

(11) Número de Publicação: PT 728200 E

(51) Classificação Internacional:

C12N 15/12 (2006.01) C07K 14/285 (2006.01)

C12N 1/21 (2006.01) A61K 39/395 (2006.01)

C07K 16/12 (2006.01)

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

(22) Data de pedido: 1994.11.07

(30) Prioridade(s): 1993.11.08 US 0148968
1993.12.29 US 0175116

(43) Data de publicação do pedido: 1996.08.28

(45) Data e BPI da concessão: 2006.08.30
012/2006

(73) Titular(es):

SANOFI PASTEUR LIMITED

1755 STEELES AVENUE WEST TORONTO,
ONTARIO M2R 3T4

CA

(72) Inventor(es):

SHEENA LOOSMORE

CA

MICHEL KLEIN

CA

YAN-PING YANG

CA

ROBIN HARKNESS

CA

ANTHONY SCHRYVERS

CA

(74) Mandatário:

ANTÓNIO JOÃO COIMBRA DA CUNHA FERREIRA
R DAS FLORES 74 4 AND 1249-235 LISBOA

PT

(54) Epígrafe: GENES DE RECEPTORES DE TRANSFERRINA DE HAEMOPHILUS

(57) Resumo:

RESUMO

"Genes de receptores de transferrina de *Haemophilus*"

Proporciona-se ácido nucleico purificado e isolado que codifica uma proteína receptora de transferrina de uma estirpe de *Haemophilus* ou um fragmento ou um análogo da proteína receptora de transferrina. A sequência de ácido nucleico pode ser utilizada para produzir péptidos isentos de contaminantes derivados de bactérias que normalmente contêm as proteínas Tbp1 ou Tbp2 para fins de diagnóstico e tratamento médico. Adicionalmente, a molécula de ácido nucleico pode ser utilizada no diagnóstico de infecção. São igualmente proporcionados Tbp1 ou Tbp2 recombinantes e métodos para a purificação das mesmas. São proporcionados vectores vivos que expressam epítomos da proteína receptora de transferrina para vacinação.

DESCRIÇÃO

"Genes de receptores de transferrina de *Haemophilus*"

CAMPO DA INVENÇÃO

A presente invenção refere-se a péptidos contendo uma sequência de aminoácidos conservada de uma proteína receptora da transferrina e a composições imunogénicas e anti-soros ou anticorpos deles derivados.

ANTECEDENTES DA INVENÇÃO

As estirpes de *Haemophilus influenzae* tipo b encapsuladas são a principal causa da meningite bacteriana e outras infecções invasivas em crianças pequenas. Contudo, as *H. influenzae* não encapsuladas ou não tipificáveis (NTHi) são responsáveis por uma ampla gama de doenças humanas incluindo otite média, epiglotite, pneumonia e traqueobronquite. As vacinas baseadas no polissacárido capsular de *H. influenzae* tipo b conjugado com o toxóide da difteria (Berkowitz *et al.*, 1987. Ao longo deste pedido são referidas várias referências entre parêntesis para descrever mais completamente o estado da arte à qual pertence a presente invenção. A informação bibliográfica completa para cada citação encontra-se no final da descrição, imediatamente antes das reivindicações. O toxóide do tétano (Classon *et al.*, 1989 e Patente US 4 496 538), ou a proteína da membrana externa de *Neisseria meningitidis* (Black *et al.*, 1991) têm sido eficazes na redução da meningite induzida por *H. influenzae* tipo b, mas não na doença induzida por NTHi (Bluestone, 1982).

A otite média é a doença mais comum no início da infância com 60-70% de todas as crianças com menos de 2 anos de idade sofrendo entre uma e três infecções nos ouvidos. A otite média crónica é responsável por deficiências auditivas, da fala e cognitivas nas crianças. As infecções por *H. influenzae* são responsáveis por cerca de 30% dos casos de otite média aguda e cerca de 60% de otite média crónica. Só nos Estados Unidos, o tratamento da otite média custa entre 1 e 2 biliões de dólares por ano para antibióticos e procedimentos cirúrgicos tais como tonsilectomias, adenoidectomias e inserção de tubos de timpanostomia. Adicionalmente, muitos dos organismos

causadores de otite média estão a tornar-se resistentes ao tratamento com antibióticos. Uma vacina profiláctica eficaz contra a otite média é portanto desejável. As estirpes não tipificáveis de *H. influenzae* são também importantes patogénios responsáveis por pneumonia nos indivíduos mais idosos e noutros indivíduos que são particularmente susceptíveis a infecções respiratórias. Existe assim uma necessidade de antigénios de *H. influenzae* que sejam úteis como componentes em preparações imunogénicas que proporcionem protecção contra os muitos serótipos de *H. influenzae*.

O ferro é um nutriente essencial para o crescimento de muitas bactérias. Vários patogénios humanos, tais como *H. influenzae*, *Branhamella catarrhalis*, *N. meningitidis*, *N. gonorrhoeae* e estirpes comensais de *Neisseria* não patogénicas, podem utilizar transferrina humana como fonte de ferro (Schryvers, 1988; Schryvers e Lee, 1989; Mickelsen e Sparling, 1981). O receptor de transferrina bacteriano (TfR) é composto por duas cadeias, Tbp1 e Tbp2. Em estirpes de *H. influenzae*, o peso molecular da Tbp1 é de aproximadamente 100 000, enquanto o peso molecular da Tbp2 é variável, variando de 60 000 a 90 000, dependendo da estirpe (Schryvers e Gray-Owen, 1992; Holland et al., 1992). Pensa-se que a expressão do receptor de transferrina de *H. influenzae* é regulada por ferro e/ou hemina (Morton et al., 1993) e o local de ligação *fur* putativo (Braun e Hantke, 1991) foi identificado a montante de *tbp2*. Esta sequência encontra-se na região promotora de genes que são regulados negativamente pelo ferro, incluindo o TfR de *N. meningitidis* (Legrain et al., 1993). O promotor é seguido pelos genes *tbp2* e *tbp1*, um arranjo encontrado noutros operões de TfR bacterianos (Legrain et al., 1993; Wilton et al., 1993). Anticorpos que bloqueiam o acesso do receptor de transferrina à sua fonte de ferro podem impedir o crescimento bacteriano. Em adição, anticorpos contra TfR que são opsonizantes ou bactericidas podem também proporcionar protecção através de mecanismos alternativos. Assim, o receptor de transferrina, seus fragmentos, suas cadeias constituintes, ou péptidos seus derivados, são candidatos a vacinas para proteger contra a doença de *H. influenzae*. Ratinhos imunizados com proteínas TfR de *N. meningitidis* em adjuvante de Freund foram protegidos da provocação homóloga e os anti-soros anti-TfR eram bactericidas

e protectores num ensaio de transferência passiva (Danve *et al.*, 1993). Porcos imunizados com Tbp2 de *A. pleuropneumoniae* recombinante foram protegidos contra a provocação homóloga mas não contra a provocação heteróloga (Rossi-Campos *et al.*, 1992). Estes resultados indicam a eficácia de vacinas baseadas em TfR na protecção contra a doença. Seria desejável proporcionar a sequência da molécula de ADN que codifica o receptor de transferrina e péptidos correspondentes a porções do receptor de transferrina e vectores contendo estas sequências, para o diagnóstico, a imunização e a produção de reagentes de diagnóstico e imunológicos.

O poliovírus é um enterovírus, um género da família *Picornaviridae*. Existem três serótipos distintos do vírus, e múltiplas estirpes em cada serótipo. As estirpes virulentas são agentes causadores de poliomielite paralítica. As estirpes atenuadas, que possuem um reduzido potencial para causar a doença paralítica, e as estirpes virulentas inactivadas, são utilizadas como vacinas. A infecção com o vírus induz imunidade mucosa, protectora, duradoura. A inoculação com vacinas de poliovírus inactivado podem também induzir uma resposta imunitária mucosa.

A estrutura do poliovírus é conhecida, e é altamente conservada entre estirpes e serótipos. As estruturas de vários outros picornavírus (vírus pertencentes a géneros da família *Picornaviridae*) foi também determinada e foi mostrado que se relacionava estreitamente com a estrutura do poliovírus. É possível expressar epítomos estranhos sobre a cápside ou os poliovírus (Murdin *et al.*, 1992) e este trabalho foi estendido a outros picornavírus. Os epítomos que foram expressos são usualmente curtos, bem definidos, epítomos contíguos, e a maioria foram expressos no interior do local I antigénico de neutralização (N_{Ag}I) de poliovírus ou no local equivalente noutros picornavírus. Este local inclui as cadeias B e C beta de ligação em ansa (a ansa BC) da proteína da cápside de poliovírus VP1. A ansa BC de VP1 é uma ansa exposta na superfície de nove aminoácidos que pode ser substituída e alongada com pelo menos vinte cinco aminoácidos heterólogos (Murdin *et al.*, 1991). Os poliovírus híbridos ou quiméricos que expressam epítomos de receptores de transferrina, que crescem até um título elevado e são imunogénicos, seriam úteis

como vacinas e como ferramentas para a produção de reagentes imunológicos.

Gray-Owen, S.D. et al. (1993), *Microbial Pathogenesis*, Vol. 14, páginas 389-398 descrevem anticorpos monoclonais específicos para a transferrina humana.

Stevenson, P. et al. (1992) *Infection & Immunity* Vol. 60, N°.6, páginas 2391-2396, descrevem um anti-soro contra TBP-2 purificado a partir de uma estirpe de *N. meningitidis* que reage cruzadamente com todos os isolados meningocócicos examinados e com TBP-2 de várias estirpes de *H. influenzae*, tipo b.

SUMÁRIO DA INVENÇÃO

A presente invenção proporciona péptidos possuindo não menos de sete aminoácidos e não mais de 150 aminoácidos e contendo uma sequência de aminoácidos que é conservada entre bactérias que produzem uma proteína receptora da transferrina compreendendo TBP-2, sendo essa sequência TBP-2 conservada:

LEGGFYG (SEQ ID NO: 85)

e contendo uma sequência de aminoácidos que é:

LEGGFYGP (SEQ ID NO: 74).

O péptido pode incluir uma sequência de aminoácidos seleccionada entre SEQ ID NO: 95 e SEQ ID NO: 74.

De acordo com outro aspecto da invenção, é proporcionada uma composição imunogénica que compreende um tal péptido como componente activo. A composição pode compreender pelo menos um péptido sintético, como aqui proporcionado, e um seu transportador farmacêuticamente aceitável. O pelo menos um componente activo produz uma resposta imunitária quando administrado a um hospedeiro.

As composições imunogénicas aqui proporcionadas podem ser formuladas na forma de uma vacina para administração *in vivo* para proteger contra doenças causadas por patogénios bacterianos que produzem receptores de transferrina. Para este fim, as composições podem ser formuladas na forma de uma preparação de micropartículas, cápsulas ou lipossomas.

Alternativamente, as composições podem ser proporcionadas em combinação com uma molécula de direccionamento para entrega a células específicas do sistema imunitário ou a superfícies mucosas. A composição imunogénica pode compreender uma pluralidade de componentes activos para proporcionar protecção contra uma doença causada por uma pluralidade de espécies de bactérias produtoras de receptores de transferrina. As composições imunogénicas podem ainda compreender um adjuvante.

As composições podem ser utilizadas num método para induzir protecção contra a infecção ou doença causadas por *Haemophilus* ou outras bactérias que produzem proteína receptora da transferrina, compreendendo o passo de administração a um hospedeiro susceptível, tal como um ser humano, de uma quantidade eficaz da composição imunogénica como atrás descrita.

De acordo com outro aspecto da invenção, é proporcionado um anti-soro ou um anticorpo específicos para qualquer uma das sequências de aminoácidos SEQ ID NO: 85, SEQ ID NO: 74 ou SEQ ID NO: 50.

Pode ser proporcionado um vector vivo para entrega do receptor de transferrina a um hospedeiro, compreendendo um vector contendo a molécula de ácido nucleico como descrita atrás. O vector pode ser seleccionado entre *Salmonella*, BCG, adenovírus, poxvírus, vaccinia e poliovírus. O vector pode ser especificamente um poliovírus e a molécula de ácido nucleico pode codificar para um fragmento de receptor de transferrina possuindo uma sequência de aminoácidos de LEGGFYGP (SEQ ID NO: 74) ou LEGGFYG (SEQ ID NO: 85). Descrevemos aqui os vectores pT7TBP2A, pT7TBP2B, pT7TBP2C e PT7TBP2D (designações ATCC N^{os}. 75931, 75932, 75933, 75934).

BREVE DESCRIÇÃO DOS DESENHOS

A presente invenção será mais compreendida a partir da seguinte descrição com referência aos desenhos, nos quais:

A Figura 1A mostra o mapa de restrição de dois clones plasmídicos (pBHT1 e pBHT2) do operão do receptor de transferrina de *Haemophilus influenzae* tipo b, estirpe DL63.

A Figura 1B mostra o mapa de restrição dos clones S-4368-3-3 e JP-901-5-3 contendo genes de TfR de *H. influenzae* tipo b, estirpe Eagan.

A Figura 1C mostra o mapa de restrição dos clones DS-712-1-3 contendo o gene do receptor de transferrina de *H. influenzae* tipo b, estirpe MinnA.

A Figura 1D mostra o mapa de restrição do clone JB-1042-7-6 contendo o gene do receptor de transferrina de *H. influenzae* não tipificável, estirpe PAK 12085.

A Figura 2 ilustra a organização e os mapas de restrição dos genes de Tbp1 e Tbp2 clonados de estirpes identificadas e a organização genética do operão de TfR com dois genes (*tbp1* e *tbp2*) em tandem formando um operão sob a regulação da transcrição de um único promotor e representa também o fragmento de ADN de 3,0 kb de pBHIT2 usado para sondar bibliotecas quanto a genes de TfR das estirpes de *Haemophilus*.

A Figura 3 mostra as sequências de nucleótidos dos genes de receptores de transferrina (SEQ ID NO: 1) e suas sequências de aminoácidos deduzidas (SEQ ID NO: 5 - Tbp1 e SEQ ID NO: 6 - Tbp2) de *H. influenzae* tipo b, estirpe DL63. As sequências de aminoácidos sublinhadas correspondem a péptidos de Tbp1 identificados por sequenciação de aminoácidos. As sequências de sinal putativas estão indicadas por sublinhado duplo e correspondem aos resíduos 1 a 17 para Tbp1 e 1 a 25 para Tbp2.

A Figura 4 mostra as sequências de nucleótidos dos genes de receptores de transferrina (SEQ ID NO: 2) e suas sequências de aminoácidos deduzidas (SEQ ID NO: 7 - Tbp1 e SEQ ID NO: 8 - Tbp2) de *H. influenzae* tipo b, estirpe Eagan. As sequências putativas -35, -10 e do local de ligação ao ribossoma estão sobrelinhadas.

A Figura 5 mostra as sequências de nucleótidos dos genes de receptores de transferrina (SEQ ID NO: 3) e suas sequências de aminoácidos deduzidas (SEQ ID NO: 9 - Tbp1 e SEQ ID NO: 10 - Tbp2) de *H. influenzae* tipo b, estirpe MinnA. As sequências

putativas -35, -10 e do local de ligação ao ribossoma estão sobrelinhadas.

A Figura 6 mostra as sequências de nucleótidos dos genes de receptores de transferrina (SEQ ID NO: 4) e suas sequências de aminoácidos deduzidas (SEQ ID NO: 11 - Tbp1 e SEQ ID NO: 12 - Tbp2) de *H. influenzae* não tipificável, estirpe PAK 12085. As sequências putativas -35, -10 e do local de ligação ao ribossoma estão sobrelinhadas.

A Figura 7 mostra as sequências de nucleótidos dos genes de receptores de transferrina (SEQ ID NO: 105) e suas sequências de aminoácidos deduzidas (SEQ ID NO: 106 - Tbp1 e SEQ ID NO: 107 - Tbp2) da estirpe de *H. influenzae* não tipificável SB33.

A Figura 8 mostra a sequência de nucleótidos do gene de Tbp2 (SEQ ID NO: 108) e a sequência de aminoácidos deduzida (SEQ ID NO: 109 - Tbp2) da estirpe não tipificável de *H. influenzae*, estirpe SB12.

A Figura 9 mostra a sequência de nucleótidos do gene de Tbp2 (SEQ ID NO: 110) e a sequência de aminoácidos deduzida (SEQ ID NO: 111 - Tbp2) da estirpe não tipificável de *H. influenzae*, estirpe SB29.

A Figura 10 mostra a sequência de nucleótidos do gene de Tbp2 (SEQ ID NO: 112) e a sequência de aminoácidos deduzida (SEQ ID NO: 113 - Tbp2) da estirpe não tipificável de *H. influenzae*, estirpe SB30.

A Figura 11 mostra a sequência de nucleótidos do gene de Tbp2 (SEQ ID NO: 114) e a sequência de aminoácidos deduzida (SEQ ID NO: 115 - Tbp2) da estirpe não tipificável de *H. influenzae*, estirpe SB32.

A Figura 12A mostra as sequências de nucleótidos das regiões promotoras e extremidade 5' dos genes *tbp2* de *H. influenzae*, das estirpes Eagan (SEQ ID NO: 116), MinnA (SEQ ID NO: 117), PAK 12085 (SEQ ID NO: 118) e SB33 (SEQ ID NO: 119). O iniciador da cadeia de codificação utilizado para

amplificar os genes *tbp2* por PCR está sublinhado (SEQ ID NO: 120).

A Figura 12B mostra a sequência de nucleótidos da região intergênica e da extremidade 5' dos genes *tbp1* de *H. influenzae*, das estirpes Eagan (SEQ ID NO: 121), MinnA (SEQ ID NO: 122), DL63 (SEQ ID NO: 123), PAK 12085 (SEQ ID NO: 124), SB12 (SEQ ID NO: 125), SB29 (SEQ ID NO: 126), SB30 (SEQ ID NO: 127), e SB32 (SEQ ID NO: 128). O iniciador da cadeia não codificante utilizado para amplificar os genes *tbp2* por PCR está sublinhado (SEQ ID NO: 129).

A Figura 13 mostra a análise em gel de agarose de genes *tbp2* amplificados por PCR das estirpes não tipificáveis de *H. influenzae*, SB12, SB29, SB30, SB32 e SB33. A pista 1 é SB33, a pista 2 é SB12, a pista 3 é SB29, a pista 4 é SB30, a pista 5 é SB32.

A Figura 14 mostra uma comparação das sequências de aminoácidos de Tbp1 de *H. influenzae*, estirpes Eagan, DL63, PAK 12085 e SB33 (SEQ ID NOS: 7, 5, 11 e 106), de *N. meningitidis*, estirpes B16B6 e M982 (SEQ ID NOS: 94 e 95), e de *N. gonorrhoeae*, estirpe FA19 (SEQ ID NO: 96).

A Figura 15 mostra uma comparação da sequência de aminoácidos de Tbp2 de *H. influenzae*, estirpes Eagan, DL63, PAK 12085, SB12, SB29, SB30 e SB32 (SEQ ID NOS: 8, 6, 12, 109, 110, 112, 114), de *N. meningitidis*, estirpes B16B6 e M982 (SEQ ID NOS: 97 e 98), de *N. gonorrhoeae*, estirpe FA19, e de *Actinobacillus pleuropneumoniae*, estirpes AP205 e AP37 (SEQ ID NOS: 99 e 100).

A Figura 16A mostra a estrutura secundária prevista da proteína Tbp1 de *H. influenzae* e a Figura 16B mostra a estrutura secundária prevista da proteína Tbp2 de *H. influenzae*.

A Figura 17 mostra o esquema de construção do plasmídeo JB-1468-29 que expressa Tbp1 de *H. influenzae* tipo b Eagan a partir de *E. coli*.

A Figura 18 mostra o esquema de construção do plasmídeo JB-1424-2-8 que expressa Tbp2 de *H. influenzae* tipo b Eagan a partir de *E. coli*.

A Figura 19 mostra os pares de oligonucleótidos (SEQ ID NOS: 130, 131) utilizados para construir o plasmídeo JB-1424-2-8.

A Figura 20 mostra a sequência dos pares de oligonucleótidos A (SEQ ID NOS: 86, 87), B (SEQ ID NOS: 88, 89), C (SEQ ID NOS: 90, 91) e D (SEQ ID NOS: 92, 93) para a construção de plasmídeos de expressão de Tbp1 e Tbp2.

A Figura 21 mostra o esquema de construção do plasmídeo JB-1600-1 que expressa Tbp2 de *H. influenzae*, estirpe SB12, a partir de *E. coli*.

A Figura 22 mostra géis de SDS-PAGE de produtos da expressão da proteína Tbp1 de *Haemophilus* tipo b Eagan, proteína Tbp2 de Eagan, e proteína Tbp2 de *H. influenzae* não tipificável SB12, a partir de *E. coli*. Na pista 1, JB-1476-2-1 (T7/Tbp1 de Eagan) a t₀; na pista 2, JB-1476-2-1 a t=4h de indução; na pista 3, marcadores de peso molecular de 200 kDa, 116 kDa, 97,4 kDa, 66 kDa, 45 kDa e 31 kDa; na pista 4, JB-1437-4-1 (T7/Tbp2 de Eagan) a t₀; na pista 5, JB-1437-4-1 a t=4h de indução; na pista 6, JB-1607-1-1 (T7/Tbp2 de SB12) a t₀; na pista 7, JB-1607-1-1 a t=4h de indução.

A Figura 23 mostra um esquema de purificação para Tbp1 e Tbp2 recombinantes expressas a partir de *E. coli*.

A Figura 24 mostra uma análise da pureza de Tbp1 e Tbp2 recombinantes purificadas através do esquema da Figura 23. A pista 1 contém marcadores de peso molecular (106, 80, 49,5, 32,5, 27,5 e 18,5 kDa), a pista 2 é lisado de células inteiras de *E. coli*. A pista 3 é corpos de inclusão solubilizados. A pista 4 é Tbp1 ou Tbp2 purificadas.

A Figura 25 mostra a imunogenicidade de rTbp1 (painel superior) e rTbp2 (painel inferior) em ratinhos.

A Figura 26 mostra a reactividade de anti-soros anti-rTbp1 de Eagan com várias estirpes de *H. influenzae* num *Western blot*. Pista 1, BL21/DE3; pista 2, SB12-EDDA; pista 3, SB12 +EDDA; pista 4, SB29 - EDDA; pista 5, SB29 +EDDA; pista 6, SB33 -EDDA; pista 7, SB33 + EDDA; pista 8, Eagan - EDDA; pista 9, Eagan +EDDA; pista 10, *B. catarrhalis* 4223 - EDDA; pista 11, *B. catarrhalis* 4223 +EDDA; pista 12, *N. meningitidis* 608 - EDDA; pista 13, *N. meningitidis* 608 + EDDA; pista 14, JB-1476-2-1 induzida que expressa Tbp1 de Eagan recombinante; pista 15, marcadores de peso molecular. As bandas específicas de ~95 kDa reagiram com os anti-soros anti-Tbp1 nas pistas 3, 4, 5, 7, 8 e 9, correspondentes às estirpes SB12, SB29, SB33 e Eagan de *H. influenzae*; as bandas de ~110 kDa nas pistas 10 e 11, correspondentes à estirpe 4223 de *B. catarrhalis*; e as bandas de ~80 kDa nas pistas 12 e 13, correspondentes a *N. meningitidis* 608.

A Figura 27 mostra a reactividade de anti-soros anti-rTbp2 de Eagan com várias estirpes de *H. influenzae* num *Western blot*. Pista 1, marcadores de peso molecular; pista 2, JB-1437-4-1 que expressa Tbp2 de Eagan recombinante, induzido; pista 3, SB12-EDDA; pista 4, SB12 +EDDA; pista 5, SB29 -EDDA; pista 6, SB29 +EDDA; pista 7, SB30 -EDDA; pista 8, SB30 +EDDA; pista 9, SB32 -EDDA; pista 10, SB33-EDDA; pista 11, SB33 +EDDA; pista 12, PAK -EDDA; pista 13, PAK +EDDA; pista 14, Eagan -EDDA; pista 15, Eagan +EDDA. As bandas específicas de 60-70 kDa eram reactivas com os anti-soros anti-Tbp2 nas pistas 3, 6, 7, 8, 13, 14 e 15, i.e. estirpes SB12, SB29, SB30, PAK e Eagan.

A Figura 28 mostra a construção dos plasmídeos pUHIT1KFH e pUHIT1KFP utilizados para produzir estirpes de *H. influenzae* que não produzem receptor de transferrina.

A Figura 29 mostra a construção de plasmídeos que codificam poliovírus quiméricos que expressam um epítipo derivado da proteína receptora da transferrina que é conservado entre bactérias que produzem proteína receptora da transferrina.

A Figura 30 é um *Western blot* que mostra a reactividade de anti-soros produzidos por imunização de coelhos com

quimeras de poliovírus que expressam um epítipo derivado da proteína receptora da transferrina que é conservado entre bactérias que produzem proteína receptora da transferrina. O painel A mostra um gel corado com Azul brilhante de Coomassie que mostra Tbp2 recombinante purificada da estirpe SB12 de *H. influenzae* expressa em *E. coli* (pista 1), Tbp2 purificada da estirpe 4223 de *Branhamella catarrhalis* (pista 2), um lisado de células inteiras da estirpe 4223 de *B. catarrhalis* limitada por ferro (pista 3), um lisado de células inteiras de *E. coli* JM109 crescida sob condições sem limitação por ferro (pista 5). O painel B mostra os resultados de um *Western blot* de um gel replicado utilizando um conjunto dos soros recolhidos no dia 27 de coelhos imunizados com PV1TBP2A (coelhos 40, 41 e 42). O painel C mostra os resultados para um conjunto de soros pré-sangria dos mesmos, que apresentava reactividade específica mínima.

Em algumas das Figuras anteriores, foram utilizadas as seguintes abreviaturas para designar locais particulares específicos de endonucleases de restrição: R, *Eco* RI; Ps, *Pst* I; H, *Hind* III; Bg, *Bgl* II; Nde, *Nde* I; Ear, *Ear* I; e Sau, *Sau*3A I.

Na Figura 28, utilizaram-se as seguintes abreviaturas para designar locais particulares específicos das endonucleases de restrição: A, *Acc* I; B *Bam* HI; E, *Eco* RI; O, *Xho* I; H, *Hind* III; Ps, *Pst* I; V, *Eco* RV; X, *Xba* I, G, *Bgl* II; S, *Sal* I; K, *Kpn* I; e S*, *Sac* I.

DESCRIÇÃO GERAL DA INVENÇÃO

Qualquer estirpe de *Haemophilus* pode ser convenientemente utilizada para proporcionar o ácido nucleico purificado e isolado que pode estar na forma de moléculas de ADN, compreendendo pelo menos uma porção do ácido nucleico que codifica para um receptor de transferrina como tipificado pelas concretizações da presente invenção. Estas estirpes estão geralmente disponíveis a partir de fontes clínicas e a partir de colecções de culturas bacterianas, tais como a American Type Culture Collection.

De acordo com um aspecto da invenção, a proteína receptora da transferrina pode ser isolada a partir de estirpes de *Haemophilus* pelos métodos descritos por Schryvers (1989), Ogunnawiwo e Schryvers (1992) e patente US 5 141 743. Embora sejam proporcionados os detalhes de um processo apropriado na patente US 5 141 743, segue-se um breve resumo de um tal processo. O isolamento do receptor de transferrina é conseguido através do isolamento de uma fracção membranar a partir de uma estirpe bacteriana que expressa actividade de ligação a transferrina e purificação do receptor de transferrina através de um método de afinidade envolvendo os passos sequenciais de pré-ligação de transferrina ao receptor de transferrina na fracção membranar, solubilização da membrana, imobilização da transferrina e separação do receptor de transferrina da transferrina imobilizada. Alternativamente, as proteínas receptoras podem ser isoladas através de uma modificação do método anterior em que o passo de pré-ligação é evitado e é incluída uma elevada concentração de sal no tampão de solubilização para permitir o isolamento directo com transferrina imobilizada como descrito em Ogunnariwo e Schryvers (1992).

No presente pedido, a expressão "receptor de transferrina" é utilizada para definir uma família de proteínas Tbp1 e/ou Tbp2 que inclui as que possuem variações nas suas sequências de aminoácidos incluindo as que ocorrem naturalmente em várias estirpes de, por exemplo, *Haemophilus*. Outras fontes bacterianas de receptor de transferrina incluem, mas não se lhes limitam, espécies de *Neisseria*, *Branhamella*, *Pasteurella* e *Actinobacillus*. Algumas, senão todas, estas bactérias, contêm tanto Tbp1 como Tbp2. As moléculas de ADN purificadas e isoladas compreendendo pelo menos uma porção que codifica para o receptor de transferrina da presente invenção incluem também as que codificam análogos funcionais de receptor de transferrina. No presente pedido, uma primeira proteína ou péptido é um "análogo funcional" de uma segunda proteína se a primeira proteína estiver imunologicamente relacionada com, e/ou tiver a mesma função que, a segunda proteína ou péptido. O análogo funcional pode ser, por exemplo, um fragmento da proteína ou um seu mutante de substituição, adição ou deleção.

Numa concretização particular, o receptor de transferrina foi isolado de *H. influenzae* tipo b, estirpe DL63, e purificado por métodos de cromatografia de afinidade, como descrito por Schryvers (1989), Ogunnariwo e Schryvers (1992) e na patente US 5 141 743. O receptor de transferrina isolado e purificado foi utilizado para gerar anti-soros anti-TfR em coelhos. O ADN cromossômico de *H. influenzae* tipo b, estirpe DL63, foi cortado mecanicamente, adicionaram-se ligantes *EcoRI*, e construiu-se uma biblioteca de expressão em λ ZAP. A biblioteca foi pesquisada com os anti-soros anti-TfR de coelho e obtiveram-se dois clones positivos (pBHIT1 e pBHIT2) que tinham mapas de restrição que se sobrepunham (Figura 1A e Figura 2). Os clones foram sequenciados e identificaram-se dois grandes quadros de leitura abertos (Figura 2). As sequências de nucleótidos dos genes de receptores de transferrina Tbp1 e Tbp2 (SEQ ID NO: 1) de *H. influenzae* DL63 e suas sequências de aminoácidos deduzidas (SEQ ID NO: 5 - Tbp1 e SEQ ID NO: 6 - Tbp2) estão apresentadas na Figura 3. A análise de sequências mostrou que o operão de TfR consiste em dois genes (Tbp1 e Tbp2) dispostos em tandem e transcritos a partir de um único promotor (como mostrado em particular na Figura 2 e na Figura 3). A proteína Tbp2 tende a variar em peso molecular dependendo da espécie enquanto a proteína Tbp1 tende a ter um peso molecular mais consistente com alguma variabilidade entre as várias bactérias que possuem genes de TfR. O peso molecular de Tbp1 está usualmente na gama de 94 a 106 000 enquanto o peso molecular de Tbp2 varia consideravelmente de 58 a 98 000.

Realizou-se a sequenciação de aminoácidos dos terminais N e fragmentos de brometo de cianogénio de receptor de transferrina de *H. influenzae* DL63. O terminal N de Tbp2 foi bloqueado mas as sequências de aminoácidos foram identificadas por sequenciação de Tbp1 e estão indicadas pelo sublinhado na sequência da proteína da Figura 3. Estas sequências peptídicas são **Glu Thr Gln Ser Ile Lys Asp Thr Lys Glu Ala Ile Ser Ser Glu Val Asp Thr** (como mostrado na Figura 3, SEQ ID NO: 101) e **Leu Gln Leu Asn Leu Glu Lys Lys Ile Gln Gln Asn Trp Leu Thr His Gln Ile Ala Phe** (como mostrado na Figura 3; SEQ ID NO: 102). A sequência de sinal de Tbp1 e a sequência de sinal putativa de Tbp2 estão indicadas por sublinhado duplo na Figura 3. A sequência de sinal putativa para Tbp1 é **Met Thr**

Lys Lys Pro Tyr Phe Arg Leu Ser Ile Ile Ser Cys Leu Leu Ile Ser Cys Tyr Val Lys Ala (SEQ ID NO: 103). A sequência de sinal putativa para Tbp2 é **Met Lys Ser Val Pro Leu Ile Ser Gly Gly Leu Ser Phe Leu Leu Ser Ala** (SEQ ID NO: 104). A sequência de aminoácidos derivada da região N-terminal de Tbp2 indica que é uma lipoproteína.

Preparou-se o ADN cromossômico de *H. influenzae* tipo b, estirpe Eagan, e geraram-se bibliotecas. A primeira biblioteca foi construída a partir de ADN parcialmente digerido com *Sau3A* I, fraccionado por tamanhos para fragmentos de ~5-10 kb, e clonado num plasmídeo à base de pUC. A segunda biblioteca foi construída a partir de fragmentos de ADN cromossômico obtidos por restrição com *Eco* RI, clonados em λ ZAP. Ambas as bibliotecas foram sondadas com um fragmento 5' do clone pBHIT como mostrado na Figura 2 e obtiveram-se os clones parciais dos genes de TfR de *H. influenzae* Eagan denominados S-4368-3-3 e JB-901-5-3. Assim, com referência à Figuras 1B e 2, ilustram-se, de acordo com aspectos adicionais da presente invenção, os clones plasmídicos S-4368-3-3 e JB-901-5-3 que codificam Tbp1 e Tbp2 de *H. influenzae* tipo b, estirpe Eagan. As sequências de ADN dos genes de Tbp1 e Tbp2 (SEQ ID NO: 2) de *H. influenzae* tipo b, estirpe Eagan, e suas sequências de aminoácidos deduzidas (SEQ ID NOS: 7 e 8) estão apresentadas na Figura 4, sendo a sequência de Tbp2 o primeiro gene no operação. Na Figura 4, as sequências putativas -35, -10 e do local de ligação ao ribossoma estão sobrelinhadas.

Preparou-se o ADN cromossômico de *H. influenzae* tipo b, estirpe MinnA, e o ADN foi parcialmente digerido com *Sau3A* I, fraccionado por tamanhos para fragmentos de 10-20 kb, e clonado no local *Bam*HI de EMBL3. A biblioteca foi sondada com o fragmento 5' do clone pBHIT (Figura 2) e obteve-se um clone de comprimento completo que codifica TfR (DS-712-1-3). Com referência às Figuras 1C e 2, ilustra-se, de acordo com aspectos adicionais da presente invenção, o clone plasmídico DS 712-1-3 que codifica Tbp1 e Tbp2 de *H. influenzae* tipo b, estirpe MinnA. As sequências de ADN de Tbp1 e Tbp2 (SEQ ID NO: 3) e suas sequências de aminoácidos deduzidas (SEQ ID NO: 9 - Tbp1 e SEQ ID NO: 10 - Tbp2) de *H. influenzae* tipo b, estirpe MinnA, estão apresentadas na Figura 5 onde a sequência de Tbp2 é a primeira no operação. Na Figura 5, as sequências putativas

-35, -10 e do local de ligação ao ribossoma estão sobrelinhadas.

Preparou-se o ADN cromossômico da *H. influenzae* não tipificável, estirpe PAK 12085. O ADN foi parcialmente digerido com *Sau3A* I, fraccionado por tamanhos para fragmentos de 10-20 kb, e clonado no local *Bam*HI de EMBL3. A biblioteca foi sondada com os fragmentos do clone pBHIT (Figura 2) e obteve-se um clone de comprimento completo que codifica Tfr (JB-1042-7-6). O mapa de restrição do clone JB-1042-7-6 está apresentado nas Figuras 1D e 2 e as sequências de nucleótidos dos genes de *Tbp1* e *Tbp2* (SEQ ID NO: 4) de *H. influenzae* PAK 12085 e suas sequências de aminoácidos deduzidas estão apresentadas na Figura 6 (SEQ ID NOS: 11, 12), com a sequência de *Tbp2* primeiro. Na Figura 6, as sequências putativas -35, -10 e do local de ligação ao ribossoma estão sobrelinhadas.

Preparou-se o ADN cromossômico a partir de *H. influenzae* não tipificável, estirpe SB33, derivada de otite média. O ADN foi parcialmente digerido com *Sau3A* I, fraccionado por tamanhos para fragmentos de 10-20 kb e clonado no local *Bam*HI de EMBL3. A biblioteca foi sondada com os fragmentos do clone pBHIT (Figura 2) e obteve-se um clone de comprimento completo que codifica Tfr (JB-1031-2-9). O mapa de restrição do clone JB-1031-2-9 está apresentado na Figura 2 e as sequências de nucleótidos dos genes de *Tbp1* e *Tbp2* (SEQ ID NO: 4) de *H. influenzae* SB33 e suas sequências de aminoácidos deduzidas estão apresentadas na Figura 7 (SEQ ID NOS: 11, 12), com a sequência de *Tbp2* primeiro. Verificou-se que o gene *tbp2* de SB33 tinha uma deleção de uma única base que resultou num desvio de enquadramento no resíduo 126 e truncagem prematura da proteína resultante no resíduo 168.

Realizou-se a amplificação por PCR dos genes *tbp2* das estirpes de NTHi derivadas de otite média SB12, SB29, SB30 e SB32 e sequenciaram-se os genes.

A sequência de nucleótidos dos genes *tbp2* das estirpes de *H. influenzae* não tipificável SB12 (SEQ ID NO: 105), SB29 (SEQ ID NO: 108), SB30 (SEQ ID NO: 110) e SB32 (SEQ ID NO: 112) estão apresentadas nas Figuras 8, 9, 10 e 11 respectivamente.

Verificou-se que todos os genes *tbp2* amplificados codificavam proteínas Tbp2 de comprimento completo, indicando que o gene *tbp2* defectivo da estirpe SB33 era atípico.

As três estirpes de *H. influenzae b* tinham todas sequências intergênicas curtas idênticas de apenas 13 pb entre *tbp2* e *tbp1*, mas as estirpes de NTHi PAK 12085 e SB33 tinham sequências intergênicas mais longas de 27 pb (Figura 12).

A estirpe SB12 tinha uma sequência intergênica de 13 pb idêntica à encontrada nas estirpes de *H. influenzae b* enquanto as estirpes SB29, SB30 e SB32 continham sequências intergênicas mais longas (27-30 pb) como encontrado nas outras estirpes de NTHi PAK 12085 e SB33 (Figura 2B). Todas as nove estirpes têm uma sequência nuclear em comum conservada de 13 pb entre os seus genes *tbp2* e *tbp1*.

Foi identificada uma sequência pentapeptídica próxima do terminal amino de Tbp1 de *H. influenzae* (Figura 12) que é similar à caixa TonB. O gene *tonB* de *H. influenzae* foi recentemente clonado e sequenciado (Jarosik *et al.*, 1994).

As sequências de aminoácidos de Tbp1 de *H. influenzae* das estirpes Eagan/Minna, DL63, estirpes PAK 12085 e SB33 são comparadas na Figura 14. As proteínas Tbp1 de Eagan e Minna são idênticas e têm 912 aminoácidos de comprimento, a de DL63 tem 914 resíduos, a de PAK 12085 tem 914 resíduos, e a de SB33 tem 911 resíduos. As proteínas Tbp1 de *H. influenzae* são altamente conservadas com 95-100% de identidade de sequências. As sequências de aminoácidos de Tbp2 de *H. influenzae*, estirpes Eagan/Minna, DL63, PAK 12085 SB12, SB29, SB30 e SB32 estão comparadas na Figura 15. As proteínas Tbp2 de Eagan e Minna são idênticas e contêm 660 aminoácidos, a de DL63 tem 644 resíduos, e a de PAK 12085 tem 654 resíduos. Há uma deleção de uma única base no gene *tbp2* de SB33 que resulta num desvio de enquadramento no resíduo 126 e truncagem prematura da proteína resultante no resíduo 168. A base que falta foi confirmada por sequenciação directa de ADN cromossômico amplificado por PCR. Com a excepção de Eagan e Minna, em que são idênticas, as sequências da proteína Tbp2 são menos conservadas com apenas 66-70% de identidade, mas existem vários segmentos curtos de sequência conservada que podem ser

identificados na Figura 15. Verificou-se que os genes *tbp2* amplificados por PCR das estirpes SB12, SB29, SB30 e SB32 codificam todas proteínas Tbp2 de comprimento completo. Havia heterogeneidade de sequência e tamanhos entre as proteínas Tbp2 deduzidas em que SB12 tinha 648 aminoácidos, SB29 tinha 631 resíduos, SB30 tinha 630 resíduos e SB32 tinha 631 resíduos.

Determinaram-se as estruturas secundárias putativas de Tbp1 e Tbp2 de Eagan (Figuras 16A e 16B). Ambas as proteínas possuem vários domínios transmembranares, com Tbp1 a atravessar a membrana 20 vezes e Tbp2 atravessando-a 12 vezes. Identificaram-se três epítomos conservados expostos na região amino-terminal de Tbp1 (DNEVTGLGK - SEQ ID NO: 43, EQVLN/DIRDLTRYD - SEQ ID NOS: 139 e 140, e GAINEIEYENVKAVEISK - SEQ ID NO: 141) e um na região C-terminal (GI/VYNLF/LNYRYVTWE - SEQ ID NOS: 142 e 143). Apenas podem ser identificadas três pequenas regiões conservadas nas proteínas Tbp2 dos patógenos humanos: CS/LGGG(G)SFD - SEQ ID NOS: 75, 144 e 145 no terminal N, LE/SGGFY/FGP - SEQ ID NOS: 74 e 146 localizadas internamente e WFGAR/K - SEQ ID NOS: 83 e 84 no terminal C.

A verificação de que a sequência de aminoácidos de Tbp2 varia entre estirpes de *Haemophilus* permite o agrupamento de *Haemophilus* em subgrupos definidos pela mesma sequência de aminoácidos de Tbp2. Esta verificação permite a selecção racional de um número mínimo de sequências de Tbp1 e/ou Tbp2 ou de péptidos sintéticos representando epítomos partilhados por estes subtipos em estirpes de *Haemophilus* para utilização em composições imunogénicas para, por exemplo, imunização contra as doenças causadas por *Haemophilus* e outras bactérias que produzem receptor de transferrina com semelhanças de sequência relativamente a Tbp1 e Tbp2 de espécies de *Haemophilus*. Assim, pode ser utilizado um número mínimo de receptor de transferrina, análogos, fragmentos e/ou péptidos, para imunizar contra muitas ou todas as estirpes de *Haemophilus* e outros patógenos bacterianos que produzem receptor de transferrina.

Adicionalmente, as sequências de aminoácidos do receptor de transferrina de uma gama de patógenos bacterianos

(*H. influenzae* tipo b, *H. influenzae* não tipificável, *Neisseria meningitidis*, *Neisseria gonorrhoeae* e *Actinobacillus (Haemophilus) pleuropneumoniae*) foram comparadas como mostrado nas Figuras 14 e 15. Esta análise revelou regiões de Tbp1 e Tbp2 que são conservadas entre todas estas bactérias. Algumas destas sequências conservadas estão contidas nos péptidos nas Tabelas 2 e 3. Em particular, as sequências DNEVTGLGK (SEQ ID: 43), EQVLNIRDLTRYPGI (SEQ ID NO: 44), EQVLNIRDLTRYPGISVVEQG RGASSGYSIRGMD (SEQ ID NO: 45), GAINEIEYENVKAVEISK (SEQ ID NO: 46) e GALAGSV (SEQ ID NO: 47) são conservadas em Tbp1 (Tabela 1 e Figura 14). As sequências conservadas particulares em Tbp2 incluem LEGGFYGP (SEQ ID NO: 74), CSGGGSFD (SEQ ID NO: 75), YVYSGL (SEQ ID NO: 76), CCSNLSYVKFG (SEQ ID NO: 77), FLLGHRT (SEQ ID NO: 78), EFNVOF (SEQ ID NO: 79), NAFTGTA (SEQ ID NO: 80), VNGAFYG (SEQ ID NO: 81), ELGGYF (SEQ ID NO: 82), VVFGAR (SEQ ID NO: 83) e VVFGAK (SEQ ID NO: 84) (Tabela 2 e Figura 15).

A identificação de sequências conservadas no interior receptor de transferrina de uma gama de patogénios bacterianos permite a selecção de um número mínimo de antigénios possuindo sequências de aminoácidos particulares (incluindo na forma de péptidos sintéticos) para imunizar contra a doença causada por patogénios que possuem receptores de transferrina. Estas bactérias, em adição às atrás mencionadas, incluem outras espécies de *Neisseria*, tais como *Neisseria gonorrhoeae*, e *Branhamella*, incluindo *Branhamella catarrhalis*. Estas sequências de aminoácidos conservadas entre muitos patogénios bacterianos permitem a produção de anticorpos específicos de TfR, incluindo anticorpos monoclonais, que reconhecem a maioria, senão todos, os receptores de transferrina. Criou-se anti-soro contra péptidos correspondentes a porções conservadas do receptor de transferrina. Estes anti-soro reconheceu o receptor de transferrina em *Branhamella catarrhalis*. Estes anti-soros são úteis para a detecção e neutralização da maioria, senão todas, as bactérias que produzem a proteína TfR e são também úteis para a imunização passiva contra as doenças causadas por esses patogénios. Ensaios de diagnóstico e kits utilizando estas sequências de aminoácidos conservadas são úteis para detectar muitas, senão todas, as bactérias que produzem receptor de transferrina.

Epítomos contendo as sequências de aminoácidos atrás mencionadas podem ser entregues a células do sistema imunitário através da utilização de péptidos sintéticos contendo estas sequências, ou através da utilização de vectores vivos que expressam estas sequências, ou através da administração directa de moléculas de ácido nucleico que codificam a sequência de aminoácidos.

Alguns péptidos contendo sequências de aminoácidos conservadas no interior das proteínas Tbp1 das estirpes Eagan, MinnA, DL63 de *H. influenzae* tipo b e da estirpe não tipificável PAK 12085 estão apresentados na Tabela 2. Criaram-se anticorpos contra alguns destes péptidos em porquinhos-da-índia (Tabela 4). Os péptidos contendo sequências de aminoácidos conservadas no interior das proteínas Tbp2 de *H. influenzae* tipo b, estirpes Eagan, Minn A, DL63 e da estirpe não tipificável PAK 12085 estão apresentados na Tabela 3. Criaram-se anticorpos contra alguns destes péptidos em porquinhos-da-índia (Tabela 4).

As sequências de codificação dos genes de Tbp1 e Tbp2 podem ser clonados em vectores de expressão apropriados para produzir proteínas recombinantes. As Tbp1 e Tbp2 recombinantes foram expressas em *E. coli* utilizando o sistema de expressão de T7. O gene *tbp1* que codifica a proteína Tbp1 de Eagan madura foi clonado em enquadramento por trás do promotor de T7 gerando o plasmídeo JB-1468-29, como mostrado na Figura 17. Quando introduzido em células BL21/DE3 e induzido com IPTG ou lactose, a proteína Tbp1 de Eagan foi expressa como mostrado na Figura 22.

O gene *tbp2* que codifica a proteína Tbp2 madura foi clonado em enquadramento por trás do promotor de T7 gerando o plasmídeo JB-1424-2-8 como mostrado na Figura 18. Quando introduzido em células de *E. coli* e induzido como acima, a proteína Tbp2 foi expressa como mostrado na Figura 22.

O gene *tbp2* da estirpe SB12 de NTHi foi amplificado por PCR. O ADN amplificado resultante contém a sequência de sinal de Tbp2 de *H. influenzae* autêntica antes da proteína madura. O gene *tbp2* de SB12 que codifica a sequência de sinal e a proteína madura foi clonado no sistema de expressão pT7-7 como

mostrado na Figura 21. Quando o plasmídeo resultante (JB-1600-1) foi introduzido em células de *E. coli* BL21/DE3 e induzido, a Tbp2 de SB12 foi expressa, como mostrado na Figura 22.

