

Memória descritiva referente à patente de invenção de BEHRINGWERKE AKTIENGESELLSCHAFT, alemã, industrial e comercial, com sede em D-3550 Marburg, República Federal Alemã, (inventores: Dr. Hans-Ulrich Bernard, Dr. Tilmann Oltersdorf, residentes na República Federal Alemã e Dr. Klaus Seedorf, residente nos E.U.A.), para "PROCESSO PARA A UTILIZAÇÃO DE UM SISTEMA DE VECTORES CASSETE PARA A EXPRESSÃO DE CODÕES ABERTOS EM CÉLULAS EUCARIÓTICAS"

Memória descritiva

A invenção refere-se a um sistema multifacetado de vectores cassette que possibilita a transcrição e translacção de codões abertos (ORF) em células eucarióticas superiores e que permite, devido a uma origem de replicação E.coli, a selecção e amplificação em células de E.coli. Os sistemas até agora conhecidos para a preparação de proteínas de fusão não permitiam na medida desejada fazer previsões sobre as proteínas de fusão obtidas, em particular no caso da expressão de um ORF num CDNA ou em exões virais ou genômicos sem sinais de iniciação de translacção em especial sem codão de início ATG.



A invenção esclarece-se com as figuras apresentadas em anexo:

A figura 1, apresenta esquematicamente o sistema de vectores cassette de acordo com a invenção, em que P é o promotor, T a iniciação de translação, L uma região de polilinguagem e S o sinal de junção de poliadenilação e em que o círculo simples simboliza o fragmento PouII-EcoRI de pBR322 com o gene resistente à ampicilina (Amp^r) e a origem de replicação (ori). O segmento assinalado na zona da região de polilinguagem representa o DNA a construir com o ORF a exprimir.

A figura 2 representa na região superior o plasmídeo p297 no qual a região 323 bp é substituída pela sequência de DNA ordenada na parte inferior, obtendo-se o plasmídeo pATG. Neste plasmídeo podem introduzir-se como se indica a título de exemplo- nos locais indicados, por um lado um fragmento 670 bp - DNA com o promotor de citomegalovirus humano (MCMV) e, por outro lado, um fragmento 57-59 bp com uma região de polilinguagem, estando as três variantes desta região de polilinguagem numa configuração tal que o codão de iniciação ATG contido na zona enquadrada esteja sempre no codão de uma sequência de DNA a introduzir numa região de polilinguagem.

A figura 3, mostra finalmente a sequência de DNA das regiões mencionadas para os três vectores de expressão PORFEX11-13 e, em baixo, as respectivas sequências de aminoácidos, isto é, a região N-terminal das proteínas codificadas por estes vectores.

Os diferentes aspectos da invenção encontram-se definidos nas reivindicações. No que se segue ilustram-se de forma mais detalhada alguns aspectos particulares da invenção.



Em primeiro lugar descreve-se a construção dos vectores de acordo com a invenção, de onde ressalta, para o especialista, a larga gama de possibilidades de variação, sem que seja necessário entrar em pormenores detalhados.

O fragmento pBR 322-PouII-EcoRI com 2295 bp (pares de bases), que representa apenas, por exemplo, uma determinada sequência de DNA de um plasmídeo de E.coli com origem de replicação e marcador de selecção, liga-se com o fragmento SV40-BglIII-PouII com 323 pb, que contém o "early-BamHI-BglIII com 848 pb que contém o sinal de junção da região de t-antigene. Esta construção completa-se com um fragmento SV40-EcoRI-BamHI com 746 bp com a região de poliadenilação "tardia". O plasmídeo p297 assim obtido (figura 2) corresponde na sua construção aos plasmídeos pSV2 conhecidos (Mulligan e Berg, Science 20g (1980) 1422-1427).

O plasmídeo corta-se com BamHI, os extremos concavos enchem-se com a ajuda de polimerase de Klenow e os extremos convexos ligam-se de novo com compostos de ligação "blunt end". A partir deste plasmídeo, com as posições BamHI eliminadas, substitui-se então o fragmento Pou-II-BglIII com o promotor SV40 por um oligonucleótido com 34 bp, que se representa na figura 2. Obtém-se assim o plasmídeo pATG, A região de 10 bp enquadrada na figura corresponde à região de início SV40-T-antigene.

