTTGTGT CGAATTTAGC CAGTTGCTTC TCCAAGATGA AAAAACCCCT GCGCAGTTTC
 4668 4676 4686 4696 4706 4716 4726 4736 4746 4756
 TGTGTGCGAA GATCCTAATC GACTTTTCCA CCCCCACAA AAGTAAATGT TTCTTTTATA CATTGCGGTG GGTAGCTAGC TCCCGAATTC TCAAAAGGACT
 4768 4776 4786 4796 4806 4816 4826 4836 4846 4856
 TAGGGACTGC ACTACATCAG AGTGTGTGCA CCTGTTTTCG TGCCTGGTTT GAAAGAAAAG ACCGGGAACT CCGCGGTTCG CCGGGAATAA TCATGCGATA
 4868 4876 4886 4896 4906 4916 4926 4936 4946 4956
 GTCCTTTGCG CTTCCAAATC GCATGTAGAG TAGACAACAG ACAGGAGGGG CAGGAAGGAT CTTTCACTGA GATGCTGTAT CTTGTTGGGT AAGTCGGATG
 4968 4976 4986 4996 5006 5016 5026 5036 5046 5056
 AAAGGGGAAT CGTATGAGAT TGGAGAGGAT GCGGAAGAGG TAACGCCCTT TGTTAACTTG TTTAATTATT ATGGGCGAGG CGAGAGGGGG AGGAATGTAT
 5068 5076 5086 5096 5106 5116 5126 5136 5146 5156
 GTGTGTGAGC CCGGCGAGAG GGAGCCATCC AGGCCAGGTA GAAATAGAGA AAGCCGAATG TTAGACAATA TGGCAGCGTA GTAGAGTAGG TAGGTAGGCA
 5168 5176 5186 5196 5206 5216 5226 5236 5246 5256
 AGTACTGCTA GCAAGAGGGA GAAAGGTAAG CTCACTCTTC GCATTCCACA CGGTAGTGTG GTCAGTTTAG ACAAAAAAAC AACTACTATA CCAATTAGTA
 5268 5276 5286 5296 5306 5316 5326 5336 5346 5356
 GACTGTGAAC TGACTTTTGG AACGGCTTTT CCGACTGCGA TTATTCGTGA GGAATCAAGG TAGGAATTTG GTCATATTTA CCGACACAGG TGGGTGATTG
 5368 5376 5386 5396 5406 5416 5426 5436 5446 5456
 CCATATGGAG TAAGAAAAGC AGATCATGGT ATCCTCAGAT ATGTTGCGGA ATTCTGTTCG CCGCAAGTTT CAGGGTGCTC TGGTGGGTTC CGGTGTGCTT
 5468 5476 5486 5496 5506 5516 5526 5536 5546 5556
 TTGCTTTGCT TCTCCCTGTT CTTGCATGTT AATAATAGCC TAGGCTGTGA GCGGAAGCTT AGGGTAGGCT TAGTGTGGA ACGTACATAT GTATCAGCTT
 5568 5576 5586 5596 5606 5616 5626 5636 5646 5656
 GACTTGGTTT AACCAAGGCA CCTGGTAGCC AGCCATACCC ACACAGCTTT TTTGATTCTC TCAGTATAGT TGTGAAAAGT GTAGGGAAGA TATGTGCTCT
 5668 5676 5686 5696 5706 5716 5726 5736 5746 5756
 GAGCAACAGC GTCTTTTCTT AGTAGTGGCG TCGGTTACTT GGTGACATT GGTATTGGA CTTTGTGCT ACACCATTGA CTACTGGAAG TCGAGTGCTA
 5768 5776 5786 5796 5806 5816 5826 5836 5846 5856
 AGGGTATGAT TTCTAGTGGT GAACACCTTT AGTTACGTAA TGTTTTCATT GCTGTTTTAC TTGAGATTTC GATTGAGAAA AAGTATTTTA ATAGCTGAAA
 5868 5876 5886 5896 5906 5916 5926 5936 5946 5956
 TCAATGTGTT ATCATTGTGA AGATGTTCTT CCTAACTCG AAAGGTATAT GAGGCTTTGG TTTCTTAGGA GAATTATTAT TCTTTTGTGA TGTGTGCTT
 5968 5976 5986 5996 6006 6016 6026 6036 6046 6056
 GTAGTTGGA AAGGTGAAGA GACAAAAGCT TAACACTTGA AATTTAGGAA AGAGCAGAAT TTGCAAAAAA AAATAAAAAA AAAATAAAAA GTTCGACTTC
 6068 6076 6086 6096 6106 6116 6126 6136 6146 6156
 TGAGCGGATA ACACTCGAGG GATCTTCATT ATGAAATCTT CTACTATATT AGCGGATGAT ACTGCTTTAA TTTCCGTGTT TATGGTGTCT CCAATTTCTA
 6168 6176 6186 6196 6206 6216 6226 6236 6246 6256
 CCGAAACTGA CATCGACGAT CTTCCAATTT CCGTTCCAGA AGAAGCCTTG ATTGGATTCA TTGACTTAAC CCGGGATGAA GTTTCCTGTT TGCCTGTTAA
 6268 6276 6286 6296 6306 6316 6326 6336 6346 6356
 TAACGGAACC CACACTGGTA TTCTATTCTT AAACACCACC ATCGCTGAAG CTGCTTTTGC TGACAAGGAT GATTGGAAGA AGCGGCTCC CATGACGACG
 6368 6376 6386 6396 6406 6416 6426 6436 6446 6456
 ACAACGCCCT TGAAGACAAQ CTGGGTTAAC TGCTCTAACA TGATCGATGA AATTATAACA CACTTAAAGC AGCCACCTTT GCTTTGCTG GACTTCAACA
 6468 6476 6486 6496 6506 6516 6526 6536 6546 6556
 ACCTCAATGG GGAAGACCAA GACATTCTGA TGGAAAAATA CTTTGAAGG CCAAACTTGC AGGCATTCAA CAGGCTGCTC AAGACTTTAC AGAAGCTATC
 6568 6576 6586 6596 6606 6616 6626 6636 6646 6656
 AGCAATTGAG AGCATTCTTA AAAATCTCCT GCCATGCTG CCGCTGGCCA CCGGCGACG CACGCGACAT CCAATCCATA TCAAGGACGG TGAATGGAAT
 6668 6676 6686 6696 6706 6716 6726 6736 6746 6756
 GAATTCCGGA GGAACCTGAC GTTCTATCTG AAAACCCCTG AGAATGCGCA GGGCTCAAGG ACCGACTTGA GCGCTCGGAT CTTTGTGATC GAAAGTGGAG
 6768 6776 6786 6796 6806 6816 6826 6836 6846 6856
 CTCGTTCTCT GGGCCTTCTC ACCACAGAGC CTCGGGACAT CAAAAACAGG AGAAGCTTGC AAACCTCTGG GTCATCTCTC ACACATTCCA GGACGAGAGG
 6868 6876 6886 6896 6906 6916 6926 6936 6946 6956
 CATTTCAOCT TTTCTGCGG CATCAGATGA ATTGTTAATT ATCTAATTC TGAATGTGAG AGCTCCCAT TGGCCTGTG CCGTTGTGTT CTCATTTTTA
 6968 6976 6986 6996 7006 7016 7026 7036 7046 7056
 TCCCATGAG ACTATTTATT TATGTATGTA TGTATTTATT TATTTATTGC CTGAGTGTG AACTGTATTT ATTTTAGCAG AGGAGCCATG TCGTGGTGTG
 7068 7076 7086 7096 7106 7116 7126 7136 7146 7156
 TCTGCAAAAA ACTCAGAGTG GGGTGGGAG CATGTTTATT TGTACCTGGA GAATTTATAC TTAGATAAGT ATGTACTTAC AGTATATTTT CTATGAGATA
 7168 7176 7186
 CTGATGTATA CATGCATGAT AATATTTAAA GCTT