As proteínas Tbp1 e Tbp2 recombinantes produzidas em *E. coli* na forma de corpos de inclusão foram purificadas através do esquema mostrado na Figura 23. As proteínas purificadas eram pelo menos cerca de 70% puras como mostrado na Figura 24. Realizaram-se estudos de imunogenicidade em ratinhos com as proteínas Tbp1 e Tbp2 recombinantes purificadas. Ambas as proteínas eliciaram uma boa resposta imunitária em ratinhos em doses de 3-10 µg (Figura 25).

Os anti-soros criados contra Tbp1 ou Tbp2 recombinantes derivadas de uma estirpe de *H. influenzae* são reactivos de forma cruzada com outras estirpes, o que os torna reagentes de diagnóstico potencialmente úteis (Figuras 26 e 27).

Os plasmídeos pUHIT1KFH e pUHITKFP mostrados na Figura 28, contêm um marcador seleccionável de resistência a antibiótico clonado no operão do receptor de transferrina e foram construídos para inactivar por inserção o operão do receptor de transferrina. Estes plasmídeos foram utilizados para transformar *Haemophilus* para gerar estirpes que não produzem receptor de transferrina Tbp1 e/ou Tbp2 como descrito no Exemplo 19. Estas estirpes são úteis como controlos negativos (uma vez que não produzem TfR) em concretizações de detecção e diagnóstico *in vitro* e *in vivo*. Espera-se também que estas estirpes sejam atenuadas para crescimento *in vivo* e sejam úteis como vacinas vivas para proporcionar protecção contra doenças causadas por *Haemophilus*.

Como descrito atrás, os epítomos de proteínas receptoras da transferrina podem ser entregues a células do sistema imunitário através da utilização de vectores vivos que expressam estas sequências de aminoácidos e o vector vivo pode ser poliovírus. Com referência à Figura 29, ilustra-se a construção de poliovírus híbridos que expressam um epítomo de proteína receptora da transferrina incluindo o epítomo conservado de Tbp2 LEGGFYGP (SEQ ID NO: 74). Estes vírus eram reconhecidos por anticorpos criados contra um péptido que

incorpora a sequência de aminoácidos LEGGFYGP (SEQ ID NO: 74) (Tabela 5) indicando que os vírus expressaram esta sequência numa forma antigenicamente reconhecível. PV1TBP2A e PV1TBP2B foram também neutralizados por anti-soros de coelho criados contra *tbp2* de *H. influenzae*, estirpe DL63, indicando que pelo menos estes dois vírus expressaram a sequência numa forma reconhecível por anticorpos criados contra a proteína. Todos os vírus eram neutralizáveis por soros anti-PV1, indicando que as alterações no local I antigénico de neutralização de polio não tinham afectado significativamente outros locais antigénicos nos vírus. Adicionalmente, o anti-soro de coelho produzido por imunização com quimera de poliovírus PV1TBP2A ou PV1TBP2B reconheceu um péptido que incorpora a sequência de aminoácidos LEGGFYGP (SEQ ID NO: 74). Isto indica que as sequências expressas por PV1TBP2A e PV1TBP2B são imunogénicas e eliciam anticorpos capazes de reconhecer a mesma sequência no contexto de um péptido sintético.

Com referência à Figura 30, o painel A mostra um gel de SDS-PAGE que mostra *tbp2* recombinante purificado de *H. influenzae*, estirpe SB12, expresso em *E. coli* (pista 1), *tbp2* de *Branhamella catarrhalis*, estirpe 4223 (pista 2), um lisado de células completas de *B. catarrhalis*, estirpe 4223 limitada por ferro (pista 3), um lisado de células completas de *E. coli* JM109 limitada por ferro (pista 4), e um lisado de células completas de *E. coli* JM109 crescidas sob condições sem limitação por ferro (pista 5). O painel B mostra os resultados de um *Western blot* de um gel replicado utilizando um conjunto de soros de coelhos imunizados com PV1TBP2A. Houve uma forte reacção com as proteínas de ligação a transferrina purificadas nas pistas 1 e 2, e com uma banda de dimensão similar na pista 3. Não houve reacção significativa com nenhuma das proteínas de *E. coli* (pistas 4 e 5). O painel C mostra os resultados para um conjunto de soros pré-sangria dos mesmos coelhos, que apresentava um mínimo de reactividade específica. Estes resultados mostram que PV1TBP2A é capaz de induzir anti-soros específicos para proteínas de ligação a transferrina de *H. influenzae* e *B. catarrhalis*, e que os anti-soros podem distinguir *B. catarrhalis* de *E. coli*, que não expressa uma proteína equivalente.

As moléculas de ADN purificadas e isoladas compreendendo pelo menos uma porção que codifica para um receptor de transferrina de uma espécie de *Haemophilus* tipificada pelas concretizações aqui descritas são vantajosas como:

- sondas de ácido nucleico para a identificação específica de estirpes de *Haemophilus in vitro* ou *in vivo*;
- os produtos codificados pelas moléculas de ADN são úteis como reagentes de diagnóstico, antigénios para a produção de anti-soros específicos de *Haemophilus*, para vacinação contra as doenças causadas por espécies de *Haemophilus* e (por exemplo) detecção de infecção por *Haemophilus*;
- péptidos correspondentes a porções do receptor de transferrina como tipificado pelas concretizações aqui descritas são vantajosos como reagentes de diagnóstico, antigénios para a produção de anti-soros específicos de *Haemophilus*, para vacinação contra as doenças causadas por espécies de *Haemophilus* e (por exemplo) para detecção de infecção por *Haemophilus*.

O receptor de transferrina codificado pelas moléculas de ácido nucleico da presente invenção, seus fragmentos e análogos, e péptidos contendo sequências correspondentes a porções do receptor de transferrina que são conservadas entre vários isolados de *Haemophilus* e outras bactérias que produzem receptor de transferrina, são úteis no diagnóstico de, e na imunização contra, doenças causadas por qualquer estirpe bacteriana que produz receptor de transferrina. Em particular, péptidos contendo as sequências LEGGFYGP são conservados nas proteínas receptoras da transferrina de muitos patogénios bacterianos que produzem receptor de transferrina e são apropriados para o diagnóstico de, e imunização contra, doenças causadas por bactérias que produzem receptor de transferrina. Estas bactérias incluem, mas não se lhes limitam, espécies de *Haemophilus*, *Neisseria* (incluindo *N. meningitidis* e *N. gonorrhoeae*) e *Branhamella* (incluindo *B. catarrhalis*).

É claramente evidente para um perito na especialidade, que as várias concretizações da presente invenção têm muitas aplicações nos campos da vacinação, diagnóstico, tratamento de, por exemplo, infecções por *Haemophilus*, e infecções com

outros patogénios bacterianos que produzem receptor de transferrina e para a produção de reagentes imunológicos. Uma outra discussão não limitante destas utilizações é adicionalmente apresentada adiante.

1. Preparação e Utilização de Vacinas

As composições imunogénicas, adequadas para utilização como vacinas, podem ser preparadas a partir de receptor de transferrina imunogénico, seus análogos e fragmentos e/ou péptidos, como aqui descrito. A vacina elicia uma resposta imunitária que produz anticorpos, incluindo anticorpos anti-receptor de transferrina e anticorpos que são opsonizantes ou bactericidas. Se o indivíduo vacinado for provocado por *Haemophilus* ou outras bactérias que produzem um receptor de transferrina, os anticorpos ligam-se ao receptor de transferrina e desse modo impedem o acesso das bactérias a uma fonte de ferro que é necessária para a sua viabilidade. Adicionalmente, os anticorpos anti-TfR opsonizantes ou bactericidas podem também proporcionar protecção por mecanismos alternativos.

As vacinas contendo péptidos são geralmente bem conhecidas na especialidade, como exemplificado pelas Patentes U.S. 4 601 903; 4 599 231; 4 599 230; e 4 596 792. As composições imunogénicas, incluindo vacinas, podem ser preparadas em formas injectáveis, como soluções líquidas ou emulsões. O receptor de transferrina, seus análogos e fragmentos e/ou péptidos podem ser misturados com excipientes farmacologicamente aceitáveis que são compatíveis com o receptor de transferrina, fragmentos análogos ou péptidos. Estes excipientes podem incluir, água, solução salina, dextrose, glicerol, etanol e suas combinações. As composições imunogénicas e vacinas podem ainda conter substâncias auxiliares tais como agentes molhantes ou emulsionantes, agentes tamponantes do pH, ou adjuvantes para aumentar a eficácia das vacinas. As composições imunogénicas e as vacinas podem ser administradas parentericamente, por injeção subcutânea ou intramuscular. Alternativamente, as composições imunogénicas formadas de acordo com a presente invenção, podem ser formuladas e entregues de maneira a invocar uma resposta imunitária em superfícies mucosas. Assim, a composição

imunogénica pode ser administrada a superfícies mucosas, por exemplo, pelas vias nasal ou oral (intragástrica). A composição imunogénica pode ser proporcionada em combinação com uma molécula de direccionamento para entrega a células específicas do sistema imunitário ou a superfícies mucosas. Algumas destas moléculas de direccionamento incluem estirpe B12 e fragmentos de toxinas bacterianas, como descrito em WO 92/17167 (Biotech Australia Pty. Ltd.), e anticorpos monoclonais, como descrito na Patente U.S. 5 194 254 (Barber *et al.*). Alternativamente, podem ser desejáveis outros modos de administração incluindo supositórios e formulações orais. Para supositórios, podem-se incluir aglutinantes e transportadores, por exemplo, polialquilenoglicóis ou triglicéridos. As formulações orais podem incluir incipientes normalmente empregues tais como, por exemplo, qualidades farmacêuticas de sacarina, celulose e carbonato de magnésio. Estas composições tomam a forma de soluções, suspensões, comprimidos, pílulas, cápsulas, formulações de libertação sustentada ou pós e contêm 10-95% do receptor de transferrina, fragmentos análogos e/ou péptidos.

As vacinas são administradas de uma maneira compatível com a formulação de dosagem, e numa quantidade tal que sejam terapeuticamente eficazes, protectoras e imunogénicas. A quantidade a administrar depende do indivíduo a tratar, incluindo, por exemplo, a capacidade do sistema imunitário do indivíduo para sintetizar anticorpos, e se necessário, para produzir uma resposta imunitária mediada por células. As quantidades precisas de ingrediente activo que é necessário administrar dependem da opinião do profissional médico. Contudo, as gamas de dosagem adequadas são prontamente determináveis por um perito na especialidade e podem ser da ordem de microgramas do receptor de transferrina, seus análogos e fragmentos e/ou péptidos. Os regimes adequados para administração inicial e doses de reforço são também variáveis, mas podem incluir uma administração inicial seguida de administrações subsequentes. A dosagem da vacina pode também depender na via de administração e variará consoante o tamanho do hospedeiro.

As molécula de ácido nucleico que codificam o receptor de transferrina da presente invenção podem também ser utilizadas

directamente para imunização por administração do ADN directamente, por exemplo por injeção para imunização genética ou por construção de um vector vivo tal como *Salmonella*, BCG, adenovírus, poxvírus, *vaccinia* ou poliovírus. Uma discussão de alguns vectores vivos que têm sido utilizados para o transporte de antigénios heterólogos para o sistema imunitário está por exemplo em O'Hagan (1992). Processos para a injeção directa de ADN em indivíduos de teste para imunização genética estão descritos em, por exemplo, Ulmer et al., 1993.

A utilização de péptidos *in vivo* pode requerer primeiro a sua modificação química pois os péptidos por si só podem não ter uma semivida suficientemente longa no soro e/ou no tecido e/ou suficiente imunogenicidade. Estes péptidos quimicamente modificados são aqui referidos como "análogos peptídicos". A expressão "análogo peptídico" estende-se a qualquer equivalente químico funcional de um péptido caracterizado pelas suas melhoradas estabilidade e/ou eficácia e imunogenicidade *in vivo* ou *in vitro* em relação à prática da invenção. A expressão "análogo peptídico" é também aqui utilizada estendendo-se a qualquer derivado de aminoácidos dos péptidos aqui descritos. Os análogos peptídicos aqui contemplados são produzidos através de procedimentos que incluem, mas não se lhes limitam, modificações nas cadeias laterais, incorporação de aminoácidos não naturais e/ou seus derivados durante a síntese peptídica e a utilização de reticulantes e outros métodos que impõem constrangimentos conformacionais nos péptidos ou seus análogos.

Os exemplos de modificações nas cadeias laterais contempladas na presente invenção incluem a modificação de grupos amino tal como por alquilação redutora por reacção com um aldeído seguida por redução com NaBH_4 ; amidação com metilacetimidato; acetilação com anidrido acético; carbamilação de grupos amino com cianato; trinitrobenzilação de grupos amino com ácido 2,4,6-trinitrobenzenossulfónico (TNBS); alquilação de grupos amino com anidrido succínico e anidrido tetra-hidroftálico; e piridoxilação de lisina com piridoxal-5'-fosfato seguida de redução com NaBH_4 .

O grupo guanidino de resíduos de arginina pode ser modificado através da formação de produtos de condensação heterocíclicos com reagentes tais como 2,3-butanodiona, fenilglioxal e glioxal.

O grupo carboxilo pode ser modificado por activação de carbodiimida por via da formação de o-acilisourea seguida de subsequente derivatização, por exemplo, numa amida correspondente.

Os grupos sulfidrilo podem ser modificados por métodos tais como carboximetilação com ácido iodoacético ou iodoacetamida; oxidação de ácido perfórmico a ácido cisteico; formação de dissulfuretos mistos com outros compostos de tiol; reacção com maleimida; anidrido maleico ou outra maleimida substituída; formação de derivados de mercúrio utilizando 4-cloromercuriobenzoato, ácido 4-cloromercuriofenilsulfónico, cloreto de fenilmercúrio, 2-cloromercúrico-4-nitrofenol e outros compostos de mercúrio; carbamilação com cianato a pH alcalino.

Os resíduos de triptofano podem ser modificados, por exemplo, por oxidação com N-bromossuccinimida ou alquilação do anel indole com brometo de 2-hidroxi-5-nitrobenzilo ou halogenetos de sulfonilo. Os resíduos de tirosina podem ser alterados por nitração com tetranitrometano para formar um derivado de 3-nitrotirosina.

A modificação do anel imidazole de um resíduo de histidina pode ser realizada por alquilação com derivados do ácido iodoacético ou N-carbetoxilação com dietilpirocarbonato.

Os exemplos de incorporação de aminoácidos não naturais e derivados durante a síntese peptídica incluem, mas não se lhes limitam, a utilização de norleucina, ácido 4-aminobutírico, ácido 4-amino-3-hidroxi-5-fenilpentanóico, ácido 6-amino-hexanóico, t-butilglicina, norvalina, fenilglicina, ornitina, sarcosina, 4-amino-3-hidroxi-6-metil-heptanóico, 2-tienilalanina e/ou isómeros D de aminoácidos.

A imunogenicidade pode ser significativamente melhorada se os antigénios forem co-administrados com adjuvantes,

vulgarmente utilizados na forma de uma solução de 0,05 a 1,0 por cento em solução salina tamponada com fosfato. Os adjuvantes aumentam a imunogenicidade de um antigénio mas não são, eles próprios, necessariamente imunogénicos. Os adjuvantes podem actuar retendo o antigénio localmente próximo do local de administração para produzir um efeito de depósito que facilita uma libertação lenta, sustentada, de antigénio nas células do sistema imunitário. Os adjuvantes podem também atrair células do sistema imunitário para um depósito de antigénio e estimular essas células para eliciar respostas imunitárias.

Os agentes imunoestimulantes ou adjuvantes têm sido utilizados há muitos anos para melhorar as respostas imunitárias do hospedeiro a, por exemplo, vacinas. Os adjuvantes intrínsecos, tais como lipopolissacáridos, normalmente são os componentes das bactérias mortas ou atenuadas utilizadas como vacinas. Os adjuvantes extrínsecos são imunomoduladores que estão tipicamente ligados de modo não covalente a antigénios e são formulados para aumentar as respostas imunitárias do hospedeiro. Assim, têm sido identificados adjuvantes que aumentam a resposta imunitária a antigénios entregues parentericamente. Alguns destes adjuvantes são contudo tóxicos, e podem causar efeitos secundários indesejáveis, o que os torna inadequados para utilização em seres humanos e muitos animais. De facto, apenas hidróxido de alumínio e fosfato de alumínio (colectivamente referidos em comum como alúmen) são rotineiramente utilizados como adjuvantes em vacinas humanas e veterinárias. A eficácia de alúmen para aumentar as respostas de anticorpos aos toxóides da difteria e do tétano está bem estabelecida e, mais recentemente, uma vacina de HBsAg foi adjuvada com alúmen. Embora a utilidade do alúmen esteja bem estabelecida para algumas aplicações, tem limitações. Por exemplo, o alúmen é ineficaz para a vacinação contra a gripe e elicia inconsistentemente uma resposta imunitária mediada por células. Os anticorpos eliciados por antigénios adjuvados por alúmen são principalmente do isótipo IgG1 no ratinho, que pode não ser óptimo para a protecção por alguns agentes de vacina.

Uma ampla gama de adjuvantes extrínsecos pode provocar respostas imunitárias potentes contra antigénios. Estes

incluem saponinas complexadas com antígenos de proteína de membrana (complexos estimulantes imunitários), polímeros Pluronic com óleo mineral, micobactérias mortas e óleo mineral, adjuvante completo de Freund, produtos bacterianos, tais como dipéptido de muramilo (MDP) e lipopolissacárido (LPS), assim como lípido A e lipossomas.

Para induzir eficientemente respostas imunitárias humorais (HIR) e imunidade mediada por células (CMI), os imunogénios são emulsionados em adjuvantes. Muitos adjuvantes são tóxicos, induzem granulomas, inflamações agudas e crónicas (adjuvante completo de Freund, FCA), citólise (saponinas e polímeros Pluronic) e pirogenicidade, artrite e uveíte anterior (LPS e MDP). Embora o FCA seja um excelente adjuvante e amplamente utilizado em investigação, não está licenciado para utilização em vacinas humanas ou veterinárias devido à sua toxicidade.

As características desejáveis de adjuvantes ideais incluem:

- (1) ausência de toxicidade;
- (2) capacidade para estimular uma resposta imunitária duradoura;
- (3) simplicidade de fabrico e estabilidade na armazenagem a longo prazo;
- (4) capacidade para elicitar tanto CMI como HIR contra antígenos administrados por várias vias, se necessário;
- (5) sinergia com outros adjuvantes;
- (6) capacidade de interagir selectivamente com populações de células apresentadoras de antígenos (APC);
- (7) capacidade para elicitar especificamente respostas imunitárias apropriadas específicas de células T_H1 ou T_H2 ; e
- (8) capacidade para aumentar selectivamente níveis de isótipos de anticorpos apropriados (por exemplo, IgA) contra antígenos.

A Patente US 4 855 283 concedida a Lockhoff *et al.* em 8 de Agosto, 1989, ensina análogos glicolipídicos incluindo N-glicosilamidas, N-glicosilureias e N-glicosilcarbamatos,

cada um substituído no resíduo de açúcar por um aminoácido, como imunomoduladores ou adjuvantes. Assim, Lockhoff *et al.* 1991, relataram que análogos N-glicolipídicos apresentando similaridades estruturais com os glicolípidos de ocorrência natural, tais como glicosfingolípidos e glicoglicerolípidos, são capazes de eliciar fortes respostas imunitárias tanto na vacina do vírus herpes *simplex* como na vacina do vírus da pseudo-raiva. Alguns glicolípidos foram sintetizados a partir de alquilaminas de cadeia longa e ácidos gordos que são ligados directamente aos açúcares através do átomo de carbono anomérico, para imitar as funções dos resíduos lipídicos de ocorrência natural.

A Patente U.S. 4 258 029 concedia a Moloney, transmitida ao presente cessionário, ensina que cloridrato de octadeciltirosina (OTH) funciona como um adjuvante quando complexado com o toxóide do tétano e vacina de vírus da poliomielite do tipo I, II e III inactivado com formalina. Também, Nixon-George *et al.* 1990, relataram que ésteres de octadecilo de aminoácidos aromáticos complexados com um antigénio de superfície da hepatite B recombinante, aumentaram as respostas imunitárias do hospedeiro contra o vírus da hepatite B.

A lipidação de péptidos sintéticos foi também utilizada para aumentar a sua imunogenicidade. Assim, Wiesmuller 1989, descreve um péptido com uma sequência homóloga a uma proteína viral da febre aftosa acoplado a um adjuvante de tripalmitil-s-gliceril-cisteinilserilserina, sendo um análogo sintético da parte N-terminal da lipoproteína de bactérias Gram-negativas. Adicionalmente, Deres *et al.* 1989, relataram a sensibilização *in vivo* de linfócitos T citotóxicos específicos do vírus com vacina de lipopéptido sintético que era constituída por péptidos sintéticos modificados derivados da nucleoproteína de vírus *influenza* por ligação a um lipopéptido, N-palmitil-s-[2,3-bis(palmitiloxi)-(2RS)-propil-[R]-cisteína (TPC).

2. Imunoensaios

O receptor de transferrina, seus análogos e fragmentos e/ou péptidos da presente invenção são úteis como imunogénios, como antigénios em imunoensaios incluindo ensaios

imunossorventes com enzimas ligadas (ELISA), RIA e outros ensaios de ligação de anticorpos sem enzima ligada ou procedimentos conhecidos na especialidade para a detecção de anticorpos antibacterianos, de *Haemophilus*, TfR e/ou péptidos. Em ensaios ELISA, o receptor de transferrina, análogos, fragmentos e/ou péptidos correspondentes a porções de proteína TfR, são imobilizados sobre uma superfície seleccionada, por exemplo uma superfície capaz de ligar proteínas ou péptidos tal como os poços de uma placa de microtítulo de poliestireno. Após lavagem para remover receptor de transferrina, análogos, fragmentos e/ou péptidos, incompletamente adsorvidos, pode ser ligada à superfície seleccionada uma proteína não específica, tal como uma solução de albumina sérica bovina (BSA) ou caseína que se sabe serem antigenicamente neutras em relação à amostra de teste. Isto permite o bloqueio de locais de adsorção não específica sobre a superfície de imobilização e reduz assim o ruído de fundo causado por ligações não específicas de anti-soros sobre a superfície. Preferivelmente, os péptidos seleccionados são das regiões conservadas da Tabela 2 ou Tabela 3 para aumentar a detecção cruzada entre espécies a menos que se pretenda detectar uma espécie bacteriana particular. Nesse caso, é seleccionado um polipéptido que é único para o TfR dessa espécie particular. Normalmente, os péptidos estão na gama de 12 resíduos até, e preferivelmente, 14 a 30 resíduos. Entenda-se contudo, que pode ser utilizada uma mistura de péptidos quer como imunogénio numa vacina quer como agente de diagnóstico. Pode haver circunstâncias em que se utiliza uma mistura de péptidos das regiões conservadas e/ou das regiões não conservadas para proporcionar protecção cruzada entre espécies e/ou diagnóstico. Neste caso, a mistura de péptidos imunogénicos é vulgarmente referida como uma preparação de "cocktail" para utilização como vacina ou agente de diagnóstico.

A superfície de imobilização é então colocada em contacto com uma amostra tal como materiais clínicos ou biológicos a testar de uma maneira condutora à formação de complexos imunitários (antigénio/anticorpo). Isto pode incluir a diluição da amostra com diluentes tais como BSA, gama-globulina bovina (BGG) e/ou solução salina tamponada com fosfato (PBS)/Tween. A amostra é então deixada a incubar durante de 2 a 4 horas, a temperaturas tais como da ordem de

25° a 37°C. Após incubação, a superfície em contacto com a amostra é lavada para remover material não imunocomplexado. O procedimento de lavagem pode incluir a lavagem com uma solução tal como PBS/Tween ou um tampão de borato.

Após a formação de imunocomplexos específicos entre a amostra de teste e o receptor de transferrina, análogos, fragmentos e/ou péptidos, ligados, e subsequente lavagem, a ocorrência, e mesmo a quantidade, de formação de imunocomplexo podem ser determinadas submetendo o imunocomplexo a um segundo anticorpo possuindo especificidade para o primeiro anticorpo. Se a amostra de teste é de origem humana, o segundo anticorpo é um anticorpo possuindo especificidade para imunoglobulinas humanas e IgG em geral. Para proporcionar meios de detecção, o segundo anticorpo pode ter uma actividade associada tal como uma actividade enzimática que irá gerar, por exemplo, um desenvolvimento de cor por incubação com um substrato cromogénico apropriado. A quantificação pode então ser conseguida medindo o grau de produção de cor utilizando, por exemplo, um espectrofotómetro no espectro do visível.

3. Utilização de Sequências como Sondas de Hibridação

As sequências de nucleótidos da presente invenção, compreendendo a sequência do gene do receptor de transferrina, permitem agora a identificação e clonagem dos genes de receptor de transferrina a partir de qualquer espécie de *Haemophilus* e outras bactérias que possuam genes de receptor de transferrina.

As sequências de nucleótidos compreendendo a sequência dos genes de receptor de transferrina da presente invenção são úteis pela sua capacidade de formar selectivamente moléculas dúplices com trechos complementares de outros genes de TfR. Dependendo da aplicação, podem ser empregues uma variedade de condições de hibridação para conseguir vários graus de selectividade da sonda em relação aos outros genes de TfR. Para um elevado grau de selectividade, utilizam-se condições relativamente rigorosas para formar as dúplices, tais como condições de baixo teor de sais e/ou temperatura elevada, tal como proporcionadas por 0,02 M a 0,15 M de NaCl a temperaturas entre cerca de 50°C e 70°C. Para algumas aplicações, são

necessárias condições de hibridação menos rigorosas tais como 0,15 M a 0,9 M de sal, a temperaturas variando de entre cerca de 20°C e 55°C. As condições de hibridação podem também ser tornadas mais rigorosas pela adição de quantidades crescentes de formamida, para desestabilizar a dúplice híbrida. Assim, as condições de hibridação particulares podem ser prontamente manipuladas, e serão geralmente um método a escolher dependendo dos resultados desejados. Em geral, as temperaturas de hibridação convenientes na presença de formamida a 50% são: 42°C para uma sonda que é 95 a 100% homóloga ao fragmento alvo, 37°C para 90 a 95% de homologia e 32°C para 85 a 90% de homologia.

Numa concretização de diagnóstico clínico, as sequências de ácido nucleico dos genes de TfR da presente invenção podem ser utilizadas em combinação com um meio apropriado, tal como um marcador, para determinar a hibridação. São conhecidos na especialidade uma ampla variedade de meios indicadores apropriados, incluindo ligandos radioactivos, enzimáticos ou outros, tais como avidina/biotina, que são capazes de proporcionar um sinal detectável. Em algumas concretizações de diagnóstico, pode ser utilizado um marcador enzimático tal como urease, fosfatase alcalina ou peroxidase, em vez de um marcador radioactivo. No caso de marcadores enzimáticos, são conhecidos substratos indicadores colorimétricos que podem ser empregues para proporcionar um meio visível ao olho humano ou espectrofotometricamente, para identificar a hibridação específica com amostras contendo sequências do gene de TfR.

As sequências de ácido nucleico dos genes de TfR da presente invenção são úteis como sondas de hibridação em hibridações em solução e em concretizações que empregam procedimentos em fase sólida. Em concretizações que envolvem procedimentos em fase sólida, o ADN (ou ARN) de teste das amostras, tais como amostras clínicas, incluindo exudados, fluidos corporais (e.g., soro, fluido amniótico, efusão do ouvido médio, esputo, fluido de lavagem broncoalveolar) ou mesmo tecidos, é adsorvido ou de outro modo fixado numa matriz ou superfície seleccionadas. O ácido nucleico de cadeia simples fixado é então submetido a hibridação específica com sondas seleccionadas compreendendo as sequências de ácido nucleico dos genes de TfR ou seus fragmentos da presente

invenção, sob as condições desejadas. As condições seleccionadas dependerão das circunstâncias particulares baseadas nos critérios particulares necessários dependendo de, por exemplo, os teores de G+C, o tipo de ácido nucleico alvo, a fonte de ácido nucleico, o tamanho da sonda de hibridação, etc. Após lavagem da superfície de hibridação de modo a remover moléculas de sonda ligadas de modo não específico, a hibridação específica é detectada, ou mesmo quantificada, por meio do marcador. Tal como com a selecção de péptidos, prefere-se seleccionar porções da sequência de ácido nucleico que são conservadas entre espécies de *Haemophilus*, tais como sequências de ácido nucleico que codificam a sequência peptídica conservada das Figuras 8, 9, 13 e 14 e particularmente enumeradas nas Tabelas 2 e 3. A sonda seleccionada pode ter pelo menos 18 pb e pode estar na gama de 30 pb a 90 pb de comprimento.

4. Expressão dos Genes do Receptor de Transferrina

Vectores plasmídicos contendo sequências de replicação e de controlo que são derivadas de espécies compatíveis com a célula hospedeira, podem ser utilizados para a expressão dos genes de receptor de transferrina em sistemas de expressão. O vector normalmente é portador de um local de replicação, assim como de sequências marcadoras que são capazes de proporcionar selecção fenotípica em células transformadas. Por exemplo, *E. coli* pode ser transformada utilizando pBR322 que contém genes para resistência a ampicilina e tetraciclina e assim proporciona um meio fácil para a identificação de células transformadas. O plasmídeo pBR322, ou outro plasmídeo microbiano ou fago, têm também que conter, ou ser modificados para conter, promotores que podem ser utilizados pela célula hospedeira para a expressão das suas próprias proteínas.

Em adição, vectores fágicos contendo sequências de replicação e de controlo que são compatíveis com o hospedeiro podem ser utilizadas como vector de transformação em relação a estes hospedeiros. Por exemplo, o fago em lambda GEM[™]-11 pode ser utilizado para preparar vectores fágicos recombinantes que podem ser utilizados para transformar células hospedeiras, tais como a *E. coli* LE392.

Os promotores normalmente utilizados na construção de ADN recombinante incluem os sistemas promotores da β -lactamase (penicilinase) e da lactose (Chang et al., 1978; Itakura et al., 1977; Goeddel et al., 1979; Goeddel et al., 1980) e outros promotores microbianos tais como o sistema promotor de T7 (Patente U.S. 4 952 496). Os detalhes relativos às sequências de nucleótidos de promotores são conhecidos, permitindo a um perito ligá-los funcionalmente a genes. O promotor particular utilizado será geralmente uma matéria a escolher dependendo dos resultados desejados. Os hospedeiros que são apropriados para a expressão dos genes de receptor de transferrina, seus fragmentos, análogos ou variantes, incluem *E. coli*, espécies de *Bacillus*, *Haemophilus*, fungos, levedura ou pode-se utilizar o sistema de expressão de baculovírus.

De acordo com a presente invenção, prefere-se preparar a proteína por métodos recombinantes, particularmente quando a proteína TfR de ocorrência natural tal como purificada a partir de uma cultura de uma espécie de *Haemophilus* pode incluir quantidades vestigiárias de materiais tóxicos ou outros contaminantes. Este problema pode ser evitado utilizando proteína TfR produzida recombinantemente em sistemas heterólogos que podem ser isolados do hospedeiro de uma maneira que minimize os contaminantes no material purificado. Os hospedeiros particularmente desejáveis para a expressão, a este respeito, incluem bactérias Gram-positivas que não possuem LPS e são portanto isentas de endotoxinas. Estes hospedeiros incluem espécies de *Bacillus* e podem ser particularmente úteis para a produção de receptor de transferrina, seus fragmentos ou análogos, não pirogênicos. Adicionalmente, os métodos de produção recombinantes permitem o fabrico de Tbp1 ou Tbp2 ou seus fragmentos, separados uns dos outros, o que é distinto das proteínas combinadas normais presentes em *Haemophilus*.

Depósitos Biológicos

Certos plasmídeos que contêm pelo menos uma porção que codifica para um receptor de transferrina de estirpes de *Haemophilus influenzae* que são aqui descritos e referidos, foram depositados na American Type Culture Collection (ATCC) localizada em Rockville, Maryland EUA, segundo os termos do

Tratado de Budapeste e antes da apresentação deste pedido. As amostras dos plasmídeos depositados ficarão disponíveis ao público após a concessão de uma patente baseada neste pedido de patente dos Estados Unidos. A invenção que se descreve e reivindica aqui não se pretende limitada em âmbito pelos plasmídeos depositados, pois a concretização depositada destina-se apenas a ilustração da invenção.

Sumário do Depósito

Clone	Designação ATCC	Data do Depósito
DS-712-1-3	75603	4 de Novembro, 1993
JB-1042-7-6	75607	4 de Novembro, 1993
JB-1424-2-8	75937	27 de Outubro, 1994
JB-1600-1	75935	27 de Outubro, 1994
JB-1468-29	75936	27 de Outubro, 1994
pT7TBP2A	75931	27 de Outubro, 1994
pT7TBP2B	75932	27 de Outubro, 1994
pT7TBP2C	75933	27 de Outubro, 1994
pT7TBP2D	75934	27 de Outubro, 1994

Estirpes de *Haemophilus*

A estirpe Eagan de Hib está disponível em Connaught Laboratories Limited, 1755 Steeles Ave. W., Willowdale, Ontario, Canadá M2R 3T4.

A estirpe MinnA de Hib foi obtida da colecção do Dr. Robert Munson, Department of Microbiology and Immunology, Washington University School of Medicine, Children's Hospital, St. Louis, Missouri 63110.

A estirpe DL63 de Hib foi obtida da colecção do Dr. Eric Hansen, Department of Microbiology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235-9048.

A PAK 12085 foi obtida da colecção do Dr. Robert Munson (*supra*).

As SB12, 29, 30, 32 e 33 foram obtidas da colecção do Dr. Stephen Barenkamp, Department of Pediatrics, School of Medicine, Saint Louis University Medical Centre, St. Louis, Missouri 63104.

EXEMPLOS

A descrição anterior descreve genericamente a presente invenção. Pode-se ter um entendimento mais completo por referência aos seguintes Exemplos específicos. Estes Exemplos são descritos somente para fins de ilustração e não se destinam a limitar o âmbito da invenção. Alterações na forma e substituição de equivalentes estão contempladas conforme as circunstâncias o possam sugerir ou tornar expedito. Embora se tenham empregue aqui termos específicos, esses termos pretendem-se num sentido descritivo e não para fins de limitação.

Os métodos de genética molecular, bioquímica das proteínas, imunologia e tecnologia das fermentações, utilizados mas não explicitamente descritos nesta descrição e nestes Exemplos, estão amplamente relatados na literatura científica e estão bem dentro das capacidades dos peritos na especialidade.

Exemplo 1

Este Exemplo ilustra a preparação de ADN cromossómico de *H. influenzae*, estirpes DL63, Eagan, MinnA e PAK 12085, e SB33.

As estirpes de *H. influenzae* foram crescidas em ágar Mueller-Hinton ou em caldo de infusão de cérebro e coração como descrito por Harkness *et al.* 1992.

A. Extracção de ADN cromossómico de *Haemophilus influenzae* tipo b DL63

O ADN cromossómico foi preparado como se segue. Duzentos e cinquenta ml de cultura foram sedimentados por centrifugação a 8 000 rpm num rotor Beckman J14 durante 15 minutos. A pelete foi lavada com 200 ml de Tris 50mM-HCl, pH 8,0, centrifugada

como anteriormente, ressuspensa em 12,5 ml de Tris 50mM-HCl, EDTA 50mM, pH 8,0 e congelada a -20°C. Depois, adicionaram-se 1,25 ml de uma solução de lisozima a 10 mg/ml em Tris 0,25M-HCl, pH 8,0, à pelete de células congelada. A pelete foi descongelada e incubada em gelo durante 45 minutos. Em seguida, adicionaram-se 2,5 ml de uma solução de proteinase K a 1mg/ml em SDS a 0,5%, EDTA 0,4M, Tris 50mM-HCl, pH 7,5 e incubou-se a mistura a 50°C durante 1 hora com mistura ocasional. O lisado foi extractado uma vez com 15 ml de fenol tamponado com Tris, depois adicionaram-se 1,5 ml de acetato de sódio 3M e 30 ml de etanol para precipitar o ADN. O ADN foi enrolado numa vareta de vidro, depois foi dissolvido em 12,5 ml de Tris 50mM-HCl, EDTA 1mM, pH 7,5 contendo ARNase A a 0,2 mg/ml por rotação durante a noite. A amostra foi extractada uma vez com um volume igual de clorofórmio, precipitada e enrolada como anteriormente. O ADN foi dissolvido em 2 ml de Tris 50mM-HCl, EDTA 1mM, pH 7,5 e armazenado a 4°C.

B. Extracção de ADN cromossómico de *Haemophilus influenzae* tipo b Eagan

Cinquenta ml de cultura foram sedimentados por centrifugação, a pelete foi ressuspensa em 25ml de TE (Tris 10mM, EDTA 1mM, pH 7,5), e utilizaram-se 2 alíquotas de 5ml para a preparação de ADN cromossómico. A cada alíquota adicionaram-se 0,6ml de sarcosilo a 10% e 0,15ml de proteinase K a 20mg/ml e incubaram-se as amostras a 37°C durante 1 hora. Extractou-se o lisado uma vez com fenol saturado com Tris e três vezes com clorofórmio:álcool isoamílico (24:1). As fases aquosas foram reunidas até um volume final de 7ml. Depois, adicionaram-se 0,7ml de acetato de sódio 3M (pH 5,2) e 4,3 ml de isopropanol para precipitar o ADN que foi enrolado, enxaguado com etanol a 70%, seco, e ressuspensa em 1 ml de água.

C. Extracção de ADN cromossómico de *Haemophilus influenzae* Eagan, MinnA, PAK 12085 e SB33

As células foram sedimentadas a partir de 50 ml de cultura por centrifugação a 5000 rpm durante 15-20 minutos, a 4°C. Ressuspendeu-se a pelete de células em 10 ml de TE (Tris

10mM-HCl, EDTA 1mM, pH 7,5), adicionaram-se pronase e SDS até concentrações finais de 500 µg/ml e 1%, respectivamente. A amostra foi incubada a 37°C durante 4 horas até se obter um lisado límpido. O lisado foi extractado uma vez com fenol saturado com Tris, uma vez com fenol saturado com Tris/clorofórmio (1:1) e uma vez com clorofórmio. A fase aquosa final foi dialisada durante 24 horas contra 2x 500 ml de NaCl 1M a 4°C, mudando o tampão uma vez, e durante 24 horas contra 2x 500 ml de TE a 4°C, mudando o tampão uma vez. O dialisado final foi dividido em alíquotas para utilização.

Exemplo 2

Este Exemplo ilustra a preparação de bibliotecas cromossómicas.

A. Biblioteca de *H. influenzae* DL63-λZAP

Cortaram-se mecanicamente 100 µg de ADN cromossómico de *H. influenzae* DL63 em TE, numa seringa de 1 ml com uma agulha de calibre 25. O ADN cortado foi tornado de extremidades lisas por adição de água até um volume final de 405 µl, 45 µl de tampão de nuclease S1 10x (NaCl 2M, NaOAc 500mM, pH 4,5, ZnSO₄ 10mM, glicerol a 5%), e 1,7 µl de nuclease S1 a 100 U/µl e incubação a 37°C durante 15 min. A amostra foi extractada uma vez com fenol/clorofórmio e uma vez com clorofórmio e adicionou-se 1 ml de etanol para precipitar o ADN. A amostra foi incubada em gelo durante 10 min ou a -20°C durante a noite e o ADN foi recolhido por centrifugação numa microcentrífuga durante 30 minutos. O ADN foi lavado com etanol a 70% e seco. Os locais *Eco* RI na sequência de ADN foram metilados utilizando procedimentos *standard*. A este ADN metilado adicionaram-se 5 µl de MgCl₂ 100mM, 8 µl de mistura de dNTP (2,5 mM de cada de dATP, dCTP, dGTP e dTTP), e 4 µl de Klenow a 5 U/µl. A mistura foi incubada a 12°C durante 30 minutos. Adicionaram-se 450 µl de STE (NaCl 0,1M, Tris 10mM-HCl, EDTA 1mM, pH 8,0), e extractou-se a mistura uma vez com fenol/clorofórmio, e uma vez com clorofórmio, antes da adição de 1 ml de etanol para precipitar o ADN. A amostra foi incubada em gelo durante 10 min ou a -20°C durante a noite. O ADN foi recolhido por centrifugação numa microcentrífuga durante 30 minutos, lavado com etanol a 70% e seco.

O ADN foi ressuspenso em 7 µl de TE e adicionaram-se à solução 14 µl de ligantes *Eco* RI fosforilados (200 ng/µl), 3 µl de tampão de ligação 10x, 3 µl de ATP 10mM e 3µl de ADN-ligase de T4 (4 U/µl). Incubou-se a amostra a 4°C durante a noite, depois incubou-se a 68°C durante 10 minutos para inactivar a ligase. Adicionaram-se à mistura 218 µl de H₂O, 45 µl de tampão Universal 10x e 7 µl de *Eco* RI a 30 U/µl. Após incubação a 37°C durante 1,5 horas, adicionaram-se 1,5 µl de EDTA 0,5M, e colocou-se a mistura em gelo.

O ADN foi fraccionado por tamanhos num gradiente de sacarose, e reuniram-se as fracções contendo ADN de 6-10 kb. O ADN reunido foi precipitado com etanol e ressuspenso em 5 µl de tampão TE. Ligaram-se 200ng de ADN de inserção durante 2-3 dias a 4°C com 1 µg de vector ZAP II num volume final de 5µl. Empacotou-se a mistura de ligação utilizando Gigapack II Gold (Stratagene) e plaqueou-se em células *E. coli* SURE em placas de NZY. Titulou-se a biblioteca, amplificou-se e armazenou-se a 4°C sob clorofórmio a 0,3%.

B. biblioteca de *H. influenzae* Eagan-pUC

ADN cromossómico preparado a partir de *H. influenzae* Eagan pelo método do Exemplo 1C foi digerido com *Sau*3A I durante 2, 5 e 10 minutos e submeteram-se amostras a electroforese num gel de agarose preparativo. Excisaram-se fatias de gel que incluíam fragmentos de ADN entre 3-10 kb de comprimento e extractou-se o ADN pelo procedimento de congelação-descongelação *standard*. ADN plasmídico de pUC 8:2 (pUC 8 com os locais para as enzimas de restrição *Bgl* II e *Xba* I adicionais no local de clonagem múltipla) foi digerido com *Bam*H I e *Bgl* II, e desfosforilado com fosfatase alcalina de vitelo (CAP). Os fragmentos de ADN de *H. influenzae* Eagan foram ligados a pUC e a mistura foi utilizada para transformar células de *E. coli* JM109.

C. Biblioteca de *H. influenzae* Eagan-λZAP

ADN cromossómico de *H. influenzae* Eagan preparado como no Exemplo 1B foi digerido com *Eco* RI e fraccionado por tamanhos num gel de agarose preparativo. Excisaram-se fatias

de gel correspondentes a fragmentos de ADN de 7-23 kb e submeteu-se o ADN a electroeluição durante a noite em tubagem de diálise contendo 3 ml de TAE (Tris 40mM-acetato, EDTA 1mM) a 14V. O ADN foi precipitado duas vezes e ressuspensão em água antes de ser ligado durante a noite a ADN de λ ZAP II digerido com *Eco* RI. A mistura de ligação foi empacotada utilizando o *kit* de empacotamento Gigapack II (Stratagene) e plaqueada em células de *E. coli* XL1-Blue. A biblioteca foi titulada, amplificada e armazenada a 4°C sob clorofórmio a 0,3%.

D. Bibliotecas de EMBL3

ADN cromossómico de *H. influenzae* MinnA (10 μ g) foi preparado como no Exemplo 1C e digerido com *Sau*3A I (40 unidades) durante 2, 4 e 6 minutos, depois foi fraccionado por tamanhos num gradiente de sacarose a 10-30% em tampão TNE (Tris 20mM-HCl, NaCl 5mM, EDTA 1mM, pH 8). Reuniram-se as fracções contendo fragmentos de ADN maiores que 5 kb e precipitaram-se. Numa segunda experiência, o ADN cromossómico (2,6 μ g) foi digerido com *Sau*3A I (4 unidades) durante 1, 2 e 3 minutos e fraccionado por tamanhos por electroforese em gel de agarose preparativo. Excisaram-se fatias de gel contendo fragmentos de ADN de 10-20 kb e extractou-se o ADN através de uma técnica *standard* de congelação/descongelação. O ADN fraccionado por tamanhos das duas experiências foi reunido para ligação com braços *Bam*H I de EMBL3 (Promega). A mistura de ligação foi empacotada utilizando o *kit* de empacotamento Gigapack II e plaqueada em células de *E. coli* LE392. A biblioteca foi titulada, depois foi amplificada e armazenada a 4°C sob clorofórmio a 0,3%.

ADN cromossómico de *H. influenzae* PAK 12085 ou SB33 preparado como no Exemplo 1C foi digerido com *Sau*3A I (0,5 unidades/10 μ g de ADN) a 37°C durante 15 minutos e fraccionado por tamanhos por electroforese em gel de agarose. Excisaram-se fatias de gel correspondentes a fragmentos de ADN de 15-23 kb e o ADN foi electroeluído durante a noite em tubagem de diálise contendo 3 ml de TAE a 14V. O ADN foi precipitado duas vezes e ressuspensão em água antes da ligação, durante a noite, a braços *Bam*H I de EMBL3 (Promega). A mistura de ligação foi empacotada utilizando o *kit* de empacotamento Lambda *in vitro* (Amersham) de acordo com as instruções do

fabricante e plaqueada em células *E. coli* NM539. A biblioteca foi titulada, depois amplificada e armazenada a 4°C na presença de clorofórmio a 0,3%.

Exemplo 3

Este Exemplo ilustra a pesquisa das bibliotecas.

A. Biblioteca de expressão de *H. influenzae* DL63-λZAP

As proteínas Tbp1 e Tbp2 foram purificadas por afinidade em transferrina humana em fase sólida (hTf). Resumidamente, preparou-se uma coluna de hTf-Sepharose de 20 ml de acordo com o protocolo do fabricante para acoplamento de ligandos de proteína a Sepharose activada por CNBr (Sigma). Lavou-se a matriz resultante com 3 volumes de coluna de Tris 50mM-HCl, NaCl 1M, guanidina 6M-HCl, pH 8,0 para remover hTf ligado não covalentemente. A coluna foi então equilibrada com Tris 50mM-HCl, pH 8,0 e o hTf ligado foi carregado com ferro utilizando 1 ml de FeCl₃ 10mg/ml em tampão contendo 100mM de cada um de citrato de sódio e bicarbonato de sódio, pH 8,6, seguidos de 2 volumes de coluna de Tris 50mM-HCl, NaCl 1M, pH 8,0. Prepararam-se membranas bacterianas totais (300 mg de proteína total) de *H. influenzae*, estirpe DL63, crescida em meios deficientes em ferro como descrito previamente (Schryvers et al., 1989). Diluíram-se as membranas para 2 mg/ml em Tris 50mM-HCl, NaCl 1M, pH 8,0 e solubilizaram-se pela adição de EDTA até 15mM e Sarcosilo NL97 até 1,5%. Após centrifugação a 40 000 x g durante 1 hora, aplicou-se o sobrenadante na coluna de hTf e lavou-se a coluna com 10 volumes de coluna de Tris 50mM-HCl, NaCl 1M, EDTA 10mM, Sarcosilo a 0,5%, pH 8,0. As proteínas receptoras foram eluídas utilizando GnHCl 2M no mesmo tampão e as fracções eluídas foram dialisadas extensamente contra tampão de bicarbonato de amónio 25mM (5 trocas de tampão), liofilizadas e armazenadas -20°C. Utilizaram-se proteínas isoladas para gerar anti-soros específicos de receptor de transferrina em coelhos New Zealand White utilizando técnicas *standard*. Resumidamente, imunizaram-se os coelhos 3 vezes subcutaneamente, com intervalos de duas semanas, utilizando adjuvante completo de Freund para a primeira injeção e adjuvante incompleto de Freund para as injeções subsequentes.