No plasmídeo pATG pode substituir-se a região desde a posição de corte SphI até à posição de corte KpnI por promotores eucarióticos desejados. O mais simples é proceder de modo a obter um fragmento com ambos os extremos convexos por meio de um promotor, por exemplo por enchimento com a ajuda de polimerase de Klenow, em que se introduz o fragmento convexo na posição de corte SmaI do vector puC18 ou puC19 e se corta depois como fragmento SphI-KpnI e se liga ao correspondente plasmídeo pATG cortado.



Escolhendo-se por exemplo o "immediate early promotor" de HCHV (Boshardt et al., Cell 41 (1985) 521-530; EP-A 0 173 177) como fragmento HincII-AvaII com 656 bp (556bp para além do "cap site") e procedendo-se como descrito anteriormente, obtém-se desse modo o vector de expressão pORFEX10.

Uma outra concretização vantajosa da invenção consiste na escolha do promotor "heat shock-protein 70" regulador de calor da Drosophila (hsp; steller, H., Dissertation 1985, Universidade de Heidelberg) na forma de um fragmento EcoRI-HindIII com 335 bp (85 bp para além do "cap-site" da mRNA). Por meio da introdução, acima descrita, no plasmídeo pATG obtém-se o plasmídeo de expressão pORFEX20.

Nos plasmídeos pORFEX10 e 20 podem introduzir-se sequências de DNA nas posições de corte BglII Hind III, que possuem um ORF no codão correcto para o codão de início ATG presente. Estes plasmídeos representam deste modo vectores de expressão de acordo com a invenção. Estes servem ainda como vectores de partida para outros vectores de acordo com a invenção, em especial para os que têm codões de leitura "desviados" ("frame shift").

Em vez do iniciador de translação SV40-T-antigene pôde também introduzir-se uma outra região correspondente, nomeadamente com vantagem por permuta do fragmento KpnI-BglII dentro do oligonucleótido sintético no plasmídeo ATG.

No lugar da região BglII-HindIII no oligonucleótido sintético no pATG pode introduzir-se o agente de polilinguagem modificado pUC18, cuja região 5'- está representada na figura 3. As sequências reproduzidas na figura 3 começam com o codão de início ATG, que está na figura 2 na zona enquadrada. As posições de corte BglII da figura 3 são idênticas à posição BglII que, na direcção 3'- é vizinha



à zona enquadrada referida. Por modificação da região na figura 3 entre as posições de corte BglIII e EcoRI obtém-se o afastamento gradual do codão de leitura em relação ao codão de início ATG.

É claro que o agente de polilinguagem pUC18 modificado é, por exemplo, uma região de polilinguagem que pode ser substituída por sequências de DNA de efeito análogo. Podem assim construir-se proteínas com um determinado terminal N-. Estas construções permitem a introdução de genes estruturais que apresentam, por exemplo, o seu próprio codão de início. Além disso podem codificar-se determinadas sequências de aminoácidos N-terminais, por exemplo para sinais de transporte (von Heijne, G., Eur. J. Biochem. 133 (1983)17-21).

As construções de acordo com a invenção não contêm qualquer codão ATG antes da inserção sintética. Deste modo garante-se um consumo muito eficiente deste triplete ATG em células eucarióticas. Assim, se se introduzir por exemplo um fragmento de DNA na posição de corte HindIII dos vectores pORFEX11 - 13, pode então dizer-se com segurança que a expressão ocorre a partir do codão ATG naquele dos três vectores em que o ORF no codão de leitura se liga com o codão ATG.

Uma condição necessária para uma expressão é que o promotor é que o promotor tenha sido construído na orientação correcta. No caso do fragmento do promotor HCMV pode verificar-se a orientação com base nas 3 posições BglI assimetricamente ordenadas e no caso do fragmento do promotor hsp pela posição de corte YhOI.

Como gene estrutural a experimentar pode utilizar-se por exemplo o gene bacteriano resistente contra a higromicina (Hm^r), bem como um outro gene que codifique para a cloranfenicol-acetiltransferase bacteriana (CAT). São



particularmente significativos os vectores de acordo com a invenção para a expressão de proteínas virais, em especial de proteína humana papillomvirus (HPV), por exemplo do tipo 6, 11,16 ou 18, que após transfecção de linhas celulares de rato ou humanas não dão transcrição ou dão apenas uma transcrição imprecisa, de modo que, por exemplo com anticorpos contra proteínas de fusão de Ecoli não se podem encontrar proteínas virais. Se se introduzir por exemplo, nos vectores de acordo com a invenção, o ORF para a proteína E7 de HPV18, obtém-se após transfecção de L-fibroblastos de rato uma expressão transiente ("transient expression") da proteína E7, que constitui cerca de 2% da quantidade total de proteína e que se pode localizar como componente do citoplasma.