FIGURA 23

10	20	30	40	50	60	70	80	90	100
TCGAATTTGC	GGGGAGAAGA	TGGATCTATG	CTAAATCTAA	ATAGGCATTT	GAATAACGAC	GACGAGTTAC	ACGACATATC	GCCATCTTTA	AATGAGCAAC
110	120	130	140	150	160	170	180	190	200
CACACTGGGA	CCTCATAGAG	GACGGGTCTC	GCTGGAGTAA	ATTTTTCAC	GGGATAATTA	AGAGGACAAG	AAGGTTCCAG	AAATCTTTAA	TGAGGTCTTT
210	220	230	240	250	260	270	280	290	300
AGTCAGAGGC	AGGAACAGCC	GTCAGGGGG	CATAAGACTA	CGGTCTATCC	CATCTGCCTC	TTCGTCCAGC	CTTGCCAACTA	GGGAGTTCTT	CAGAGGACTA
310	320	330	340	350	360	370	380	390	400
GAGGCTCAA	ACGAAATTAT	TGACAGCCTA	GACATCAATA	GTCATACAAC	AGAAAGCGAC	CACCCAACTT	TGGCTGATAA	TAGCGTATAA	ACAATGCTA
410	420	430	440	450	460	470	480	490	500
CTTTGTACGT	TCAAAATACA	ATGCAGTAGA	TATATTTATG	CATATTACAT	ATAATACATA	TCACATAGGA	AGCAACAGGC	GGGTTGGACT	TTTAATTTC
510	520	530	540	550	560	570	580	590	600
GAGGACCGCG	AATCCTTAGA	TCACACCCAA	TCCCCCAGAA	GTGATCCGUC	ACACACCCATA	GCTTCAAAAT	GTTTCTACTC	CTTTTACT	CTTCCAGATT
610	620	630	640	650	660	670	680	690	700
TTCTCGGACT	CCGCGCATCG	CCGTACCAC	TCAAAACACC	CAAGCACAGC	ATACTAAATT	TCCCTCTTTT	CTTCCCTCTAG	GGTGTCGTTA	ATTACCCGTA
710	720	730	740	750	760	770	780	790	800
CTAAAGGTTT	GGAAAAGAAA	AAAGAGACCG	CCTCGTTTCT	TTTTCTTCT	CGAAAAGGC	AATAAAAATT	TTTATCACGT	TTCTTTTCT	TGAAAATTTT
810	820	830	840	850	860	870	880	890	900
TTTTTTTGAT	TTTTTTCTCT	TTCGATGACC	TCCCATTTGAT	ATTTAAGTTA	ATAAACGGTC	TTCAATTCT	CAAGTTTCAG	TTTCATTTT	CTTGTCTAT
910	920	930	940	950	960	970	980	990	1000
TACAACTTTT	TTTACTTCTT	GCTCATTAGA	AAGAAAGCAT	AGCAATCTAG	TCGACAGATC	TCTCGAGTGC	TTTTGTGCGC	GTATGTTTAT	GTATGTACCT
1010	1020	1030	1040	1050	1060	1070	1080	1090	1100
CTCTCTCTAT	TTCTATTTT	AAACGACCTT	CTCAATATAA	TAAAAATAAT	AAAGTATTTT	TAAGGAAAAG	ACGTGTTTAA	GCACTGACTT	TATCTACTTT
1110									
TTGTACGTCT	AGA								

FIGURA 24