A biblioteca de DL63- λ ZAP foi plaqueada em células de *E. coli* SURE e transferiram-se as placas fágicas para membranas de nitrocelulose que tinham sido pré-embebidas em IPTG 10mM para induzir a expressão a partir do promotor lacZ de pBluescript. Bloquearam-se os filtros utilizando leite desnatado a 0,5% em Tris 50mM-HCl, NaCl 150mM, pH 7,5, antes de serem sondados com os anti-soros anti-TfR policlonais e IgG anti-coelho de cabra conjugado com peroxidase de rábano. Purificaram-se as placas fágicas através de 3 ciclos de pesquisa e recuperaram-se os plasmídeos pBluescript recombinantes (pBHIT1 e pBHIT2; Figuras 1A e 2) pelo procedimento de excisão *in vivo* (Short et al., 1988).

B. Bibliotecas que não de expressão de Eagan, MinnA e PAK 12085

(i) Pesquisa da biblioteca de *H. influenzae* Eagan-pUC

Realizaram-se levantamentos de colónias para nitrocelulose utilizando técnicas *standard* e sondaram-se os filtros com a sonda 5'pBHIT2 do gene do receptor de transferrina ilustrada na Figura 2. A sonda foi marcada com digoxigenina (dig, Boehringer Mannheim) seguindo as especificações do fabricante. Submeteram-se vários clones putativos a *dot blot* sobre nitrocelulose e submeteram-se a um segundo ciclo de pesquisa utilizando a mesma sonda 5'pBHIT2. Analisaram-se os putativos do segundo ciclo por mapeamento com enzimas de restrição e seleccionou-se o clone S-4368-3-3 (Figura 1B, Figura 2) para análise da sequência.

(ii) Pesquisa da biblioteca de *H. influenzae* Eagan- λ ZAP

Plaqueou-se a biblioteca de fagos utilizando técnicas *standard* em células XLI Blue (Stratagene) utilizando placas de LB e uma sobrecamada de agarose a 0,7%. Levantaram-se as placas fágicas para nitrocelulose utilizando protocolos *standard* e cozeram-se os filtros a 80°C, durante 2 horas, sob vácuo, para fixar o ADN. A sonda 5'pBHIT2 do gene do receptor de transferrina (Figura 2) foi marcada com digoxigenina e os filtros foram pré-hibridados durante 4 horas a 42°C, e depois hibridados com a sonda marcada a 42°C, durante a noite. Lavaram-se os filtros a 68°C e após autorradiografia, seleccionaram-se várias placas fágicas para um segundo ciclo

de pesquisa. A excisão *in vivo* de ADN fagemídico dos putativos do segundo ciclo foi realizada de acordo com protocolos proporcionados com o sistema λ ZAP (Promega). Obtiveram-se quatro clones com inserções *Eco* RI de ~2,5 kb idênticas dos quais o JB-901-5-3 da Figura B, Figura 2 é um exemplo. As placas fágicas putativas foram também amplificadas e o ADN fágico foi purificado a partir de 500 ml de cultura. Excisou-se o ADN inserido por digestão com *Xba* I e clonou-se em pUC 8:2 (pUC 8 contendo locais *Bgl* II e *Xba* I adicionais no seu local de clonagem múltipla) que tinha sido digerido com *Xba* I e desfosforilado. O clone JB-911-3-2 (Figura 17) contém a metade 3' do operão de TfR de *H. influenzae* Eagan.

(iii) Pesquisa de bibliotecas de EMBL 3

A biblioteca de *H. influenzae* MinnA foi plaqueada sobre células LE392 em placas de NZCYM utilizando agarose a 0,7% no topo em NZCYM na forma de sobrecamada. Realizaram-se levantamentos de placas fágicas para filtros de nitrocelulose seguindo os procedimentos *standard* e processaram-se os filtros e sondaram-se com a sonda 5.'pBHIT2 (Figura 2) marcada com digoxigenina. Plaquearam-se as placas fágicas putativas e submeteram-se a um segundo e um terceiro ciclos de pesquisa utilizando os mesmos procedimentos. Preparou-se ADN fágico a partir de 500 ml de cultura utilizando técnicas *standard*, excisou-se o ADN inserido por digestão com *Sal* I e clonou-se em pUC para gerar o clone DS-712-1-3 (Figuras 1C e 2).

Plaqueou-se a biblioteca de *H. influenzae* PAK 12085 em células LE392 sobre placas de NZCYM utilizando agarose a 0,7% em NZCYM na forma de sobrecamada. Levantaram-se placas fágicas para nitrocelulose e processaram-se os filtros e sondaram-se com a sonda 5'pBHIT2 marcada com digoxigenina (Figura 2). Plaquearam-se as placas fágicas putativas e submeteram-se a um segundo ciclo de pesquisa utilizando o mesmo procedimentos. Preparou-se o ADN fágico a partir de culturas de 500 ml por técnicas *standard*, excisou-se o ADN inserido por digestão com *Sal* I, e clonou-se em pUC para gerar o clone JB-1042-7-6 (Figura 1D e 2).

A biblioteca de *H. influenzae* SB33 foi plaqueada em células LE392 sobre placas de NZCYM utilizando agarose a 0,7% em NZCYM na forma de sobrecamada. Levantaram-se placas fágicas

para nitrocelulose e processaram-se os filtros e sondaram-se com a sonda 5'pBHIT2 marcada com digoxigenina (Figura 2). Plaquearam-se as placas fágicas putativas e submeteram-se a um segundo ciclo de pesquisa utilizando os mesmos procedimentos. Preparou-se o ADN fágico a partir de culturas de 500 ml por técnicas *standard*, excisou-se a inserção de ADN por digestão com *Sal* I, e clonou-se em pUC para gerar o clone JB-1031-2-9 (Figura 2).

Exemplo 4

Este Exemplo ilustra a sequenciação dos genes de *Tbp1* e *Tbp2* do operão de *TfR*.

ADN plasmídico dos clones pBHIT 1, pBHIT 2, S-4368-3-3, JB-901-5-3, DS-712-1-3, JB-1042-7-6 e JB-1031-2-9 foi preparado utilizando técnicas *standard*. Sintetizaram-se iniciadores de sequenciação oligonucleotídicos de 17-25 bases de comprimento no sintetizador de ADN ABI modelo 380B e purificaram-se por cromatografia utilizando cartuchos OPC obtidos de Applied Biosystems Inc., e utilizaram-se de acordo com as recomendações dos fabricantes. Sequenciaram-se as amostras utilizando o sequenciador de ADN ABI modelo 370A e a química do terminador corante de acordo com os protocolos do fabricante. A sequência do operão de *TfR* da estirpe DL63 está ilustrada na Figura 3, a da estirpe Eagan na Figura 4, a da estirpe MinnA na Figura 5, a de PAK 12085 na Figura 6 e a de SB33 na Figura 7.

Exemplo 5

Este Exemplo ilustra a amplificação por PCR dos genes *tbp2* de *H. influenzae* não tipificável estirpes SB12, SB29, SB30 e SB32.

O ADN cromossômico de *H. influenzae* não tipificável, estirpes SB12, SB29, SB30 e SB32 foi preparado como descrito atrás. Os genes de *TfR* estão dispostos na forma de um operão com *tbp2* seguido de *tbp1* (vejam-se as Figuras 12A e 12B). Sintetizaram-se oligonucleótidos para a extremidade 5' do *tbp2* e o complemento inverso da extremidade 5' das sequências de codificação de *tbp1*. Os iniciadores eram: GGATCCATATGAAATCTGT

ACCTCTTATCTCTGGT (SEQ ID NO: 120) correspondente a MKSVPLISGS (SEQ ID NO: 147) da sequência de comando de Tbp2 e TCTAGAAGCTTTTTTAGTCATTTTATAGTATTCCAT (SEQ ID NO: 137) que é o complemento inverso da sequência de comando MTKK (SEQ ID NO: 138) de Tbp1 e uma parte da sequência intergénica (Figuras 12A e 12B). A amplificação por PCR foi realizada em tampão contendo Tris 10mM/HCl, pH 8,3, cloreto de potássio 50 mM e cloreto de magnésio 1,5 mM. Cada mistura reaccional de 100 µl continha 5 ng de ADN cromossómico, 1 µg de cada iniciador, 5 unidades de ADN-polimerase amplitaq (Perkin Elmer Cetus) e 0,4 mM dos dNTP (Perkin Elmer Cetus). As condições dos ciclos foram de 25 ciclos de 94°C durante 1,0 min, 45°C durante 2,0 min e 72°C durante 1,5 min. Amplificaram-se fragmentos específicos de 2 kb para cada amostra (Figura 13). Utilizou-se ADN de SB33 como controlo positivo (Pista 1). O ADN cromossómico utilizado para a amplificação do gene de Tbp2 era, na pista 1, SB33; na pista 2, SB12; na pista 3, SB29; na pista 4, SB30; e na pista 5, SB32. Os fragmentos foram clonados no vector de clonagem TA (Invitrogen) e determinaram-se as suas sequências de nucleótidos. As sequências de ácido nucleico de Tbp2 das estirpes SB12 (SEQ ID NO: 108), SB29 (SEQ ID NO: 110), SB30 (SEQ ID NO: 112) e SB32 (SEQ ID NO: 114) estão apresentadas nas Figuras 8, 9 10 e 11 respectivamente.

Exemplo 6

Este Exemplo ilustra a comparação das sequências de aminoácidos de transferrina, a identificação de epítomos potencialmente expostos de proteínas receptoras da transferrina por análise da estrutura secundária.

Com referência à Figura 14, mostra-se uma comparação da sequência de aminoácidos de Tbp1 de *H. influenzae* tipo b Eagan, DL63, *H. influenzae* não tipificável, estirpes PAK 12085 e SB33, *N. meningitidis*, estirpes B16B6 e M982 (Legrain et al., 1993) e *N. gonorrhoeae* FA19 (Cornelissen et al., 1992). Esta análise revelou regiões de Tbp1 que são conservadas entre todas estas bactérias.

Com referência à Figura 15, mostra-se uma comparação da sequência de aminoácidos de Tbp2 de *H. influenzae* tipo b,

estirpes Eagan, DL63, *H. influenzae* não tipificável PAK 12085, SB12, SB29, SB30 e SB32, *N. meningitidis*, estirpes B16B6 e M982, *N. gonorrhoeae* FA19 e *Actinobacillus (Haemophilus) pleuropneumoniae* (Gerlach et al., 1992) 205 e 37. Esta análise revelou regiões de Tbp2 que são conservadas entre todas estas bactérias.

As análises da estrutura secundária das proteínas foram realizadas utilizando os algoritmos de Chou e Fasman (1978) e fizeram-se as representações gráficas de hidrofiliicidade/hidrofobicidade utilizando o algoritmo de Hopp (1986). Os valores eram derivados das médias de janelas heptapeptídicas e estão representados no gráfico no ponto médio de cada fragmento. A Figura 16A ilustra a estrutura secundária prevista de Tbp1 de *H. influenzae* tipo b Eagan e a Figura 16B ilustra a estrutura secundária prevista de Tbp2 de *H. influenzae* tipo b Eagan. As estruturas secundárias previstas representadas nas Figuras 16A e 16B foram obtidas utilizando os procedimentos atrás descritos. Contudo, os inventores ainda não foram capazes de verificar que a estrutura secundária é correctamente representada por essas Figuras.

Os epítomos conservados das proteínas Tbp1 e Tbp2 de várias bactérias diferentes foram identificados por alinhamento de sequências como mostrado nas Figuras 14 e 15, respectivamente. Alguns destes epítomos conservados incluem:

TBP1	DNEVTGLGK	SEQ ID NO:43
TBP1	EQVLNIRLTRYDPGI	SEQ ID NO:44
TBP1	GAINEIEYENVKAVEISK	SEQ ID NO:45
TBP1	GALAGSV	SEQ ID NO:46
TBP2	LEGGFYGP	SEQ ID NO:74
TBP2	CSGGGSFD	SEQ ID NO:75
TBP2	YVYSGL	SEQ ID NO:76
TBP2	CCSNLSYVKFG	SEQ ID NO:77
TBP2	FLLGHRT	SEQ ID NO:78
TBP2	EFNVDF	SEQ ID NO:79
TBP2	NAFTGTA	SEQ ID NO:80
TBP2	VNGAFYG	SEQ ID NO:81
TBP2	LEGGYF	SEQ ID NO:82
TBP2	VVFGAR	SEQ ID NO:83

Adicionalmente, em combinação com as estruturas secundárias previstas, identificaram-se quatro epítomos conservados expostos em Tbp1 e identificaram-se dois em Tbp2. Estes são:

Tbp1	DNEVTGLGK	SEQ ID NO:43
Tbp1	EQVLN/DIRDLTRYD	SEQ ID NOS: 139 e 140
Tbp1	GAINEIEYENVKAVEISK	SEQ ID NO:141
Tbp1	GI/VYNLF/LNYRYVTWE	SEQ ID NOS:142 e 143
Tbp2	CS/LGGG(G)SFD	SEQ ID NOS: 75, 144 e 145
Tbp2	LE/SGGFY/FGP	SEQ ID NOS: 74 e 146

As proteínas, polipéptidos ou péptidos contendo as sequências de aminoácidos conservadas atrás mencionadas são particularmente úteis como meio de detecção em concretizações de diagnóstico e como imunogénios para detectar ou proteger de doenças causadas por bactérias que produzem a proteína receptora da transferrina. Para imunização, as sequências de aminoácidos particularmente indicadas podem ser apresentadas ao sistema imunitário na forma de proteínas ou péptidos ou pode-se utilizar um veículo de entrega vivo, tal como *Salmonella*, BCG, adenovírus, poxvírus, vaccinia ou poliovírus.

Exemplo 7

Este Exemplo ilustra a construção do plasmídeo JB-1468-29 que expressa Tbp1 de Eagan a partir de *E.coli*.

Os plasmídeos S-4368-3-3 (Figuras 1B e 2) e JB-911-3-2 (Figura 17) contêm as partes a 5' e 3' do gene *tbp1* de Eagan, respectivamente. A Figura 17 ilustra o esquema de construção para o plasmídeo JB-1468-29. As sequências de oligonucleótidos utilizadas na construção de JB-1468-29 estão mostradas na Figura 20, (SEQ ID NOS: 86 e 87). O plasmídeo JB-1468-29 foi introduzido em *E. coli*, estirpe BL21/DE3, por electroporação, para gerar a estirpe JB-1476-2-1.

Cresceu-se JB-1476-2-1 em meio YT e induziu-se com IPTG seguindo protocolos *standard*. Para a preparação de Tbp1 para estudos de imunogenicidade e outros estudos, cresceu-se a estirpe JB-1476-2-1 durante a noite em meio NZCYM contendo glucose a 3%. Adicionou-se um inóculo de 1:40 a meio NZCYM

fresco sem glucose, e cresceu-se a cultura até $A_{578}=0,3$. Adicionou-se lactose até 1% e induziu-se a cultura durante 4 horas. A análise SDS-PAGE de lisados de células inteiras de JB-1476-2-1 está mostrada na Figura 22. Na pista 1, JB-1476-2-1 (T7/Tbp1 de Eagan) a t_0 ; na pista 2, JB-1476-2-1 a $t=4h$ de indução; na pista 3, marcadores de peso molecular de 200 kDa, 116 kDa, 97,4 kDa, 66 kDa, 45 kDa e 31 kDa; na pista 4, JB-1437-4-1 (T7/Tbp2 de Eagan) a t_0 ; na pista 5, JB-1437-4-1 a $t=4h$ de indução; na pista 6, JB-1607-1-1 (T7/Tbp2 de SB12) a t_0 ; na pista 7, JB-1607-1-1 a $t=4h$ de indução.

Exemplo 8

Este Exemplo ilustra a construção do plasmídeo JB-1424-2-8 que expressa Tbp2 de Eagan a partir de *E. coli*.

Com referência à Figura 18, mostra-se o plasmídeo S-4368-3-3 que contém a totalidade do gene *tbp2* de *H. influenzae* tipo b Eagan. A Figura 18 ilustra o plasmídeo JB-1424-2-8 e a Figura 19 mostra os oligonucleótidos utilizados. O plasmídeo JB-1424-2-8 foi introduzido em *E. coli* estirpe BL21/DE3 por electroporação para gerar *E. coli*, estirpe JB-1437-4-1. Por indução com IPTG ou lactose, a Tbp2 foi expressa pela *E. coli* JB-1437-4-1 como mostrado na Figura 22. Na pista 1, JB-1476-2-1 (T7/Tbp1 de Eagan) a t_0 ; na pista 2, JB-1476-2-1 a $t=4h$ de indução; na pista 3, marcadores de peso molecular de 200 kDa, 116 kDa, 97,4 kDa, 66 kDa, 45 kDa e 31 kDa; na pista 4, JB-1437-4-1 (T7/Tbp2 de Eagan) a t_0 ; na pista 5, JB-1437-4-1 a $t=4h$ de indução; na pista 6, JB-1607-1-1 (T7/Tbp2 de SB12) a t_0 ; na pista 7, JB-1607-1-1 a $t=4h$ de indução.

Exemplo 9

Este Exemplo ilustra a construção de plasmídeos que codificam uma sequência de comando de lipoproteína antes da sequência de Tbp2.

Os oligonucleótidos utilizados para a construção de plasmídeos com sequências de comando de lipoproteínas derivadas de *E. coli* lpp (SEQ ID NOS: 88 e 89), rlpB (SEQ ID

NOS: 90 e 91), e pal (SEQ ID NOS: 92 e 93) antes de Tbp2 estão mostrados na Figura 20. Os plasmídeos construídos e as estirpes geradas correspondentes estão ilustrados na Tabela 1 adiante.

Exemplo 10

Este Exemplo ilustra a construção do plasmídeo JB-1600-1 que expressa Tbp2 de SB12 a partir de *E. coli*.

O plasmídeo DS-1047-1-2 (Figura 21) contém o gene *tbp2* de SB12 amplificado por PCR. O gene *tbp2* foi excisado na forma de um fragmento de restrição *Nde* I a *EcoR* I e inserido no vector de expressão pT7-7 para gerar o plasmídeo JB-1600-1. A electroporação em células BL21/DE3 originou a *E. coli*, da estirpe JB-1607-1-1 que expressa Tbp2 de SB12. Por indução com IPTG ou lactose, a Tbp2 de SB12 foi expressa, como mostrado na Figura 22. Na pista 1, JB-1476-2-1 (T7/Tbp1 de Eagan) a t_0 ; na pista 2, JB-1476-2-1 a $t=4h$ de indução; na pista 3, marcadores de peso molecular de 200 kDa, 116 kDa, 97,4 kDa, 66 kDa, 45 kDa e 31 kDa; na pista 4, JB-1437-4-1 (T7/Tbp2 de Eagan) a t_0 ; na pista 5, JB-1437-4-1 a $t=4h$ de indução; na pista 6, JB-1607-1-1 (T7/Tbp2 de SB12) a t_0 ; na pista 7, JB-1607-1-1 a $t=4h$ de indução.

Exemplo 11

Este Exemplo ilustra a extracção e a purificação de Tbp1 e Tbp2.

O esquema de purificação para Tbp1 e Tbp2 é mostrado na Figura 23. Ambas as proteínas recombinantes são expressas na forma de corpos de inclusão em *E.coli* e os esquemas de purificação são idênticos. Células de uma cultura de 500 ml, preparada como descrito no Exemplo 7 para Tbp1 e no Exemplo 8 para Tbp2, foram ressuspensas em 50 ml de Tris 50mM-HCl, pH 8,0, e rompidas com ultra-sons (3x 10 min, círculo de trabalho de 70%). O extracto foi centrifugado a 20 000 x g durante 30 min e o sobrenadante resultante que continha >95% das proteínas solúveis de *E. coli* foi rejeitado.

A pelete remanescente (Figura 23, PPT₁) foi adicionalmente extractada em 50 ml de Tris 50 mM, pH 8,0 contendo Triton X-100 a 0,5% e EDTA 10 mM. Após centrifugação a 20 000 x g durante 30 min, o sobrenadante contendo proteínas solúveis residuais e a maioria das proteínas de membrana, foi rejeitado. A pelete resultante (Figura 23, PPT₂) obtida após a extracção anterior, continha os corpos de inclusão. As proteínas Tbp1 e Tbp2 foram solubilizadas em Tris 50 mM, pH 8,0, contendo SDS a 0,1% e DTT 5 mM. Após centrifugação, o sobrenadante resultante foi adicionalmente purificado numa coluna de filtração de gel Superdex 200 equilibrada em Tris 50 mM, pH 8,0, contendo SDS a 0,1% e DTT 5mM. As fracções foram analisadas por SDS-PAGE e as que continham Tbp1 ou Tbp2 purificadas foram dialisadas durante a noite a 4°C contra Tris 50 mM, pH 8,0, e depois centrifugadas a 20 000 x g durante 30 min. A proteína permaneceu solúvel nestas condições e a Tbp1 e a Tbp2 purificadas foram armazenadas a -20°C.

A análise SDS-PAGE do processo de purificação é mostrada na Figura 24. Nas pistas 1, marcadores de peso molecular de proteínas pré-corados (106, 80, 49,5, 32,5, 27,5, 18,5 kDa); nas pistas 2, lisados de células inteiras de *E.coli*; nas pistas 3, corpos de inclusão solubilizados; nas pistas 4, Tbp1 ou Tbp2 purificadas.

Exemplo 12

Este Exemplo ilustra os estudos de imunogenicidade de Tbp1 e Tbp2 recombinantes em ratinhos.

Grupos de cinco ratinhos Balb/c foram injectados subcutaneamente (s.c.) nos dias 1, 29 e 43 com rTbp1 ou rTbp2 purificadas (1 µg a 10 µg), preparadas como descrito no Exemplo 11, na presença de AlPO₄ (1,5 mg por dose). Colheram-se amostras de sangue nos dias 14, 28, 42 e 56 para análise dos títulos de anticorpos anti-rTbp1 e anti-rTbp2 por EIA. Os resultados dos estudos de imunogenicidade são ilustrados na Figura 25.

Exemplo 13

Este Exemplo ilustra o desenvolvimento dos EIA para determinação de anticorpos anti-rTbpl e anti-rTbp2 em soros de ratinhos.

Os títulos de anticorpos anti-rTbpl e anti-rTbp2 foram determinados essencialmente como descrito por Panezutti et al. (1993). Revestiram-se os poços de microtítulo com 0,5 µg de rTbpl ou rTbp2, preparadas como descrito no Exemplo 11, durante 16 h à temperatura ambiente, depois bloquearam-se com 0,1% (p/v) de BSA em PBS. Os soros foram diluídos em série, adicionados aos poços, depois incubados durante uma hora à temperatura ambiente. Utilizaram-se fragmentos F(ab')₂ purificados por afinidade de anticorpo IgG de cabra anti-ratinho (específico para Fc) conjugados com peroxidase de rábano como segundo anticorpo. As reacções foram desenvolvidas utilizando tetrametilbenzidina (TMB/H₂O₂) e a absorvância foi medida a 450 nm (utilizando 540 nm como comprimento de onda de referência) num leitor de microplacas Flow Multiskan MCC. O título reactivo de um anti-soro foi definido como o recíproco da diluição que apresenta consistentemente um aumento de duas vezes na absorvância relativamente ao que é obtido com a amostra de soro pré-imune.

Exemplo 14

Este Exemplo ilustra a reactividade cruzada de anti-soros anti-Tbpl, produzidos por imunização com Tbpl de Eagan recombinante, com várias estirpes de *H. influenzae*.

Lisados de células inteiras de estirpes de *H. influenzae* crescidas em meios BHI suplementados com NAD e heme (Harkness et al., 1992) ± EDDA foram separados por gel de SDS-PAGE, transferidos para membrana de nitrocelulose e sondados com anti-soros anti-Tbpl de porquinho-da-índia criados contra Tbpl de Eagan recombinante purificada (Figura 26). Na pista 1, BL21/DE3; na pista 2, SB12-EDDA; na pista 3, SB12 +EDDA; na pista 4, SB29 - EDDA; na pista 5, SB29 +EDDA; na pista 6, SB33 -EDDA; na pista 7, SB33 + EDDA; na pista 8, Eagan -EDDA; na pista 9, Eagan +EDDA; na pista 10, *B. catarrhalis* 4223 - EDDA; na pista 11, *B. catarrhalis* 4223 +EDDA; na pista 12,

N. meningitidis 608 - EDDA; na pista 13, *N. meningitidis* 608 + EDDA; na pista 14, JB-1476-2-1 induzida expressando Tbp1 de Eagan recombinante; na pista 5, marcadores de peso molecular. As bandas específicas de ~95 kDa reagiram com os anti-soros anti-Tbp1 nas pistas 3, 4, 5, 7, 8 e 9, correspondentes a *H. influenzae*, estirpes SB12, SB29, SB33 e Eagan; as bandas de ~110 kDa nas pistas 10 e 11, correspondentes a *B. catarrhalis*, estirpe 4223; e as bandas de ~80 kDa nas pistas 12 e 13, correspondentes a *N. meningitidis* 608.

Exemplo 15

Este Exemplo ilustra a reatividade cruzada de anti-soros anti-Tbp2, produzidos por imunização com Tbp2 de Eagan recombinante, com várias estirpes de *H. influenzae*.

Lisados de células inteiras de estirpes de *H. influenzae* crescidas em meios BHI suplementados com NAD e heme (Harkness *et al.*, 1992) ± EDDA foram separados num gel de SDS-PAGE, transferidos para membrana de nitrocelulose e sondados com anti-soros anti-Tbp2 de porquinho-da-índia criados contra Tbp2 de Eagan recombinante purificada (Figura 27). Na pista 1, marcadores de peso molecular; na pista 2, JB-1437-4-1 induzida expressando Tbp2 de Eagan recombinante; na pista 3, SB12-EDDA; na pista 4, SB12 +EDDA; na pista 5, SB29 -EDDA; na pista 6, SB29 +EDDA; na pista 7, SB30 -EDDA; na pista 8, SB30 +EDDA; na pista 9, SB32 -EDDA; na pista 10, SB33-EDDA; na pista 11, SB33 +EDDA; na pista 12, PAK -EDDA; na pista 13, PAK +EDDA; na pista 14, Eagan -EDDA; na pista 15, Eagan +EDDA. As bandas específicas de cerca de 60-70 kDa eram reativas com os anti-soros anti-Tbp2 nas pistas 3, 6, 7, 8, 13, 14 e 15, correspondentes às estirpes SB12, SB29, SB30, PAK e Eagan de *Haemophilus*.

Exemplo 16

Este Exemplo ilustra a síntese de péptidos sintéticos correspondentes a regiões conservadas em Tbp2 e Tbp1.

As sequências de aminoácidos deduzidas de Tbp1 e Tbp2 foram comparadas como mostrado na Figuras 14 e 15, respectivamente. Esta comparação identificou regiões de

conservação de sequência de aminoácidos no interior do receptor de transferrina descrito atrás e, como mostrado nas Tabelas 2 e 3, foram sintetizados péptidos contendo uma porção do receptor de transferrina. Esta síntese pode ser efectuada por expressão num hospedeiro adequado de vectores recombinantes contendo ácido nucleico que codifica os referidos péptidos ou por síntese peptídica *standard*.

Resumidamente, os péptidos foram sintetizados utilizando um sintetizador de péptidos ABI 430A e química do t-Boc optimizada utilizando as condições recomendadas pelo fabricante, e os péptidos foram clivados da resina utilizando ácido fluorídrico (HF). Os péptidos foram purificados por cromatografia líquida de alta resolução de fase inversa (RP-HPLC) numa coluna Vydac C4 semi-preparativa (1 x 30 cm) utilizando um gradiente de acetonitrilo de 15 a 55% em ácido trifluoroacético a 0,1% (TFA), desenvolvida durante 40 minutos a um caudal de 2 ml/minuto. Todos os péptidos sintéticos utilizados em estudos bioquímicos e imunológicos eram >95% puros como avaliado por HPLC analítica. Realizaram-se análises da composição em aminoácidos num sistema Waters Pico-Tag e eram concordantes com as composições teóricas.

Exemplo 17

Este Exemplo ilustra a imunogenicidade de péptidos sintéticos em animais de teste.

Porquinhos-da-índia foram imunizados intramuscularmente com 100 µg de péptido, preparado como descrito no Exemplo 16, emulsionado em adjuvante completo de Freund no dia 0 seguido de reforços nos dias + 14 e + 28 utilizando a mesma quantidade de péptido emulsionado em adjuvante incompleto de Freund. Obtiveram-se amostras de soros no dia 42 + e determinaram-se os títulos de anticorpos por ensaio imunossorvente com enzima ligada (ELISA). Resumidamente, revestiram-se poços de microtítulo (Nunc-Immunoplate, Nunc, Dinamarca) com 500 ng de qualquer péptido particular em 50 µL de tampão de revestimento (Na₂CO₃ 15 mM, NaHCO₃ 35 mM, pH 9,6) durante 16 horas à temperatura ambiente. As placas foram então bloqueadas com 0,1% (p/v) de BSA em tampão salino de fosfato (PBS) durante 30 minutos à temperatura ambiente. Os anti-soros foram diluídos

em série, adicionados aos poços e incubados durante 1 hora à temperatura ambiente. Após remoção dos anti-soros, lavaram-se as placas cinco vezes com PBS contendo Tween-20 a 0,1% (p/v) e BSA a 0,1% (p/v). F(ab')₂ de anticorpos IgG de cabra anti-porquinho-da-índia conjugados com peroxidase de rábano (Jackson ImmunoResearch Labs Inc., PA) foram diluídos (1/8 000) com tampão de lavagem, e adicionados a placas de microtítulo. Após 1 hora de incubação à temperatura ambiente, lavaram-se as placas cinco vezes com o tampão de lavagem. Desenvolveram-se as placas utilizando o substrato tetrametilbenzidina (TMB) em H₂O₂ (ADI, Toronto), parou-se a reacção com H₂SO₄ 1N e mediu-se a densidade óptica a 450 nm utilizando um Titretrek Multiskan II (Flow Labs., Virginia). Incluíram-se dois péptidos irrelevantes de 32 resíduos de aminoácido como controlos negativos nestes ELISA. Os ensaios foram realizados em triplicado, e o título reactivo de cada anti-soro foi definido como a diluição que apresentava consistentemente um aumento de 2 vezes no valor da absorvância relativamente aos obtidos com os controlos negativos. Os anti-soros criados em porquinhos-da-índia eram monoespecíficos para o péptido utilizado para imunização. Os títulos dos soros obtidos após a imunização estão apresentados na Tabela 4.

Os péptidos da presente invenção compreendem cópias únicas de quaisquer dos apresentados nas Tabelas 2 e 3 ou péptidos contendo múltiplas cópias de seus análogos. Um péptido pode ainda compreender múltiplos péptidos diferentes seleccionados entre os mostrados nas Tabelas 2 e 3 ou seus análogos e incluem moléculas transportadoras adequadas. Prefere-se que sejam utilizados péptidos de regiões conservadas para desenvolver os anticorpos porque um imunoensaio ou outro tipo de ensaio de ligação pode então ser utilizado para detectar várias espécies de *Haemophilus*. As Tabelas 2 e 3 apresentam portanto várias outras regiões conservadas do receptor de transferrina para identificar outros péptidos que seriam úteis em diagnóstico, imunização e tratamento médico.

Exemplo 18

Este Exemplo descreve a capacidade de anti-soro criado contra péptidos correspondentes a porções conservadas do

receptor de transferrina, para reconhecer o receptor de transferrina de *Branhamella catarrhalis*.

Imunizaram-se porquinhos-da-índia com péptido, correspondente a porções conservadas do receptor de transferrina, e os anti-soros obtidos estão descritos no Exemplo 17. Submeteu-se um extracto de células inteiras de *Branhamella catarrhalis* a immunoblot com o anti-soro específico do péptido que reconheceu especificamente o receptor de transferrina desta bactéria. O anti-soro anti-péptido de um porquinho-da-índia imunizado com o péptido N-terminal de Tbp2 e o péptido TBP2-25 reconheceu especificamente a proteína Tbp2 de *Branhamella catarrhalis* e a Tbp2 recombinante expressa pelo clone plasmídico pBHIT2 em *E. coli*. O clone pBHIT2 expressa uma versão truncada de Tbp2 que começa no aminoácido 80. (i.e. NKKFYSG SEQ ID NO:105). Portanto, a proteína Tbp2 de pBHIT2 apenas pode ser reconhecida por anticorpos criados contra o segundo epítipo LEGGFYGP (TBP2-25). Esta análise mostra que péptidos correspondentes a sequências conservadas entre receptores de transferrina são úteis na detecção da maioria, senão todas, as bactérias que produzem receptor de transferrina, e como componentes em composições imunogénicas, incluindo vacinas para produzir uma resposta imunitária contra o receptor de transferrina e proteger contra doenças causadas por essas bactérias.

Os soros destes coelhos foram testados por ELISA contra um péptido que incorpora a sequência LEGGFYGP (SEQ ID NO:74) ou contra Tbp2 de *H. influenzae*, estirpe DL63. Placas de ELISA foram revestidas com o péptido ou a proteína e depois bloqueadas com leite desnatado a 5%. Incubaram-se diluições em série de duas vezes de soros em solução salina tamponada com fosfato, tween-20 a 0,05%, e leite seco a 1%, nas placas, durante duas horas a 37°C, após o que as placas foram lavadas cinco vezes com solução salina tamponada com fosfato com 0,05% de tween-20. Sondaram-se as placas lavadas com um IgG de burro anti-coelho conjugado com peroxidase de rábano (HRPO) durante 30 minutos à temperatura ambiente, e depois lavaram-se cinco vezes com solução salina tamponada com fosfato com 0,05% de tween-20. O substrato da HRPO foi adicionado a todos os poços durante 30 minutos à temperatura ambiente no escuro, depois

parou-se o desenvolvimento de cor pela adição de 50 μ l de ácido sulfúrico 1M. A cor foi medida por determinação da absorvância a 450nm.

Exemplo 19

Este Exemplo ilustra a produção de estirpes de *H. influenzae* que não produzem receptor de transferrina.

Um fragmento *Eco* RI de 2,55 da inserção de pBHIT1 foi subclonado no local *Eco* RI de pUC4K, resultando a remoção da cassette de resistência a canamicina (kan) Tn903 deste vector (pUHIT1; Figura 28). Este passo de subclonagem facilitou a subsequente inserção de um fragmento *Hinc*II ou *Pst*I de pUC4K contendo a cassette kan nos locais *Hind* III ou *Pst* I de pUHIT1 pois ambos são locais únicos nesta construção, para produzir pUHIT1KFH e pUHIT1KFP, Figura 28. Após digestão com *Eco* RI para remover as sequências génicas interrompidas, introduziram-se as construções no genoma de *H. influenzae* do tipo selvagem por transformação utilizando meios M-IV como descrito previamente (Barcak *et al.*, 1991) e seleccionaram-se os transformantes em ágar BHINH contendo 20 μ g/ml de canamicina.

Exemplo 20

Este Exemplo ilustra a construção de poliovírus que expressam um epítopo de um receptor de transferrina.

Um clone de ADNc das bases 1175 a 2956 do genoma do poliovírus do tipo 1, estirpe Mahoney (PV1-M) foi cortado com as enzimas de restrição *Sau* I e *Hind* III. Estas enzimas excisam um fragmento contendo as bases 2754 a 2786, que codifica os aminoácidos 1094 a 1102 de PV1-M, como mostrado na Figura 29. No presente pedido, utilizamos o código de quatro dígitos para os aminoácidos do poliovírus; por exemplo, 1095 é o aminoácido 95 da proteína da cápside VP1. Novos clones híbridos de ADNc que codificam sequências de aminoácidos tanto do poliovírus como de receptores de transferrina foram construídos substituindo o fragmento excisado por oligonucleótidos sintéticos que codificam para aminoácidos de Tbp2 de *H. influenzae*. Os novos clones de ADNc híbridos foram

cortados com as enzimas de restrição *Nhe* I e *Sna*B I, que excisam um fragmento híbrido, incluindo as sequências de ADN do receptor de transferrina, das bases 2471 a 2956 de poliovírus. Um clone de ADNc, por exemplo pT7XLD ou pT7CMCB, deste genoma completo de PV1-M foi cortado com *Nhe*I e *Sna*BI para excisar um fragmento das bases 2471 a 2956 de poliovírus. Este foi então substituído por um fragmento híbrido incluindo as sequências de ADN do receptor de transferrina para produzir um clone de ADNc híbrido do genoma de PV1-M com as bases 2754 a 2786 substituídas por bases que codificam uma ansa BC híbrida incluindo aminoácidos do receptor de transferrina, como mostrado na Figura 29.

O plasmídeo pT7XLD e os clones derivados de pT7XLD, tais como pT7CMCB, contêm uma sequência promotora para a enzima ARN-polimerase de T7 na extremidade 5' do ADNc de PV1-M. Os transcritos de ARN do ADNc de PV1-M, incluindo quaisquer bases que codificam aminoácidos do receptor de transferrina, foram preparados utilizando ARN-polimerase de T7 como descrito por van der Werf et al. A transfecção de células Vero com estes transcritos de ARN produziu quatro vírus híbridos viáveis, designados PV1TBP2A, PV1TBP2B, PV1TBP2C e PV1TBP2D. A transfecção com transcritos de pT7CMCB originou um poliovírus do tipo selvagem derivado por transfecção designado por PV1XLD (Figura 29).

As características antigénicas de PV1TBP2A, PV1TBP2B, PV1TBP2C e PV1TBP2D estão apresentadas na Tabela 5. Todos foram neutralizados por anti-soros de porquinho-da-índia criados contra um péptido incorporando a sequência LEGGFYGP (SEQ ID NO: 74), o que indica que os vírus expressaram esta sequência numa forma antigenicamente reconhecível. Para produzir os anti-soros, imunizaram-se IM porquinhos-da-índia fêmea com um volume de 500 µl contendo 200 µg de péptido formulado em fosfato de alumínio (3 mg/ml). Os animais foram imunizados nos dias 1, 14, 28 e 42 e sangrados nos dias 0, 28, 42 e 56. Os soros eram do dia 56 de sangria. PV1TBP2A e PV1TBP2B foram também neutralizados por anti-soros de coelho criados contra Tbp2 de *H. influenzae*, estirpe DL63, indicando que pelo menos estes dois vírus expressaram a sequência numa forma reconhecível por anticorpos criados contra a proteína. Todos os vírus eram neutralizáveis por soros anti-PV1, o que

indica que as alterações no local I antigénico de neutralização de polio não tinham afectado significativamente outros locais antigénicos nos vírus.

Exemplo 21

Este Exemplo ilustra a utilização de poliovírus híbridos para induzir títulos elevados de anti-soros contra Tbp2.

Inocularam-se coelhos com PV1TBP2A purificado por CsCl (coelhos #40, 41, 42). Note-se que, embora os vírus utilizados estivessem vivos, o poliovírus não replica em coelhos e que qualquer resposta observada é efectivamente a resposta a um antigénio inactivado. No dia 1, os coelhos foram inoculados com 1 ug de vírus em adjuvante completo de Freund subcutaneamente no dorso, e, no dia 14, os coelhos receberam um reforço com 1 ug de vírus em adjuvante incompleto de Freund inoculado subcutaneamente no dorso. Os coelhos foram sangrados no dia 0 (pré-sangria) e no dia 27. A dose de vírus por inoculação foi de $2,5 \times 10^8$ pfu, que se determinou a partir dos valores de A_{260} ser de aproximadamente $3,0 \times 10^{11}$ viriões. Isto equivalia a 0,5 pmol de vírus ou 30 pmol do epítipo LEGGFYG (SEQ ID NO: 74), pois cada virião expressa 60 cópias do epítipo.

SUMÁRIO DA DESCRIÇÃO

Em resumo desta descrição, a presente invenção proporciona moléculas de ADN purificadas e isoladas contendo genes de receptores de transferrina, as sequências destes genes de receptores de transferrina e as sequências de aminoácidos suas derivadas. A invenção também proporciona péptidos correspondentes a porções do receptor de transferrina. Os genes, as sequências de ADN, as proteínas e péptidos recombinantes, são úteis para o diagnóstico, a imunização e a produção de reagentes de diagnóstico e imunológicos. Vacinas baseadas em Tbp1 e/ou Tbp2 recombinantes expressas, suas porções, ou péptidos derivados das sequências proporcionadas, podem ser preparados para a prevenção de doenças causadas por patogénios bacterianos que produzem receptor de transferrina.

TABELA 1

comando	1° resíduo	plasmídeo	estirpe
<i>E. coli</i> lpp	Cys	JB-1360-1R-10	JB-1407-1-1
<i>E. coli</i> lpp	Ser	JB-1366-1R-7	JB-1407-3-1
<i>E. coli</i> pal	Cys	JB-1360-3-10	JB-1407-2-1
<i>E. coli</i> pal	Ser	JB-1366-3R-5	JB-1407-4-4
<i>E. coli</i> rlpB	Cys	JB-1399-1	JB-1437-1-1
<i>E. coli</i> rlpB	Ser	JB-1378-7	JB-1407-5-1

TABELA 2

PÉPTIDOS Tbp1 ANTIGÉNICOS PREVISTOS

PÉPTIDO	RESÍDUOS ¹	SEQUÊNCIAS	SEQ ID NO:
TBP1-N	1-36	AETQSIKDTKEAISSEVDTQSTEDSELETISVTAEK	13
TBP1-2	31-66	SVTAEKVRDRKDNEVTGLGKIIKTSESISREQVLNI	14
TBP1-3	59-94	SREQVLNIRDLTRYPGISVVEQGRGASSGYSIRGM	15
TBP1-4	88-123	GYSIRGMDRNRVALLVDGLPQTQSYWQSPLVARSG	16
TBP1-5	117-152	PLVARSGYGTGAINIEIYENVKAVEISKGGSSEYG	17
TBP1-6	147-182	SSSEYGNAGALAGSVTFQSKSAADILEGDKSWGQTK	18
TBP1-7	179-214	GIQTKNAYSSKNKGFTTHSLAVAGKQGGFEGVAIYTH	19
TBP1-8	208-243	GVAIYTHRNSIETQVHKDALKGVQSYDRFIATTEDQ	20
TBP1-9	236-271	IATTEDQSAYFVMQDECLDGYDKCKTSPKRPATLST	21
TBP1-10	266-301	PATLSTQRETVSVDYTGANRIKPNPMKYESQSWFL	22
TBP1-11	293-328	YESQSWFLRGGYHFSEQHYIGGIFFTQQKFDIRDM	23
TBP1-12	322-357	KFDIRDMTFPAYLRPTEDKDLQSRPFYQDYGAYQ	24
TBP1-13	352-387	DYGAYQHIGDGRGVKYASGLYFDEHHRKQRVGIEYI	25
TBP1-14	383-418	GIEYIYENKNKAGIIDKAVLSANQQNIILDSYMRHT	26
TBP1-15	412-447	DSYMRHTHCSLYPNPSKNCRPTLDKPYSYYHSDRNV	27
TBP1-16	443-478	SDRNVYKEKHNLQLNLEKKIQNWLTHTQIAFNLGF	28
TBP1-17	469-504	THQIAFNLGFDDFTSALQHKDYLTRRVIATASSISE	29
TBP1-M	498-534	TASSISEKRGARRNGLQSSPYLYPTPKAELVGGDLC	30
TBP1-19	528-563	LVGGDLCNYQGKSSNYSCKVRLIKGKNYYFAARNN	31
TBP1-20	558-593	FAARNNMALGKYVDLGLGMRYDVSRTKANESTISVG	32
TBP1-21	588-623	STISVGKFKNFSWNTGIVIKPTEWLDLSYRLSTGFR	33
TBP1-22	618-653	LSTGFRNPSFAEMYGWRYGGKDTDVYIGKFKPETS	34
TBP1-23	648-683	KPETS RNQEFGLALKGDFGNIEISHFSNAYRNLI	35
TBP1-24	677-712	YRNLI AFAEELSKNGTTGKGNYGYHNAQNAKLGVN	36

PÉPTIDOS Tbp1 ANTIGÉNICOS PREVISTOS

PÉPTIDO	RESÍDUOS ¹	SEQUÊNCIAS	SEQ ID NO:
TBP1-25	706-741	AKLVGVNITAQLDFNGLWKRIPTYGWYATFAYNRVKV	37
TBP1-26	735-770	AYNRVKVKDQKINAGLASVSSYLFDIAIQPSRYIIGL	38
TBP1-27	764-799	SRYIIIGLDYDHPSTWGIKTMFTQSKAKSQNELLGK	39
TBP1-28	794-829	NELLGKRALGNNRNVKSTRKLTRAHWILDVSGYYM	40
TBP1-29	825-854	SGYYMVNRSILFRLGVYNLLNYRYVTWEAV	41
TBP1-30	843-865	LLNYRYVTWEAVRQTAQGAEFDI	42
TBP1-31	42-50	DNEVTGLGK	43
TBP1-32	61-76	EQVLNIRDLDTRYDPI	44
TBP1-33	61-95	EQVLNIRDLDTRYDPGISVVEQGRGASSGYSIRGMD	45
TBP1-34	128-146	GAINEIEYENVKAVEISKG	46
TBP1-35	155-161	GALAGSV	47
TBP1-1	1-14	AETQSIKDTKEAISC ²	48

1. Número do resíduo da sequência de Tbp1 de *H. influenzae* tipo b, estirpe Egan (como mostrado na Figura 8).
2. Cisteína adicionada para facilitar o acoplamento a uma proteína transportadora, por exemplo KLH.