Em comparação com as células Hela, que contêm 10 a 50 cópias endrogénicas de HPV18 (Schwarz et al., Nature 314 (1985) 111-114), obtém-se uma expressão superior por um factor de cerca de 100.

Nos exemplos seguintes esclarece-se mais em pormenor a invenção

Exemplo 1

A partir do plasmídeo pSV2CAT (Gorman et al., Mol. Cell. Biol. 2 (1982) 1044-1051) corta-se o fragmento com o gene CAT, por meio de HindIII e HpaI, que no sentido ascendente contém mais 36 nucleótido e no sentido descendente apresenta um segmento genético com o sinal de junção SV40 (que corresponde ao segmento BglII-HpaI de p297).

O vector pORFX10 abre-se com HindIII e HpaI e o fragmento respectivo liga-se ao gene CAT, obtendo-se o plasmídeo pORFEX10-CAT. Este possui, antes do codão de início CAT-ATG, um outro codão ATG, mas um outro codão de leitura. Este plasmídeo activa por isso apenas uma expressão reduzida pelo factor 20.



O plasmídeo pORFEX10-CAT corta-se por isso com KpnI e HindIII e o fragmento grande assim obtido liga-se com o oligonucleótido sintético AGCTTGTAC na presença de dATP. Obtém-se assim o plasmídeo pORFEX10-CAT- -ATG. Este plasmídeo activa em comparação com pSV2CAT- uma expressão de CAT mais elevada cerca de um factor 5.

TABELA 1

Plasmídeo	% de cloranfenicol transformado
-	0
pSV2CAT	12,4
pORFEX10-CAT	3,0
pORFEX10-CAT- ATG	60,9

Exemplo 2

Bernard et al., Exp. Cell Res. 158 (1985) 237-243, descreveram fragmentos BamHI-HindIII que contêm o gene bacteriano resistente à higromicina (Hm^r) em todos os três codões de leitura. Estes fragmentos introduzem-se como fragmentos BamHI-HindIII no pORFEX10, que se cortou com HindIII e BamHI. A transfecção em células-L dá origem, após 2 semanas e selecção contra a higromicina apenas no caso do clono pORFEX10-Hm24, a colónias bem formadas que correspondem ao plasmídeo de partida pHMR272 (Bernard et al., referência anterior). As poucas e pequenas colónias no caso do clono pORFEX10-Hm19 devem atribuir-se a um segundo codão ATG exterior ao codão de leitura. Os resultados apresentam-se na Tabela 2.

TABELA 2

Plasmídeo	Número de colônias
-	0
pHMR272	290
pORFEX10	0
pORFEX10-Hm16	0
pORFEX10-Hm19	11 (pequeno)
pORFEX10-Hm24	220
pORFEX10-Hm24*	0

* orientação invertida

Exemplo 3

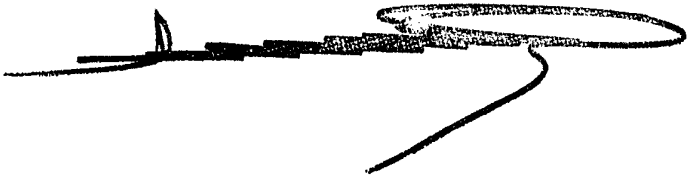
No registo de patente alemã
P 36 25 257.3 propõem-se construções de vectores semelhantes,
nas quais uma região de polilinguagem está deslocada por pas-
sos, de uma posição de nucleótido, de modo que o gene a expri-
mir aí introduzido fica automaticamente numa destas três or-
denações no codão de leitura correcto. Introduzindo-se no
vector pEX10-mer o ORF para a proteína HPV18-E7, o vector
obtido codifica para uma proteína de fusão na qual, na posi-
ção de transição entre a proteína MS2 e a proteína HPV18-E7
se encontra a seguinte sequência de DNA (cadeia codificante):

CTG AAT TCC GGG GGA TCC GTC

EcoRI

BamHI

Na sequência de EcoRI o triplete AAT
está no mesmo codão de leitura queem pORFEX12, mas não em pORF
EX11 ou 13. O mesmo é válido para o triplete TCC na posição
de corte BamHI.



Uma vez que ao cortar-se com BamHI se obtêm o mesmo extremo convexo como com BglII, podem ligar-se os vectores pORFEX10 e pORFEX20 depois de cortados com BglII com as sequências cortadas com BamHI, estando as últimas também ordenadas em códões de leitura correctos.