TABELA 3

PÉPTIDOS Tbp2 ANTIGÉNICOS CONSERVADOS PREVISTOS

PÉPTIDO	RESÍDUOS ¹	SEQUÊNCIAS	SEQ ID NO:
TBP2-1	18-31	CSGGGSFDDVDNVS	49
TBP2-2	231-261	LEGGFYGPKEELGFRFLAGDKKVFVFSK	50
TBP2-3	358-380	TVGKKTYQVEACCSNLSYVKFGM	51
TBP2-4	527-549	ATVKGAFYGPKEELGGYFTYNG	52
TBP2-5	1-36	MKLAALNLFDRNKPSLLNEDSYMIFSSRSTIEEDV	53
TBP2-6	29-64	STIEEDVKNDNQNGEHPIDSIVDPAPNSNENRHG	54
TBP2-7	57-92	SNENRHGQKYVYSGLYYIQSWSLRDLNKKFYSGY	55
TBP2-8	85-120	KKFYSGYGYAYYFGNTTASALPVGGVATYKGTWS	56
TBP2-9	113-148	TYKGTWSFITAAENGKNYELLRNSSGGGQAYSRRSA	57
TBP2-10	141-176	AYSRRSATPEDIDLDRKTGLTSEFTVNFGTKKLTG	58
TBP2-11	169-204	GTKKLTGGLYNNLRETDANKSQNRTHKLYDLEADV	59
TBP2-12	197-232	YDLEADVHSNRFRGKVKPTKESSEHPFTSEGT	60
TBP2-13	225-260	FTSEGTLEGGFYGPEGQELGGKFLAHDKKVLGVFS	61

PÉPTIDOS Tbp2 ANTIGÉNICOS CONSERVADOS PREVISTOS

PÉPTIDO	RESÍDUOS ¹	SEQUÊNCIAS	SEQ ID NO:
TBP2-14	253-288	KVLGVFSAKEQQETSENKKLPKETLIDGKLTTFKT	62
TBP2-15	281-316	KLTTFKTTNATANATTDATTSTTASTKTDTTNAT	63
TBP2-16	309-344	DTTNNATANTENFTTKDIPSLGEADYLLIDNYPVP	64
TBP2-17	337-372	IDNYPVPLFPESGDFISSKHHTVGKKTYQVEACCS	65
TBP2-M	360-406	CSNLSYVKFGMYEAPPKEEEKEKEKDKDKEKEKQA	66
TBP2-19	393-428	KEKDKDKEKEKQATTSIKTYQFLLGLRTPSSEIP	67
TBP2-20	421-456	TPSSEIPKEGSAKYHGNWFGYISDGETSYSASGDK	68
TBP2-21	449-484	YSASGDKERSKNAVAEFNVNFAEKTTLTGELKRHDT	69
TBP2-22	477-512	ELKRHDTQNPVFKINATFQSGKNDFGTATAKDLA	70
TBP2-23	505-540	ATAKDLAIDGKNTQGTSKVNFTATVNGAFYGPAT	71
TBP2-24	533-559	FYGPATELGGYFTYNGNPTDKNSS	72
TBP2-C	553-581	CPTDKNSSSNSEKARAAVVFAGKKQQVETTK	73
TBP2-25	231-238	LEGGFYGP	74
TBP2-26	18-25	CSGGGSFD	75
TBP2-27	130-134	YVYSGL	76
TBP2-28	345-355	CCSNLSYVKFG	77
TBP2-29	401-407	FLLGHRT	78
TBP2-30	450-456	EFNVDF	79
TBP2-31	485-491	NAFTGTA	80
TBP2-32	516-522	VNGAFYG	81
TBP2-33	527-532	ELGGYF	82
TBP2-34	562-566	VVFGAR	83
TBP2-35	562-568	VVFGAK	84
TBP2-36	231-237	LEGGFYG	85

1. Número do resíduo da sequência de Tbp2 de *H. influenzae* tipo B, estirpe Eagan (como mostrado na Figura 9).

TABELA 4

Respostas de anticorpos de porquinho-da-índia a péptidos de Tbp1 e Tbp2

PÉPTIDO	SEQ ID	SEQUÊNCIAS	TÍTULO
TBP1-N	13	AETQSIKDTKEAISSEVDTQSTEDSELETISVTAEK	500
TBP1-M	30	TASSISEKRGGEARRNGLQSSPYLYPTPKAELVGGDLG	1562500
TBP1-1	48	AETQSIKDTKEAISC	<100
TBP2-1	49	CSGGGSFDVDNVSN	2500
TBP2-2	50	LEGGFYGPKGEELGFRFLAGDKKVFVFSK	12500
TBP2-3	51	TVGKKTYQVEACCSNLSYVKFGM	62500
TBP2-4	52	ATVKGAIFYGPKASELGGYFTYNG	<100
TBP2-M	66	CSNLSYVKFGMYEAPPKEEEEKEKEKDKDKEKEKQA	62500
TBP2-C	73	CPTDKNSSSNSEKARAAVVFAGKKQVETTK	312500

TABELA 5

Actividade de neutralização de soros anti-Tbp2 e anti-péptido contra vírus híbridos polio/Tbp2

Soros ^a	Título de Neutralização v. Vírus ^b				
	PVITBP2A	PVITBP2B	PVITBP2C	PVITBP2D	PV1XLD
Rb @PV1	>409600	25844	20480	16763	>40960
Rb 516 D0	<4	<4	<4	<4	<4
Rb 516 D42	40	20	<4	<4	<4
conjunto de GP561, 562, D0	<4	<4	<4	<4	<4
GP 561 D56	>2048	>2048	>2048	1164	<4
GP 562 D56	>2048	>2048	25	10	<4
Conjunto de GP558, 559, 560, D56	<4	<4	<4	<4	<4

^a Rb @PV1 é o conjunto de soros imunes de coelho criados contra PV1XLD. O coelho (Rb) 516 foi imunizado com três doses sucessivas de 3 µg de proteína 2 de ligação a transferrina de *H. influenzae* DL63 recombinante nos dias 1, 14 e 28. O soro foi recolhido nos dias 0 (D0) e 42 (D42). Os porquinhos-da-índia foram imunizados com quatro doses sucessivas de 200µg de péptido nos dias 1, 14, 28 e 42. Os soros foram recolhidos no dia 0 (D0) e no dia 56 (D56). Os porquinhos-da-índia (GP) 561 e 562 receberam um péptido contendo a sequência LEGGFYGP (SEQ ID NO:74). Os porquinhos-da-índia (GP) 558, 559 e 556 receberam um péptido de controlo com uma sequência não relacionada.

^b O título é o inverso da diluição de soro que origina 50% do ponto final num ensaio de neutralização de vírus versus 100 TCID₅₀ de vírus.

TABELA 6
Títulos de IgG específica do péptido de coelhos imunizados com
PV1TBP2A ou PV1TBP2B

COELHO (ANTIGÉNIO)	TÍTULOS DE IgG ESPECÍFICA DO PÉPTIDO ^a	
	PRÉ-SANGRIA	DIA 27
40 (PV1TBP2A)	<20	640
41 (PV1TBP2A)	<20	640
42 (PV1TBP2A)	<20	2560
43 (PV1TBP2B)	<20	160
44 (PV1TBP2B)	<20	1280
45 (PV1TBP2B)	<20	1280
10 (PV1 Mahoney)		<20 ^b
11 (PV1 Mahoney)		<20

^a Os títulos são o recíproco da maior diluição de soros que origina uma A₄₅₀ de pelo menos duas vezes o valor do ruído de fundo. O valor do ruído de fundo era a A₄₅₀ média de poços ensaiados na ausência de soros de coelho.

^b Os títulos para os coelhos 10 e 11 referem-se a soros colhidos no dia 42 após três imunizações com PV1 Mahoney. Os coelhos 10 e 11 foram imunizados como os coelhos 40 a 45, excepto administrando uma dose de reforço adicional no dia 28.

LISTAGEM DE REFERÊNCIAS

- Barcak *et al.*, (1991) *Methods Enzymol.* **204**: 321-342.
- Berkowitz *et al.*, (1987) *J. Pediatr.* **110**:509.
- Black *et al.*, (1991) *Pediatr. Infect. Dis. J.* **10**:97.
- Bluestone, N. (1982) *Engl. J. Med.* **306**:1399.
- Chang *et al.*, (1978) *Nature* **375**:615.
- Chou, *et al.*, (1978). *Annual Reviews of Biochemistry* **47**, 251-276.
- Claesson *et al.*, (1989) *J. Pediatr.* **114**:97.
- Cornelissen *et al.*, (1992) *J. Bacteriol.* **174**:5788
- Danve, *et al.*, (1993). *Vaccine* **11**, 1214-1220.
- Deres *et al.*, (1989) *Nature* **342**:651.
- Gerlach, *et al.*, (1992) *Infect. Immun.* **608**:325
- Goeddel *et al.*, (1979) *Nature* **281**:544.

- Goeddel et al., (1980) *Nucl. Acids Res.* **8**:4057
- Harkness et al., (1992) *J. Bacteriol.* **174**:2425.
- Holland et al., (1992) *Infect. Immun.* **60**:2986.
- Hopp, T.P. (1986) *Journal of Immunological Methods* **88**, 1-18.
- Itakura et al., (1977) *Science* **198**:1056.
- Jarosik et al., (1994). *Infection and Immunity* **62**, 2470-2477.
- Legrain et al., (1993). *Gene* **130**:73
- Lockhoff et al., (1991) *Chem. Int. Ed. Engl.* **30**:1611.
- Mickelsen e Sparling, (1981) *Infect. Immun.* **33**:555.
- Morton et al., (1993) *Infection and Immunity* **61**, 4033-4037.
- Murdin et al., (1992) *J. Gen. Viral* **73**: 607.
- Murdin et al., (1991) *Microbial Pathogenesis* **10**:27.
- Nixon-George et al., (1990) *J. Immunol.* **14**:4798.
- Ogunnariwo, e Schryvers, (1992) *Avian Dis.* **36**:655.
- O'Hagan (1992) *Clin Pharmacokinet.* **22**:1.
- Panezutti et al., (1993) *Infection and Immunity* **61**, 1867-1872.
- Roosi-Campos et al., (1992) *Vaccine* **10**, 512-518.
- Schryvers, (1988) *Molec. Microbiol.* **2**:467.
- Schryvers e Lee, (1989) *Can. J. Microbiol.* **35**:409.
- Schryvers e Gray-Owen, (1992) *J. Infect. Dis.* **165** suppl.1: S103.
- Schryvers (1989) *Med. Microbiol.* **29**:121.
- Short et al., (1988) *Nucl. Acids Res.* **16**:7583.
- Ulmer et al., (1993) *Curr. Opinion Invest. Drugs.* **2 (9)**: 983-989.
- Van der Werf et al., (1986) *Proc. Natl. Acad. Sci.* **83**: 2330.
- Weismuller et al., (1989) *Vaccine* **8**:29.
- Wilton et al., (1993) *FEMS Microbiology Letters* **107**, 59-66.
- Patente U.S. 4 855 283

Patente U.S. 4 258 029

Patente U.S. 4 496 538

Patente U.S. 4 599 230

Patente U.S. 4 599 231

Patente U.S. 4 601 903

Patente U.S. 5 141 743

Patente U.S. 4 596 792

Patente U.S. 4 952 496

Patente U.S. 5 194 254

WO 92/17167

Lisboa,

REIVINDICAÇÕES

1. Péptido possuindo não menos de sete aminoácidos e não mais de 150 aminoácidos e contendo uma sequência de aminoácidos que é conservada entre bactérias que produzem uma proteína receptora da transferrina compreendendo TBP-2, sendo a referida sequência conservada de TBP-2:

LEGGFYG (SEQ ID NO: 85).

2. Péptido de acordo com a reivindicação 1 e contendo uma sequência de aminoácidos que é:

LEGGFYGP (SEQ ID NO: 74).

3. Péptido de acordo com a reivindicação 1, em que o referido péptido é:

LEGGFYGP (SEQ ID NO: 74) ou

LEGGFYG (SEQ ID NO: 85).

4. Péptido de acordo com a reivindicação 1, em que o referido péptido é:

LEGGFYGPKGEELGFRFLAGDKKVFGVFSAK (SEQ ID NO: 50).

5. Composição imunogénica, compreendendo como componente activo um péptido de acordo com qualquer reivindicação anterior.

6. Composição imunogénica de acordo com a reivindicação 5, formulada na forma de uma vacina para administração *in vivo* para proteger contra doenças causadas por patogénios bacterianos que produzem proteínas receptoras da transferrina.

7. Composição imunogénica de acordo com a reivindicação 6, formulada na forma de uma composição de micropartículas, cápsulas ou lipossomas.

8. Composição imunogénica de acordo com qualquer uma das reivindicações 5 a 7, compreendendo uma pluralidade dos referidos componentes activos para proporcionar protecção contra uma doença causada por uma pluralidade de espécies de bactérias produtoras de proteínas receptoras da transferrina.

9. Anti-soro ou anticorpo específicos para uma sequência de aminoácidos

LEGGFYG (SEQ ID NO: 85),

LEGGFYGP (SEQ ID NO: 74), ou

LEGGFYGPKGEELGFRFLAGDKKVFGVFSAK (SEQ ID NO: 50).

Lisboa,

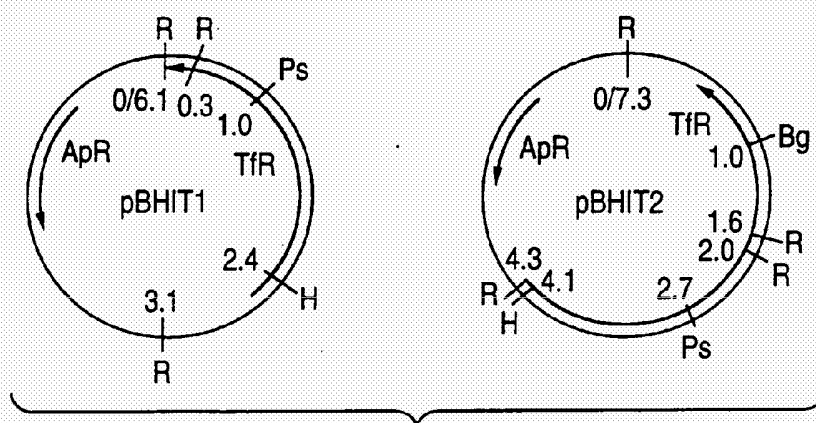


FIG.1A.

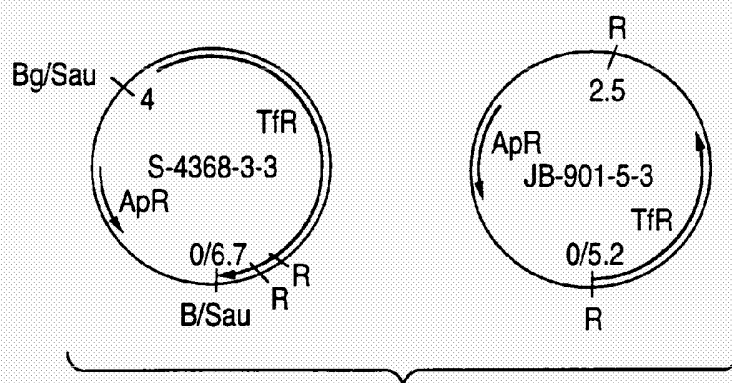


FIG.1B.

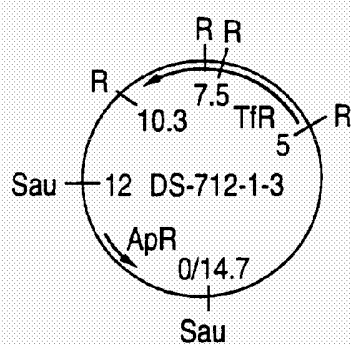


FIG.1C.

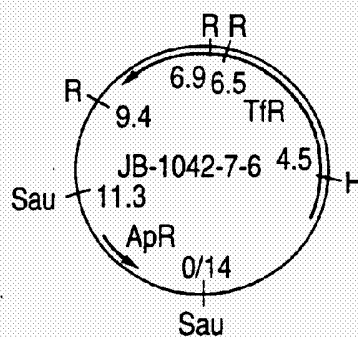


FIG.1D.

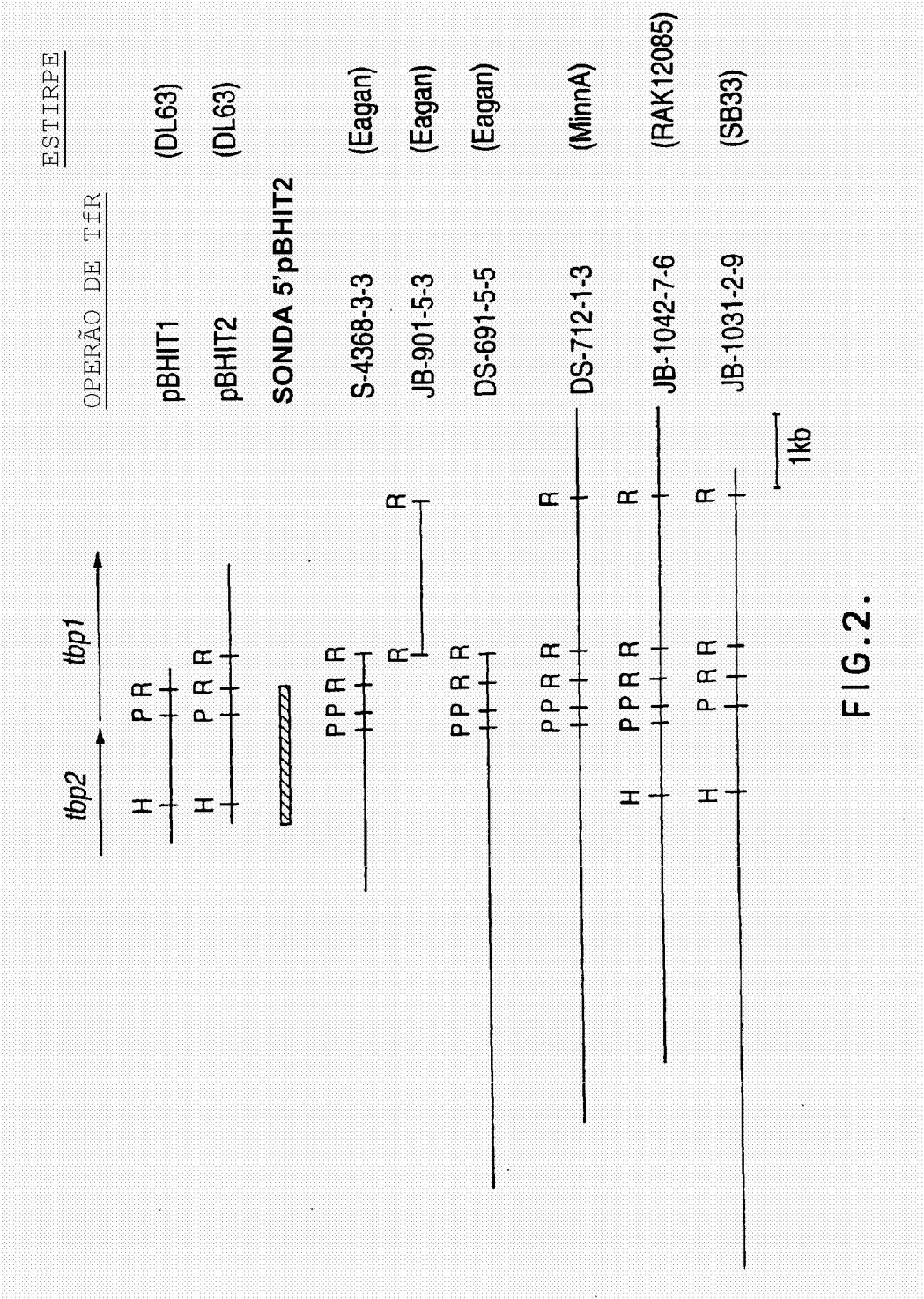


FIG.2.

FIG. 3 A.

TATAACTCA ATG AAA TCT GTA CCT CTT ATC TCT GGT GGA CTT TCC TTT
Met Lys Ser Val Pro Leu Ile Ser Gly Gly Leu Ser Phe
 1 5 10

TTA CTA AGT GCT TGT AGC GGA GGG GGG TCT TTT GAT GTA GAT AAC GTC
Leu Leu Ser Ala Cys Ser Gly Gly Gly Ser Phe Asp Val Asp Asn Val
 15 20 25

TCT AAT ACC CCC TCT TCT AAA CCA CGT TAT CAA GAC GAT ACT TCA AGT
Ser Asn Thr Pro Ser Ser Lys Pro Arg Tyr Gln Asp Asp Thr Ser Ser
 30 35 40 45

TCA AGA ACA AAA TCT AAA TTG GAA AAG TTG TCC ATT OCT TCT TTA GGG
Ser Arg Thr Lys Ser Lys Leu Glu Lys Leu Ser Ile Pro Ser Leu Gly
 50 55 60

GGA GGG ATG AAG TTA GCG GCT CTG AAT CTT TTT GAT AGG AAC AAA CCT
Gly Gly Met Lys Leu Ala Ala Leu Asn Leu Phe Asp Arg Asn Lys Pro
 65 70 75

AGT CTC TTA AAT GAA GAT AGC TAT ATG ATA TTT TCC TCA CGT TCT ACG
Ser Leu Leu Asn Glu Asp Ser Tyr Met Ile Phe Ser Ser Arg Ser Thr
 80 85 90

FIG. 3B.

```

ATT GAA GAG GAT GTT AAA AAT GAC AAT CAA AAC GGC GAG CAC CCT ATT
Ile Glu Glu Asp Val Lys Asn Asp Asn Gln Asn Gly Glu His Pro Ile
95
GAC TCA ATA GTC GAT CCT AGA GCA CCA AAT TCA AAC GAA AAT CGT CAT
Asp Ser Ile Val Asp Pro Arg Ala Pro Asn Ser Asn Glu Asn Arg His
110
GGA CAA AAA TAT GTA TAT TCA GGG CTT TAT TAT ATT CAA TCG TGG AGT
Gly Gln Lys Tyr Val Tyr Ser Gly Leu Tyr Tyr Ile Gln Ser Trp Ser
130
CTA AGA GAT TTA CCA AAT AAA AAG TTT TAT TCA GGT TAC TAT GGA TAT
Leu Arg Asp Leu Pro Asn Lys Lys Phe Tyr Ser Gly Tyr Tyr Gly Tyr
145
GGG TAT TAC TTT GGC AAT ACA ACT GCC TCT GCA TTA OCT GTA GGT GGC
Ala Tyr Tyr Phe Gly Asn Thr Thr Ala Ser Ala Leu Pro Val Gly Gly
160
GTA GCA ACG TAT AAA GGA ACT TGG AGC TTC ATC ACC GCA GCT GAA AAT
Val Ala Thr Tyr Lys Gly Thr Trp Ser Phe Ile Thr Ala Ala Glu Asn
175

```

FIG.3C.

```

GGC AAG AAT TAT GAA TTG TTA AGA AAT TCT GGT GGC GGT CAA GCT TAT
Gly Lys Asn Tyr Glu Leu Leu Arg Asn Ser Gly Gly Gln Ala Tyr
190                                     200                205

TCT CGA CGT AGT GCT ACT ACT CCA GAA GAT ATT GAT TTA GAT CGT AAG ACG
Ser Arg Arg Ser Ala Thr Pro Glu Asp Ile Asp Leu Asp Arg Lys Thr
210                                     215                220

GGC TTA ACA AGT GAA TTT ACT GTC AAT TTT GGT ACA AAA AAG CTC ACT
Gly Leu Thr Ser Glu Phe Thr Val Asn Phe Gly Thr Lys Lys Leu Thr
225                                     230                235

GGA GGA CTT TAT TAT AAT TTA CGT GAA ACA GAT GCT AAT AAA TCA CAA
Gly Gly Leu Tyr Tyr Asn Leu Arg Glu Thr Asp Ala Asn Lys Ser Gln
240                                     245                250

AAT AGA ACA CAT AAA CTC TAC GAT CTA GAA GCT GAT GTT CAT AGC AAC
Asn Arg Thr His Lys Leu Tyr Asp Leu Glu Ala Asp Val His Ser Asn
255                                     260                265

CGA TTC AGG GGT AAA GTA AAG CCA ACC AAA AAA GAG TCT TCT GAA GAA
Arg Phe Arg Gly Lys Val Lys Pro Thr Lys Lys Glu Ser Ser Glu Glu
270                                     275                280                285

```

FIG.3D.

```

CAT CCC TTT ACC AGC GAG GGA ACA TTA GAA GGT GGT TTT TAC GGG CCT
His Pro Phe Thr Ser Glu Gly Thr Leu Glu Gly Phe Tyr Gly Pro
290 295 300

GAG GGT CAA GAA TTA GGA GGA AAG TTT TTA GCT CAC GAC AAA AAA GTT
Glu Gly Gln Glu Leu Gly Gly Lys Phe Leu Ala His Asp Lys Lys Val
305 310 315

TTG GGG GTA TTT AGT GCC AAA GAA CAG CAA GAA ACG TCA GAA AAC AAA
Leu Gly Val Phe Ser Ala Lys Glu Gln Gln Glu Thr Ser Glu Asn Lys
320 325 330

AAA TTA CCC AAA GAA ACC TTA ATT GAT GGC AAG CTA ACT ACT TTT AAA
Lys Leu Pro Lys Glu Thr Leu Ile Asp Gly Lys Leu Thr Thr Phe Lys
335 340 345

ACA ACC AAT GCA ACA GCC AAT GCA ACA ACC GAT GCA ACA ACC AGT ACA
Thr Thr Asn Ala Thr Ala Asn Ala Thr Thr Asp Ala Thr Thr Ser Thr
350 355 360 365

ACA GCC AGT ACA AAA ACC GAT ACA ACA ACC AAT GCA ACA GCC AAT ACA
Thr Ala Ser Thr Lys Thr Asp Thr Thr Thr Asn Ala Thr Ala Asn Thr
370 375 380

```

FIG. 3E.

GAA AAC TTT ACG ACA AAA GAT ATA CCA AGT TTG GGT GAA GCT GAT TAT
 Glu Asn Phe Thr Thr Lys Asp Ile Pro Ser Leu Gly Glu Ala Asp Tyr 395
 385

CTT TTA ATT GAT AAT TAC CCT GTT CCT CTT TTC CCT GAG AGT GGT GAT
 Leu Leu Ile Asp Asn Tyr Pro Val Pro Leu Phe Pro Glu Ser Gly Asp 410
 400

TTC ATA AGT AGT AAG CAC CAT ACT GTA CGA AAG AAA ACC TAT CAA GTA
 Phe Ile Ser Ser Lys His His Thr Val Gly Lys Lys Thr Tyr Gln Val 425
 415

GAA GCA TGT TGC AGT AAT CTA AGC TAT GTA AAA TTT GGT ATG TAT TAT
 Glu Ala Cys Cys Ser Ser Asn Leu Ser Tyr Val Lys Phe Gly Met Tyr Tyr 445
 430

GAA GCC CCA CCT AAA GAA GAA GAA AAA GAA AAA GAA GAC AAA GAC
 Glu Ala Pro Pro Lys Glu Glu Glu Lys Glu Lys Lys Asp Lys Asp 460
 450

AAA GAA AAA GAA AAA CAA CCG ACA ACA TCT ATC AAG ACT TAT TAT CAA
 Lys Glu Lys Lys Glu Lys Gln Ala Thr Thr Ser Ile Lys Thr Tyr Tyr Gln 475
 465

FIG. 3F.

```

TTC TTA TTA GGT CTC CGT ACT CCC AGT TCT GAA ATA CCT AAA GAA GGA
Phe Leu Leu Gly Leu Arg Thr Pro Ser Ser Glu Ile Pro Lys Glu Gly
480
AGT GCA AAA TAT CAT GGT AAT TGG TTT GGT TAT ATT AGT GAT GCC GAG
Ser Ala Lys Tyr His Gly Asn Trp Phe Gly Tyr Ile Ser Asp Gly Glu
495
ACA TCT TAC TCC GCC AGT GGT GAT AAG GAA CGC AGT AAA AAT GCT GTC
Thr Ser Tyr Ser Ala Ser Gly Asp Lys Glu Arg Ser Lys Asn Ala Val
510
GCC GAG TTT AAT GTA AAT TTT GCC GAG AAA ACA TTA ACA GGC GAA TTA
Ala Glu Phe Asn Val Asn Phe Ala Glu Lys Thr Leu Thr Gly Glu Leu
530
AAA CGA CAC GAT ACT CAA AAT CCC GTA TTT AAA ATT AAT GCA ACC TTT
Lys Arg His Asp Thr Gln Asn Pro Val Phe Lys Ile Asn Ala Thr Phe
545
CAA AGT GGT AAG AAT GAC TTC ACT GGT ACA GCA ACC GCA AAA GAT TTA
Gln Ser Gly Lys Asn Asp Phe Thr Gly Thr Ala Thr Ala Lys Asp Leu
560

```

FIG. 36.

```

GCA ATA GAT GGT AAA AAT ACA CAA GGC ACA TCT AAA GTC AAT TTC ACG
Ala Ile Asp Gly Lys Asn Thr Gln Gly Thr Ser Lys Val Asn Phe Thr
575                                     580                                     585

GCA ACA GTA AAC GGG GCA TTT TAT GGT CCG CAC GCT ACA GAA TTA GGC
Ala Thr Val Asn Gly Ala Phe Tyr Gly Pro His Ala Thr Glu Leu Gly
590                                     595                                     600                                     605

GGT TAT TTC ACC TAT AAC GGA AAC AAT CCT ACA GAT AAA AAT TCA TCA
Gly Tyr Phe Thr Tyr Asn Gly Asn Asn Pro Thr Asp Lys Asn Ser Ser
610                                     615                                     620

TCC AAT TCA GAA AAG GCA AGA GCT GGC GTT GTG TTT GGA GCT AAA AAA
Ser Asn Ser Glu Lys Ala Arg Ala Val Val Phe Gly Ala Lys Lys
625                                     630                                     635

CAA CAA GTA GAA ACA ACC AA GTAATGGAAT ACTAAA A ATG ACT AAA AAA
Gln Gln Val Glu Thr Thr Lys
640                                     645                                     645
Met Thr Lys Lys

CCC TAT TTT GGC CTA AGT AAT AAT TCT TGT CTT TTA ATT TCA TCC TAT
Pro Tyr Phe Arg Leu Ser Ile Ile Ser Cys Leu Leu Ile Ser Cys Tyr
650                                     655                                     660

```

FIG. 3H.

GTA AAA GCA	GAA ACT CAA AGT ATA AAA GAT ACA AAA GAA OCT ATA TCA	
<u>Val Lys Ala Glu Thr Gln Ser Ile Lys Asp Thr Lys Glu Ala Ile Ser</u>		
665	670	675 680
TCT GAA GTG GAC ACT CAA AGT ACA GAA GAT TCA GAA TTA GAA ACT ATC		
<u>Ser Glu Val Asp Thr Gln Ser Thr Glu Asp Ser Glu Leu Glu Thr Ile</u>		
	685	690 695
TCA GTC ACT GCA GAA AAA GTT AGA GAT CGT AAA GAT AAT GAA GTA ACT		
<u>Ser Val Thr Ala Glu Lys Val Arg Asp Arg Lys Asp Asn Glu Val Thr</u>		
	700	705 710
GGA CTT GGC AAA ATT ATA AAA ACT AGT GAA AGT ATC AGC CGA GAA CAA		
<u>Gly Leu Gly Lys Ile Ile Lys Thr Ser Glu Ser Ile Ser Arg Glu Gln</u>		
	715	720 725
GTA TTA AAT ATT CGT GAT CTA ACA CGC TAT GAT CCA GCG ATT TCA GTT		
<u>Val Leu Asn Ile Arg Asp Leu Thr Arg Tyr Asp Pro Gly Ile Ser Val</u>		
	730	735 740
GTA GAA CAA GGT CGC CGT GCA AGT TCT CGA TAT TCT ATT CGT CGT ATG		
<u>Val Glu Gln Gly Arg Gly Ala Ser Ser Gly Tyr Ser Ile Arg Gly Met</u>		
	745	750 755 760

FIG.31.

GAC AGA AAT AGA GTT GCT TTA TTA GTA GAT GGT TTA CCT CAA ACG CAA
 Asp Arg Asn Arg Val Ala Leu Leu Val Asp Gly Leu Pro Gln Thr Gln
 765 770 775
 TCT TAT GTA GTG CAA AGC OCT TTA GTT GCT OCT TCA GGA TAT TCT GGC
 Ser Tyr Val Val Gln Ser Pro Leu Val Ala Arg Ser Gly Tyr Ser Gly
 780 785 790
 ACT GGT GCA ATT AAT GAA ATT GAA TAT GAA AAT GTA AAG GCC GTC GAA
 Thr Gly Ala Ile Asn Glu Ile Glu Tyr Glu Asn Val Lys Ala Val Glu
 795 800 805
 ATA AGC AAG GGG GGG AGT TCT TCT GAG TAT GGT AAT GGA GCA CTA GCT
 Ile Ser Lys Gly Gly Ser Ser Ser Glu Tyr Gly Asn Gly Ala Leu Ala
 810 815 820
 GGT TCT GTA ACA TTT CAA AGC AAA TCA GCA GCC GAT ATC TTA GAA GGA
 Gly Ser Val Thr Phe Gln Ser Lys Ser Ala Ala Asp Ile Leu Glu Gly
 825 830 835 840
 GAC AAA TCA TGG GGA ATT CAA ACT AAA AAT GCT TAT TCA AGC AAA AAT
 Asp Lys Ser Trp Gly Ile Gln Thr Lys Asn Ala Tyr Ser Ser Lys Asn
 845 850 855

FIG. 3J.

```

AAA GGC TTT ACC CAT TCT TTA GCT GTA GCA GGA AAA CAA GGT GGA TTT
Lys Gly Phe Thr His Ser Leu Ala Val Ala Gly Lys Gln Gly Gly Phe
860                                     870

GAA GGG GTC GCC ATT TAC ACT CAC CGA AAT TCA ATT GAA ACC CAA GTC
Glu Gly Val Ala Ile Tyr Thr His Arg Asn Ser Ile Glu Thr Gln Val
875                                     885

CAT AAA GAT GCA TTA AAA GGC GTG CAA AGT TAT GAT CGA TTC ATC GCC
His Lys Asp Ala Leu Lys Gly Val Gln Ser Tyr Asp Arg Phe Ile Ala
890                                     900

ACA ACA GAG GAT CAA TCT GCA TAC TTT GTG ATG CAA GAT GAG TGT CTA
Thr Thr Glu Asp Gln Ser Ala Tyr Phe Val Met Gln Asp Glu Cys Leu
905                                     915

GAT GGT TAT GAC AAG TGT AAA ACT TCA CCC AAA CGA CCT GCG ACT TTA
Asp Gly Tyr Asp Lys Cys Lys Thr Ser Pro Lys Arg Pro Ala Thr Leu
920                                     935

TCC ACC CAA AGA GAA ACC GTA ACC GTT TCA GAT TAT ACC GGG GCT AAC
Ser Thr Gln Arg Glu Thr Val Ser Val Ser Asp Tyr Thr Gly Ala Asn
940                                     950

```

FIG. 3K.

CGT ATC AAA CCT AAT CCA ATG AAA TAT GAA AGC CAG TCT TCG TTT TTA
 Arg Ile Lys Pro Asn Pro Met Lys Tyr Glu Ser Gln Ser Trp Phe Leu
 955 960 965

AGA GGA GGT TAT CAT TTT TCT GAA CAA CAC TAT ATT GGT GGT ATT TTT
 Arg Gly Gly Tyr His Phe Ser Glu Gln His Tyr Ile Gly Gly Ile Phe
 970 975 980

GAA TTC ACA CAA CAA AAA TTT GAT ATC CGT GAT ATG ACA TTT CCC GCT
 Glu Phe Thr Gln Gln Lys Phe Asp Ile Arg Asp Met Thr Phe Pro Ala
 985 990 995 1000

TAT TTA AGG CCA ACA GAA GAC AAG GAT TTA CAA AGT CGC CCT TTT TAT
 Tyr Leu Arg Pro Thr Glu Asp Lys Asp Leu Gln Ser Arg Pro Phe Tyr
 1005 1010 1015

CCA AAG CAA GAT TAT GGT GCA TAT CAA CAT ATT GGT GAT GGC AGA GGC
 Pro Lys Gln Asp Tyr Gly Ala Tyr Gln His Ile Gly Asp Gly Arg Gly
 1020 1025 1030

GTT AAA TAT GCA AGT GGG CTT TAT TTC GAT GAA CAC CAT AGA AAA CAG
 Val Lys Tyr Ala Ser Gly Leu Tyr Phe Asp Glu His His Arg Lys Gln
 1035 1040 1045

FIG.3L.

```

CGT GTA GGT ATT GAA TAT ATT TAC GAA AAT AAG AAC AAA GCG GGC ATC
Arg Val Gly Ile Glu Tyr Ile Tyr Glu Asn Lys Asn Lys Ala Gly Ile
1050
ATT GAC AAA GCG GIG TTA AGT GCT AAT CAA CAA ACA TCA TAC TTG ACA
Ile Asp Lys Ala Val Leu Ser Ala Asn Gln Gln Thr Ser Tyr Leu Thr
1065
GTT ATA TGC GAC ATA CCG ATT GCA GTC TTT ATC CAT AAT CCA AGT AAG
Val Ile Cys Asp Ile Arg Ile Ala Val Phe Ile His Asn Pro Ser Lys
1085
AAT TGC CCG CCA ACA CTT GAT AAA CCT TAT TCA TAC TAT CAT TCT GAT
Asn Cys Arg Pro Thr Leu Asp Lys Pro Tyr Ser Tyr Tyr His Ser Asp
1100
AGA AAT GTT TAT AAA GAA AAA CAT AAC ATG TTG CAA TTG AAT TTA GAG
Arg Asn Val Tyr Lys Glu Lys His Asn Met Leu Gln Leu Asn Leu Glu
1115
AAA AAA ATT CAA CAA AAT TGG CTT ACT CAT CAA ATT GCC TTC AAT CTT
Lys Lys Ile Gln Gln Asn Trp Leu Thr His Gln Ile Ala Phe Asn Leu
1130

```

FIG. 3M.

```

GGT TTT GAT GAC TTT ACT TCC GCA CTT CAG CAT AAA GAT TAT TTA ACT
Gly Phe Asp Asp Phe Thr Ser Ala Leu Gln His Lys Asp Tyr Leu Thr
1145                               1150                               1155                               1160

CGA CGT GTT ATC GCT ACG GCA AGT AGT ATT TCA GAG AAA CGT GGT GAA
Arg Arg Val Ile Ala Thr Ala Ser Ser Ile Ser Glu Lys Arg Gly Glu
1165                               1170                               1175

GCA AGA ACA AAT GGT TTA CAA TCA AGT CCT TAC TTA TAC CCA ACA CCA
Ala Arg Arg Asn Gly Leu Gln Ser Ser Pro Tyr Leu Tyr Pro Thr Pro
1180                               1185                               1190

AAA GCA GAG TTG GTA GCA GCA GAT CTT TGT AAT TAT CAA GGT AAG TCC
Lys Ala Glu Leu Val Gly Gly Asp Leu Cys Asn Tyr Gln Gly Lys Ser
1195                               1200                               1205

TCT AAT TAC AGT GAC TGT AAA GIG CCG TTA ATT AAA GGG AAA AAT TAT
Ser Asn Tyr Ser Asp Cys Lys Val Arg Leu Ile Lys Gly Lys Asn Tyr
1210                               1215                               1220

TAT TTC GCA GCA CGC AAT AAT ATG GCA TTA GGG AAA TAC GTT GAT TTA
Tyr Phe Ala Ala Arg Asn Asn Met Ala Leu Gly Lys Tyr Val Asp Leu
1225                               1230                               1235                               1240

```


FIG. 3N.

GGT TTA GGT ATG AGG TAT GAC GTA TCT CGT ACA AAA GCT AAT GAA TCA
 Gly Leu Gly Met Arg Tyr Asp Val Ser Arg Thr Lys Ala Asn Glu Ser
 1245 1250 1255

ACT ATT AGT GTT GGT AAA TTT AAA AAT TTC TCT TGG AAT ACT GGT ATT
 Thr Ile Ser Val Gly Lys Phe Lys Asn Phe Ser Trp Asn Thr Gly Ile
 1260 1265 1270

GTC ATA AAA CCA ACG GAA TGG CTT GAT CTT TCT TAT CGC CTT TCT ACT
 Val Ile Lys Pro Thr Glu Trp Leu Asp Leu Ser Tyr Arg Leu Ser Thr
 1275 1280 1285

CGA TTT ACA AAT CCT AGT TTT GCT GAA ATG TAT GGT TGG CGG TAT GGT
 Gly Phe Arg Asn Pro Ser Phe Ala Glu Met Tyr Gly Trp Arg Tyr Gly
 1290 1295 1300

GGC AAG GAT ACC GAT GTT TAT ATA GGT AAA TTT AAG CCT GAA ACA TCT
 Gly Lys Asp Thr Asp Val Tyr Ile Gly Lys Phe Lys Pro Glu Thr Ser
 1305 1310 1315 1320

CGT AAC CAA GAG TTT GGT CTC GCT CTA AAA GGG GAT TTT GGT AAT ATT
 Arg Asn Gln Glu Phe Gly Leu Ala Leu Lys Gly Asp Phe Gly Asn Ile
 1325 1330 1335

FIG.30.

GAG ATC AGT CAT TTT AGT AAT GCT TAT CGA AAT CTT ATC GCC TTT GCT
 Glu Ile Ser His Phe Ser Asn Ala Tyr Arg Asn Leu Ile Ala Phe Ala
 1340 1345 1350

CAA GAA CTT AGT AAA AAT GGA ACT ACT GGA AAG GCC AAT TAT CGA TAT
 Glu Glu Leu Ser Lys Asn Gly Thr Thr Gly Lys Gly Asn Tyr Gly Tyr
 1355 1360 1365

CAT AAT GCA CAA AAT GCA AAA TTA GTT GGC GTA AAT ATA ACT GCG CAA
 His Asn Ala Gln Asn Ala Lys Leu Val Gly Val Asn Ile Thr Ala Gln
 1370 1375 1380

TTA GAT TTT AAT GGT TTA TGG AAA CGT ATT CCC TAC GGT TGG TAT GCA
 Leu Asp Phe Asn Gly Leu Trp Lys Arg Ile Pro Tyr Gly Trp Tyr Ala
 1385 1390 1395 1400

ACA TTT GCT TAT AAC CGA GTA AAA GTT AAA GAT CAA AAA ATC AAT GCT
 Thr Phe Ala Tyr Asn Arg Val Lys Val Lys Asp Gln Lys Ile Asn Ala
 1405 1410 1415

GGT TTA GCT TOC GTA AGC AGT TAT TTA TTT GAT GCC ATT CAG CCC AGC
 Gly Leu Ala Ser Val Ser Ser Tyr Leu Phe Asp Ala Ile Gln Pro Ser
 1420 1425 1430

FIG. 3P.

CGT TAT ATC ATT GGT TTA GGC TAT GAT CAT CCA AGT AAT ACT TCG GCA
Arg Tyr Ile Ile Gly Leu Gly Tyr Asp His Pro Ser Asn Thr Trp Gly
1435 1440 1445

ATT AAG ACA ATG TTT ACT CAA TCA AAA GCA AAA TCT CAA AAT GAA TTG
Ile Lys Thr Met Phe Thr Gln Ser Lys Ala Lys Ser Gln Asn Glu Leu
1450 1455 1460

CTA GGA AAA CGT GCA TTG GGT AAC AAT TCA AGG AAT GTA AAA TCA ACA
Leu Gly Lys Arg Ala Leu Gly Asn Asn Ser Arg Asn Val Lys Ser Thr
1465 1470 1475 1480

AGA AAA CTT ACT CCG GCA TGG CAT ATC TTA GAT GTA TCG GGT TAT TAC
Arg Lys Leu Thr Arg Ala Trp His Ile Leu Asp Val Ser Gly Tyr Tyr
1485 1490 1495

ATG GTG AAT AGA AGT ATT TTG TTC CGA TTA GGA GTA TAT AAT TTA TTA
Met Val Asn Arg Ser Ile Leu Phe Arg Leu Gly Val Tyr Asn Leu Leu
1500 1505 1510

AAC TAT CGC TAT GTC ACT TGG GAA GCG GTG CGT CAA ACA GCA CAA GGT
Asn Tyr Arg Tyr Val Thr Trp Glu Ala Val Arg Gln Thr Ala Gln Gly
1515 1520 1525

FIG. 3 Q.

GCG GTC AAT CAA CAT CAA AAT GTT GGT AAC TAT ACT CGC TAC GCA GCA
Ala Val Asn Gln His Gln Asn Val Gly Asn Tyr Thr Arg Tyr Ala Ala
1530 1535 1540

TCA CGA CGA AAC TAT ACC TTA ACA TTA GAA ATG AAA TTC TAA
Ser Gly Arg Asn Tyr Thr Leu Thr Leu Glu Met Lys Phe
1545 1550 1555

FIG. 4 A.

GCCCAAGCTA CATTGGTTAA TCATAAGCCT ATAAATGATA AGAAGAAAT TTGTTTACG
-35
CCATTTTTCA TATTTTATCC ATGGAAGTTAA AAAACTCTAA CTTCACATTAA TTACAAAAA
-10 RBS
AGATCAATTA TGGCAATTAT TATCAATTTT GTATCAGTAT ATAATTCT ATG AAA TCT
Met Lys Ser
1
GTA CCT CTT ATC TCT GGT GGA CTT TCC TTT TTA CTA AGT GCT TGT AGC
5 Val Pro Leu Ile Ser Gly Gly Leu Ser Phe Leu Leu Ser Ala Cys Ser
10
GGA GGG GGG TCT TTT GAT GTA GAT AAC GTC TCT AAT ACC CCC TCT TCT
20 Gly Gly Gly Ser Phe Asp Val Asp Asn Val Ser Asn Thr Pro Ser Ser
25 30 35
AAA CCA CGT TAT CAA GAC GAT ACC TCG AAT CAA AGA AAA AAA TCT AAT
40 Lys Pro Arg Tyr Gln Asp Asp Thr Ser Asn Gln Arg Lys Lys Ser Asn
45 50
TTG AAA AAG TTG TIC ATT OCT TCT TTA GGA GGA GGG ATG AAA TTG GTG
55 Leu Lys Lys Leu Phe Ile Pro Ser Leu Gly Gly Gly Met Lys Leu Val
60 65

FIG.4B.

```

GCT CAG AAT CTT CGT AAT AAA GAA CCT AGT TTC TTA AAT GAA GAT
Ala Gln Asn Leu Arg Gly Asn Lys Glu Pro Ser Phe Leu Asn Glu Asp
70 75 80

GAC TAT ATA TCA TAT TTT TCC TCA CTT TCT ACG ATT GAA AAG GAT GTT
Asp Tyr Ile Ser Tyr Phe Ser Ser Leu Ser Thr Ile Glu Lys Asp Val
85 90 95

AAA GAT AAC AAT AAA AAC GGG GCG GAC CTT ATT GGC TCA ATA GAC GAG
Lys Asp Asn Asn Lys Asn Gly Ala Asp Leu Ile Gly Ser Ile Asp Glu
100 105 110 115

CCT AGT ACA ACA AAT CCA CCC GAA AAG CAT CAT GGA CAA AAA TAT GTA
Pro Ser Thr Thr Asn Pro Pro Glu Lys His His Gly Gln Lys Tyr Val
120 125 130

TAT TCA GGG CTT TAT TAT ACT CCA TCG TGG AGT TTA AAC GAT TCT AAA
Tyr Ser Gly Leu Tyr Tyr Thr Pro Ser Trp Ser Leu Asn Asp Ser Lys
135 140 145

AAC AAG TTT TAT TTA CGT TAC TAT GGA TAT GCG TTT TAT TAT GGT AAT
Asn Lys Phe Tyr Leu Gly Tyr Tyr Gly Tyr Ala Phe Tyr Tyr Gly Asn
150 155 160

```

FIG.4C.

```

AAA ACT GCA ACA AAC TTG CCA GTA AAC GGT GTA GCT AAA TAC AAA GGA
Lys Thr Ala Thr Asn Leu Pro Val Asn Gly Val Ala Lys Tyr Lys Gly
165 170 175

ACT TGG GAT TTC ATC ACT GCA ACT AAA AAT GGC AAA CGT TAT OCT TTG
Thr Trp Asp Phe Ile Thr Ala Thr Lys Asn Gly Lys Arg Tyr Pro Leu
180 185 190 195

TTA AGT AAT GGC AGT CAC GCT TAT TAT CGA CGT AGT GCA ATT CCA GAA
Leu Ser Asn Gly Ser His Ala Tyr Tyr Arg Arg Ser Ala Ile Pro Glu
200 205 210

GAT ATT GAT TTA GAA AAT GAT TCA AAG AAT GGT GAT ATA GGC TTA ATA
Asp Ile Asp Leu Glu Asn Asp Ser Lys Asn Gly Asp Ile Gly Leu Ile
215 220 225

AGT GAA TTT AGT GCA GAT TTT GGG ACT AAA AAA CTG ACA GGA CAA CTG
Ser Glu Phe Ser Ala Asp Phe Gly Thr Lys Lys Leu Thr Gly Gln Leu
230 235 240

TCT TAC ACC AAA AGA AAA ACT AAT AAT CAA CCA TAT GAA AAG AAA AAA
Ser Tyr Thr Lys Arg Lys Thr Asn Asn Gln Pro Tyr Glu Lys Lys Lys
245 250 255

```

FIG. 4D.

```

CTC TAT GAT ATA GAT GCC GAT ATT TAT AGT AAT AGA TTC AGG GGT ACA
Leu Tyr Asp Ile Asp Ala Asp Ile Tyr Ser Asn Arg Phe Arg Gly Thr
260                               270                               275

GTA AAG CCA ACC GAA AAA GAT TCT GAA GAA CAT CCC TTT ACC AGC GAG
Val Lys Pro Thr Glu Lys Asp Ser Glu Glu His Pro Phe Thr Ser Glu
280                               285                               290

GGA ACA TTA GAA GGT GGT TTT TAT GGG CCT AAT GCT GAA GAA CTA GGG
Gly Thr Leu Glu Gly Gly Phe Tyr Glu Pro Asn Ala Glu Glu Leu Gly
295                               300                               305

GGG AAA TTT TTA GCT ACG GAT AAC CGA GTT TTT GGG GTA TTT AGT GCC
Gly Lys Phe Leu Ala Thr Asp Asn Arg Val Phe Gly Val Phe Ser Ala
310                               315                               320

AAA GAA ACG GAA GAA ACA AAA AAG GAA GCG TTA TCC AAG GAA ACC TTA
Lys Glu Thr Glu Glu Thr Lys Lys Glu Ala Leu Ser Lys Glu Thr Leu
325                               330                               335

ATT GAT GGC AAG CTA ATT ACT TTC TCT ACT AAA AAA ACC GAT GCA AAA
Ile Asp Gly Lys Leu Ile Thr Phe Ser Thr Lys Lys Thr Asp Ala Lys
340                               345                               355

```


FIG.4E.

```

ACC AAT GCA ACA ACC AGT ACC GCA GCT AAT ACA ACA ACC GAT ACA ACC
Thr Asn Ala Thr Thr Ser Thr Ala Ala Asn Thr Thr Thr Asp Thr Thr
360
GCC AAT ACA ATA ACC GAT GAA AAA AAC TTT AAG ACG GAA GAT ATA TCA
Ala Asn Thr Ile Thr Asp Glu Lys Asn Phe Lys Thr Glu Asp Ile Ser
375
AGT TTT GGT GAA GCT GAT TAT CTG TTA ATT GAC AAA TAT CCT ATT CCA
Ser Phe Gly Glu Ala Ala Asp Tyr Leu Leu Ile Asp Lys Tyr Pro Ile Pro
390
CTT TTA CCT GAT AAA AAT ACT AAT GAT TTC ATA AGT AGT AAG CAT CAT
Leu Leu Pro Asp Lys Asn Thr Asn Asp Phe Ile Ser Ser Lys His His
405
ACT GTA GGA AAT AAA CGC TAT AAA GTG GAA GCA TGT TGC AGT AAT CTA
Thr Val Gly Asn Lys Lys Arg Tyr Lys Val Glu Ala Cys Cys Ser Asn Leu
420
AGC TAT GTG AAA TTT GGT ATG TAT TAT GAA GAC CCA CTT AAA GAA AAA
Ser Tyr Val Lys Phe Gly Met Tyr Tyr Glu Asp Pro Leu Lys Glu Lys
440

```

FIG.4F.

```

GAA ACA GAA ACA GAA ACA GAA ACA GAA AAA GAA AAA GAA AAA
Glu Thr Glu Thr Glu Thr Glu Thr Glu Lys Asp Lys Glu Lys Glu Lys
455                                     460                                     465

GAA AAA GAC AAA GAC AAA GAA AAA CAA ACG GCG GCA ACG ACC AAC ACT
Glu Lys Asp Lys Asp Lys Glu Lys Glu Thr Ala Ala Thr Thr Asn Thr
470                                     475                                     480

TAT TAT CAA TTC TTA TTA GGT CAC CGT ACT CCC AAG GAC GAC ATA CCT
Tyr Tyr Gln Phe Leu Leu Leu Gly His Arg Thr Pro Lys Asp Asp Ile Pro
485                                     490                                     495

AAA ACA GGA AGT GCA AAA TAT CAT GGT AGT TGG TTT GGT TAT ATT ACT
Lys Thr Gly Ser Ala Lys Tyr His Gly Ser Trp Phe Gly Tyr Ile Thr
500                                     505                                     510                                     515

GAC GGT AAG ACA TCT TAC TCC CCC AGT GGT GAT AAG AAA CCG GAT AAA
Asp Gly Lys Thr Ser Tyr Ser Pro Ser Gly Asp Lys Lys Arg Asp Lys
520                                     525                                     530

AAT GCT GTC GCC GAG TTT AAT GTT GAT TTT GCC GAG AAA AAG CTA ACA
Asn Ala Val Ala Glu Phe Asn Val Asp Phe Ala Glu Lys Lys Leu Thr
535                                     540                                     545

```

FIG. 4 G.

```

GGC GAA TTA AAA CGA CAC GAT ACT GGA AAT CCC GTA TTT AGT ATT GAG
Gly Glu Leu Lys Arg His Asp Thr Gly Asn Pro Val Phe Ser Ile Glu
550
560
GCA AAC TTT AAT AAT AGT AGT AGT AAT GGC TTC ACT GGT ACA GCA ACC GCA
Ala Asn Phe Asn Asn Ser Ser Asn Ala Phe Thr Gly Thr Ala Thr Ala
565
570
575
ACA AAT TTT GTA ATA GAT GGT AAA AAT AGT CAA AAT AAA AAT ACC CCA
Thr Asn Phe Val Ile Asp Gly Lys Asn Ser Gln Asn Lys Asn Thr Pro
580
585
590
595
ATT AAT ATT ACA ACT AAA GTA AAC GGG GCA TTT TAT GGA CCT AAG GCT
Ile Asn Ile Thr Thr Lys Val Asn Gly Ala Phe Tyr Gly Pro Lys Ala
600
605
610
TCT GAA TTA GGC GGT TAT TTC ACT TAT AAC GGA AAT TCT ACA GCT ACA
Ser Glu Leu Gly Tyr Phe Thr Tyr Asn Gly Asn Ser Thr Ala Thr
615
620
625
AAT TCT GAA AGT TCC TCA ACC GTA TCT TCA TCA TCC AAT TCA AAA AAT
Asn Ser Glu Ser Ser Thr Val Ser Ser Ser Ser Asn Ser Lys Asn
630
635
640

```

FIG. 4 H.

GCA AGA GCT GCA GTT GTC TTT GGT GCG AGA CAA CAA GTA GAA ACA ACC
645 Ala Arg Ala Ala Val Val Phe Gly Ala Arg Gln Gln Val Glu Thr Thr 655

AAA TAATGGAATA CTAAAA ATG ACT AAA AAA CCC TAT TTT CGC CTA AGT
Lys Met Thr Lys Lys Pro Tyr Phe Arg Leu Ser 660 670

ATT ATT TCT TGT CTT TTA ATT TCA TGC TAT GTA AAA GCA GAA ACT CAA
675 Ile Ile Ser Cys Leu Leu Ile Ser Cys Tyr Val Lys Ala Glu Thr Gln 685

AGT ATA AAA GAT ACA AAA GAA GCT ATA TCA TCT GAA GTG GAC ACT CAA
690 Ser Ile Lys Asp Thr Lys Glu Ala Ile Ser Ser Glu Val Asp Thr Gln 700

AGT ACA GAA GAT TCA GAA TTA GAA ACT ATC TCA GTC ACT GCA GAA AAA
705 Ser Thr Glu Asp Ser Glu Leu Glu Thr Thr Ile Ser Val Thr Ala Glu Lys 715

ATA AGA GAT CGT AAA GAT AAT GAA GTA ACT GGA CTT GGC AAA ATT ATC
720 Ile Arg Asp Arg Lys Asp Asn Glu Val Thr Gly Leu Gly Lys Ile Ile 730

FIG. 41.

```

AAA ACT AGT GAA AGT ATC AGC CGA GAA CAA GTA TTA AAT ATT CGT GAT
Lys Thr Ser Glu Ser Ile Ser Arg Glu Gln Val Leu Asn Ile Arg Asp 750
735
CTA ACA CGC TAT GAT CCA GCG ATT TCA GTT GTA GAA CAA GGT CGC GGT
Leu Thr Arg Tyr Asp Pro Gly Ile Ser Val Val Glu Gln Gly Arg Gly 765
755
GCA AGT TCT GGA TAT TCT ATT CGT CGT ATG GAC AGA AAT ACA GTT GCT
Ala Ser Ser Gly Tyr Ser Ile Arg Gly Met Asp Arg Asn Arg Val Ala 780
770
TTA TTA GTA GAT GGT TTA CCT CAA ACG CAA TCT TAT GTA GTG CAA AGC
Leu Leu Val Asp Gly Leu Pro Gln Thr Gln Ser Tyr Val Val Gln Ser 795
785
CCT TTA GTT GCT CGT TCA GGA TAT TCT GCG ACT GGT GCA ATT AAT GAA
Pro Leu Val Ala Arg Ser Gly Tyr Ser Gly Thr Gly Ala Ile Asn Glu 810
800
ATT GAA TAT GAA AAT GTA AAG GCC GTC GAA ATA AGC AAG GCG GCG AGT
Ile Glu Tyr Glu Asn Val Lys Ala Val Glu Ile Ser Lys Gly Gly Ser 830
815

```

FIG. 4 J.

```

TCT TCT GAG TAT GGT AAT GGA GCA CTA GCT GGT TCT GTA ACA TTT CAA
Ser Ser Glu Tyr Gly Asn Gly Ala Leu Ala Gly Ser Val Thr Phe Gln
835
AGC AAA TCA GCA GGC GAT ATC TTA GAA GGA GAC AAA TCA TGG GGA ATT
Ser Lys Ser Ala Ala Asp Ile Leu Glu Gly Asp Lys Ser Trp Gly Ile
850
CAA ACT AAA AAT GCT TAT TCA AGC AAA AAT AAA GGC TTT ACC CAT TCT
Gln Thr Lys Asn Ala Tyr Ser Ser Lys Asn Lys Gly Phe Thr His Ser
865
TTA GCT GTA GCA GGA AAA CAA GGT GGA TTT GAA GGG CTA GCC ATT TAC
Leu Ala Val Ala Gly Lys Gln Gly Gly Phe Glu Gly Leu Ala Ile Tyr
880
ACT CAA CGA AAT TCA ATT GAA ACC CAA GTC CAT AAA GAT GCA TTA AAA
Thr Gln Arg Asn Ser Ile Glu Thr Thr Gln Val His Lys Asp Ala Leu Lys
895
GGC GTA CAA AGT TAT GAT CGA TTA ATC GCC ACA ACA GAT AAA TCT TCA
Gly Val Gln Ser Tyr Asp Arg Leu Ile Ala Thr Thr Asp Lys Ser Ser
915

```

FIG.4K.

```

GCA TAC TTT GTG ATA CAA GGT GAG TGT CCA AAT GGT GAT GAC AAG TGT
Gly Tyr Phe Val Ile Gln Gly Glu Cys Pro Asn Gly Asp Asp Lys Cys
930 935 940

GCA GCC AAG CCA CCT GCG ACT TTA TCC ACC CAA AGC GAA ACC GTA AGC
Ala Ala Lys Pro Pro Ala Thr Leu Ser Thr Gln Ser Glu Thr Val Ser
945 950 955

GTT TCA GAT TAT ACG GCG GCT AAC GGT ATC AAA CCT AAT CCA ATG AAA
Val Ser Asp Tyr Thr Gly Ala Asn Arg Ile Lys Pro Asn Pro Met Lys
960 965 970

TAT GAA AGC CAG TCT TGG TTT TTA AGA GGA GGG TAT CAT TTT TCT GAA
Tyr Glu Ser Ser Gln Ser Trp Phe Leu Arg Gly Gly Tyr His Phe Ser Glu
975 980 985 990

CAA CAT TAT ATT GGT GGT ATT TTT GAA TTC ACA CAA CAA AAA TTT CAT
Gln His Tyr Ile Gly Gly Ile Phe Glu Phe Thr Gln Gln Lys Phe Asp
995 1000 1005

ATC CGT GAT ATG ACA TTT CCC GCT TAT TTA AGC CCA ACA GAA AGA CCG
Ile Arg Asp Met Thr Phe Pro Ala Tyr Leu Ser Pro Thr Glu Arg Arg
1010 1015 1020

```

FIG.4L.

```

GAT GAT AGT AGT CGT TCT TTT TAT CCA ATG CAA GAT CAT GGT GCA TAT
Asp Asp Ser Ser Arg Ser Phe Tyr Pro Met Gln Asp His Gly Ala Tyr
1025
CAA CAT ATT GAG GAT GGC AGA GGC GTT AAA TAT GCA AGT GGG CTT TAT
Gln His Ile Glu Asp Gly Arg Gly Val Lys Tyr Ala Ser Gly Leu Tyr
1040
TTC GAT GAA CAC CAT ACA AAA CAG CGT GTA GGT ATT GAA TAT ATT TAC
Phe Asp Glu His His Arg Lys Gln Arg Val Gly Ile Glu Tyr Ile Tyr
1055
GAA AAT AAG AAC AAA GCG GGC ATC ATT GAC AAA GCA GTG TTA AGT GCT
Glu Asn Lys Asn Lys Ala Gly Ile Ile Asp Lys Ala Val Leu Ser Ala
1075
AAT CAA CAA AAC ATC ATA CTT GAC AGT TAT ATG CGA CAT ACG CAT TGC
Asn Gln Gln Asn Ile Ile Leu Asp Ser Tyr Met Arg His Thr His Cys
1090
AGT CTT TAT OCT AAT CCA AGT AAG AAT TGC CGC CCA ACA CTT GAT AAA
Ser Leu Tyr Pro Asn Pro Ser Lys Asn Cys Arg Pro Thr Leu Asp Lys
1105

```


FIG. 4 M.

```

OCT TAT TCA TAC TAT CGT TCT GAT AGA AAT GGT TAT AAA GAA AAA CAT
Pro Tyr Ser Tyr Tyr Arg Ser Asp Arg Asn Val Tyr Lys Glu Lys His
1120                                     1130

AAT ATG TTG CAA TTG AAT TTA GAG AAA AAA AAT CAA AAT TGG CTT
Asn Met Leu Gln Leu Asn Leu Glu Lys Lys Ile Gln Gln Asn Trp Leu
1135                                     1145

ACT CAT CAA ATT GTC TTC AAT CTT GGT TTT GAT GAC TTT ACT TCA GCG
Thr His Gln Ile Val Phe Asn Leu Gly Phe Asp Asp Phe Thr Ser Ala
1155                                     1165

CTT CAG CAT AAA GAT TAT TTA ACT CGA CGT GGT ATC GCT ACG GCA GAT
Leu Gln His Lys Asp Tyr Leu Thr Arg Arg Val Ile Ala Thr Ala Asp
1170                                     1180

AGT ATT CCA AGG AAA CCT GGT GAA ACT GGT AAA CCA AGA AAT GGT TTG
Ser Ile Pro Arg Lys Pro Gly Glu Thr Gly Lys Pro Arg Asn Gly Leu
1185                                     1195

CAA TCA CAA CCT TAC TTA TAC CCA AAA CCA GAG CCA TAT TTT GCA GGA
Gln Ser Gln Pro Tyr Leu Tyr Pro Lys Pro Glu Pro Tyr Phe Ala Gly
1200                                     1205                                     1210

```

FIG. 4N.

[illegible]

FIG.40.

```

TTT TCT GAA ATG TAT GGT TGG CCG TAT GGT GGC AAG AAT GAC GAG GTT
Phe Ser Glu Met Tyr Gly Trp Arg Tyr Gly Gly Lys Asn Asp Glu Val
1315                               1320                               1325

TAT GTA GGT AAA TTT AAG CCT GAA ACA TCT CGT AAC CAA GAG TTT GGT
Tyr Val Gly Lys Phe Lys Pro Glu Thr Ser Arg Asn Gln Glu Phe Gly
1330                               1335                               1340

CTC GCT CTA AAA GGG GAT TTT GGT AAT ATT GAG ATC AGT CAT TTT AGT
Leu Ala Leu Lys Gly Asp Phe Gly Asn Ile Glu Ile Ser His Phe Ser
1345                               1350                               1355

AAT GCT TAT CGA AAT CTT ATC GCC TTT GCT GAA GAA CTT AGT AAA AAT
Asn Ala Tyr Arg Asn Leu Ile Ala Phe Ala Glu Glu Leu Ser Lys Asn
1360                               1365                               1370

CGA ACT CGA AAG GGC AAT TAT CGA TAT CAT AAT GCA CAA AAT GCA AAA
Gly Thr Gly Lys Gly Asn Tyr Gly Tyr His Asn Ala Gln Asn Ala Lys
1375                               1380                               1385                               1390

TTA GTT GGC GTA AAT ATA ACT GCA CAA TTA GAT TTT AAT GGT TTA TGG
Leu Val Gly Val Asn Ile Thr Ala Gln Leu Asp Phe Asn Gly Leu Trp
1395                               1400                               1405

```

FIG. 4 P.

```

AAA CGT ATT CCC TAC GGT TGG TAT GCA ACA TTT GCT TAT AAC CAA GTA
Lys Arg Ile Pro Tyr Gly Trp Tyr Ala Thr Phe Ala Tyr Asn Gln Val
1410
AAA GTT AAA GAT CAA AAA ATC AAT GCT GGT TTA GCC TCC GTA AGC AGT
Lys Val Lys Asp Gln Lys Ile Asn Ala Gly Leu Ala Ser Val Ser Ser
1425
TAT TTA TTT GAT GCC ATT CAG CCC ACC GGT TAT ATC ATT GGT TTA GGC
Tyr Leu Phe Asp Ala Ile Gln Pro Ser Arg Tyr Ile Ile Gly Leu Gly
1440
TAT GAT CAT CCA AGT AAT ACT TGG GGA ATT AAT ACA ATG TTT ACT CAA
Tyr Asp His Pro Ser Asn Thr Trp Gly Ile Asn Thr Met Phe Thr Gln
1455
TCA AAA GCA AAA TCT CAA AAT GAA TTG CTA GGA AAA CGT GCA TTA GGT
Ser Lys Ala Lys Ser Gln Asn Glu Leu Leu Gly Lys Arg Ala Leu Gly
1475
AAC AAT TCA AGG GAT GTA AAA TCA ACA AGA AAA CTT ACT CGG GCA TGG
Asn Asn Ser Arg Asp Val Lys Ser Thr Arg Lys Leu Thr Arg Ala Trp
1490

```

FIG. 4 Q.

CAT ATC TTA GAT GTA TCG GGT TAT TAC ATG GCG AAT AAA AAT ATT ATG
 His Ile Leu Asp Val Ser Gly Tyr Tyr Met Ala Asn Lys Asn Ile Met
 1505 1510 1515

CTT CGA TTA GGG ATA TAT AAT TTA TTC AAC TAT CGC TAT GTT ACT TGG
 Leu Arg Leu Gly Ile Tyr Asn Leu Phe Asn Tyr Arg Tyr Val Thr Trp
 1520 1525 1530

GAA GCG GTG CGT CAA ACA GCA CAA GGT GCG GTC AAT CAA CAT CAA AAT
 Glu Ala Val Arg Gln Thr Ala Gln Gly Ala Val Asn Gln His Gln Asn
 1535 1540 1545 1550

GTT GGT AGC TAT ACT CGC TAC GCA GCA TCA GGA CGA AAC TAT ACC TTA
 Val Gly Ser Tyr Thr Arg Tyr Ala Ala Ser Gly Arg Asn Tyr Thr Leu
 1555 1560 1565

ACA TTA GAA ATG AAA TTC TAAATTAAAA TCGGCCAGAT GGA CTAGATA
 Thr Leu Glu Met Lys Phe
 1570

TGCTATATCT ATACCTTACT GGCGCATCTT TTTCIGTTCT ATAACTTGCT TAAGTGAAAA

ACCAAACTTG GATTTTTTAC AAGATCTTTT CACACATTTA TTG

FIG. 5 A.

ATTGTGTTTA CGCAATTTT CATATTTTAT CCATGAACCTT AAAAAACTCT AACTTGACAT -35
TATTACAAAA AAAGAICAAT AATGCGAATT ATTATCAATT TTGTATGAGT ATATAATTCT -10
RBS
ATG AAA TCT GTA CCT CTT ATC TCT GGT GGA CTT TCC TTT TTA CTA AGT
Met Lys Ser Val Pro Leu Ile Ser Gly Gly Leu Ser Phe Leu Leu Ser 15
1 5 10
GCT TGT AGC GGA GGG GGT TCT TTT GAT GTA GAT AAC GTC TCT AAT ACC
Ala Cys Ser Gly Gly Gly Ser Phe Asp Val Asp Asn Val Ser Asn Thr 20
25 30
CCC TCT TCT AAA CCA CGT TAT CAA GAC GAT ACC TCG AAT CAA AGA AAA
Pro Ser Ser Lys Pro Arg Tyr Gln Asp Asp Thr Ser Asn Gln Arg Lys 35
40 45
AAA TCT AAT TTG AAA AAG TTG TTC ATT OCT TCT TTA GGA GGA GGG ATG
Lys Ser Asn Leu Lys Lys Leu Phe Ile Pro Ser Leu Gly Gly Gly Met 50
55 60
AAA TTG GTG GCT CAG AAT CTT CGT AAT AAA GAA CCT AGT TTC TTA
Lys Leu Val Ala Gln Asn Leu Arg Gly Asn Lys Glu Pro Ser Phe Leu 65
70 75 80

FIG. 5B.

```

AAT GAA GAT GAC TAT ATA TCA TAT TTT TCC TCA CTT TCT ACG ATT GAA
Asn Glu Asp Asp Tyr Ile Ser Tyr Phe Ser Ser Leu Ser Thr Ile Glu
85                                     90                                     95

AAG GAT GTT AAA GAT AAC AAT AAA AAC GCG GCG GAC CTT ATT GGC TCA
Lys Asp Val Lys Asp Asn Asn Lys Asn Gly Ala Asp Leu Ile Gly Ser
100                                     105                                     110

ATA GAC GAG CCT AGT ACA ACA AAT CCA CCC GAA AAG CAT CAT GCA CAA
Ile Asp Glu Pro Ser Thr Thr Asn Pro Pro Glu Lys His His Gly Gln
115                                     120                                     125

AAA TAT GTA TAT TCA GCG CTT TAT TAT ACT CCA TCG TCG AGT TTA AAC
Lys Tyr Val Tyr Ser Gly Leu Tyr Tyr Thr Pro Ser Trp Ser Leu Asn
130                                     135                                     140

GAT TCT AAA AAC AAG TTT TAT TTA GGT TAC TAT GGA TAT GCG TTT TAT
Asp Ser Lys Asn Lys Phe Tyr Leu Gly Tyr Tyr Gly Tyr Ala Phe Tyr
145                                     150                                     155                                     160

TAT GGT AAT AAA ACT GCA ACA AAC TTG CCA GTA AAC GGT GTA GCT AAA
Tyr Gly Asn Lys Thr Ala Thr Asn Leu Pro Val Asn Gly Val Ala Lys
165                                     170                                     175

```

FIG.5C.

```

TAC AAA GGA ACT TGG GAT TTC ATC ACT GCA ACT AAA AAT GGC AAA CGT
Tyr Lys Gly Thr Trp Asp Phe Ile Thr Ala Thr Lys Asn Gly Lys Arg
180
TAT CCT TTG TTA AGT AAT GGC AGT CAC GCT TAT TAT CGA CGT AGT GCA
Tyr Pro Leu Leu Ser Asn Gly Ser His Ala Tyr Tyr Arg Arg Ser Ala
195
ATT CCA GAA GAT ATT GAT TTA GAA AAT GAT TCA AAG AAT GGT GAT ATA
Ile Pro Glu Asp Ile Asp Leu Glu Asn Asp Ser Lys Asn Gly Asp Ile
210
GGC TTA ATA AGT GAA TTT AGT GCA GAT TTT GGG ACT AAA AAA CTG ACA
Gly Leu Ile Ser Glu Phe Ser Ala Asp Phe Gly Thr Lys Lys Leu Thr
225
GGA CAA CTG TCT TAC ACC AAA AGA AAA ACT AAT AAT CAA CCA TAT GAA
Gly Gln Leu Ser Tyr Thr Lys Arg Lys Thr Asn Asn Gln Pro Tyr Glu
245
AAG AAA AAA CTC TAT GAT ATA GAT GCC GAT ATT TAT AGT AAT AGA TTC
Lys Lys Lys Leu Tyr Asp Ile Asp Ala Asp Ile Tyr Ser Asn Arg Phe
260

```


FIG. 5D.

AGG GGT ACA GTA AAG CCA ACC GAA AAA GAT TCT GAA GAA CAT CCC TTT
 Arg Gly Thr Val Lys Pro Thr Glu Lys Asp Ser Glu Glu His Pro Phe
 275 280 285

ACC AGC GAG GGA ACA TTA GAA GGT GGT TTT TAT GCG CCT AAT GCT GAA
 Thr Ser Glu Gly Thr Leu Glu Gly Gly Phe Tyr Gly Pro Asn Ala Glu
 290 295 300

GAA CTA GGG GGG AAA TTT TTA GCT ACG GAT AAC CGA GTT TTT GCG GTA
 Glu Leu Gly Gly Lys Phe Leu Ala Thr Asp Asn Arg Val Phe Gly Val
 305 310 315 320

TTT AGT GCC AAA GAA ACG GAA GAA ACA AAA AAG GAA GCG TTA TCC AAG
 Phe Ser Ala Lys Glu Thr Glu Glu Thr Lys Lys Glu Ala Leu Ser Lys
 320 325 330 335

GAA ACC TTA ATT GAT GGC AAG CTA ATT ACT TTC TCT ACT AAA AAA ACC
 Glu Thr Leu Ile Asp Gly Lys Leu Ile Thr Phe Ser Thr Lys Lys Thr
 340 345 350

GAT GCA AAA ACC AAT GCA ACA ACC AGT ACC GCA GCT AAT ACA ACA ACC
 Asp Ala Lys Thr Asn Ala Thr Thr Ser Thr Ala Ala Asn Thr Thr Thr
 355 360 365

FIG. 5E.

```

GAT ACA ACC GCC AAT ACA ATA ACC GAT GAA AAA AAC TTT AAG ACG GAA
Asp Thr Thr Ala Asn Thr Ile Thr Asp Glu Lys Asn Phe Lys Thr Glu
370                                     380

GAT ATA TCA AGT TTT GGT GAA GCT GAT TAT CTG TTA ATT GAC AAA TAT
Asp Ile Ser Ser Phe Gly Glu Ala Asp Tyr Leu Leu Ile Asp Lys Tyr
385                                     390                                     400

CCT ATT CCA CTT TTA OCT GAT AAA AAT ACT AAT GAT TTC ATA AGT AGT
Pro Ile Pro Leu Leu Pro Asp Lys Asn Thr Asn Asp Phe Ile Ser Ser
405                                     410                                     415

AAG CAT CAT ACT GTA GGA AAT AAA CGC TAT AAA GTG GAA GCA TGT TGC
Lys His His Thr Val Gly Asn Lys Arg Tyr Lys Val Glu Ala Cys Cys
420                                     425                                     430

AGT AAT CTA AGC TAT GTG AAA TTT GGT AUG TAT TAT GAA GAC CCA CTT
Ser Asn Leu Ser Tyr Val Lys Phe Gly Met Tyr Tyr Glu Asp Pro Leu
435                                     440                                     445

AAA GAA AAA GAA ACA GAA ACA GAA ACA GAA ACA GAA AAA GAC AAA GAA
Lys Glu Lys Glu Thr Glu Thr Glu Thr Glu Thr Glu Lys Asp Lys Glu
450                                     455                                     460

```

FIG.5F.

```

AAA GAA AAA GAA AAA GAC AAA GAC AAA GAA AAA CAA ACG GCG GCA ACG
Lys Glu Lys Lys Glu Lys Asp Lys Asp Lys Glu Lys Gln Thr Ala Ala Thr
465                               470                               475                               480

ACC AAC ACT TAT TAT CAA TTC TTA TTA GGT CAC CGT ACT CCC AAG GAC
Thr Asn Thr Tyr Tyr Gln Phe Leu Leu Gly His Arg Thr Pro Lys Asp
485                               490                               495

GAC ATA CCT AAA ACA GGA AGT GCA AAA TAT CAT GGT AGT TGG TTT GGT
Asp Ile Pro Lys Thr Gly Ser Ala Lys Tyr His Gly Ser Trp Phe Gly
500                               505                               510

TAT ATT ACT GAC GGT AAG ACA TCT TAC TCC CCC AGT GGT GAT AAG AAA
Tyr Ile Thr Asp Gly Lys Thr Ser Tyr Ser Pro Ser Gly Asp Lys Lys
515                               520                               525

CGC GAT AAA AAT GCT GTC GCC GAG TTT AAT GTT GAT TTT GCC GAG AAA
Arg Asp Lys Asn Ala Val Ala Glu Phe Asn Val Asp Phe Ala Glu Lys
530                               535                               540

AAG CTA ACA GGC GAA TTA AAA CGA CAC GAT ACT CGA AAT CCC GTA TTT
Lys Leu Thr Gly Glu Leu Lys Arg His Asp Thr Gly Asn Pro Val Phe
545                               550                               555                               560

```

FIG. 56.

```

AGT ATT GAG GCA AAC TTT AAT AAT AGT AGT AGT GGC TTC ACT GGT ACA
Ser Ile Glu Ala Asn Phe Asn Asn Ser Ser Asn Ala Phe Thr Gly Thr
565 570 575

GCA ACC GCA ACA AAT TTT GTA ATA GAT GGT AAA AAT AGT CAA AAT AAA
Ala Thr Ala Thr Asn Phe Val Ile Asp Gly Lys Asn Ser Gln Asn Lys
580 585 590

AAT ACC CCA ATT AAT ATT ACA ACT AAA GTA AAC GGG GCA TTT TAT GCA
Asn Thr Pro Ile Asn Ile Thr Thr Lys Val Asn Gly Ala Phe Tyr Gly
595 600 605

CCT AAG GCT TCT GAA TTA GGC GGT TAT TTC ACT TAT AAC GGA AAT TCT
Pro Lys Ala Ser Glu Leu Gly Gly Tyr Phe Thr Tyr Asn Gly Asn Ser
610 615 620

ACA GCT ACA AAT TCT GAA AGT TCC TCA ACC GTA TCT TCA TCA TCC AAT
Thr Ala Thr Asn Ser Glu Ser Ser Ser Thr Val Ser Ser Ser Asn
625 630 635 640

TCA AAA AAT GCA ACA GCT GCA GTT GTC TTT GGT GCG AGA CAA CAA GTA
Ser Lys Asn Ala Arg Ala Ala Val Val Phe Gly Ala Arg Gln Gln Val
645 650 655

```

FIG. 5H.

```

GAA ACA ACC AAA TAA TGA ATA CTAAA ATG ACT AAA AAA CCC TAT TTT
Glu Thr Thr Lys                               Met Thr Lys Lys Pro Tyr Phe
660
CGC CTA AGT ATT ATT TCT TGT CTT TTA ATT TCA TGC TAT GTA AAA GCA
Arg Leu Ser Ile Ile Ser Cys Leu Leu Ile Ser Cys Tyr Val Lys Ala
670
GAA ACT CAA AGT ATA AAA GAT ACA AAA GAA GCT ATA TCA TCT GAA GTG
Glu Thr Gln Ser Ile Lys Asp Thr Lys Glu Ala Ile Ser Ser Glu Val
685
GAC ACT CAA AGT ACA GAA GAT TCA GAA TTA GAA ACT ATC TCA GTC ACT
Asp Thr Gln Ser Thr Glu Asp Ser Glu Leu Glu Thr Ile Ser Val Thr
700
GCA GAA AAA ATA AGA GAT CGT AAA GAT AAT GAA GTA ACT GCA CTT GGC
Ala Glu Lys Ile Arg Asp Arg Lys Asp Asn Glu Val Thr Gly Leu Gly
720
AAA ATT ATC AAA ACT AGT GAA AGT ATC AGC CGA GAA CAA GTA TTA AAT
Lys Ile Ile Lys Thr Ser Glu Ser Ile Ser Arg Glu Gln Val Leu Asn
735

```

FIG. 51.

```

ATT CGT GAT CTA ACA CGC TAT GAT CCA GCG ATT TCA GTT GTA GAA CAA
Ile Arg Asp Leu Thr Arg Tyr Asp Pro Gly Ile Ser Val Val Glu Gln
750
GGT CGC GGT GCA AGT TCT CGA TAT TCT ATT CGT GGT ATG GAC AGA AAT
Gly Arg Gly Ala Ser Ser Gly Tyr Ser Ile Arg Gly Met Asp Arg Asn
765
AGA GTT GCT TTA TTA GTA GAT GGT TTA CCT CAA ACG CAA TCT TAT GTA
Arg Val Ala Leu Leu Val Asp Gly Leu Pro Gln Thr Ser Tyr Val
780
GTG CAA AGC CCT TTA GTT GCT CGT TCA CGA TAT TCT GGC ACT GGT GCA
Val Gln Ser Pro Leu Val Ala Arg Ser Gly Tyr Ser Gly Thr Gly Ala
800
ATT AAT GAA ATT GAA TAT GAA AAT GTA AAG GCC GTC GAA ATA AGC AAG
Ile Asn Glu Ile Glu Tyr Glu Asn Val Lys Ala Val Glu Ile Ser Lys
815
GGG GGG AGT TCT TCT GAG TAT GGT AAT GGA GCA CTA GCT GGT TCT GTA
Gly Gly Ser Ser Ser Glu Tyr Gly Asn Gly Ala Leu Ala Gly Ser Val
830

```

FIG. 5 J.

ACA TTT CAA AGC AAA TCA GCA GCC GAT ATC TTA GAA GGA GAC AAA TCA
 Thr Phe Gln Ser Lys Ser Ala Ala Asp Ile Leu Glu Gly Asp Lys Ser
 845 850 855
 TGG GGA ATT CAA ACT AAA AAT GCT TAT TCA AGC AAA AAT AAA GGC TTT
 Trp Gly Ile Gln Thr Lys Asn Ala Tyr Ser Ser Lys Asn Lys Gly Phe
 860 865 870 875
 ACC CAT TCT TTA GCT GTA GCA GGA AAA CAA GGT GGA TTT GAA GGG CTA
 Thr His Ser Leu Ala Val Ala Gly Lys Gln Gly Gly Phe Glu Gly Leu
 880 885 890
 GCC ATT TAC ACT CAA CGA AAT TCA ATT GAA ACC CAA GTC CAT AAA GAT
 Ala Ile Tyr Thr Gln Arg Asn Ser Ile Glu Thr Gln Val His Lys Asp
 895 900 905
 GCA TTA AAA GGC GTA CAA AGT TAT GAT CGA TTA ATC GCC ACA ACA GAT
 Ala Leu Lys Gly Val Gln Ser Tyr Asp Arg Leu Ile Ala Thr Thr Asp
 910 915 920
 AAA TCT TCA GGA TAC TTT GTG ATA CAA GGT GAG TGT CCA AAT GGT GAT
 Lys Ser Ser Gly Tyr Phe Val Ile Gln Gly Glu Cys Pro Asn Gly Asp
 925 930 935

FIG. 5K.

GAC AAG TGT CCA GCC AAG CCA CCT CCG ACT TTA TCC ACC CAA AGC GAA
 Asp Lys Cys Ala Ala Lys Pro Pro Ala Thr Leu Ser Thr Gln Ser Glu
 940 945 950 955

ACC GTA AGC GTT TCA GAT TAT TAT ACG GCG GCT AAC CGT ATC AAA CCT AAT
 Thr Val Ser Val Ser Asp Tyr Thr Gly Ala Asn Arg Ile Lys Pro Asn
 960 965 970

CCA ATG AAA TAT GAA AGC CAG TCT TGG TTT TTA AGA GGA GGG TAT CAT
 Pro Met Lys Tyr Glu Ser Gln Ser Trp Phe Leu Arg Gly Gly Tyr His
 975 980 985

TTT TCT GAA CAA CAT TAT ATT GGT GGT ATT TTT GAA TTC ACA CAA CAA
 Phe Ser Glu Gln His Tyr Ile Gly Gly Ile Phe Glu Phe Thr Gln Gln
 990 995 1000

AAA TTT GAT ATC CGT GAT ATG ACA TTT CCC GCT TAT TTA AGC CCA ACA
 Lys Phe Asp Ile Arg Asp Met Thr Phe Pro Ala Tyr Leu Ser Pro Thr
 1005 1010 1015

GAA AGA CCG GAT GAT AGT AGT CGT TCT TTT TAT CCA ATG CAA GAT CAT
 Glu Arg Arg Asp Asp Ser Ser Arg Ser Phe Tyr Pro Met Gln Asp His
 1020 1025 1030 1035

FIG. 5L.

GGT GCA TAT CAA CAT ATT GAG GAT GGC AGA GGC GTT AAA TAT GCA AGT
Gly Ala Tyr Gln His Ile Glu Asp Gly Arg Gly Val Lys Tyr Ala Ser
1040 1045 1050

GGG CTT TAT TTC GAT GAA CAC CAT AGA AAA CAG CGT GTA GGT ATT GAA
Gly Leu Tyr Phe Asp Glu His His Arg Lys Gln Arg Val Gly Ile Glu
1055 1060 1065

TAT ATT TAC GAA AAT AAG AAC AAA GCG GGC ATC ATT GAC AAA GCA GTC
Tyr Ile Tyr Glu Asn Lys Asn Lys Ala Gly Ile Ile Asp Lys Ala Val
1070 1075 1080

TTA AGT GCT AAT CAA CAA AAC ATC ATA CTT GAC AGT TAT ATG CGA CAT
Leu Ser Ala Asn Gln Gln Asn Ile Ile Leu Asp Ser Tyr Met Arg His
1085 1090 1095

ACG CAT TGC AGT CTT TAT CCT AAT CCA AGT AAG AAT TGC CGC CCA ACA
Thr His Cys Ser Leu Tyr Pro Asn Pro Ser Lys Asn Cys Arg Pro Thr
1100 1105 1110 1115

CTT GAT AAA CCT TAT TCA TAC TAT CGT TCT GAT AGA AAT GTT TAT AAA
Leu Asp Lys Pro Tyr Ser Tyr Tyr Arg Ser Asp Arg Asn Val Tyr Lys
1120 1125 1130

FIG. 5M.

```

GAA AAA CAT AAT ATG CAA TTG AAT TTA GAG AAA AAA ATT CAA CAA
Glu Lys His Asn Met Leu Gln Leu Asn Leu Glu Lys Lys Ile Gln Gln
1135                               1140                               1145

AAT TGG CTT ACT CAT CAA AAT GTC TTC AAT CTT GGT TTT GAT GAC TTT
Asn Trp Leu Thr His Gln Ile Val Phe Asn Leu Gly Phe Asp Asp Phe
1150                               1155                               1160

ACT TCA GCG CTT CAG CAT AAA GAT TAT TTA ACT CGA CGT GGT ATC GCT
Thr Ser Ala Leu Gln His Lys Asp Tyr Leu Thr Arg Arg Val Ile Ala
1165                               1170                               1175

ACG GCA GAT AGT ATT CCA AGG AAA CCT GGT GAA ACT GGT AAA CCA AGA
Thr Ala Asp Ser Ile Pro Arg Lys Pro Gly Glu Thr Gly Lys Pro Arg
1180                               1185                               1190                               1195

AAT GGT TTG CAA TCA CAA CCT TAC TTA TAC CCA AAA CCA GAG CCA TAT
Asn Gly Leu Gln Ser Gln Pro Tyr Leu Tyr Pro Lys Pro Glu Pro Tyr
1200                               1205                               1210

TTT GCA GGA CAA GAT CAT TGT AAT TAT CAA GGT AGC TCC TCT AAT TAC
Phe Ala Gly Gln Asp His Cys Asn Tyr Gln Gly Ser Ser Ser Asn Tyr
1215                               1220                               1225

```

FIG. 5N.

```

AGA GAC TGT AAA GTG CCG TTA ATT AAA GCG AAA AAT TAT TAT TTC GCA
Arg Asp Cys Lys Val Arg Leu Ile Lys Gly Lys Asn Tyr Tyr Phe Ala
1230                                     1235
GCA CCG AAT AAT ATG GCA TTA GCG AAA TAC GTT GAT TTA GGT TTA GGT
Ala Arg Asn Asn Met Ala Leu Gly Lys Tyr Val Asp Leu Gly Leu Gly
1245                                     1250
ATT CCG TAT GAC GTA TCT CGT ACA AAA GCT AAT GAA TCA ACT ATT AGT
Ile Arg Tyr Asp Val Ser Arg Thr Lys Lys Ala Asn Glu Ser Thr Ile Ser
1260                                     1265
GTT GGT AAA TTT AAA AAT TTC TCT TCG AAT ACT GGT ATT GTC ATA AAA
Val Gly Lys Phe Lys Asn Phe Ser Trp Asn Thr Gly Ile Val Ile Lys
1280                                     1285
CCA ACG GAA TCG CTT GAT CTT TCT TAT CCG CTT TCT ACT GGA TTT AGA
Pro Thr Glu Trp Leu Asp Leu Ser Tyr Arg Leu Ser Thr Gly Phe Arg
1295                                     1300
AAT CCT AGT TTT TCT GAA ATG TAT GGT TGG CCG TAT GGT GGC AAG AAT
Asn Pro Ser Phe Ser Glu Met Tyr Gly Trp Arg Tyr Gly Gly Lys Asn
1310                                     1315

```

FIG. 50.

GAC GAG GTT TAT GTA GGT AAA TTT AAG CCT GAA ACA TCT CGT AAC CAA
 Asp Glu Val Tyr Val Gly Lys Phe Lys Pro Glu Thr Ser Arg Asn Gln
 1325 1330 1335
 GAG TTT GGT CTC GCT CTA AAA GGG GAT TTT GGT AAT ATT GAG ATC AGT
 Glu Phe Gly Leu Ala Leu Lys Gly Asp Phe Gly Asn Ile Glu Ile Ser
 1340 1345 1350 1355
 CAT TTT AGT AAT GCT TAT CGA AAT CTT ATC GCC TTT GCT GAA GAA CTT
 His Phe Ser Asn Ala Tyr Arg Asn Leu Ile Ala Phe Ala Glu Glu Leu
 1360 1365 1370
 AGT AAA AAT CGA ACT CGA AAG GGC AAT TAT TAT GCA TAT CAT AAT GCA CAA
 Ser Lys Asn Gly Thr Gly Lys Gly Asn Tyr Gly Tyr His Asn Ala Gln
 1375 1380 1385
 AAT GCA AAA TTA GTT GGC GTA AAT ATA ACT GCA CAA TTA GAT TTT AAT
 Asn Ala Lys Leu Val Gly Val Asn Ile Thr Ala Gln Leu Asp Phe Asn
 1390 1395 1400
 GGT TTA TGG AAA CGT ATT CCC TAC GGT TCG TAT GCA ACA TTT GCT TAT
 Gly Leu Trp Lys Arg Ile Pro Tyr Gly Trp Tyr Ala Thr Phe Ala Tyr
 1405 1410 1415

FIG. 5P.

AAC CAA GTA AAA GTT AAA GAT CAA AAA ATC AAT GCT GGT TTA GCC TCC
Asn Gln Val Lys Val Lys Asp Gln Lys Ile Asn Ala Gly Leu Ala Ser
1420 1425 1430 1435

GTA AGC AGT TAT TTA TTT GAT GCC ATT CAG CCC AGC GGT TAT ATC ATT
Val Ser Ser Tyr Leu Phe Asp Ala Ile Gln Pro Ser Arg Tyr Ile Ile
1440 1445 1450

GGT TTA GGC TAT GAT CAT CCA AGT AAT ACT TGG GCA ATT AAT ACA ATG
Gly Leu Gly Tyr Asp His Pro Ser Asn Thr Trp Gly Ile Asn Thr Met
1455 1460 1465

TTT ACT CAA TCA AAA GCA AAA TCT CAA AAT GAA TTG CTA GGA AAA CGT
Phe Thr Gln Ser Lys Ala Lys Ser Gln Asn Glu Leu Leu Gly Lys Arg
1470 1475 1480

GCA TTA GGT AAC AAT TCA AGG GAT GTA AAA TCA ACA AGA AAA CTT ACT
Ala Leu Gly Asn Asn Ser Arg Asp Val Lys Ser Thr Arg Lys Leu Thr
1485 1490 1495

CGG GCA TGG CAT ATC TTA GAT GTA TCG GGT TAT TAC ATG GCG AAT AAA
Arg Ala Trp His Ile Leu Asp Val Ser Gly Tyr Tyr Met Ala Asn Lys
1500 1505 1510 1515

FIG. 5Q.

AAT ATT ATG CTT CGA TTA GCG ATA TAT AAT TTA TTC AAC TAT CGC TAT
Asn Ile Met Leu Arg Leu Gly Ile Tyr Asn Leu Phe Asn Tyr Arg Tyr 1530
1520

GTT ACT TCG GAA GCG GTG CGT CAA ACA GCA CAA CGT GCG GTC AAT CAA
Val Thr Trp Glu Ala Val Arg Gln Thr Ala Gln Gly Ala Val Asn Gln 1545
1535

CAT CAA AAT GTT GGT AGC TAT ACT CGC TAC GCA GCA TCA GGA CGA AAC
His Gln Asn Val Gly Ser Tyr Thr Arg Tyr Ala Ala Ser Gly Arg Asn 1560
1550

TAT ACC TTA ACA TTA GAA ATG AAA TTC TAAATTAAAA TGGGCCAGAT
Tyr Thr Leu Thr Leu Glu Met Lys Phe 1570
1565

GGACTAGATA TGCTATATCT ATACCTTACT GGCGCATCTT TTCTGTGTCT ATAATCTGCT

TAAGTGA AAA ACCAAACTTG GATTTTATAC AAGATCTTTT CACACATTTA TTGTAAATC

TCCGACAATT TTGACCG

FIG. 6 A.