Transfectando-se células-L com os vectores assim obtidos, obtêm-se uma "transient expression" da proteína de fusão codificada, que abrange a maior parte da proteína HPV18-E7. Após lisagem das células e análise "western blot" com anticorpos, que são mencionados no registo de patente referido atrás, encontra-se a proteína esperada nas células que foram infectadas com pORFEX10, 12 e 20 (no último caso após "heat shock"), mas não nas células que foram infectadas com pORFEX11 e 13 ou nas células que continham construções com a orientação do promotor HCMV invertida.

Em células Hela com 10 a 50 cópias endogénicas de HPV18 encontrou-se, por meio de anticorpos de coelho contra a proteína de fusão HPV18-E7, uma pequena banda para uma proteína 12,5 kDa, isto é, da ordem de grandeza esperada. Uma análise quantitativa por meio de contagem de cintilação das bandas E7 mostrou que as células-L infectadas, com 2% em média de proteína E7 em relação ao total de proteína celular, continham uma quantidade 100 vezes superior à das células Hela, que apenas tinham 0,01% do total de proteína celular como proteína E7.

Por meio de "immunestaining" pôde mostrar-se que a proteína E7 se localiza no citoplasma.

Condições do processo:

Cultivaram-se fibroblastos LTK- de rato num meio Eagle modificado de acordo com Dulbecco (DMEM)



suplementado com soro de vitela fetal a 10%, tampão de HEPES 10 mM (pM 7,4) e 100 ug/ml de penicilina/estreptomicina. Para a transfecção "transiente" de DNA colocaram-se $3 \cdot 10^6$ células numa cápsula de 10 cm de diâmetro, 24 horas mais tarde lavaram-se 2 vezes com DMEM e adicionaram-se 10 µg de DNA-plasmídeo em 2 ml de DMEM, que continha 200 µg/ml de DMEM-Dextran (0,5 mDa) e 0,1M de cloroquina. Após 5 horas de incubação a 37°C lavaram-se as células 2 vezes com DMEM e cultivaram-se depois, durante 48 horas, em DMEM com o suplemento acima mencionado. No caso em que se introduziu um plasmídeo com o promotor hsp, efectuou-se um "heat shock" de 5 horas a 42°C, após o tempo de incubação.

A análise das células transfectadas, por meio de um teste de CAT, efectuou-se com 5 µg de proteína, que se trata com cloranfenicol- ^{14}C de 62,5 uCi (Gorman et al., referência acima indicada).

Para a imunocitoquímica colocaram-se as células 24 horas após a utilização do DNA em recipientes de vidro tratados em autoclave. 24 horas depois lavaram-se os recipientes com uma solução de cloreto de sódio tamponizada com fosfato. As células fixaram-se durante 5 minutos em metanol gelado e durante 2 minutos em acetona. A imunocitoquímica efectuou-se com antisoro de coelho contra a proteína de fusão MS2-HPV18-E7 (registro de patente alemão P 36 26 257.3; Seedorf, K., Disertação, Universidade de Heidelberg, 1986) e com um sistema de detecção usual no mercado. Os recipientes incubaram-se primeiro durante 1 hora com os anticorpos referidos numa diluição de 1:50 a 1:200, lavaram-se 3 vezes com uma solução de cloreto de sódio tamponizada com fosfato, incubaram-se durante 1 hora com um anticorpo biotinil-anti-coelho, lavaram-se novamente e incubaram-se durante 1 hora com um complexo de estreptavidina-biotinil-peroxidase (diluição 1:100). Como substrato para a reacção corada utilizou-se 3-amina-9-etilcarbazole a 0,02% em tampão de acetato 50 mM (pM 5,2 com peróxido de hidrogénio a 0,93%. Todas as incubações se efectuaram a 37°C

R E I V I N D I C A Ç Õ E S

- 1ª -

Processo de engenharia genética em que se utiliza um sistema de vectores cassette, contendo