AAAATTCGGT AATGATAACC CTATAAATGA TAAAGAGAGAA AGTTGTTTTTAT GGCATTTTTT
CATATTTTAT CCATGAACTT AAAAAATTCCT AAGTTGACAT TATTACAAAA AAAGAACAAT
AATCGAATT ATTATCAATT TTGTATAAGT ATTAATTCCT ATG AAA TCT GTA CCT
Met Lys Ser Val Pro
1 5

CTT ATC ACT CGT GGA CTT TCC TTT TTA CTA AGC GCT TGT AGC GCG GGA
Leu Ile Thr Gly Gly Leu Ser Phe Leu Leu Ser Ala Cys Ser Gly Gly
10 15 20

GGT GGT TCT TTT GAT GTA GAT GAC GTC TCT AAT OCC TOC TCT TCT AAA
Gly Gly Ser Phe Asp Val Asp Asp Val Ser Asn Pro Ser Ser Ser Lys
25 30 35

CCA CGT TAT CAA GAC GAT ACC TCG AAT CAA AGA ACA AAA TCT GAT TTG
Pro Arg Tyr Gln Asp Asp Thr Ser Asn Gln Arg Thr Lys Ser Asp Leu
40 45 50

GAA AAG TTG TTC ATT OCT TCT TTA GCG GGA GCG ATG AAG TTA GTG GCT
Glu Lys Leu Phe Ile Pro Ser Leu Gly Gly Gly Met Lys Leu Val Ala
55 60 65

FIG.6B.

```

CAA AAT TTT ATT GGT GCT AGA GAA CCT AGT TTC TTA AAT GAA GAT GGC
Gln Asn Phe Ile Gly Ala Arg Glu Pro Ser Phe Leu Asn Glu Asp Gly
70 75 80 85

TAT ATG ATA TTT TCC TCA CTT TCT ACG ATT GAA GAG GAT GTT GAA AAA
Tyr Met Ile Phe Ser Ser Leu Ser Thr Ile Glu Glu Asp Val Glu Lys
90 95 100

GTT AAA AAT AAC AAT AAA AAC GGG GGG AGG CTT ATT GGC TCA ATT GAG
Val Lys Asn Asn Lys Lys Asn Gly Gly Arg Leu Ile Gly Ser Ile Glu
105 110 115

GAA CCT AAT GGA ACA TCA CAA AAT TCT AAT TCA CAA GAA TAC GTT TAT
Glu Pro Asn Gly Thr Ser Gln Asn Ser Asn Ser Gln Glu Tyr Val Tyr
120 125 130

TCT GGT TTG TAT TAT ATC GAT GAT AGT TCG CGT GAT TAT AAG AAG GAA GAG
Ser Gly Leu Tyr Tyr Ile Asp Ser Tyr Arg Asp Tyr Lys Lys Glu Glu
135 140 145

CAA AAA GCT TAT ACT GGC TAT TAT GGT TAT GCA TTT TAT TAT GGT AAT
Gln Lys Ala Tyr Thr Gly Tyr Tyr Gly Tyr Ala Phe Tyr Tyr Gly Asn
150 155 160 165

```


FIG. 6C.

```

GAA ACT GCA AAA AAC TTG CCA GTA AAA GGT GTA GCT AAA TAC AAA GGA
Glu Thr Ala Lys Asn Leu Pro Val Lys Gly Val Ala Lys Tyr Lys Gly
170 175 180

ACG TGG AAC TTC ATC ACT GCA ACT GAA AAT GGC AAA CGT TAT TCT TTG
Thr Trp Asn Phe Ile Thr Ala Thr Glu Asn Gly Lys Arg Tyr Ser Leu
185 190 195

TTC AGT AAT TCT ATC GGT CAA GCT TAT TCC AGA CGC AGC GCT ATT TCA
Phe Ser Asn Ser Ile Gly Gln Ala Tyr Ser Arg Arg Ser Ala Ile Ser
200 205 210

GAA GAT ATC TAT AAT TTA GAA AAC GGT GAC GCG GGC TTA ATA AGT GAA
Glu Asp Ile Tyr Asn Leu Glu Asn Gly Asp Ala Gly Leu Ile Ser Glu
215 220 225

TTT AGT GTA GAT TTT GGT AAG AAA GAG CTC ACT CGA GAA CTT TAT TAT
Phe Ser Val Asp Phe Gly Lys Lys Glu Leu Thr Gly Glu Leu Tyr Tyr
230 235 240 245

AAT GAA AGG AAA ACA AGT GTT AAT GAA TCA CAA AAT ACA ACA CAT AAA
Asn Glu Arg Lys Thr Ser Val Asn Glu Ser Gln Asn Thr Thr His Lys
250 255 260

```

FIG. 6D.

CTC TAC ACT CTA GAA GCT AAA GIG TAT AGC AAC CGA TTC AGA GGT AAA
 Leu Tyr Thr Leu Glu Ala Lys Val Tyr Ser Asn Arg Phe Arg Gly Lys 275
 265
 GTA AAG CCA ACC AAA ACA AAG TCT GAA GAT CAT CCC TTT ACC AGC GAG
 Val Lys Pro Thr Lys Thr Lys Ser Glu Asp His Pro Phe Thr Ser Glu 290
 280
 GGA ACA TTA GAA GGT GGT TTT TAT GGG CCT AAT GCT GAA GAA CTA GGG
 Gly Thr Leu Glu Gly Gly Phe Tyr Gly Pro Asn Ala Glu Glu Leu Gly 305
 295
 GGA AAG TTT TTA GCT AAC GAC GAA AAA GGT TTT GGG GTA TTT AGT GCC
 Gly Lys Phe Leu Ala Asn Asp Glu Lys Val Phe Gly Val Phe Ser Ala 325
 310
 AAA GAA GAC CCA CAA AAC CCA GAA AAC CAA AAA TTA TCC ACA GAA ACC
 Lys Glu Asp Pro Gln Asn Pro Glu Asn Gln Lys Leu Ser Thr Glu Thr 340
 330
 TTA ATT GAT GGC AAG CTA ATT ACT TTT AAA AGA ACT GAT GCA ACA ACC
 Leu Ile Asp Gly Lys Leu Ile Thr Phe Lys Arg Thr Asp Ala Thr Thr 355
 345

FIG.6E.

```

AAT GCA ACA ACC GAT GCA AAA ACC AGT GCA ACA ACC GAT GCA ACC AGT
Asn Ala Thr Thr Asp Ala Lys Thr Ser Ala Thr Thr Asp Ala Thr Ser
360
ACA ACA GCC AAT AAA AAA ACC GAT GCA GAA AAC TTT AAG ACG GAA GAT
Thr Thr Ala Asn Lys Lys Thr Asp Ala Glu Asn Phe Lys Thr Glu Asp
375
ATA CCA AGT TTT GGT GAA GCT GAT TAC CTT TTA ATT GGC AAT CAG CCT
Ile Pro Ser Phe Gly Glu Ala Asp Tyr Leu Leu Ile Gly Asn Gln Pro
390
ATT CCT CTT TTA CCT GAA AAA AAT ACT GAT GAT TTC ATA AGT AGT AAG
Ile Pro Leu Leu Pro Glu Lys Asn Thr Asp Asp Phe Ile Ser Ser Lys
410
CAC CAT ACG GTA GGA GGT AAA ACC TAT AAA GTA GAA GCA TGT TGC AAG
His His Thr Val Gly Gly Lys Thr Tyr Lys Val Glu Ala Cys Cys Lys
425
AAT CTA ACG TAT GIG AAA TTT GGT ATG TAT TAT GAG GAT AAA GAT AAG
Asn Leu Ser Tyr Val Lys Phe Gly Met Tyr Tyr Glu Asp Lys Asp Lys
440

```

FIG. 6F.

GAC AAC AAA AAT GAA ACA GAC AAA GAA AAA GGC AAA GAA AAA CCA ACG
Asp Asn Lys Asn Glu Thr Asp Lys Glu Lys Gly Lys Glu Lys Pro Thr
455 460 465

ACG ACA ACA TCT ATC AAC ACT TAT TAT CAA TTC TTA TTA GGT CTC CGT
Thr Thr Thr Ser Ile Asn Thr Tyr Tyr Gln Phe Leu Leu Gly Leu Arg
470 475 480 485

ACT CCC AAG GAC GAA ATA CCT AAA GAA GGA AGT GCA AAA TAT CAT GGT
Thr Pro Lys Asp Glu Ile Pro Lys Glu Gly Ser Ala Lys Tyr His Gly
490 495 500

AAT TGG TTT GGT TAT ATT AGT GAT GGC GAG ACA TCT TAC TCC GCC AGT
Asn Trp Phe Gly Tyr Ile Ser Asp Gly Glu Thr Ser Tyr Ser Ala Ser
505 510 515

GGT GAT AAG GAA CGC AGT AAA AAT GCT GTC GCC GAG TTT GAT GTA AGT
Gly Asp Lys Glu Arg Ser Lys Asn Ala Val Ala Glu Phe Asp Val Ser
520 525 530

TTT GCC AAT AAA ACA TTA ACA GGC GAA TTA AAA CGA CAC GAT AAT CGA
Phe Ala Asn Lys Thr Leu Thr Gly Glu Leu Lys Arg His Asp Asn Gly
535 540 545

FIG.66.

```

AAT ACC GTA TTT AAA ATT AAT GCA GAA TTA AAT GGT AGT AAT GAC TTC
Asn Thr Val Phe Lys Ile Asn Ala Glu Leu Asn Gly Ser Asn Asp Phe 565
550
ACT GGT ACA GCA ACC GCA ACA AAT TTT GTA ATA GAT GGT AAC AAT AGT
Thr Gly Thr Ala Thr Ala Thr Asn Phe Val Ile Asp Gly Asn Asn Ser 580
570
CAA ACT TCA AAT GCC AAA ATT AAT ATT ACA ACT AAA GTA AAT GGG GCA
Gln Thr Ser Asn Ala Lys Ile Asn Ile Thr Thr Lys Val Asn Gly Ala 595
585
TTT TAT GGA CCT AAG GCT TCT GAA TTA GGA GGG TAT TTC ACC TAT AAC
Phe Tyr Gly Pro Lys Ala Ser Glu Leu Gly Gly Tyr Phe Thr Tyr Asn 610
600
GGA AAA AAT CCT ACA GCT ACA AAT TCT GAA AGT TCC TCA ACC GTA CCT
Gly Lys Asn Pro Thr Ala Thr Asn Ser Glu Ser Ser Thr Val Pro 625
615
TCA CCA CCC AAT TCA CCA AAT GCA AGC GCT GCA GTT GTC TTT GGT GCT
Ser Pro Pro Asn Ser Pro Asn Ala Ser Ala Ala Val Val Phe Gly Ala 645
630
635
640

```

FIG. 6H.

AAA AAA CAA GTA GAA ACA ACC AAC AAG TAAACAAC CAAGTAATGG
 Lys Lys Gln Val Glu Thr Thr Asn Lys 650

AATACTAAAA ATG ACT AAA AAA CCC TAT TTT CGC CTA AGT ATT ATT TCT
 Met Thr Lys Lys Lys Pro Tyr Phe Arg Leu Ser Ile Ile Ser 655
 660 665

TGT CTT TTA ATT TCA TGC TAT GTA AAA GCA GAA ACT CAA AGT ATA AAA
 Cys Leu Leu Ile Ser Cys Tyr Val Lys Ala Glu Thr Gln Ser Ile Lys 670
 675 680

GAT ACA AAA GAA GCT ATA TCA TCT GAA GTG GAC ACT CAA AGT ACA GAA
 Asp Thr Lys Lys Glu Ala Ile Ser Ser Glu Val Asp Thr Gln Ser Thr Glu 685
 690 695

GAT TCA GAA TTA GAA ACT ATC TCA GTC ACT GCA GAA AAA ATA AGA GAT
 Asp Ser Glu Leu Glu Thr Thr Ile Ser Val Thr Ala Glu Lys Ile Arg Asp 700
 705 710 715

CGT AAA GAT AAT GAA GTA ACT GGA CTT GGC AAA ATT ATC AAA ACT AGT
 Arg Lys Asp Asn Glu Val Thr Gly Leu Gly Lys Ile Ile Lys Thr Ser 720
 725 730

FIG. 61.

GAA AGT ATC AGC CGA GAA CAA GTA TTA AAT ATT CGT GAT CTA ACA CGC
 Glu Ser Ile Ser Arg Glu Gln Val Leu Asn Ile Arg Asp Leu Thr Arg
 735 740 745
 TAT GAT CCA GGC ATT TCA GTT GTA GAA CAA GGC CGT GGT GCA AGT TCT
 Tyr Asp Pro Gly Ile Ser Val Val Glu Gln Gly Arg Gly Ala Ser Ser
 750 755 760
 GGA TAT TCT ATT CGT CGT ATG GAC AGA AAT AGA GTT GCT TTA TTA GTA
 Gly Tyr Ser Ile Arg Gly Met Asp Arg Asn Arg Val Ala Leu Leu Val
 765 770 775
 GAT GGT TTA CCT CAA ACG CAA TCT TAT GTA GTG CAA AGC CCT TTA GTT
 Asp Gly Leu Pro Gln Thr Gln Ser Tyr Val Val Gln Ser Pro Leu Val
 780 785 790 795
 GCT CGT TCA GGA TAT TCT GGC ACT GGT GCA ATT AAT GAA ATT GAA TAT
 Ala Arg Ser Gly Tyr Ser Gly Thr Gly Ala Ile Asn Glu Ile Glu Tyr
 800 805 810
 GAA AAT GTA AAG GCC GTC GAA ATA AGC AAG GCG GCG AGT TCT TCT GAG
 Glu Asn Val Lys Ala Val Glu Ile Ser Lys Gly Gly Ser Ser Ser Glu
 815 820 825

FIG.6J.

TAT GGT AAT GGA GCA CTA GCT GGT TCT GTA ACA TTT CAA AGC AAA TCA
 Tyr Gly Asn Gly Ala Leu Ala Gly Ser Val Thr Phe Gln Ser Lys Ser
 830 840

GCA GCC GAT ATC TTA GAA GGA GAC AAA TCA TGG CGA ATT CAA ACT AAA
 Ala Ala Asp Ile Leu Glu Gly Asp Lys Ser Trp Gly Ile Gln Thr Lys
 845 855

AAT GCT TAT TCA AGC AAA AAT AAA GGC TTT ACC CAT TCT TTA GCT GTA
 Asn Ala Tyr Ser Ser Lys Asn Lys Gly Phe Thr His Ser Leu Ala Val
 860 875

GCT GGA AAA CAA GGG CGA TTT GAC GGG GTC GCC AAT TAT ACT CAA CGA
 Ala Gly Lys Gln Gly Gly Phe Asp Gly Val Ala Ile Tyr Thr Gln Arg
 880 890

AAT TCA ATT GAA ACC CAA GTC CAT AAA GAT GCA TTA AAA GGC GTA CAA
 Asn Ser Ile Glu Thr Gln Val His Lys Asp Ala Leu Lys Gly Val Gln
 895 905

AGT TAT CAT CGA TTA ATC GCC AAA CCA GAG GAT CAA TCT GCA TAC TTT
 Ser Tyr His Arg Leu Ile Ala Lys Pro Glu Asp Gln Ser Ala Tyr Phe
 910 920

FIG. 6K.

```

GIG ATG CAA GAT GAG TGT CCA AAG CCA GAT GAT TAT AAC AGT TGT TTA
Val Met Gln Asp Glu Cys Pro Lys Pro Asp Asp Tyr Asn Ser Cys Leu
925                                     930                                     935

CCT TTC GCC AAA CGA CCT GCG ATT TTA TCC TCC CAA AGA GAA ACC GTA
Pro Phe Ala Lys Arg Pro Ala Ile Leu Ser Ser Gln Arg Glu Thr Val
940                                     945                                     950                                     955

AGC GTT TCA GAT TAT ACG GCG GCT AAC CGT ATC AAA CCT AAT CCA ATG
Ser Val Ser Asp Tyr Thr Gly Ala Asn Arg Ile Lys Pro Asn Pro Met
960                                     965                                     970

AAA TAT GAA AGC CAG TCT TGG TTT TTA AGA GGA GGG TAT CAT TTT TCT
Lys Tyr Glu Ser Gln Ser Trp Phe Leu Arg Gly Gly Tyr His Phe Ser
975                                     980                                     985

GAA CAA CAT TAT ATT GGT GGT ATT TTT GAA TTC ACA CAA CAA AAA TTT
Glu Gln His Tyr Ile Gly Gly Ile Phe Glu Phe Thr Gln Gln Lys Phe
990                                     995                                     1000

GAT ATC CGT GAT ATG ACA TTT CCC GCT TAT TTA AGA TCA ACA GAA AAA
Asp Ile Arg Asp Met Thr Phe Pro Ala Tyr Leu Arg Ser Thr Glu Lys
1005                                     1010                                     1015

```

FIG. 6L.

```

CGG GAT GAT AGC AGT GGC TCT TTT TAT CCA AAG CAA GAT TAT GGT GCA
Arg Asp Asp Ser Ser Gly Ser Phe Tyr Pro Lys Gln Asp Tyr Gly Ala
1020                                     1030 1035

TAT CAA CGT ATT GAG GAT GGC CGA GGC GTT AAC TAT GCA AGT GGG CTT
Tyr Gln Arg Ile Glu Asp Gly Arg Gly Val Asn Tyr Ala Ser Gly Leu
1040                                     1045 1050

TAT TTC GAT GAA CAC CAT AGA AAA CAG CGT GTA GGT ATT GAA TAT ATT
Tyr Phe Asp Glu His His Arg Lys Gln Arg Val Gly Ile Glu Tyr Ile
1055                                     1060 1065

TAC GAA AAT AAG AAC AAA GCG GCG ATC ATT GAC AAA GCA GTG TTA AGT
Tyr Glu Asn Lys Asn Lys Ala Gly Ile Ile Asp Lys Ala Val Leu Ser
1070                                     1075 1080

GCT AAT CAA CAA AAC ATC ATA CTT GAC AGT TAT ATG CAA CAT ACG CAT
Ala Asn Gln Gln Asn Ile Ile Leu Asp Ser Tyr Met Gln His Thr His
1085                                     1090 1095

TGC AGT CTT TAT CCT AAT CCA AGT AAG AAT TGC CGC CCA ACA CGT GAT
Cys Ser Leu Tyr Pro Asn Pro Ser Lys Asn Cys Arg Pro Thr Arg Asp
1100                                     1105 1110 1115

```

FIG. 6M.

```

AAA CCT TAT TCA TAC TAT CAT TCT GAT AGA AAT GTT TAT AAA GAA AAA
Lys Pro Tyr Ser Tyr Tyr His Ser Asp Arg Asn Val Tyr Lys Glu Lys
1120
CAT AAT ATG TTG CAA TTG AAT TTA GAG AAA AAA ATT CAA CAA AAT TGG
His Asn Met Leu Gln Leu Asn Leu Glu Lys Lys Ile Gln Gln Asn Trp
1135
CTT ACT CAT CAA ATT GTC TTC AAT CTT GGT TTT CAT GAC TTT ACT TCA
Leu Thr His Gln Ile Val Phe Asn Leu Gly Phe Asp Asp Phe Thr Ser
1150
GCG CTT CAG CAT AAA GAT TAT TTA ACT CGA CGT GTT ACC GCT ACG GCA
Ala Leu Gln His Lys Asp Tyr Leu Thr Arg Arg Val Thr Ala Thr Ala
1165
AAG AGT ATT TCA GAG AAA GCT AAT GAA ACA AGA AGA AAT GGT TAC AAA
Lys Ser Ile Ser Glu Lys Ala Asn Glu Thr Arg Arg Asn Gly Tyr Lys
1180
AAA CAA CCT TAC TTA TAC CCA AAA CCA ACA GTA GGT TTT GTA GTA CAA
Lys Gln Pro Tyr Leu Tyr Pro Lys Pro Thr Val Gly Phe Val Val Gln
1200

```

FIG. 6N.

GAT CAT TGT GAT TAT AAA GGT AAC TCC TCT AAT TAC AGA GAC TGT AAA
Asp His Cys Asp Tyr Lys Gly Asn Ser Ser Asn Tyr Arg Asp Cys Lys
1215 1220 1225

GTG CCG TTA ATT AAA CCG AAA AAT TAT TAT TTC GCA GCA CCG AAT AAT
Val Arg Leu Ile Lys Gly Lys Asn Tyr Tyr Phe Ala Ala Arg Asn Asn
1230 1235 1240

ATG GCA TTA GGG AAA TAC GTT GAT TTA GGT TTA GGT ATT CCG TAT GAC
Met Ala Leu Gly Lys Tyr Val Asp Leu Gly Leu Gly Ile Arg Tyr Asp
1245 1250 1255

GTA TCT CGC ACA AAA GCT AAT GAA TCA ACT ATT AGT GTT GGT AAA TTT
Val Ser Arg Thr Lys Ala Asn Glu Ser Thr Ile Ser Val Gly Lys Phe
1260 1265 1270 1275

AAA AAT TTC TCT TCG AAT ACT GGT ATT GTC ATA AAA CCA ACG GAA TGG
Lys Asn Phe Ser Trp Asn Thr Gly Ile Val Ile Lys Pro Thr Glu Trp
1280 1285 1290

CTT GAT CTT TCT TAT CCG CTT TCT ACT GGA TTT AGA AAT CCT AGT TTT
Leu Asp Leu Ser Tyr Arg Leu Ser Thr Gly Phe Arg Asn Pro Ser Phe
1295 1300 1305

FIG.60.

```

GCT GAA ATG TAT GGT TGG CGG TAT GGT GGC AAT AAT AGC GAG GTT TAT
Ala Glu Met Tyr Gly Trp Arg Tyr Gly Gly Asn Asn Ser Glu Val Tyr
1310
GTA GGT AAA TTT AAG CCT GAA ACA TCT CGT AAC CAA GAG TTT GGT CTC
Val Gly Lys Phe Lys Pro Glu Thr Ser Arg Asn Gln Glu Phe Gly Leu
1325
GCT CTA AAA GGG GAT TTT GGT AAT ATT GAG ATC AGT CAT TTT AGT AAT
Ala Leu Lys Gly Asp Phe Gly Asn Ile Glu Ile Ser His Phe Ser Asn
1340
GCT TAT CGA AAT CTT ATC GCC TTT GCT GAA GAA CTT AAT AAA AAT GGA
Ala Tyr Arg Asn Leu Ile Ala Phe Ala Glu Glu Leu Asn Lys Asn Gly
1360
ACT GGA AAG GCC AAT TAT GGA TAT CAT AAT GCA CAA AAT GCA AAA TTA
Thr Gly Lys Ala Asn Tyr Gly Tyr His Asn Ala Gln Asn Ala Lys Leu
1375
GTT GGC GTA AAT ATA ACT GCG CAA TTA GAT TTT AAT GGT TTA TGG AAA
Val Gly Val Asn Ile Thr Ala Gln Leu Asp Phe Asn Gly Leu Trp Lys
1390

```

FIG. 6P.

CGT ATT CCC TAC GGT TGG TAT GCA ACA TTT GCT TAT AAC CGA GTA AAA
 Arg Ile Pro Tyr Gly Trp Tyr Ala Thr Phe Ala Tyr Asn Arg Val Lys
 1405 1410 1415
 GGT AAA GAT CAA AAA ATC AAT GCT GGT TTG GCC TCC GTA AGC AGT TAT
 Val Lys Asp Gln Lys Ile Asn Ala Gly Leu Ala Ser Val Ser Ser Tyr
 1420 1425 1430 1435
 TTA TTT GAT GCC ATT CAG CCC AGC CGT TAT ATC ATT GGT TTA GGC TAT
 Leu Phe Asp Ala Ile Gln Pro Ser Arg Tyr Ile Ile Gly Leu Gly Tyr
 1440 1445 1450
 GAT CAT CCA AGT AAT ACT TGG GGA ATT AAT ACA AUG TTT ACT CAA TCA
 Asp His Pro Ser Asn Thr Trp Gly Ile Asn Thr Met Phe Thr Gln Ser
 1455 1460 1465
 AAA GCA AAA TCT CAA AAT GAA TTG CTA GGA AAA CGT GCA TTG GGT AAC
 Lys Ala Lys Ser Gln Asn Glu Leu Leu Gly Lys Arg Ala Leu Gly Asn
 1470 1475 1480
 AAT TCA AGG GAT GTA AAA TCA ACA AGA AAA CTT ACT CGG GCA TGG CAT
 Asn Ser Arg Asp Val Lys Ser Thr Arg Lys Leu Thr Arg Ala Trp His
 1485 1490 1495

FIG. 6 Q.

ATC TTA GAT GTA TCG GGT TAT TAC ATG GCG AAT AAA AAT ATT ATG CTT
 Ile Leu Asp Val Ser Gly Tyr Tyr Met Ala Asn Lys Asn Ile Met Leu
 1500 1505 1510 1515

CGA TTA GGG ATA TAT AAT TTA TTC AAC TAT CCG TAT GTT ACT TCG GAA
 Arg Leu Gly Ile Tyr Asn Leu Phe Asn Tyr Arg Tyr Val Thr Trp Glu
 1520 1525 1530

GCG GTG CGT CAA ACA GCA CAA GGT GCG GTC AAT CAA CAT CAA AAT GTT
 Ala Val Arg Gln Thr Ala Gln Gly Ala Val Asn Gln His Gln Asn Val
 1535 1540 1545

GGT AGC TAT ACT CCG TAC GCA GCA TCA GGA CGA AAC TAT ACC TTA ACA
 Gly Ser Tyr Thr Arg Tyr Ala Ala Ser Gly Arg Asn Tyr Thr Leu Thr
 1550 1555 1560

TTA GAA ATG AAA TTCTAATTA AATGGCCA GATGGACTAG ACATGCTATA
 Leu Glu Met Lys
 1565

TCATATACCTT ACTGGCGCAT CTTTTCTGT TCATATATCT GGTAAAGIGA AAAACCAAC

TTGGATTTTT TAGAAGATCT TTCCAGGCAT TTATGTGAAA ATCTCCGACA ATTTTACCG

CACCTTTCIC TATTACAAAA ACAATAAGGA TCCTTTGTIG AATCTCTCA

FIG. 7 A.

```

CAACATCTGC CCAAGCTATA TTGGTTAATG ATAAGCCTAT TAATGATPAAG CCTATTAAATG
ATAACAAAGA AATTGTGTTT ACGCCATTTT TCATATTTTA TCCATGAACT TAAAAAATTC
TAAGTTGACA TTATTACAAA AAAAGAACAA TAATGCGAAT TATTATCAAT TTGTATPAAG
AATATAATTC T ATG AAA TCT GTA CCT CTT ATC TCT TGT GGA CTT TCC TTT
Met Lys Ser Val Pro Leu Ile Ser Gly Gly Leu Ser Phe
1      5      10
TTA TTA AGT GCT TGT AGC GGA GGA GGG TCT TTT GAT GTA GAT AAC GTC
Leu Leu Ser Ala Cys Ser Gly Gly Gly Ser Phe Asp Val Asp Asn Val
15      20      25
TCT AAT CCC TCC TCT TCT AAA CCA CGT TAT CAA GAC GAT ACC TCG AAT
Ser Asn Pro Ser Ser Lys Pro Arg Tyr Gln Asp Asp Thr Ser Asn
30      35      40      45
CAA AGA ACA AAA TCT GAT TTG CAA AAG TTG TCC ATT CCT TCT TTA GGG
Gln Arg Thr Lys Ser Asp Leu Gln Lys Leu Ser Ile Pro Ser Leu Gly
50      55      60

```


FIG. 7B.

```

GGA GGG ATG AAG TTA GTG GCT CAG AAT CTT CTT GGT AAG AAA GAA CCT
Gly Gly Met Lys Leu Val Ala Gln Asn Leu Leu Gly Lys Glu Pro
65 70 75

AGT CTC TTA AAT AAT GAA GAT GGC TAT ATG ATA TTT TCC TCA CTT TCT
Ser Leu Leu Asn Asn Glu Asp Gly Tyr Met Ile Phe Ser Ser Leu Ser
80 85 90

ACG ATT GAA GAG GAT GTT ACA AAA GAA AAT AAA TCT CAG GAA CCC ACT
Thr Ile Glu Glu Asp Val Thr Lys Lys Glu Asn Lys Ser Gln Glu Pro Thr
95 100 105

ATT GGC TCA ATA GAC GAG CCT AGC AAA ACA AAT TCA CCC CAA AAT CAT
Ile Gly Ser Ile Asp Glu Pro Ser Lys Thr Asn Ser Pro Gln Asn His
110 115 120 125

CAT GGC AAT ATG TAT ATT CGG GTC TTT ATT ATA TTC AAT CGT GGC GTA
His Gly Asn Met Tyr Ile Arg Val Phe Ile Ile Phe Asn Arg Gly Val
130 135 140

ATT CCT CAA ATG GCA AGT TTT ATT CAG GTT ACT ATG GAT ATG CGT ATT
Ile Pro Gln Met Ala Ser Phe Ile Gln Val Thr Met Asp Met Arg Ile
145 150 155

```

FIG.7C.

```
ACT TTG GCA AGC AAA CAG CCA CTA CAT TAC CTG TAGATGGCGA AGCAACGTAT
Thr Leu Ala Ser Lys Gln Pro Leu His Tyr Leu
160                                     165
AAAGGAACTT GGCACCTTCAT CACCGCAACT GAAATGGCA AAAAGTATTC TTGTTCAGT
AATGATAGCG GTCAAGCTTA TCGCAGAGT AGTGCATTC CAGACATAT TGAATTAGAA
AAAAATGATT CAACTAATGG TGACAAGGC TTAAATAGTG AATTAGTGT CAATTTGGT
ACAAAAAAGC TCACTGGAAA ACTTTATTAT AATGAAAGAG AAACAGAACT TAATAATCA
AAAGATAGAA AACATPACT CTACAATCTA GAAGCTGAAG TGTATAGTAA CCGATTCAGG
GGTACAGTAA AGCCAACCGA AAAAGATTCT ACAGATCATC CCTTTACCAG CGAGGGTACA
TTAGAAGGTG GTTTTATGG GCCTAAAGGT GAAGAACTAG GAGGAAAGTT TTTAGCTGGC
GATATAAAAG TTTTGGGGT ATTTAGTGCC AAAGAAACGG AAGAAACAAA AAAGAAACGG
TTATCCAAGG AAACCTTAAT TGATGGCAAG CTACTACTTT TTAAACAAC CAATGCAACA
ACCAATGCAA CAGCCAATGC AACAAACAGT ACAACAGCCA GTACAACAAC CGATGCAGAA
```

FIG. 7 D.

AACITTTACGA CGAAAGATAT ACCAAGTTTT GGGAAGCTG AITTACCTTTT AATGTATAAT
TACCTGTTC CTCITTTTAC TGAGAGTGGT GATTTCATAA GTAGTAAGCA CCATACGTGA
GGAAAGAAAA CCTATCAAGT AGAAGCATGT TGCAGTAATC TAAGCTATGT GAAATTTGGT
ATGTTTTATG AAGACCCACT TAAAGAAGAA AAGGACAAG AAAAAGAAGA AGACAAGAA
AAACAAAACG CGGCAACGAC CAACACITAT TATCAATTCT TATTAGGTCT CCGTACTGOC
AGTTCIGAAA TTCCTPAAAT GGGAAAGTG GAATATGCG GTAAATTTGGT TGGTTATATT
AGTGATGGA CGACATCTTA CTCGCCAGT GGIGATPAGG AACGCAATTA AAATGCTOOC
GCCGATTTTA ATGTTGATTT TGICAATAAA AAGCTAACAG CCACATTAAA ACGACACGAT
AATCGAAATA CCGTATTTAG TATTGAGGA AACITTPACA GIGGGAATGA CTCACITGGT
AAAGCAACCG CAAAAGATTT AGTAATAGAT GGTAAAGTA CACAAGCCAC ATCTAAGTC
AATTTCACGG CAACAGTAAA AGGGGCATTT TATGACCTG ATGCTTCIGA ATTAGGCGGT
TATTTCACCT ATAACGGAAA AAATCTTACA GCTACAAATT CCCCACCGT ATCTTCACCA

FIG. 7 E.

TCCAATTCAG CAAATGCTCG TCGTCCCGTT GGTGTTGGAG CTAATAAACA AGTAGACACA

ACCAACAAGT AGAAAAAACC AATAAATGGA ATACTAAA ATG ACT AAA AAA CCC
Met Thr Lys Lys Pro
170

TAT TTT CGC CTA AGT ATT ATT TCT TGT CTT TTA ATT TCA TGC TAT GTA
Tyr Phe Arg Leu Ser Ile Ile Ser Cys Leu Leu Ile Ser Cys Tyr Val
175 180 185

AAA GCA GAA ACT CAA AGT ATA AAA GAT ACA AAA GAA CCT ATA TCA TCT
Lys Ala Glu Thr Gln Ser Ile Lys Asp Thr Lys Glu Ala Ile Ser Ser
190 195 200 205

GAA GTG GAC ACT CAA AGT ACA GAA GAT TCA GAA TTA GAA ACT ATC TCA
Glu Val Asp Thr Gln Ser Thr Glu Asp Ser Glu Leu Glu Thr Ile Ser
210 215 220

GTC ACT GCA GAA AAA ATA AGA GAT CGT AAA GAT AAT GAA GTA ACT CGA
Val Thr Ala Glu Lys Ile Arg Asp Arg Lys Asp Asn Glu Val Thr Gly
225 230 235

CTT GGC AAA ATT ATA AAA ACG AGT GAA AGT ATC AGC CGA GAA CAA GTA
Leu Gly Lys Ile Ile Lys Thr Ser Glu Ser Ile Ser Arg Glu Gln Val
240 245 250

FIG. 7F.

```

TTA AAT ATT CGT GAT CTA ACA CGC TAT GAT CCA GGC ATT TCA GTT GTA
Leu Asn Ile Arg Asp Leu Thr Arg Tyr Asp Pro Gly Ile Ser Val Val
255                               260                               265

GAA CAA GGT CGC GGT GCA AGT TCT GGA TAT TCT ATT CGT GGT ATG GAC
Glu Gln Gly Arg Gly Ala Ser Ser Gly Tyr Ser Ile Arg Gly Met Asp
270                               275                               280                               285

AGA AAT AGA GGT GCT TTA TTA GTA GAT GGT TTA CCT CAA ACG CAA TCT
Arg Asn Arg Val Ala Leu Leu Val Asp Gly Leu Pro Gln Thr Gln Ser
290                               295                               300

TAT GTA GTG CAA AGC CCT TTA GTT GCT CGT TCA GGA TAT TCT GGC ACT
Tyr Val Val Gln Ser Pro Leu Val Ala Arg Ser Gly Tyr Ser Gly Thr
305                               310                               315

CGT GCA ATT AAT GAA ATT GAA TAT GAA AAT GTA AAG GCC GTC GAA ATA
Gly Ala Ile Asn Glu Ile Glu Tyr Glu Asn Val Lys Ala Val Glu Ile
320                               325                               330

AGC AAG GGG AGT TCT TCT GAG TAT GGT AAT GGA GCA CTA GCT GGT
Ser Lys Gly Gly Ser Ser Ser Glu Tyr Gly Asn Gly Ala Leu Ala Gly
335                               340                               345

```

FIG. 7 G.

```

TCT GTA ACA TTT CAA AGC AAA TOC GCA GCC GAT ATC TTA GAA GGA GAC
Ser Val Thr Phe Gln Ser Lys Ser Ala Ala Asp Ile Leu Glu Gly Asp
350                                     360                                     365

AAA TCA TGG GGA ATT CAA ACT AAA AAT GCT TAT TCA AGC AAA AAT AAA
Lys Ser Trp Gly Ile Gln Thr Lys Asn Ala Tyr Ser Ser Lys Asn Lys
370                                     375                                     380

GGC TTT ACC CAT TCT TTA GCT GTA GCA GGA AAA CAA GGT GGA TTT GAA
Gly Phe Thr His Ser Leu Ala Val Ala Gly Lys Gln Gly Gly Phe Glu
385                                     390                                     395

GGG GTC GCC ATT TAC ACT CAA CGA AAT TOG GAG GAA ACC CAA GTC CAT
Gly Val Ala Ile Tyr Thr Gln Arg Asn Ser Glu Glu Thr Gln Val His
400                                     405                                     410

AAA GAT GCA TTA AAA GGC GTA CAA AGT TAT GAG CGA TTC ATC GCC ACA
Lys Asp Ala Leu Lys Gly Val Gln Ser Tyr Glu Arg Phe Ile Ala Thr
415                                     420                                     425

ACA GAT AAA TCT TCA CGA TAC TTT GTG ATA CAA GGT GAG TGT CCA AAT
Thr Asp Lys Ser Ser Gly Tyr Phe Val Ile Gln Gly Glu Cys Pro Asn
430                                     435                                     440                                     445

```

FIG. 7H.

GGT GAT GAC AAG TGT GCA GCC AAA CCA OCT GCA AAG TTA TCC CCC CAA
Gly Asp Asp Lys Cys Ala Ala Lys Pro Pro Ala Lys Leu Ser Pro Gln 450 460

AGC GAA ACC GTA AGC GTT TCA GAT TAT ACG GGG GCT AAC CGT ATC AAA
Ser Glu Thr Val Ser Val Ser Asp Tyr Thr Gly Ala Asn Arg Ile Lys 465 475

CCT AAT CCA ATG AAA TAT GAA AGC CAG TCT TGG TTT TTA AGA GGA GGG
Pro Asn Pro Met Lys Tyr Glu Ser Gln Ser Trp Phe Leu Arg Gly Gly 480 490

TAT CAT TTT TCT GAA CAA CAC TAT ATT GGT GGT ATT TTT GAA TTC ACA
Tyr His Phe Ser Glu Gln His Tyr Ile Gly Gly Ile Phe Glu Phe Thr 495 505

CAA CAA AAA TTT GAT ATC CGT GAT ATG ACA TTT CCC GCT TAT TTA AGA
Gln Gln Lys Phe Asp Ile Arg Asp Met Thr Phe Pro Ala Tyr Leu Arg 510 525

TCA ACA GAA AAA CGG GAT GAT AGA ACT GGC OCT TTT TAT CCA AAG CAA
Ser Thr Glu Lys Arg Asp Asp Arg Thr Gly Pro Phe Tyr Pro Lys Gln 530 540

FIG. 71.

GAT TAT GGT GCA TAT CAA CGT ATT GAG GAT GGC CGC GGT AAC TAT
 Asp Tyr Gly Ala Tyr Gln Arg Ile Glu Asp Gly Arg Gly Val Asn Tyr
 545 550 555
 GCA AGT GGG CTT TAT TTC GAT GAA CAC CAT AGA AAA CAG CGT GTA GGT
 Ala Ser Gly Leu Tyr Phe Asp Glu His His Arg Lys Gln Arg Val Gly
 560 565 570
 ATT GAA TAT ATT TAC GAA AAT AAG AAC AAA GCG GGC ATC ATT GAC AAA
 Ile Glu Tyr Ile Tyr Glu Asn Lys Asn Lys Ala Gly Ile Ile Asp Lys
 575 580 585
 GCA GTG TTA AGT GCT AAT CAA CAA AAC ATC ATA CTT GAC AGT TAT ATG
 Ala Val Leu Ser Ala Asn Gln Gln Asn Ile Ile Leu Asp Ser Tyr Met
 590 595 600 605
 CGA CAT ACG CAT TGC AGT CTT TAT CCT AAT CCA AGT AAG AAT TGC CCG
 Arg His Thr His Cys Ser Ser Leu Tyr Pro Asn Pro Ser Lys Asn Cys Arg
 610 615 620
 CCG ACA CTT GAT AAA CCT TAT TCA TAC TAT CGT TCT GAT AGA AAT GTT
 Pro Thr Leu Asp Lys Pro Tyr Ser Tyr Tyr Arg Ser Asp Arg Asn Val
 625 630 635

FIG. 7 J.

```

TAT AAA CAA AAA CAT AAT ATG TTG CAA TTG AAT TTA GAG AAA AAA ATT
Tyr Lys Glu Lys His Asn Met Leu Gln Leu Asn Leu Glu Lys Lys Ile
640 645 650

CAA CAA AAT TGG CTT ACT CAT CAA AAT GTC TTC AAT CTT GGT TTT GAT
Gln Gln Asn Trp Leu Thr His Gln Ile Val Phe Asn Leu Gly Phe Asp
655 660 665

GAC TTT ACT TCA GCG CTT CAG CAT AAA GAT TAT TTA ACT CGA CGT GTT
Asp Phe Thr Ser Ala Leu Gln His Lys Asp Tyr Leu Thr Arg Arg Val
670 675 680 685

ACC GCT ACG GCA AAT AAT ATT TCA GCG ACA GGT GCT GGT AAA CGA AGA
Thr Ala Thr Ala Asn Ile Ile Ser Gly Thr Val Ala Gly Lys Arg Arg
690 695 700

AAT GGT TAC GAA AAA CAA CCT TAC TTA TAC TCA AAA CCA AAA GTA GAT
Asn Gly Tyr Glu Lys Lys Gln Pro Tyr Leu Tyr Ser Lys Pro Lys Val Asp
705 710 715

TTT GTA GGA CAA GAT CAT TGT AAT TAT AAA GGT AGC TCC TCT AAT TAC
Phe Val Gly Gln Asp His Cys Asn Tyr Lys Gly Ser Ser Ser Asn Tyr
720 725 730

```

FIG. 7K.

```

AGC GAC TGT AAA GTG CCG TTA ATT AAA GCG AAA AAT TAT TAT TTC GCA
Ser Asp Cys Lys Val Arg Leu Ile Lys Gly Lys Asn Tyr Tyr Phe Ala
735                               740                               745

GCA GCG AAT AAT ATG GCA TTA GCG AAA TAC ATT GAT TTA GGT TTA GGT
Ala Arg Asn Asn Met Ala Leu Gly Lys Tyr Ile Asp Leu Gly Leu Gly
750                               755                               760                               765

ATT CCG TAT GAC GTA TCT CGT ACA AAA GCT AAT GAA TCA ACT ATT AGT
Ile Arg Tyr Asp Val Ser Arg Thr Lys Ala Asn Glu Ser Thr Ile Ser
770                               775                               780

GTT GGT AAA TTT AAA AAT TTC TCT TCG AAT ACT GGT ATT GTC ATA AAA
Val Gly Lys Phe Lys Asn Phe Ser Trp Asn Thr Gly Ile Val Ile Lys
785                               790                               795

CCA ACG GAA TCG CTT GAT CTT TCT TAT CCG CTT TCT ACT GGA TTT AGA
Pro Thr Glu Trp Leu Asp Leu Ser Tyr Arg Leu Ser Thr Gly Phe Arg
800                               805                               810

AAT CCT AGT TTT GCT GAA ATG TAT TAT GGT TCG CCG TAT GGT GGC AAT AAT
Asn Pro Ser Phe Ala Glu Met Tyr Gly Trp Arg Tyr Gly Gly Asn Asn
815                               820                               825

```

FIG. 7L.

```

AGC GAT GTT TAT GTA GGT AAA TTT AAG CCT GAA ACA TCT CGT AAC CAA
Ser Asp Val Tyr Val Gly Lys Phe Lys Pro Glu Thr Ser Arg Asn Gln
830                               840                               845

GAG TTT GGT CTC GCT CTA AAA GGG GAT TTT GGT AAT ATT GAG ATC AGT
Glu Phe Gly Leu Ala Leu Lys Lys Gly Asp Phe Gly Asn Ile Glu Ile Ser
850                               855                               860

CAT TTT AGT AAT GCT TAT CGA AAT CTT ATC GCC TTT GCT GAA GAA CTT
His Phe Ser Asn Ala Tyr Arg Asn Leu Ile Ala Phe Ala Glu Glu Leu
865                               870                               875

AGT AAA AAT GGA ACT ACT CGA AAG GGC AAT TAT GGA TAT CAT AAT GCA
Ser Lys Asn Gly Thr Thr Gly Lys Gly Asn Tyr Gly Tyr His Asn Ala
880                               885                               890

CAA AAT GCA AAA TTA GTT GGC GTA AAT ATA ACT GCG CAA TTA GAT TTT
Gln Asn Ala Lys Leu Val Gly Val Asn Ile Thr Ala Gln Leu Asp Phe
895                               900                               905

AAT GGT TTA TGG AAA CGT ATT OCC TAC GGT TGG TAT GCA ACA TTT GCT
Asn Gly Leu Trp Lys Arg Ile Pro Tyr Gly Trp Tyr Ala Thr Phe Ala
910                               915                               920                               925

```

FIG. 7 M.

```

TAT AAC CGA GTA AAA GTT AAA GAT CAA AAA ATC AAT GCT GGT TTG GCC
Tyr Asn Arg Val Lys Val Lys Asp Gln Lys Ile Asn Ala Gly Leu Ala
930
935
TCC GTA AGC AGT TAT TTA TTT GAT GCC ATT CAG CCC AGC GGT TAT ATC
Ser Val Ser Ser Tyr Leu Phe Asp Ala Ile Gln Pro Ser Arg Tyr Ile
945
950
ATT GGT TTA GGC TAT GAT CAT CCA AGT AAT ACT TGG GGA ATT AAT ACA
Ile Gly Leu Gly Tyr Asp His Pro Ser Asn Thr Trp Gly Ile Asn Thr
960
965
ATG TTT ACT CAA TCA AAA GCA AAA TCT CAA AAT GAA TTG CTA GGA CAA
Met Phe Thr Gln Ser Lys Ala Lys Ser Gln Asn Glu Leu Leu Gly Gln
975
980
CGT GCA TTG GGT AAC AAT TCA AGG AAT GTA AAA TCA ACA AGA AAA CTT
Arg Ala Leu Gly Asn Asn Ser Arg Asn Val Lys Ser Thr Arg Lys Leu
990
1000
ACT CGG GCA TGG CAT ATC TTA GAT GTA TCG GGT TAT TAC ATG GCG AAT
Thr Arg Ala Trp His Ile Leu Asp Val Ser Gly Tyr Tyr Met Ala Asn
1010
1015
1020

```

FIG. 7 N.

AAA AAT AIT ATG CTT CGA TTA GCG ATA TAT AAT TTA TTC AAC TAT CGC
 Lys Asn Ile Met Leu Arg Leu Gly Ile Tyr Asn Leu Phe Asn Tyr Arg
 1025 1030 1035

TAT GTT ACT TCG GAA GCG GTG CGT CAA ACA GCA CAA GGT GCG GTC AAT
 Tyr Val Thr Trp Glu Ala Val Arg Gln Thr Ala Gln Gly Ala Val Asn
 1040 1045 1050

CAA CAT CAA AAT GTT GGT AGC TAT ACT CGC TAC GCA GCA TCA GGA CGA
 Gln His Gln Asn Val Gly Ser Tyr Thr Arg Tyr Ala Ala Ser Gly Arg
 1055 1060 1065

AAC TAT ACC TTA ACA TTA GAA ATG AAA TTC TAAATTAAAA TCGGCCAGAT
 Asn Tyr Thr Leu Thr Leu Glu Met Lys Phe
 1070 1075

GGACTAGATA TGCATATATCT ATACCTTACT GCGGCAICTT TTTCIGTICT ATAAICTCT

TAAGTCAAAA ACCAAACTTG GATTTTTTAC AAGATCTTTT CAOCCATTTA TTGTAAATTC

TCCGACAATT TTTCACGCAC TTTCCTCTAT TACAAAAACA ATAAGCATCC TTITGIGACT

CTCTCAATCT TTGCAAGIT GCCTGTACAA CTCAGATCA AGTTTCAGCC AGCATCTTA

GGCAGTGGG TTGGGC

FIG. 8 A.

```

AT ATG AAA TCT GTA CCT CTT ATC TCT GGT GGA CTT TCC TTT TTA TTA
Met Lys Ser Val Pro Leu Ile Ser Gly Gly Leu Ser Phe Leu Leu
1      5      10      15

AGT GCT TGT AGC GGG GGA GGT GGT TCT TTT GAT GTA GAT GAC GTC TCT
Ser Ala Cys Ser Gly Gly Gly Ser Phe Asp Val Asp Asp Val Ser
20      25      30

AAT CCC TCC TCT TCT AAA CCA CGT TAT CAA GAC GAT ACT TCA AGT TCA
Asn Pro Ser Ser Ser Lys Pro Arg Tyr Gln Asp Asp Thr Ser Ser Ser
35      40      45

AGA ACA AAA TCT AAA TTG GAA AAT TTG TCC ATT CCT TCT TTA GGG GGA
Arg Thr Lys Ser Ser Lys Leu Glu Asn Leu Ser Ile Pro Ser Leu Gly Gly
50      55      60

GGG ATG AAG TTA GTG CCT CAG AAT CTT CGT GAT AGG ACA AAA CCT AGT
Gly Met Lys Leu Val Ala Gln Asn Leu Arg Asp Arg Thr Lys Pro Ser
65      70      75

CTC TTA AAT GAA GAT GAC TAT ATG ATA TTT TCC TCA CTT TCA ACG ATT
Leu Leu Asn Glu Asp Asp Tyr Met Ile Phe Ser Ser Leu Ser Thr Ile
80      85      90      95

```

FIG. 8B.

AAA GCT GAT GGT GAA AAA GAA AAT AAA CAC TAT ACA AGT CCA GGT GGC
 Lys Ala Asp Val Glu Lys Glu Lys Asn Lys His Tyr Thr Ser Pro Val Gly
 100 105 110

TCA ATA GAC GAG CCT AGT ACA ACA AAT CCA AAA GAA AAT GAT CAT GGA
 Ser Ile Asp Glu Pro Ser Thr Thr Asn Pro Lys Glu Asn Asp His Gly
 115 120 125

CAA AGA TAT GTA TAT TCA GGA CTT TAT TAT ATT CCA TCG TGG AAT TTA
 Gln Arg Tyr Val Tyr Ser Gly Leu Tyr Tyr Ile Pro Ser Trp Asn Leu
 130 135 140

AAC GAT CTT AAA AAT AAC AAG TAT TAT TAT TCT GGT TAC TAT GGA TAT
 Asn Asp Leu Lys Asn Asn Lys Tyr Tyr Tyr Ser Gly Tyr Tyr Gly Tyr
 145 150 155

GCG TAT TAC TTT GGC AAG CAA ACA GCC ACT ACA TTA CCT GTA AAT GGC
 Ala Tyr Tyr Phe Gly Lys Glu Thr Thr Ala Thr Thr Leu Pro Val Asn Gly
 160 165 170 175

AAA GTA ACG TAT AAA GGA ACT TGG AGC TTC ATC ACC GCA GCT GAA AAT
 Lys Val Thr Tyr Lys Gly Thr Trp Ser Phe Ile Thr Ala Ala Glu Asn
 180 185 190

FIG.8C.

GGC AAA AGG TAT CCT TTG TTA AGT AAT GGC AGT CAA GCT TAT TTT CCA
 Gly Lys Arg Tyr Pro Leu Ser Asn Gly Ser Gln Ala Tyr Phe Arg
 195 200 205

CGT AGT GCA ATT CCA GAA GAT ATT GAT TTA GAA GTT AAA AAT GAT GAG
 Arg Ser Ala Ile Pro Glu Asp Ile Asp Leu Glu Val Lys Asn Asp Glu
 210 215 220

AAT AGA GAA AAA GGG CTA GTG AGT GAA TTT AGT CCA GAT TTT GGG ACT
 Asn Arg Glu Lys Gly Leu Val Ser Glu Phe Ser Ala Asp Phe Gly Thr
 225 230 235

AAA AAA CTG ACA GGA GGA CTG TTT TAC ACC AAA AGA CAA ACT CAT ATT
 Lys Lys Leu Thr Gly Gly Leu Phe Tyr Thr Lys Arg Gln Thr His Ile
 240 245 250 255

CAA AAC CAT GAA AAG AAA AAA CTC TAT GAT ATA GAT GCC CAT ATT TAT
 Gln Asn His Glu Lys Lys Lys Leu Tyr Asp Ile Asp Ala His Ile Tyr
 260 265 270

AGT AAT AGA TTC AGA GGT AAA GTA AAT CCT ACC CAA AAA GAT TCT AAA
 Ser Asn Arg Phe Arg Gly Lys Val Asn Pro Thr Gln Lys Asp Ser Lys
 275 280 285

FIG. 8D.

```

GAA CAT CCC TTT ACC AGC GAG GGA ACA TTA GAA GGT GGT TTT TAC GGG
Glu His Pro Phe Thr Ser Glu Gly Thr Leu Glu Gly Phe Tyr Gly
290
300
CCT GAA GGT CAA GAA TTA GGA GGA AAG TTT TTA GCT GGC GAC AAA AAA
Pro Glu Gly Gln Glu Leu Glu Gly Lys Phe Leu Ala Gly Asp Lys Lys
305
310
315
GTT TTT GGG GTA TTT AGT GCC AAA GGA ACG GAA GAA AAC AAA AAA TTA
Val Phe Gly Val Phe Ser Ala Lys Gly Thr Glu Glu Asn Lys Lys Leu
320
325
330
335
CCC AAA GAA ACC TTA ATT GAT GGC AAG CTA ACT ACT TTC TCT ACT AAA
Pro Lys Glu Thr Leu Ile Asp Gly Lys Leu Thr Thr Phe Ser Thr Lys
340
345
350
ACA ACC GAT GCA AAA ACC AAT GCA ACA GCC AAT GCA ACA ACC AGT ACC
Thr Thr Asp Ala Lys Thr Asn Ala Thr Ala Asn Ala Thr Thr Ser Thr
355
360
365
GCA GCC AAT ACA ACA ACC GAT ACA ACA GCC AAT ACA ATA ACC GAT GCA
Ala Ala Asn Thr Thr Thr Asp Thr Thr Ala Asn Thr Ile Thr Asp Ala
370
375
380

```

FIG. 8E.

GAA AAC TTT AAG ACG AAA GAT ATA TCA AGT TTT GGT GAA GCT GAT TAC
 Glu Asn Phe Lys Thr Lys Asp Ile Ser Ser Phe Gly Glu Ala Asp Tyr
 385 390 395

CTT TTA ATT GAT AAT TAC OCT GTT OCT CTT TTA OCT GAG AGT GGT GAT
 Leu Leu Ile Asp Asn Tyr Pro Val Pro Leu Leu Pro Glu Ser Gly Asp
 400 405 410 415

TTC ATA AGT AGT AAG CAC CAT ACT GTA GGA AAG AAA ACC TAT CAA GTA
 Phe Ile Ser Ser Lys His Thr Val Gly Lys Lys Thr Tyr Gln Val
 420 425 430

AAA GCA TGT TGC AGT AAT CTA AGC TAT GTG AAA TTT GGT ATG TAT TAT
 Lys Ala Cys Cys Ser Asn Leu Ser Tyr Val Lys Phe Gly Met Tyr Tyr
 435 440 445

GAA GTC CCA CCT AAA GAA GAA GAA GAA GAA GAA GAA GAA GAA GAA
 Glu Val Pro Pro Lys Glu Glu Glu Lys Asp Lys Glu Lys Lys Glu Lys
 450 455 460

GAA AAA GAA AAA CAA GCG ACA AAT CTA TCG AAC ACT TAT TAT CAA TTC
 Glu Lys Glu Lys Gln Ala Thr Asn Leu Ser Asn Thr Tyr Tyr Gln Phe
 465 470 475

FIG.8F.

```

TTA TTA GGT CTC GGT ACT CCC AGT TCT GAA ATT CCT AAA GGA GGA AGT
Leu Leu Gly Leu Arg Thr Pro Ser Ser Glu Ile Pro Lys Gly Gly Ser
480                                     490                                     495

GCA AAA TAT CTC GGT AGT TGG TTT GGT TAT CTG AGC GAT GGT TCA ACA
Ala Lys Tyr Leu Gly Ser Trp Phe Gly Tyr Leu Ser Asp Gly Ser Thr
500                                     505                                     510

TCT TAC TCC CCC AGT GGT GAT AAG AAA CCG GAG AAC AAT GCT CTC GCC
Ser Tyr Ser Pro Ser Gly Asp Lys Lys Arg Glu Asn Asn Ala Leu Ala
515                                     520                                     525

GAG TTT AAT GTA AAT TTT GTC GAT AAA ACA TTA AAA GGC CAA TTA ATA
Glu Phe Asn Val Asn Phe Val Asp Lys Thr Leu Lys Gly Gln Leu Ile
530                                     535                                     540

CGA CAC GAT AAT CAA AAT ACC GTT TTT ACA ATT GAT GCA ACC TTT AAA
Arg His Asp Asn Gln Asn Thr Val Phe Thr Ile Asp Ala Thr Phe Lys
545                                     550                                     555

GGT GGT AAG AAT AAC TTC ACT GGT ACA GCA ACC GCA AAC AAT GTA GCG
Gly Gly Lys Asn Asn Phe Thr Gly Thr Ala Thr Ala Asn Asn Val Ala
560                                     565                                     570                                     575

```

FIG. 8G.

ATT GAT CCC CAA AGT ACA CAA GGC ACA TCT AAC GTC AAT TTC ACG GCA
Ile Asp Pro Gln Ser Thr Gln Gly Thr Ser Asn Val Asn Phe Thr Ala
580 585
ACA GTA AAT GCG GCA TTT TAT GCG CCG AAC GCT ACA GAA TTA GGC GGT
Thr Val Asn Gly Ala Phe Tyr Gly Pro Asn Ala Thr Glu Leu Gly Gly
595 600
TAT TTC ACC TAT AAC GGA AAT CCT ACA GAT AAA AGT TOC TCA ACC GTA
Tyr Phe Thr Tyr Asn Gly Asn Pro Thr Asp Lys Ser Ser Thr Val
610 615
CCT TCA TCA TCC AAT TCA AAA AAT GCA AGA GCT GCA GTT GTC TTT GGT
Pro Ser Ser Ser Asn Ser Lys Asn Ala Arg Ala Ala Val Val Phe Gly
625 630 635
GCG AGA CAA CAA GTA GAA ACA ACC AAA TAATGGAATA CTAAAAATGA
Ala Arg Gln Gln Val Glu Thr Thr Lys
640 645
CTAAAAAAGC TTCTAGAAGC CGAATTTC

FIG. 9 A.

GAATTGGCT TGGATCCAT ATG AAA TCT GTA CCT CTT ATC TCT GGT GGA CTT
Met Lys Ser Val Pro Leu Ile Ser Gly Gly Leu 10
1 5

TCC TTT TTA CTA AGT GCT TGT AGC GGA GGG GGG TCT TTT GAT GTA GAT
Ser Phe Leu Leu Ser Ala Cys Ser Gly Gly Gly Ser Phe Asp Val Asp 25
15 20

AAC GTC TCT AAT OCA TCC TCT TCT AAA CCA CGT TAT CAA GAC GAT ACT
Asn Val Ser Asn Pro Ser Ser Lys Pro Arg Tyr Gln Asp Asp Thr 40
30 35

TCA AGT TCA AGA ACA AAA TCT AAT TTG AAA AAG TTG TCC ATT CCT TCT
Ser Ser Arg Thr Lys Ser Asn Leu Lys Lys Leu Ser Ile Pro Ser 55
45 50

TTA GGG GGA GGG ATG AAG TTA GTG GCT CAG AAT CTT AGT GAT AAG AAC
Leu Gly Gly Gly Met Lys Leu Val Ala Gln Asn Leu Ser Asp Lys Asn 75
60 65 70

AAA CCT AGT CTC TTA AAT GAA GAT GAC TAT ATA TCA TAT TTT TCC TCA
Lys Pro Ser Leu Leu Asn Glu Asp Asp Tyr Ile Ser Tyr Phe Ser Ser 90
80 85

FIG. 9B.

```

CTT TCT ACA ATT CAA GAT GAT GAT AAA AAA GAA AAT AAA CGC CAT ACA
Leu Ser Thr Ile Gln Asp Asp Val Lys Lys Glu Asn Lys Arg His Thr
95 100 105

AAT CCA GTT GGC TCA ATA GAC GAG CCT AAC GCA ACA AAT CCA CCC GAA
Asn Pro Val Gly Ser Ile Asp Glu Pro Asn Ala Thr Asn Pro Pro Glu
110 115 120

AAG CAT CAT GGA CAA AGA TAT GTA TAT TCA GGG CTT TAT TAT ATT CCA
Lys His His Gly Gln Arg Tyr Val Tyr Ser Gly Leu Tyr Tyr Ile Pro
125 130 135

TCG TGG AGT CAT TCC TCA AAT GGC AAG CTT TAT TTA GGT TAC TAT GGA
Ser Trp Ser His Ser Ser Asn Gly Lys Leu Tyr Leu Gly Tyr Tyr Gly
140 145 150 155

TAT GCG TTT TAT TAT GGT AAT AAA ACT GCA ACA AAC TTG CCA GTA AGC
Tyr Ala Phe Tyr Tyr Gly Asn Lys Thr Ala Thr Asn Leu Pro Val Ser
160 165 170

GGC ATA GCT AAA TAC AAA GGA ACT TGG GAT TTT ATT ACT GCA ACT AAA
Gly Ile Ala Lys Tyr Lys Gly Thr Trp Asp Phe Ile Thr Ala Thr Lys
175 180 185

```

FIG. 9C.

AAAT GGC CAA CGT TAT TCT TTA TTT GGT AGC GCT TTT CGA GCT TAT AAT	
Asn Gly Gln Arg Tyr Ser Leu Phe Gly Ser Ala Phe Gly Ala Tyr Asn	190 195 200
AGA CGC AGT GCT ATT TCA GAA GAT ATA GAT AAT TTA GAA AAT AAT CTA	
Arg Arg Ser Ala Ile Ser Glu Asp Ile Asp Asn Leu Glu Asn Asn Leu	205 210 215
AAG AAT GGT GCG CGA TTA ACT AGT GAA TTT ACT GTC AAT TTT GGT ACG	
Lys Asn Gly Ala Gly Leu Thr Ser Glu Phe Thr Val Asn Phe Gly Thr	220 225 230 235
AAA AAG CTC ACT CGA AAA CTT TAT TAT AAT GAA AGG GAA ACA AAT CTT	
Lys Lys Leu Thr Gly Lys Leu Tyr Tyr Asn Glu Arg Glu Thr Asn Leu	240 245 250
AAT AAA TTA CAA AAG AGA AAA CAT GAA CTC TAT GAT ATA GAT GCC GAT	
Asn Lys Leu Gln Lys Arg Lys His Glu Leu Tyr Asp Ile Asp Ala Asp	255 260 265
ATT TAT AGT AAT AGA TTC AGA GGT AAA GTA AAG OCA ACA ACC CAA AAA	
Ile Tyr Ser Asn Arg Phe Arg Gly Lys Val Lys Pro Thr Thr Gln Lys	270 275 280

FIG. 9D.

```

GAT TCT CAA GAA CAT CCC TTT ACC ACC GAG GGA ACA TTA GAA GGT GGT
Asp Ser Gln Glu His Pro Phe Thr Ser Glu Gly Thr Leu Glu Gly Gly
285 290 295

TTT TAT GGG CCT AAC GGT GAA GAA TTA GGA GGA AAG TTT TTA GGT GGC
Phe Tyr Gly Pro Asn Gly Glu Glu Leu Gly Gly Lys Phe Leu Ala Gly
300 305 310 315

GAT AAC CGA GTT TTT GGG GTA TTT AGT GGC AAA GAA GAA ACA AAA
Asp Asn Arg Val Phe Gly Val Phe Ser Ala Lys Glu Glu Thr Lys
320 325 330

GAC AAA AAA TTA TCC AGA GAA ACC TTA ATT GAT GGC AAG CTA ATT ACT
Asp Lys Lys Leu Ser Arg Glu Thr Leu Ile Asp Gly Lys Leu Ile Thr
335 340 345

TTT AAA AGA ACT GAT GCA ACA ACC AAT ACA GCA GCC AAT GCA AAA ACC
Phe Lys Arg Thr Asp Ala Thr Thr Asn Thr Ala Ala Asn Ala Lys Thr
350 355 360

GAT GAA AAA AAC TTT ACG ACG AAA GAT ATA CCA AGT TTT GGT GAA GCT
Asp Glu Lys Lys Asn Phe Thr Thr Lys Asp Ile Pro Ser Phe Gly Glu Ala
365 370 375

```


FIG.9E.

```

GAT TAC CTT TTA ATT GAT AAT TAC CCT GTT CCT CTT TTC OCT GAA GAA
Asp Tyr Leu Leu Ile Asp Asn Tyr Pro Val Pro Leu Phe Pro Glu Glu
380                                     385                                     390                                     395

AAT ACT AAT GAT TTC ATA ACT AGT AGG CAC CAT AAG GTA GGA GAT AAA
Asn Thr Asn Asp Phe Ile Thr Ser Arg His His Lys Val Gly Asp Lys
400                                     405                                     410

ACC TAT AAA GTA GAA GCA TGT TGC AAG AAT CTA AGC TAT GTG AAA TTT
Thr Tyr Lys Val Glu Glu Ala Cys Cys Lys Asn Leu Ser Tyr Val Lys Phe
415                                     420                                     425

GGT ATG TAT TAT GAA GAC CCA TTA AAT GGA GAA AAT GGC AAA GAA AAA
Gly Met Tyr Tyr Glu Asp Pro Leu Asn Gly Glu Asn Gly Lys Glu Lys
430                                     435                                     440

GAA AAA GAA GAA AAA GAC AAA GAA GAA AAA CAA GCG ACA ACA TCT ATC
Glu Lys Glu Lys Glu Lys Asp Lys Glu Lys Glu Lys Gln Ala Thr Thr Ser Ile
445                                     450                                     455

AAG ACT TAT TAT CAA TTC TTA TTA GGT CAC CGT ACT GCC AAG GCC GAC
Lys Thr Tyr Tyr Gln Phe Leu Leu Gly His Arg Thr Ala Lys Ala Asp
460                                     465                                     470                                     475

```

FIG.9F.

ATA CCT GCA ACG GGA AAC GTG AAA TAT CCG GGT AAT TCG TTT GGT TAT
 Ile Pro Ala Thr Gly Asn Val Lys Tyr Arg Gly Asn Trp Phe Gly Tyr 480
 485
 AAT GGT GAT GAC AAG ACA TCT TAC TCC ACT ACT GGA GAT AAA AAT GCT
 Ile Gly Asp Asp Lys Thr Ser Tyr Ser Thr Thr Gly Asp Lys Asn Ala 495
 500
 GTC GCC GAG TTT GAT GTA AAT TTT GCC GAT AAA ACA TTA ACA GGC ACA
 Val Ala Glu Phe Asp Val Asn Phe Ala Asp Lys Thr Leu Thr Gly Thr 510
 515
 TTA AAA CGA CAC GAT AAT GGA AAT CCC GTA TTT ACA ATT AAT GCA AGC
 Leu Lys Arg His Asp Asn Gly Asn Pro Val Phe Thr Ile Asn Ala Ser 525
 530
 TTT CAA AGT GGT AAG AAT GAC TTC ACT GGT ACA GCA ACC GCA AAC AAT
 Phe Gln Ser Gly Lys Asn Asp Phe Thr Gly Thr Ala Thr Ala Asn Asn 540
 545
 GTA GCG ATT GAT CCC CAA AAT ACA CAA ACC ACA TCT AGA GTC AAT TTC
 Val Ala Ile Asp Pro Gln Asn Thr Gln Thr Thr Ser Arg Val Asn Phe 555
 560
 565
 570

FIG.96.

ACG GCA ACA GTA AAC GGG GCA TTT TAT GGA OCT AAG GCT ACA GAA TTA
Thr Ala Thr Val Asn Gly Ala Phe Tyr Gly Pro Lys Ala Thr Glu Leu
575 580 585

GGC GGT TAT TTC ACT TAT AAC GGA AAC AAT OCT ACA GAT AAA AAT TCC
Gly Gly Tyr Phe Thr Tyr Asn Gly Asn Asn Pro Thr Asp Lys Asn Ser
590 595 600

TCA ACC GTT TCA CCA TCC AAT TCA GCA AAT GCT CGT GCT GCC GTT GIG
Ser Thr Val Ser Pro Ser Asn Ser Ala Asn Ala Arg Ala Val Val
605 610 615

TTT GGC GCT AAA AAA CAA GTA GAA ACA ACC AAC AAG TAAAAACAAC
Phe Gly Ala Lys Lys Lys Gln Val Glu Thr Thr Asn Lys
620 625 630

CAAGTAATGG AATACTAAAA ATGACTAAAA AAGCTTCTAG AAAGCCGAAT TC

FIG.10A.

```

ATG AAA TCT GTA CCT CTT ATC TCT GGT GGA CTT TCC CTT TTA TTA AGT
Met Lys Ser Val Pro Leu Ile Ser Gly Gly Leu Ser Leu Leu Ser
1      5      10      15

GCT TGT AGC GGG GGA GGT GGT TCT TTT GAT GTA GAT GAC GTC TCT AAT
Ala Cys Ser Gly Gly Gly Ser Phe Asp Val Asp Asp Val Ser Asn
20     25     30

CCC TCC TCT TCT AAA CCA CGT TAT CAA GAC GAT ACC TCG AGT CAA AGA
Pro Ser Ser Ser Lys Pro Arg Tyr Gln Asp Asp Thr Ser Ser Gln Arg
35     40     45

ACA AAA TCT AAT TTG GAA AAG TTG TCC ATT CCT TCT TTA GGA GGA GGG
Thr Lys Ser Asn Leu Glu Lys Leu Ser Ile Pro Ser Leu Gly Gly Gly
50     55     60

ATG AAA TTG GTG GCT CAG AAT CTG AGT GGT AAT AAA GAA CCT AGT TTC
Met Lys Leu Val Ala Gln Asn Leu Ser Gly Asn Lys Glu Pro Ser Phe
65     70     75     80

TTA AAT GGA AAT GAC TAT ATG ATA TTT TCC TCA CGT TCT ACG ATT AAA
Leu Asn Gly Asn Asp Tyr Met Ile Phe Ser Ser Arg Ser Thr Ile Lys
85     90     95

```

FIG.10B.

GAT GAT GTT GAA AAT AAC AAT ACA AAC GGG GGG GAC TAT ATT GGC TCA
 Asp Asp Val Glu Asn Asn Asn Thr Asn Gly Gly Asp Tyr Ile Gly Ser
 100 105 110

ATA GAC GAG CCT AGT ACA ACA AAT CCA CTC GAA AAG CAT CAT GGA CAA
 Ile Asp Glu Pro Ser Thr Thr Asn Pro Leu Glu Lys His His Gly Gln
 115 120 125

ACG TAT GTA TAT TCA GGG CTT TAT TAT ATT CAA TCG TGG AGT CTA AGA
 Arg Tyr Val Tyr Ser Gly Leu Tyr Tyr Ile Gln Ser Trp Ser Leu Arg
 130 135 140

GAT TTA CCA AAG AAG TTT TAT TCA GGT TAC TAT GGA TAT GCG TAT TAC
 Asp Leu Pro Lys Lys Phe Tyr Ser Gly Tyr Tyr Gly Tyr Ala Tyr Tyr
 145 150 155 160

TTT GGC AAG GAA ACA GCC ACT ACA TTA CCT GTA AAT GGC GAA GCA ACG
 Phe Gly Lys Glu Thr Ala Thr Thr Leu Pro Val Asn Gly Glu Ala Thr
 165 170 175

TAT AAA GGA ACT TGG GAT TTC ATC ACT GCA ACT AGA AAT GGC AAA AGT
 Tyr Lys Gly Thr Trp Asp Phe Ile Thr Ala Thr Arg Asn Gly Lys Ser
 180 185 190

FIG.10C.

```

TAT TCT TTG TTA AGT AAT AAC CGA CAA GCT TAT TCC AAA CGT AGT GCA
Tyr Ser Leu Leu Ser Asn Asn Arg Gln Ala Tyr Ser Lys Arg Ser Ala
195                                     200 205

ATT CCA GAA GAC AIT GAT TTA GAA AAT GAT CCA AAG AAT GGT GAG ACG
Ile Pro Glu Asp Ile Asp Leu Glu Asn Asp Pro Lys Asn Gly Glu Thr
210                                     215 220

AGA TTA ACT AGT GAA TTT ACT GTG AAT TTT GGT ACG AAA AAG CTC ACA
Arg Leu Thr Ser Glu Phe Thr Val Asn Phe Gly Thr Lys Lys Leu Thr
225                                     230 235 240

GGT GGA CTT TAT TAC CAT TTA CGT AAA ACA AAT GCT AAT GAA AAC CAA
Gly Gly Leu Tyr Tyr His Leu Arg Lys Thr Asn Ala Asn Glu Asn Gln
245                                     250 255

AAT AGA AAA CAT AAA CTC TAC AAT CTA GAA GCT GAT GTG TAT AGC AAC
Asn Arg Lys His Lys Leu Tyr Asn Leu Glu Ala Asp Val Tyr Ser Asn
260                                     265 270

CGA TTC AGA GGT AAA GTA AAG CCA ACC AAA GAG TCT TCT GAA GAA CAT
Arg Phe Arg Gly Lys Val Lys Pro Thr Lys Glu Ser Ser Glu Glu His
275                                     280 285

```

FIG.10D.

```

CCC TTT ACC AGC GAG GGA ACA TTA GAA GGT GGT TTT TAT GGG CCT AAT
Pro Phe Thr Ser Glu Gly Thr Leu Glu Gly Phe Tyr Gly Pro Asn
290                                     300
GCT GAA GAA CTA GGG GGA AAA TTT TTA GCT AGC GAT AAA AAA GTT TTT
Ala Glu Glu Leu Glu Gly Lys Phe Leu Ala Ser Asp Lys Lys Val Phe
305                                     315
GGG GTA TTT AGT GCC AAA GAA CAG CAA GAA ACG GAA GAA AAC AAA AAA
Gly Val Phe Ser Ala Lys Glu Gln Gln Glu Thr Glu Glu Asn Lys Lys
325                                     330
TTA CTC AAA GAA ACC TTA ATT GAT GGC AAG CTA ACT ACT TTC TCT ACT
Leu Leu Lys Glu Thr Leu Ile Asp Gly Lys Leu Thr Thr Phe Ser Thr
340                                     350
AAA AAA ACC AAT GCA ACA ACC GAT GCA ACA ACC AGT ACA ACA ACC AGT
Lys Lys Thr Asn Ala Thr Thr Asp Ala Thr Thr Ser Thr Thr Thr Ser
355                                     365
ACA GCA ACC AAT GCA ACA GCC GAT GCA GAA AAC TTT ACG ACA AAA GAT
Thr Ala Thr Asn Ala Thr Ala Asp Ala Glu Asn Phe Thr Thr Lys Asp
370                                     375

```

FIG.10E.

ATA TCA AGT TTT GGT GAA GCT GAT TAT CTT TTA ATT GAT AAT TAC CCT
 Ile Ser Ser Phe Gly Glu Ala Asp Tyr Leu Leu Ile Asp Asn Tyr Pro
 385 390 395 400
 GTT CCT CTT TTA CCT GAA AAT ACT AAT GAT TTC ATA AGC AGT AAG CAC
 Val Pro Leu Leu Pro Glu Asn Thr Asn Asp Phe Ile Ser Ser Lys His
 405 410 415
 CAT GAG GTA GGA GGT AAA CAC TAT AAA GTG GAA GCA TGT TGC AAG AAT
 His Glu Val Gly Gly Lys Lys His Tyr Lys Val Glu Ala Cys Cys Lys Asn
 420 425 430
 CTA AGC TAT GTG AAA TTT GGT ATA TAT TAT GAG CAT AAT GAG AAG AAC
 Leu Ser Tyr Val Lys Phe Gly Ile Tyr Tyr Glu Asp Asn Glu Lys Asn
 435 440 445
 ACC AAA ATT GAA ACA GAA CAA TAC CAC CAA TTT TTG TTA GGT CTC CGT
 Thr Lys Ile Glu Thr Glu Gln Tyr His Gln Phe Leu Leu Gly Leu Arg
 450 455 460
 ACT CCC AGT TCT CAA ATT CCT GCA ACG GGA AAC GTG AAA TAT CGC GGT
 Thr Pro Ser Ser Gln Ile Pro Ala Thr Gly Asn Val Lys Tyr Arg Gly
 465 470 475 480

FIG.10F.

```

AGT TGG TTT GGT TAT ATT GGT GAT GAC AAG ACA TCT TAC TCC ACT ACT
Ser Trp Phe Gly Tyr Ile Gly Asp Asp Lys Thr Ser Tyr Ser Thr Thr
485                               490                               495

GGA GAT AAA AAT GCT CTC GCC GAG TTT GAT GTA AAT TTT ACC GAT AAA
Gly Asp Lys Asn Ala Leu Ala Glu Phe Asp Val Asn Phe Thr Asp Lys
500                               505                               510

AAG CTA ACA GGC GAA TTA AAA CGA GCC GAT AAT CAA AAT ACC GTA TTT
Lys Leu Thr Gly Glu Leu Lys Arg Ala Asp Asn Gln Asn Thr Val Phe
515                               520                               525

AGA ATT AAT GCA GAC TTT AAA AAT AAT GAT AAT GCC TTC AAA GGT ACA
Arg Ile Asn Ala Asp Phe Lys Asn Asn Asp Asn Ala Phe Lys Gly Thr
530                               535                               540

GCA ACC GCA GAA AAT TTT GTA ATA GAT GGT AAC AAT AGT CAA ACT GGA
Ala Thr Ala Glu Asn Phe Val Ile Asp Gly Asn Asn Ser Gln Thr Gly
545                               550                               555                               560

AAT ACC CAA ATT AAT ATT AAA ACT GAA GTA AAT GGG GCA TTT TAT GGT
Asn Thr Gln Ile Asn Ile Lys Thr Glu Val Asn Gly Ala Phe Tyr Gly
565                               570                               575

```

FIG.10G.

CCG AAC GCT ACA GAA TTA GGC GGT TAT TTC ACT TAT AAC GGA AAA AAT
Pro Asn Ala Thr Glu Leu Gly Tyr Phe Thr Tyr Asn Gly Lys Asn
580 585 590

CCT ACA GAT AAA AAT TCT GAA AGT TCC TCA ACC GTA CCT TCA CCA CCC
Pro Thr Asp Lys Asn Ser Glu Ser Ser Ser Thr Val Pro Ser Pro Pro
595 600 605

AAT TCA CCA AAT GCA AGA GCT GCA GTT GTC TTT GGT GCT AAA AAA CAA
Asn Ser Pro Asn Ala Arg Ala Val Val Phe Gly Ala Lys Lys Gln
610 615 620

GTA GAA AAA AAC AAC AAG TAAAAACAAC CAAGTAATGG AATACTAAAA
Val Glu Lys Asn Asn Lys
625 630

ATGACTAAAA AAGCTTCTAG AAGCCGAATT C

FIG.11A.

```

ATG AAA TCT GTA CCT CTT ATC TCT GGT GGA CTT TCC TTT TTA CTA AGT
Met Lys Ser Val Pro Leu Ile Ser Gly Gly Leu Ser Phe Leu Leu Ser
1      5      10      15

GCT TGT AGC GGA GGG GGG TCT TTT GAT GTA GAT AAC GTC TCT AAT ACC
Ala Cys Ser Gly Gly Gly Ser Phe Asp Val Asp Asn Val Ser Asn Thr
20     25     30

CCC TCT TCT AAA CCA CGT TAT CAA GAC GAT ACC TCG AAT CAA AGA ACA
Pro Ser Ser Lys Pro Arg Tyr Gln Asp Asp Thr Ser Asn Gln Arg Thr
35     40     45

AAA TCT AAA TTG GAA AAG TTG TCC ATT CCT TCT TTA GGA GGA GGG ATG
Lys Ser Lys Leu Glu Lys Leu Ser Ile Pro Ser Leu Gly Gly Met
50     55     60

AAG TTA GTT GTG CAA AAT TTT GCT GCT GCT AAA GAA OCT AGT TTC TTA
Lys Leu Val Val Gln Asn Phe Ala Gly Ala Lys Glu Pro Ser Phe Leu
65     70     75     80

AAT GAA AAT GAC TAT ATA TCA TAT TTT TCC TCA CTT TCT ATG ATT AAA
Asn Glu Asn Asp Tyr Ile Ser Tyr Phe Ser Ser Leu Ser Met Ile Lys
85     90     95

```

FIG.11B.

GAT GAT GTT GAA AAT AAC AAT AAA AAT AAG GAT ACT CCA ATT GGC TCA
Asp Asp Val Glu Asn Asn Lys Asn Lys Asp Thr Pro Ile Gly Ser
100 105 110

ATA GAC GAG CCT AGA GCA CCA AAT TCA AAC GAA AAT CAT CAA AAT CAT
Ile Asp Glu Pro Arg Ala Pro Asn Ser Asn Glu Asn His Gln Asn His
115 120 125

CAT GGA CAG CAA TAT GTA TAT TCG GGT CTT TAT TAT ATT CCA TCG TCG
His Gly Gln Gln Tyr Val Tyr Ser Gly Leu Tyr Tyr Ile Pro Ser Trp
130 135 140

CGT CTA ATA AAT TTA CCA AAT AAG TTT TAT TCA GGT TAC TAT GGA TAT
Arg Leu Ile Asn Leu Pro Asn Lys Phe Tyr Ser Gly Tyr Tyr Gly Tyr
145 150 155 160

GCG TAT TAC TTT GGC AAG CAA ACT GCC ACT ACA TTA CCT GTA AAT GGC
Ala Tyr Tyr Phe Gly Lys Gln Thr Ala Thr Thr Leu Pro Val Asn Gly
165 170 175

GAA GCA ACG TAT AAA GGA ACT TCG AGC TTC ACC GCA ACT GAA AGA
Glu Ala Thr Tyr Lys Gly Thr Trp Ser Phe Ile Thr Ala Thr Glu Arg
180 185 190

FIG.11C.

```

GGC AAA AAT TAT TCT TTG TTC AAT AAT AGA GGT CAA GCT TAT TCT CGA
Gly Lys Asn Tyr Ser Leu Phe Asn Asn Arg Gly Gln Ala Tyr Ser Arg
195 200 205

CGT AGT CCT ACT CCA GGA GAT ATT GAT TTA GAA AAC GGT GAC GCA GGC
Arg Ser Ala Thr Pro Gly Asp Ile Asp Leu Glu Asn Gly Asp Ala Gly
210 215 220

TTA ACA AGT GAA TTT ACT GTC AAT TTT GGT ACA AAA AAG CTC ACT CGA
Leu Thr Ser Glu Phe Thr Val Asn Phe Gly Thr Lys Lys Leu Thr Gly
225 230 235

GAA CCT TAT TAT AAT GAA AGG GAA ACA AAT CTT AAT CAA TCA AAA GAT
Glu Pro Tyr Tyr Asn Glu Arg Glu Thr Asn Leu Asn Gln Ser Lys Asp
240 245 250 255

AGA AAA CAT AAA CTC TAC GAT CTA GAA GCT GAT GTG TAT AGC AAC CGA
Arg Lys His Lys Leu Tyr Asp Leu Glu Ala Asp Val Tyr Ser Asn Arg
260 265 270

TTC AGA GGT ACA GTA AAG CCA ACC AAA AAA GAG TCT TCT GAA GAA CAT
Phe Arg Gly Thr Val Lys Pro Thr Lys Lys Glu Ser Ser Glu Glu His
275 280 285

```

FIG.11D.

```

CCC TTT ACC AGC GAG GGA ACA TTA GAA GGT GGT TTT TAT GGG CCT AAT
Pro Phe Thr Ser Glu Gly Thr Leu Glu Gly Phe Tyr Gly Pro Asn
290                                     300

GCT GAA GAA CTA GGG GGA AAA TTT TTA GCT AGC GAT AAA AAA GTT TTT
Ala Glu Glu Leu Glu Gly Lys Phe Leu Ala Ser Asp Lys Lys Val Phe
305                                     315

GGG GTA TTT AGT GGC AAA GAA ACG GAA GAA AAA CCA AAA TTA CCC AAA
Gly Val Phe Ser Ala Lys Glu Thr Glu Glu Lys Pro Lys Leu Pro Lys
325                                     335

GAA ACC TTA ATT GAT GGC AAG CTA ACT ACT TTC TCT AAA ACA ACC GAT
Glu Thr Leu Ile Asp Gly Lys Leu Thr Thr Phe Ser Lys Thr Thr Asp
340                                     350

ACA ACA ACC AAT AAA ACA ACC AGT GCA AAA ACC AAT ACA GAA AAC TTT
Thr Thr Thr Asn Lys Thr Thr Ser Ala Lys Thr Asn Thr Glu Asn Phe
355                                     365

ACG ACA AAA GAT ATA CCA AGT TTT GGT GAA GCT GAT TAT CTT TTA ATT
Thr Thr Lys Asp Ile Pro Ser Phe Gly Glu Ala Asp Tyr Leu Leu Ile
370                                     380

```

FIG.11E.

```

GAT AAT TAC CCT ATT CCG CTT TTA CCT GAG AGT GGT GAT TTC ATA AGT
Asp Asn Tyr Pro Ile Pro Leu Leu Pro Glu Ser Gly Asp Phe Ile Ser
385                                     395                                     400

AGT AAG CAC CAT GAG GTA GGA GGT AAA CCG TAT AAA GTG GAA GCA TGT
Ser Lys His His Glu Val Gly Gly Lys Arg Tyr Lys Val Glu Ala Cys
405                                     410                                     415

TGC AAG AAT CTA TGC TAT GTG AAA TTT GGT ATG TAT TAT GAG GAT AAA
Cys Lys Asn Leu Cys Tyr Val Lys Phe Gly Met Tyr Tyr Glu Asp Lys
420                                     425                                     430

GAG AAC AAC AAA AAT GAA ACA GAC AAA GAA AAA GAA AAA CAA ACG ACA
Glu Asn Asn Lys Lys Asn Glu Thr Asp Lys Glu Lys Glu Lys Gln Thr Thr
435                                     440                                     445

ACA TCT ATC AAG ACT TAT TAT CAA TTC TTA TTA GGT CTC CGG ACT CCC
Thr Ser Ile Lys Thr Tyr Tyr Gln Phe Leu Leu Gly Leu Arg Thr Pro
450                                     455                                     460

AGT TCT GAA ATT CCT AAA ATG GGA AAC GTG ACA TAT CGC GGT AGT TGG
Ser Ser Glu Ile Pro Lys Met Gly Asn Val Thr Tyr Arg Gly Ser Trp
465                                     470                                     475                                     480

```

FIG.11F.

```

TTT GGT TAT ATT GGT GAT GAC AAG ACA TCT TAC TCC GCT ACA GGA GAT
Phe Gly Tyr Ile Gly Asp Asp Lys Thr Ser Tyr Ser Ala Thr Gly Asp
485
AAA CGA CAA GAT AAA AAT GCT CCC GCC GAG TTT AAT GCT GAT TTT AAC
Lys Arg Gln Asp Lys Asn Ala Pro Ala Glu Phe Asn Ala Asp Phe Asn
500
AAT AAA AAG CTA ACA GGC ACA TCA AAA CGA CAC GAT AAT CAA AAT CCC
Asn Lys Lys Leu Thr Gly Thr Ser Lys Arg His Asp Asn Gln Asn Pro
515
GTG TTT AAC ATT AAG GCA ACC TTT CAA AAT GGT CGG AAT GAC TTT GAA
Val Phe Asn Ile Lys Ala Thr Phe Gln Asn Gly Arg Asn Asp Phe Glu
530
GGT ACA GCA ACC GCA GAA AAT TTT GTA ATA GAT GGT AAA GAT AGT CAA
Gly Thr Ala Thr Ala Glu Asn Phe Val Ile Asp Gly Lys Asp Ser Gln
545
CGA AAT ACC CCA ATT AAT ATT ACA ACT AAA GTA AAC GGG CCA TTT TAT
Gly Asn Thr Pro Ile Asn Ile Thr Thr Lys Val Asn Gly Ala Phe Tyr
565

```


FIG.11G.

GGG CCT GAT GCT TCT GAA TTA GGC GGT TAT TTC ACC TAT AAC GGA AAA
Gly Pro Asp Ala Ser Glu Leu Gly Gly Tyr Phe Thr Tyr Asn Gly Lys
580 585 590

GAC ACT ATA ACT AAA AAT ACT GAA AGT TCC TCA ACC GTA CCT TCA CCA
Asp Thr Ile Thr Lys Asn Thr Glu Ser Ser Thr Val Pro Ser Pro
595 600 605

CCC AAT TCA CCA AAT GCA AGA GCT GCA GTT GTG TTT GGA GCT AAA AAA
Pro Asn Ser Pro Asn Ala Arg Ala Val Val Phe Gly Ala Lys Lys
610 615 620

CAA GTA GAA ACA ACC AAC AAG TAGAAAAAAA CAAATAATGG AATACTAAAA
Gln Val Glu Thr Thr Asn Lys
625 630

ATGACTAAAA AAGCTTCTAG AAAGCCGAAT TC

FIG. 12A

...
...

-35
TCTAACTTGACATTATCAAAAAAGATCAATAATGCGAATTATTATCAATTTGTATGAG...RBS...
TCTAACTTGACATTATCAAAAAAGATCAATAATGCGAATTATTATCAATTTGTATGAG...
TCTAAGTTGACATTATCAAAAAAGAACAATAATGCGAATTATTATCAATTTGTATAAG...
TCTAAGTTGACATTATCAAAAAAGAACAATAATCCGAATTATTATCAATTTGTATAAG...

local de ligação a fur ATAATGATAATCATTATC
.....
...5'GGATCCAT \ATGAAATCTGTACCCTCTTATCTCTGGT 3'
...
... M K S V P L I S G
... TATATAATTCTATGAAATCTGTACCCTCTTATCTCTGGT EAGAN
... TATATAATTCTATGAAATCTGTACCCTCTTATCTCTGGT MINNA
... TAT-TAATTCTATGAAATCTGTACCCTCTTATCTCTGGT PAK
... AATATAATTCTATGAAATCTGTACCCTCTTATCTCTGGT SB33

FIG. 12B

3' <i>tdp2</i>	5' <i>tdp1</i>	
GTAGAAACAACCAATAA-----TGGAACTACTAAAAATGACTAAAAAACCCCTATTTTCGCCTAAGT	GTAGAAACAACCAATAA-----TGGAACTACTAAAAATGACTAAAAAACCCCTATTTTCGCCTAAGT	EAGAN
GTAGAAACAACCAATAA-----TGGAACTACTAAAAATGACTAAAAAACCCCTATTTTCGCCTAAGT	GTAGAAACAACCAATAA-----TGGAACTACTAAAAATGACTAAAAAACCCCTATTTTCGCCTAAGT	MINNA
GTAGAAACAACCAAGTAA-----TGGAACTACTAAAAATGACTAAAAAACCCCTATTTTCGCCTAAGT	GTAGAAACAACCAAGTAA-----TGGAACTACTAAAAATGACTAAAAAACCCCTATTTTCGCCTAAGT	DL63
GTAGAAACAACCAACCAAGTAAACCAACCAAGTAAATGGAATACTAAAAATGACTAAAAAACCCCTATTTTCGCCTAAGT	GTAGAAACAACCAACCAAGTAAATGGAATACTAAAAATGACTAAAAAACCCCTATTTTCGCCTAAGT	PAK
3' TACCTTATGATTTTACIGATTTTT \CGAAGAATC 5'		
GTAGAAACAACCAATAA-----TGGAACTACTAAAAATGACTAAAAAA	GTAGAAACAACCAATAA-----TGGAACTACTAAAAATGACTAAAAAA	SB12
GTAGAAACAACCAACCAAGTAAAAACCAACCAAGTAAATGGAATACTAAAAATGACTAAAAAA	GTAGAAACAACCAACCAAGTAAAAACCAACCAAGTAAATGGAATACTAAAAATGACTAAAAAA	SB29
GTAGAAAAAACCAACTAGTAAAAAACCAACCAAGTAAATGGAATACTAAAAATGACTAAAAAA	GTAGAAAAAACCAACTAGTAAAAAACCAACCAAGTAAATGGAATACTAAAAATGACTAAAAAA	SB30
GTAGAAACAACCAACCAAGTAAAAAACCAACCAAGTAAATGGAATACTAAAAATGACTAAAAAA	GTAGAAACAACCAACCAAGTAAAAAACCAACCAAGTAAATGGAATACTAAAAATGACTAAAAAA	SB32



FIG.13.

FIG. 14C.

G.	A.	VQD	VR.	RWA	V	A.	YRS.	HSEKSV	T	THR	L.	A	V	L.	FT	M.	T.	A.	L.	A.	A.	N.	SD.	SB33		
G.	A.	V	D	VR.	RWA	V	A	L.	YRS.	HSDG	SV	T	THR	L.	A.	L.	AD.	T.	A.	L.	A.	S.	ESK	TLDL	B16B6	
S.	A.	V	D	VR.	RWA	V	A	L.	YRS.	HSDG	SV	T	THR	L.	A.	L.	AD.	T.	A.	L.	A.	S.	---	VQ	RAV	M982
S.	A.	V	D	VR.	RWA	V	A	L.	YRS.	HSDG	SV	T	THR	L.	A.	L.	AD.	T.	A.	L.	A.	S.	---	KI	RAV	FA19

[illegible]

TSYLFDIAIQPSRYIIGLYGDHPSTNTGINTWFTQSKAKSQNELLGKRALGNNRSD-VKSTRKLTRAWHILDVSGYMANKNIMLRGLYINLFNRYVYTW
K.....N-.....V.RS.LF..V..L.....
Q.....N-.....
 V.....VL.....DGI.....Y.....VD.....SQ..L.GNANAK.AASRR..P.YVT.....NIK.HLT.A.V..L.....
 H.....VV.....Q.EGK.V.G.L.Y.....EIT..S..L.GNSRNT.A.ARR..P.Y.V.....TIK.HFT.A.V..L.....
 Q.H.....VV.S..Q.EGK.V.G.L.Y.....EIT..S..L.GNSRNT.A.ARR..P.Y.V.....TVK.HFT.A.V..L.H.....
 EAGAN
 DL63
 PAK
 SB33
 B16B6
 M982
 FA19

MAVRQFAQGVNQHNVGSYTRYAASGRNYTLTLEMKF*	EAGAN
.....N.....*	DL63
.....*	PAK
.....*	SB33
N.....G.....K.....V.N.....P.....FS.....*	B16B6
N.....G.....G.K.....V.N.....P.....FS.....*	M982
N.....A.....K.....V.N.....P.....FS.....*	FA19

FIG.15B.

YGNKTATNLPVNGVAKYGTWDFITATKNGKRYPLLSNGS---HAYYRRSAIPEDIDLENDKNGDI-GLISEFSADFQTKKLTGOLSYTKRKT-----N
 F..T..SA...G...T...S...AE...N.E..R.SGGG-Q..S...T...DRKT-----T...TVN...G.Y.NL.E.DAN---K
 ...E..K...K...N...E...S.F.SIG-Q..S...S...YNLENGDA---V...K.E...E.Y.NE...SVN---E
 F.KQ...T...KVT...S...AE...Q..F...VKNENREK..V...Q...G.F...Q...H
 ...S.I...S...Q..S.FGSF---G..N...S...NLENNKNGA-.T...TVN...K.Y.NE.E...N
 F.KE...T...E.T...R...S.S...NR---Q..SK...P...ETR-.T...TVN...G.Y.HL..MAN---E
 F.KQ...T...E.T...S...ER..N.S.FN.RG--Q..S...T.G...A-.T...TVN...EPY.NE.E.N---L
 KPLP-SEKITYK.TWD.VTDAMEKQRFGL--GSAAGFKSGALSALKEGVLRNQAAS--SGHT.F..MT...EV..SD.TIK.T.YRNN.I.QNNSENKQ
 POLPASGRVIYK..WHFVTDTKKGQDFREIIQPSKKQGDYRFGSGDSEYVSNKNESTLKDHEGY-FT.NLEV..N...K.IRNNASLNNTNNDK
 QOLPASEAVIYK..WHFVTDTKQKQKFNDILETSKQGDYRFGSGDDEGETSNRT.SNLD.HEGY-FT.N.KV..NN...K.IRNNKVINTAASDG-
 Y..SP.KE...QLLT.T.S...TSNANLNNEEGRPNYLN---DD..TKFIGRVRGLVSG.A.PAKH-KYT.Q.EV..A...M..KJ.-D.E...
 L.VTPSKE..KGK.IS...VSNINLERIDGKDTSGDGRNVSATISITETVNR.HKVGE.L.N.EVKGVAHSSEFAVDFDNKKLTGSLYRNGYINRNK
 AP37

NO--PYEKKLYDIDADIYSNRRFGTVKPEKD-SEHPFTSEGT-LEGGFYGNPAEELGGKFLATNDRVFGVFSAKETEETKKEA-LSKETLIDGLITFFS
 K---NRTH---LE.VH...K...K.ES...EGQ...H.KK.L...QQ..SENKK.P...T..K
 S---NTTH---TLE.KV...K...KTK-.D...N.EK...DPQNPENQK..T...K
 ---NH...H...K.N.Q.-K...EGQ...G.KK...G..N---P...T...K
 NKLOQR.HE...K...TQKD.Q...G...S.KK...QQ..BENKK.L...T...K
 ---NR.H...NLE.V...K...KES-...S.KK...KP---P...T...
 -SKDR.H...LE.V...K.ES...I.DSDS...KG..A...SN.K.AA.G..QKDKKDG.NAAGPA.E-----
 IK--TTRYTIQATLHGNRFKKGKALAAD,GATNG--I.DSDS...V.DSSS.S...F.QG...FR..SD.QK.AV.G...TKDKLENG.AA.GS.GAAASGGAAG
 HT--TQYYSLDAQ.TGNRENGTATA.D.KENET-KL--V.DSSS.S...F.QG...FR..SD.QK.AV.G...TKDKLENG.AA.GS.GAAASGGAAG
 YT--TEYYSILDATLRGNRF.GKAIA.D..NTGGTKL--VFDSSS.S...F.QG...FR..SD.QK.AV.G...TKDSTANGNAPAASSG-----
 IY--TV---NA..RGNRFTGAATASD.NKG.GE.YNF..SADSQS...K..MA..V.N.KSL.A...-----
 A.----VT.R.S.E...AG...KA.A--AGD---IFTDSNY...K..MA..FTNNKSL.A.A.-----
 EAGAN
 DL63
 PAK
 SB12
 SB29
 SB30
 SB32
 B16B6
 M982
 FA19
 AP205
 AP37

TKKTDKAT---NATSTAAATTTDTTANTITDEKNFKTEDISSFGADYLLIDKY-----PIPLLPDKNTNDFI
 TNAT.NATT--D...T.S.K...T.ATANTE..T.K.P.L...N-----V..F.--ESG...
 RTDATTNATT--D.K..ATTDA.S...KK..AE...P...GNO-----E..D...
 .T...NATA.....AE...K...N-----V...--ESG...
 EAGAN
 DL63
 PAK
 SB12

SUBSTITUTE SHEET

FIG.15C.