- a) um segmento de um plasmídeo de E.coli com um início de replicação activo na E.coli e um marcador de selecção,
- b) segmentos permutáveis, activos em células eucarióticas superiores, nomeadamente
 - b¹) um promotor,
 - b²) um iniciador de translação
 - b³) um agente de polilincagem,
 - b⁴) um sinal de junção e
 - b⁵) um sinal de poliadenilação,

podendo também dois destes segmentos permutáveis apresentar-se como unidades permutáveis em conjunto, caracterizado por se ligar o fragmento pBr 322-PvuII-EcoRI de 2295 pb (pares de bases) com o fragmento SV40-Bgl II-Pvu II de 323 pb e em seguida com o fragmento SV40-Bam HI-Bgl II de 848 pb, completando-se pelo fragmento SV40-EcoRI-Bam HI de 746 pb, obtendo-se o plasmídeo p 297, e se cortar este com Bam HI, se encherem os extremos convexos com polimerase de Klenow e se ligarem de novo os extremos concavos e se substituir, no plasmídeo obtido, o fragmento Pvu II-Bgl II por um oligonucleótido com 34 pb, obtendo-se o plasmídeo p ATG com as sequências de DNA (cadeia codificante)

ATG GAT AGA TCT (1)

ATG GAT AGA TCT C (2)

ATG GAT AGA TCT CC (3)

em que

o tripleto ATG no extremo 5' das sequências (1) (2) e (3) constitui o único codão de início de translação, colocado funcionalmente depois do promotor, e em que se liga às sequências (1), (2) e (3) do agente de polilincagem b³.

- 2ª -

Processo de acordo com a reivindicação 1, caracterizado por o segmento do plasmídeo de E.coli ser o fragmento 2295 bp pBr 322 Pvu II- EcoRI.

- 3ª -

Processo de acordo com as reivindicações 1 ou 2, caracterizado por o promotor ser o promotor hsp da proteína-70 Drosophila- "heat Shock".

A requerente declara que o primeiro pedido desta patente foi apresentado na República Federal Alemã, em 24 de Janeiro de 1987, sob o nº P 37 02 116.8.

Lisboa, 21 de Janeiro de 1988

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



R E S U M O

"PROCESSO PARA A UTILIZAÇÃO DE UM SISTEMA DE VECTORES CASSETE PARA A EXPRESSÃO DE CODÕES ABERTOS EM CÉLULAS EUCARIÓTICAS"

A invenção refere-se a um processo de engenharia genética em que se utiliza um sistema de vectores cassette, contendo

- a) um segmento de um plasmídeo de E.coli com um início de replicação activo na E.coli e um marcador de selecção,
- b) segmentos permutáveis, activos em células eucarióticas superiores, nomeadamente

b¹) um promotor,

b²) um iniciador de translação,

b³) um agente de polilincagem,

b⁴) um sinal de junção e

b⁵) um sinal de poliadenilação,

podendo também dois destes segmentos permutáveis apresentar-se como unidades permutáveis em conjunto,

que compreende ligar-se o fragmento pBr 322-PvuII-EcoRI de 2295 pb (pares de bases) com o fragmento SV40-Bgl II-Pvu II de 323 pb e em seguida com o fragmento SV40-Bam HI-Bgl II de 848 pb, completando-se pelo fragmento SV40-EcoRI-Bam HI de 746 pb, obtendo-se o plasmídeo p 297,

e se cortar este com Bam HI, se encherem os extremos convexos com polimerase de Klenow e se ligarem de novo os extremos côncavos e se substituir, no plasmídeo obtido, o fragmento Pvu II-Bgl II por um oligonucleótido com 34 pb, obtendo-se o plasmídeo pATG com as sequências de DNA (cadeia codificante)

ATG GAT AGA TCT (1)

ATG GAT AGA TCT C (2)

ATG GAT AGA TCT CC (3)

em que

o triplete ATG no extremo 5' das sequências (1), (2) e (3)



constitui o único codão de início de translacção, colocado funcionalmente depois do promotor, e que se liga às sequências (1), (2) e (3) do agente de polilincagem b^3 .

FIG.1

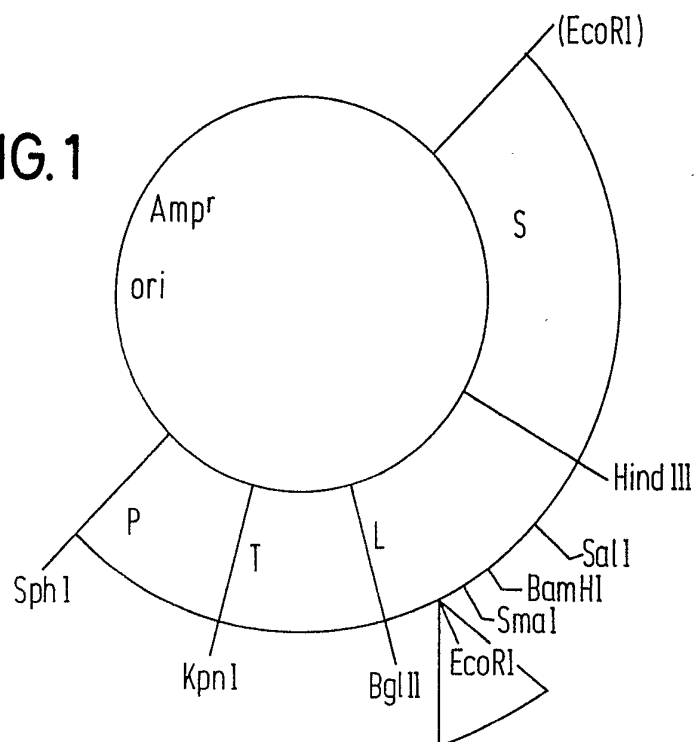
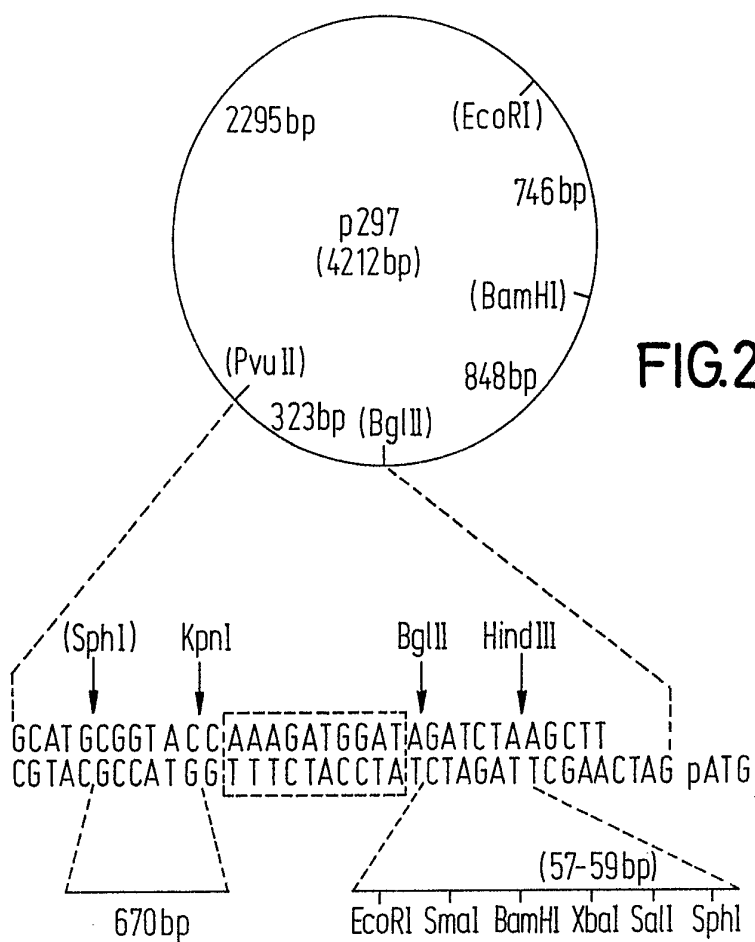


FIG.2



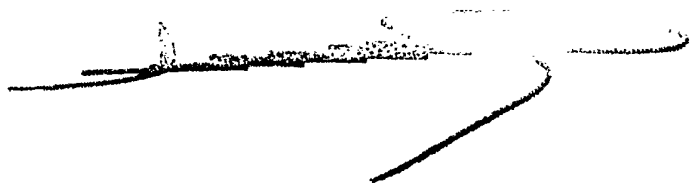


FIG. 3

pORFEX11 BglII EcoRI BamHI HindIII
ATG GAT AGA TCT GGA ATT CGA GCT CGG TAC CCG GGG ATC CTT
Met-Asp-Arg-Ser-Gly-Ile-Arg-Ala-Arg-Tyr-Pro-Gly-Ile-Leu-

pORFEX12 BglII *EcoRI BamHI HindIII
ATG GAT AGA TCT CGG AAT TCG AGC TCG GTA CCC GGG GAT CCT
Met-Asp-Arg-Ser-Arg-Asn-Ser-Ser-Ser-Val-Pro-Gly-Asp-Pro-

pORFEX13 BglII **EcoRI BamHI HindIII
ATG GAT AGA TCT CCG GAA TTC GAG CTC GGT ACC CCG GGA TCC
Met-Asp-Arg-Ser-Pro-Glu-Phe-Glu-Leu-Gly-Thr-Arg-Gly-Ser-