.....T-----A.AK.....T.K.P.....N-----V.F.EE.....SB29
N.T-----D.....S.ATNATA,AE..T.K.....N-----V.....SB30
KTTD...NK.TSAK.NTE..T.K.P.....N-----ESG.....SB32
VID.YRI.GEEFKEQIDSGDVKKLLVD.VELS..PSEGNKAA-----PQHEIE.....B16B6
 .S-----SENS--KL..VLD.VEL.LNDKKIKNLDNFSNAAQLVVD.IMIP..PKDSESGNTQADKCKNGGTEFTFRKFEHTPESDKDAQAGTQTNGAQTASN M982
 PGAATMPSET--RL..VLD.VEL.PDGKEIKNLDNFSNATRLVVD.IMIP..PT--ESGNGQADKCKNGGTDFTYTTTPESDKDKTKAQTGAGGMQTASG FA19
 -----HNGSNVN.VRIIDASKI.LT..SISELNN..D.SV.I.....GKKIKLAGSG.T AP205
 -----SENG--ET..E-----RIIDA.KI.LTQ.NAKELNN..D.SV.I.....GQKI.LAGVN.K AP37

EAGAN
 DL63
 PAK
 SB12
 SB29
 SB30
 SB32
 B16B6
 M982
 FA19
 AP205
 AP37

.....K.T--Q.....A.P.--E.K.K.KD.....ATTSIK.....L..SSE.....E.....N...
G.T--K.....K.....KDKDN.NETDKEGKEKPT.TTSI.....L..E.....E.....N...
K.T--Q.K.....V.P.E.....ATNLS.....L..SSE.....G.....L...
R.K.D.T--K.....K.....NG.NG.....E.....A.A..A--NV.R.N...
E.G.H--K.....I.....N.N.KIE.EQ.H...L..SSQ..A--NV.R...
E.G.--K.C.....KENNKN.T.E.....TTSIK.....L..SSE.....M.NVT.R...
ONGVKAT-----V.....D.MS..KLSKEN-----DDM..Q.V..VS.VAARTEAN..R.T.Y.
GDTNGKT.T--E.V...N.L.Y..LTRKNS.....SAWQAGNSSQADAKTEQVEQSM..Q.E.DEKE.--TDQNVV.R...Y.
GVNGGQVGTKT...QV...N.L.Y..LLTREN-----NNSVMQAVKNSS.AD.K.KQIE.SMFLOGERT.ENKIPQEQGIV.L.F.Y.
HTIEING.T--MVAV.....E.M...QLW-----QQAEGGKPENNSL..Q.E..AT.KM.--G.NY..I.T.D-
 N...TVEING.T--MVAV.....E.M...QLW-----Q.EGKQVQKDNLSL..Q.E..AT.KM.A--G.NY..V.T.D-

EAGAN
 DL63
 PAK
 SB12
 SB29
 SB30
 SB32

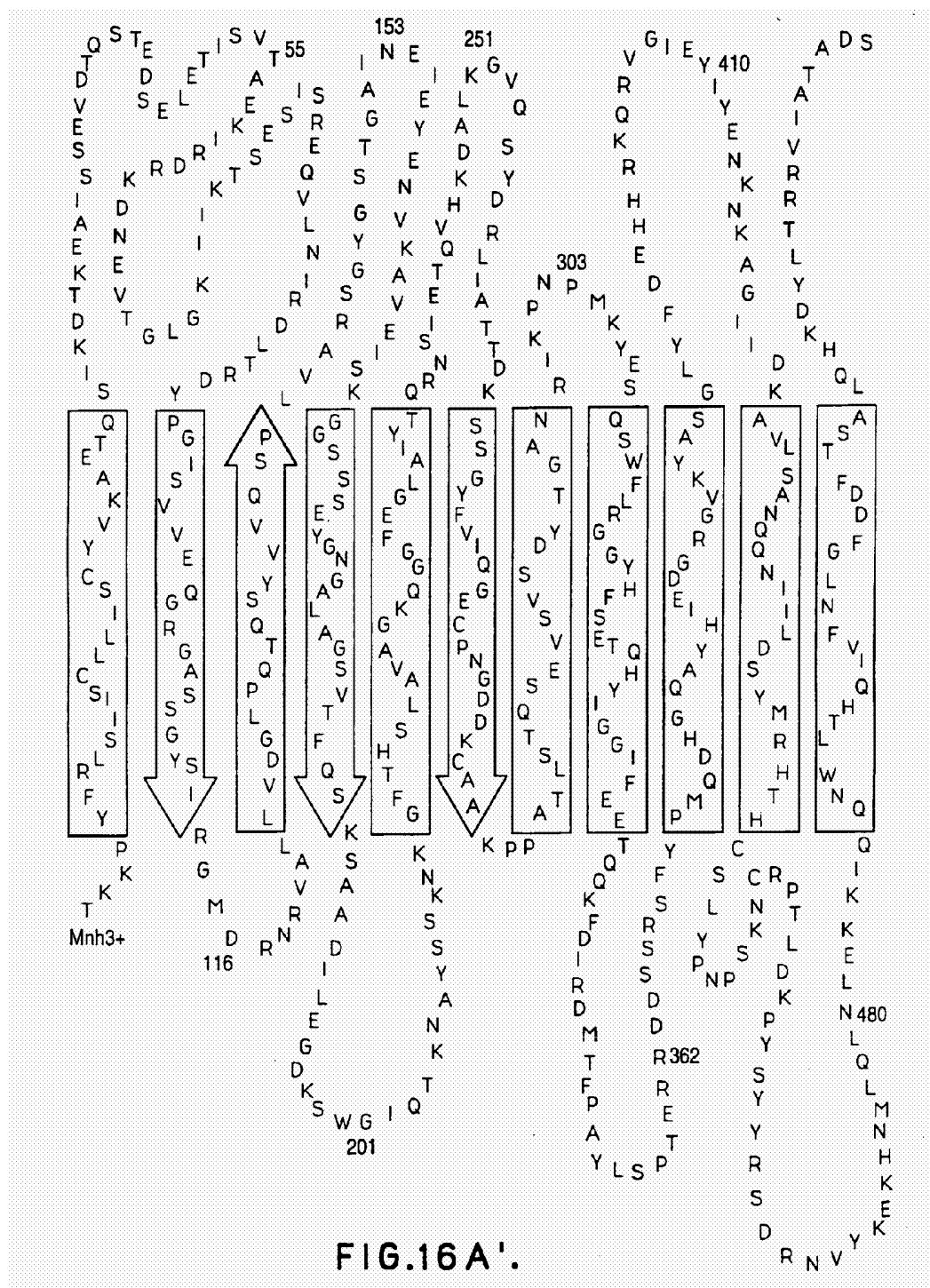
.....S.E....A....E.S.....N...T.....Q-.....K.N.T.QSGK.D.....KD--LA.....T.GTSKVNFTA...
S.E....A....E.S.....D.S.N.T.....N--T.K.N.EL.G.-D.....N..TS.AK.....
 ..LS..S.....EN..L...N.VD.T.KGQ.I...NQ--T..T.D.T.KGK.N.....N--VA..POSTQGTSNVNFTA...N.
 ..G.D....TT-----D.N.D.T..T...N--T.N.S.QSGK.D.....N--VA..PQ.T.TTSKVNFTA...
 ..G.D....TT-----L..D.N.TD.....A.NQ--T.R.N.D.K.ND..K...E--N..TG..Q..K.E.....N.
 ..G.D....AT..RQ....P....A..NN.....TS.....NQ--N.K.T.Q.GR.D.E.....E--D.G--D.

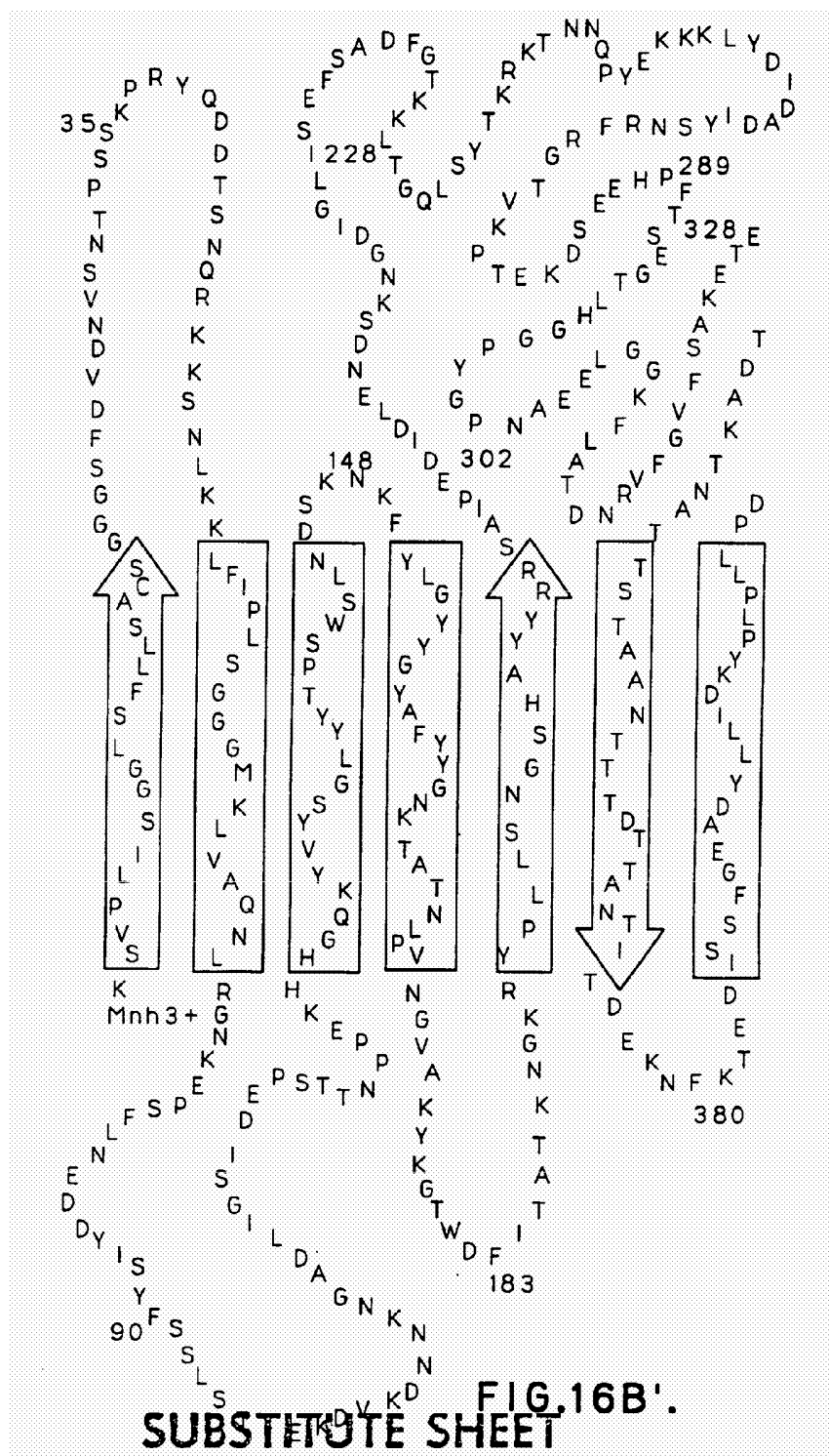
FIG.15D.

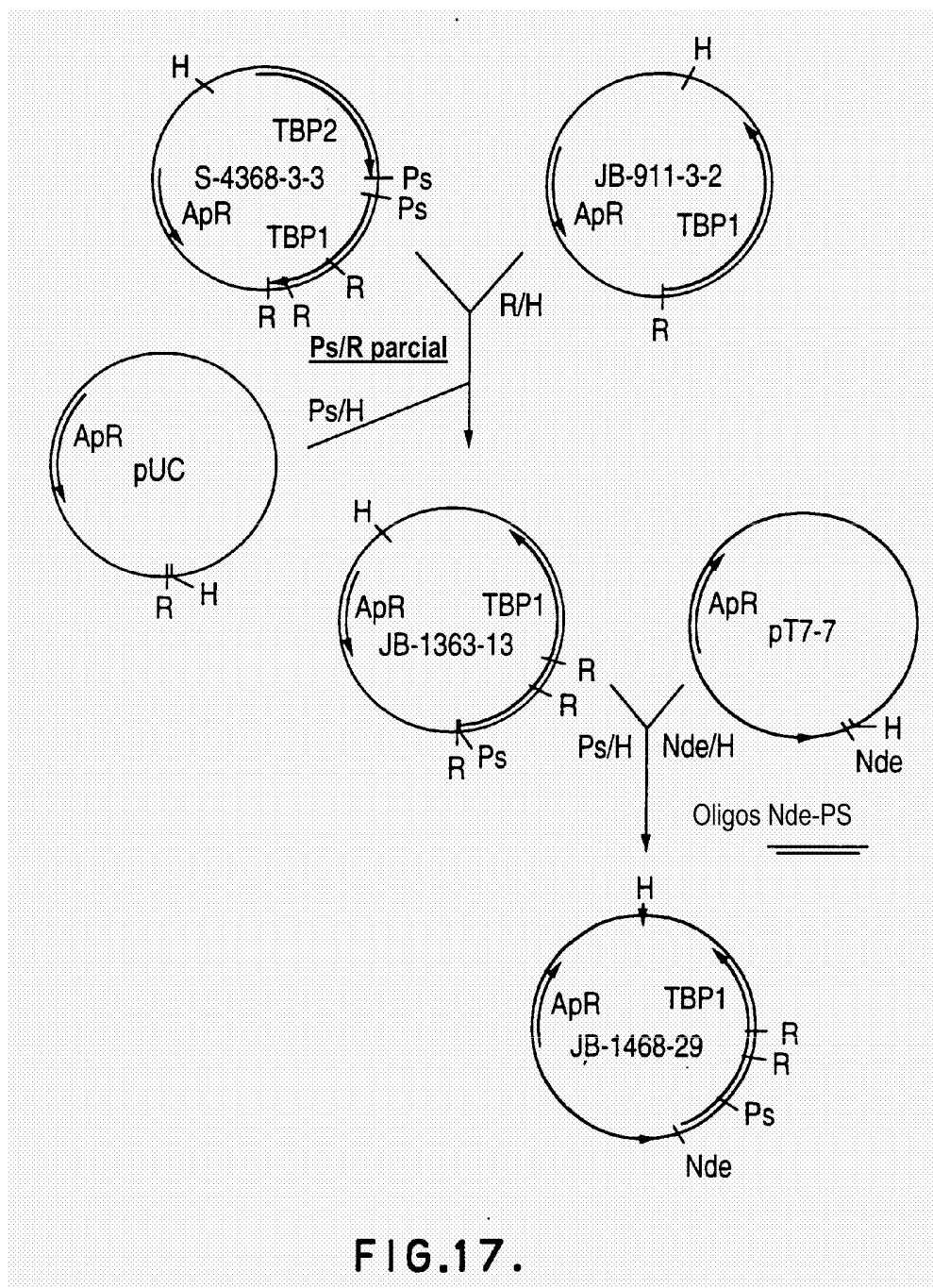
..AN.-..W.GEASNOEGG.-R...D...ST...IS.T.TAK.RT.S.A.T.T.MIKD--G.S.V.KTGENG.AL.PQ.TG.SHYTHI-EAT.S.G...KN. B16B6
H.AN.-..W.GNASD.EGG.-R...T.N.D..I..K.TAENRO-AQT.T..GMIQG--G.E...KTAESG.DL.Q..TTRTPKAYITDA..K.G..... M982
R.AN.-..W.GKASNAT.G.-R.K.T.N.DR.EI..T.TAENRS-EAT.T.D.MIEG--G.K...KTG.DG.AP.QN..TVTHKVVHIANAE.Q.G...N. FA19
AQVSKENNVATA.DD.KSGYRT..D...GN.N.S.K.LFDKN.V...TVD.KIDG--G...K.KTSDEG.AL.SGS.RYE.VKF.DVA--S.G...T. AP205
ALVSKG.NWIAEA.NN.ESGYRT..D.N.SD..VN.K.-FDKG.V...TVD.TI.G--G.I.S.KTSDSG.AL.AGS..HG.AVFSDI-.....G...T. AP37

..LGGYFTYNGN-STATNSESSSTVSSSSNSKNARA...VFGAR-QQVETT-K* EAGAN
..NPTDRN...EK.....* DL63
..KNP.....P.PP.P.S.....KK.....N.* PAK
..PTDK...P.....* SB12
..NPTDRN...P..A.....KK...N.* SB29
..KNP.DK.....P.PP.P.....KK...KNN.* SB30
..KDTITK.T.....P.PP.P.....KK.....N.* SB32
..M..S.SFP..APEGKQE-----K.S...KR..LVQ* B16B6
E...W.A.P.DKQ.EKAT-----AT..DGNSASS.T...KR..PVQ* M982
..W.A.P..EQ.KNA-----E.GNGNSASS.T...KR.KLVK* FA19
..Q.HHKSENGSVGA-----K...KK* AP205
..Q.HHKSDNGSVGA-----K-R.I.K* AP37

SECRET







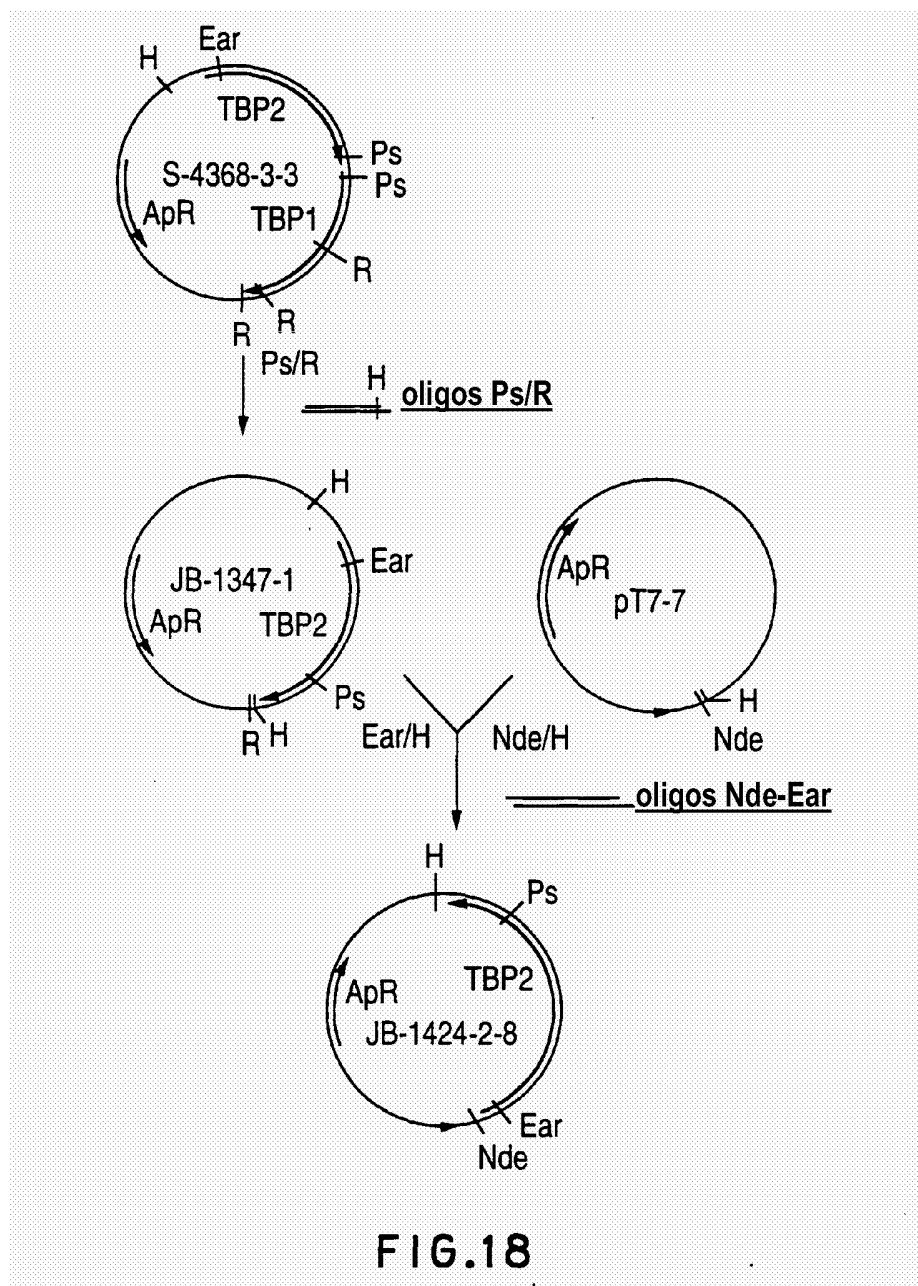


FIG.19.

Oligonucleótidos para expressar TBP2 sem sequência de sinal

Nde I	Ear I
TATGTTCTGGTGGTTCTTTCGACGTTGACAACGTTTCTAACACTCCCCTTCT	
ACACAAGACCAACCAAGAAAGCTGCAACTGTTGCAAAGATTGTGAGGGAGAAGATT	

O codão de iniciação ATG está sublinhado

A cisteína TGT da proteína madura está duplamente sublinhada

FIG.20A.

Sequência de pares de oligonucleótidos (A, B, C e D) para a construção de plasmídeos de expressão de TBP1 e TBP2

Par de oligonucleótidos A (Seq. ID 86 e 87) para unir o promotor de T7 e o gene de TBP1 de Eagan

Nde I
TATGGAAGCTCAAAGTATAAAAGATACAAAAGAAGCTATATCATCTGAAGT...
ACCTTTGAGTTTCATATTTTCTATGTTTCTTCGATATAGTACTTCA...
...GGACACTCAAAGTACAGAAGATTCAGAATTAGAAACTATCTCAGTCACTGCA
...CCTGTGAGTTTCATGTCTTCTAAGTCTTAATCTTTGATAGAGTCAGTG
Pst I

Par de oligonucleótidos B (Seq. ID 88 e 89) para unir o promotor de T7 e os genes de TBP2 de Eagan através do comando lpp de E. coli

Nde I
TATGAAAGCTACTAAAGTGTCTGGGTGCTGTATCTCTGGGTTCCTCACTCTG...
ACTTTCGATGATTTGACCAAGACCCACGACAATAGGACCCCAAGGTGAGAC...
Ear I
...CTGGCTGGT~~TTGT~~AGCGGAGGTGGTGTGTTTGTGATGTAGATAACGTCCTAAATACCCCTCTTCT
...GACCCACCAACATCGCCCTCCACCAACAAACTACATCTATTGCAGAGATTATGGGGGAGAAGATT

FIG.20B.

Par de oligonucleótidos C (Seq. ID 90 e 91) para unir o promotor de T7 e os genes de TBP2 de Eagan através do comando rlp B de *E. coli*

Nde I

TATGCGATATCTGGCAACATTGTTGTTATCTCTGGCGGTGTTAATCACCGGTG...
ACGCTATAGACCGTGTAAACAACAATAGAGACCGCCACAATTAGTGGCGAC...

Ear I

...GTTGTAGCGGAGGTGGTCTTTTGTGATGTAGATAACGTCTCTAATACCCCTCTTCT
...CAACATCGCCTCCACCAGAAACTACATCTATTGCAGAGATTATGGGGAGAAGATTT

Par de oligonucleótidos D (Seq. ID 92 e 93) para unir o promotor de T7 e os genes de TBP2 de Eagan através do comando pal de *E. coli*

Nde I

TATGCAACTGAACAAAGTGCTGAAAGGGCTGATGATTGCTCTGCCTGTTATGGCAA...
ACGTTGACTTGTTTCACGACTTCCCGACTACTAACGAGACGGACAATACCGTT...

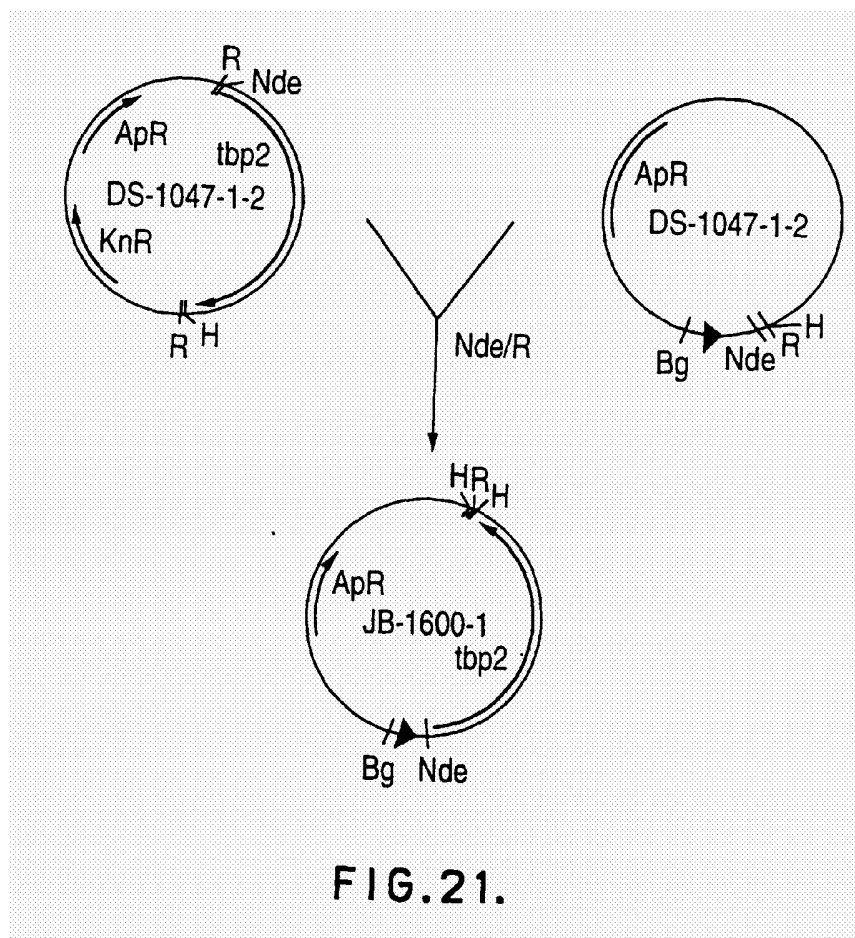
Ear I

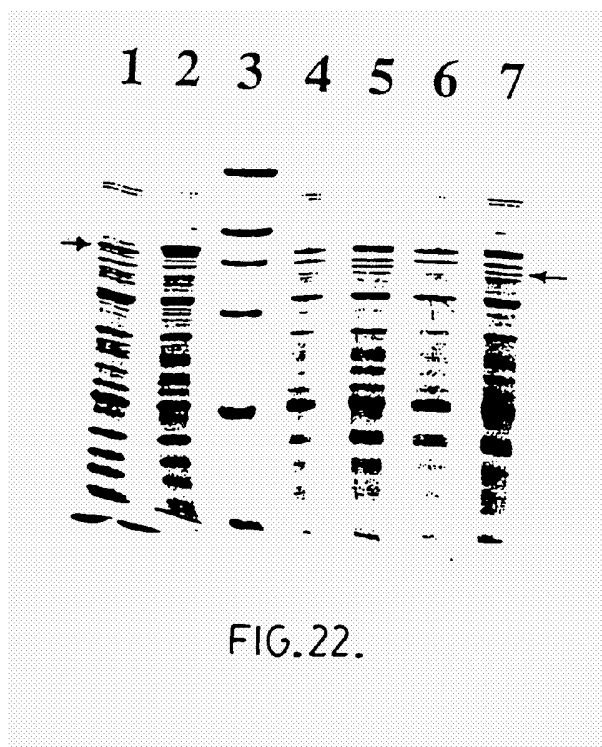
...TTGCTGGTTGTAGCGGAGGTGGTCTTTTGTGATGTAGATAACGTCTCTAATACCCCTCTTCT
...AACGACCAACATCGCCTCCTACCAGAAACTACATCTATTGCAGAGATTATGGGGAGAAGATTT

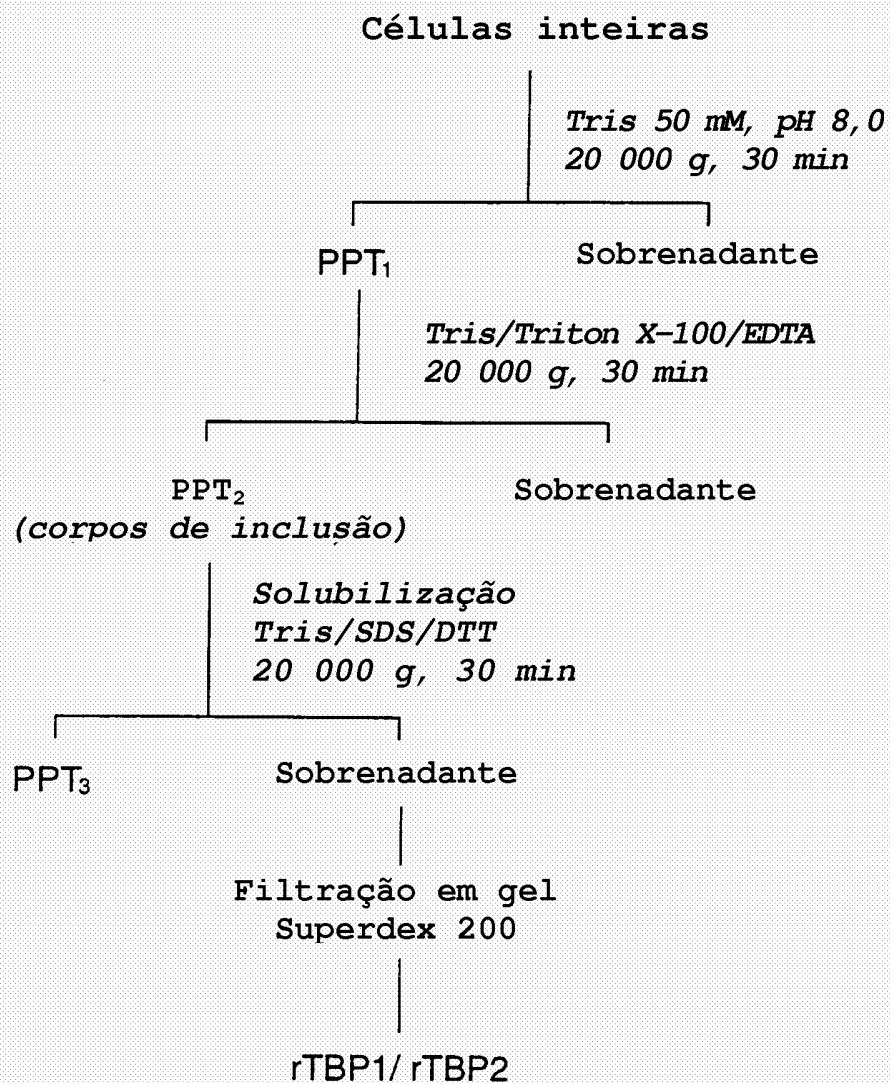
Os códons de iniciação ATG estão sublinhados

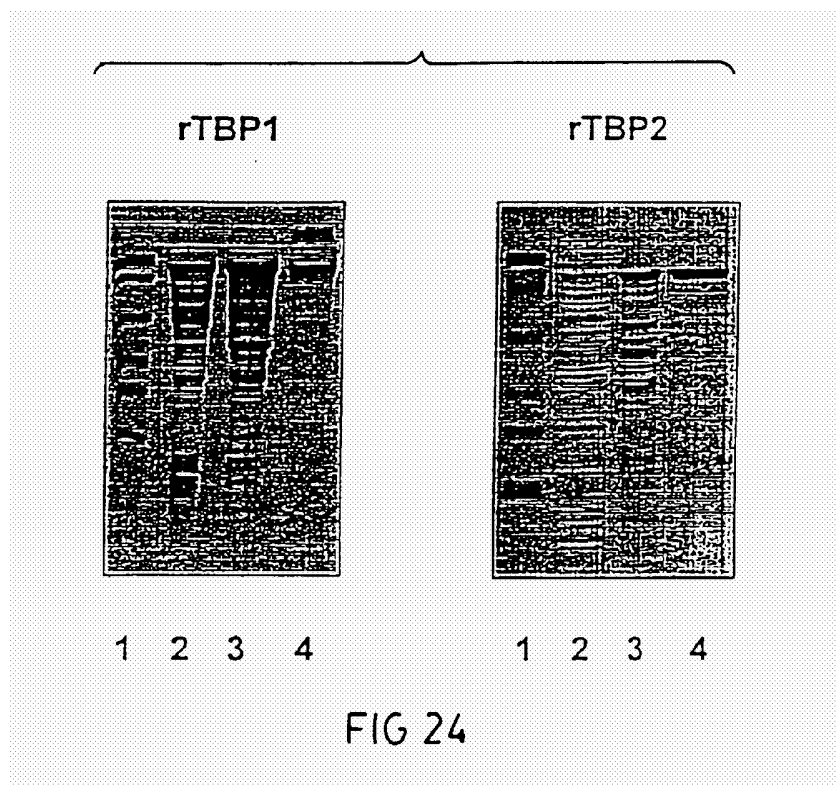
TGT, que codifica para Cys, de lipoproteínas, estão duplamente sublinhados

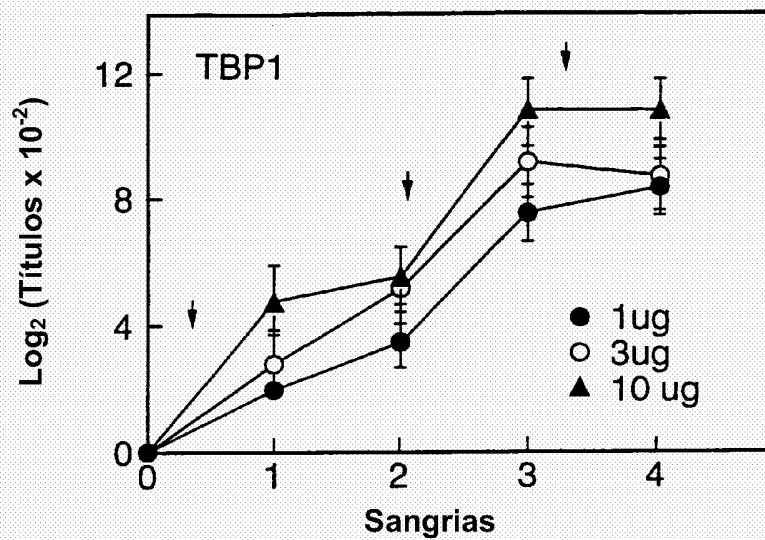
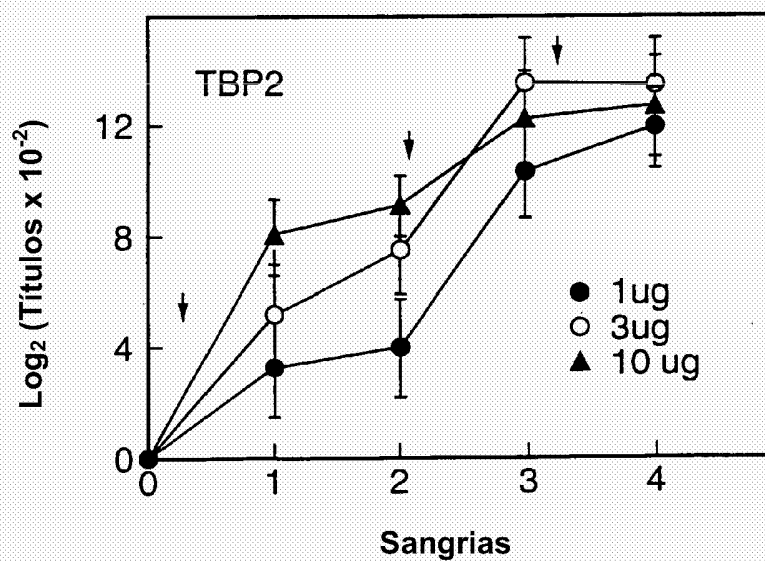
O TGT pode estar substituído por TCC, que codifica para Ser, para segregar não lipoproteínas

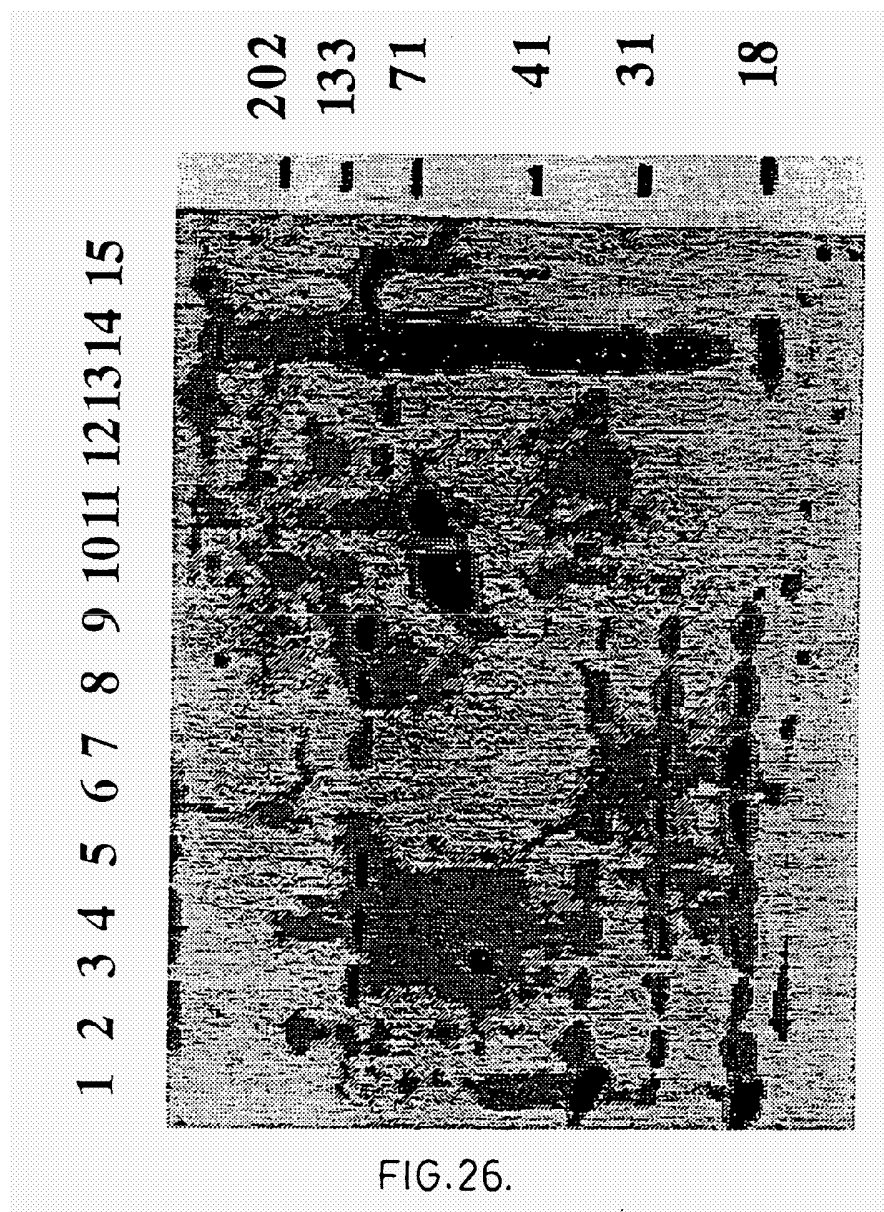


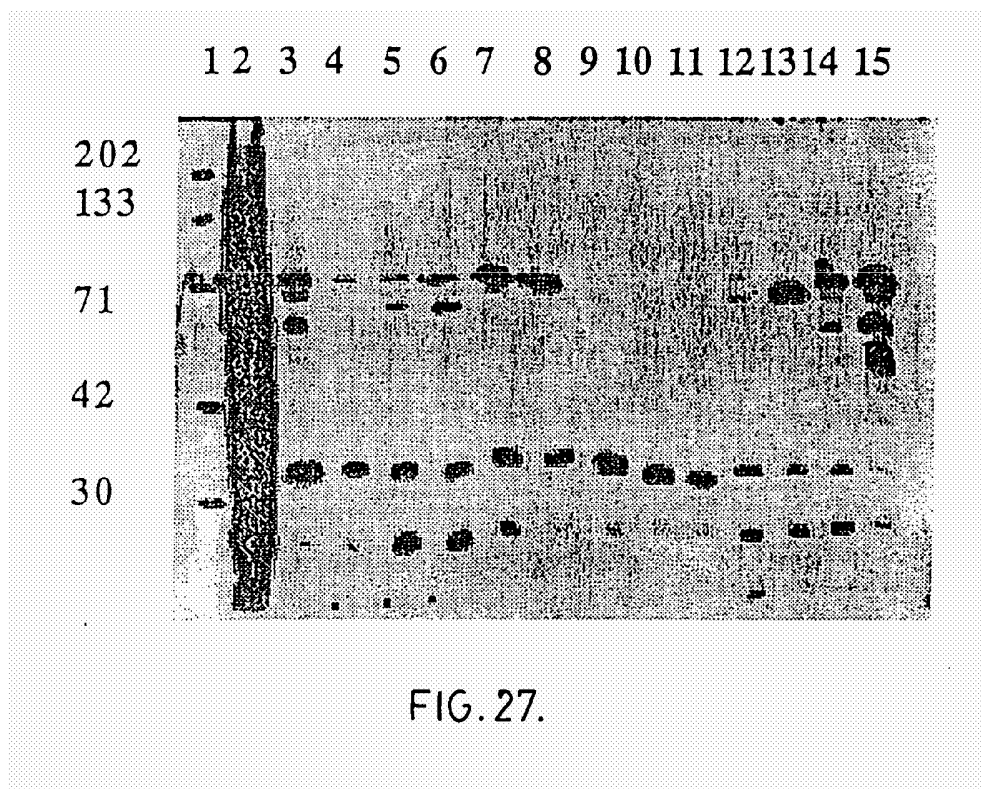


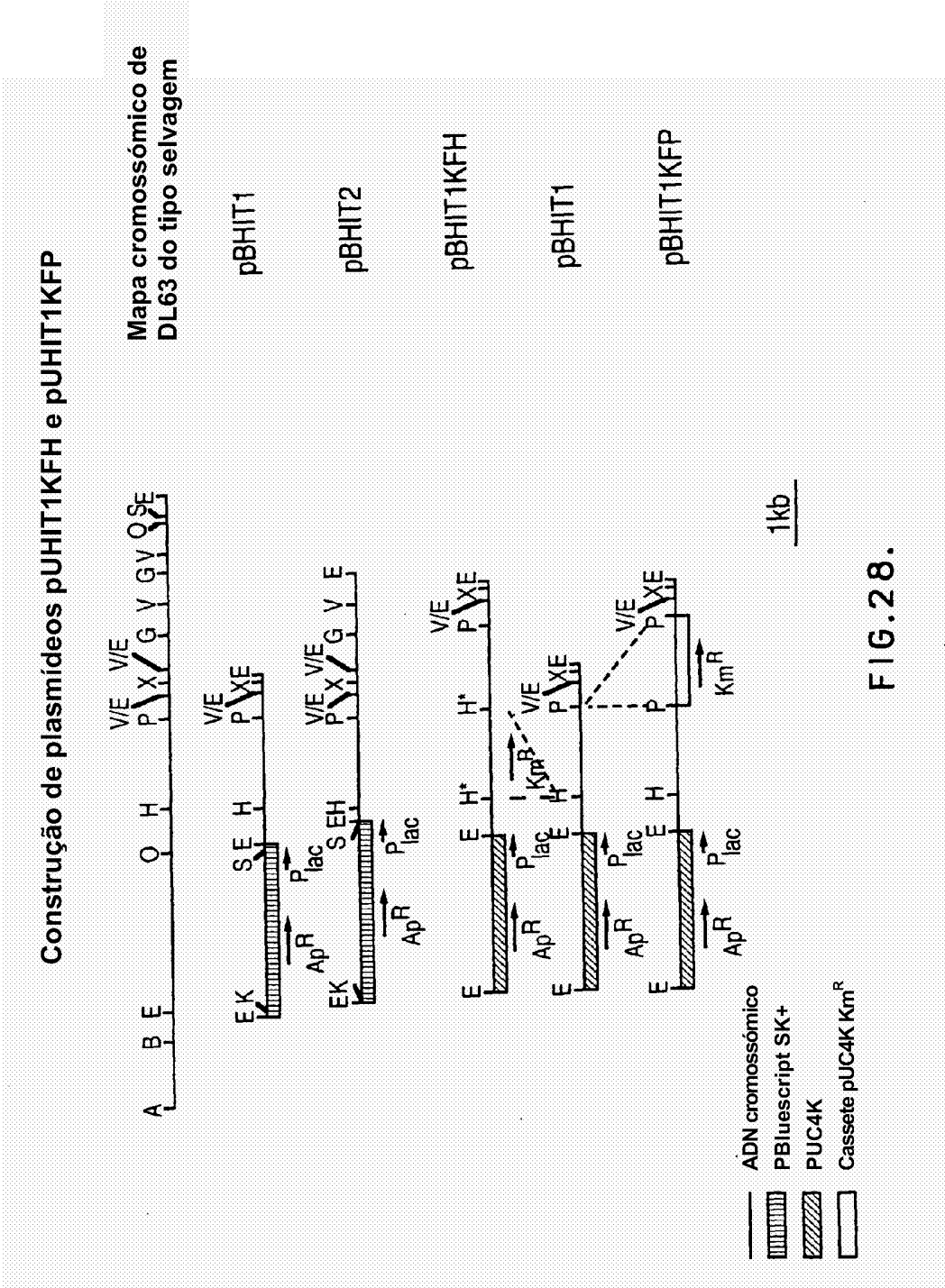
PURIFICAÇÃO DE rTBP1/rTBP2 DE *E. coli***FIG.23.**



Cinética da Resposta de Anticorpos a TBP1/TBP2 em Ratinhos**FIG.25 A.****FIG.25 B.**







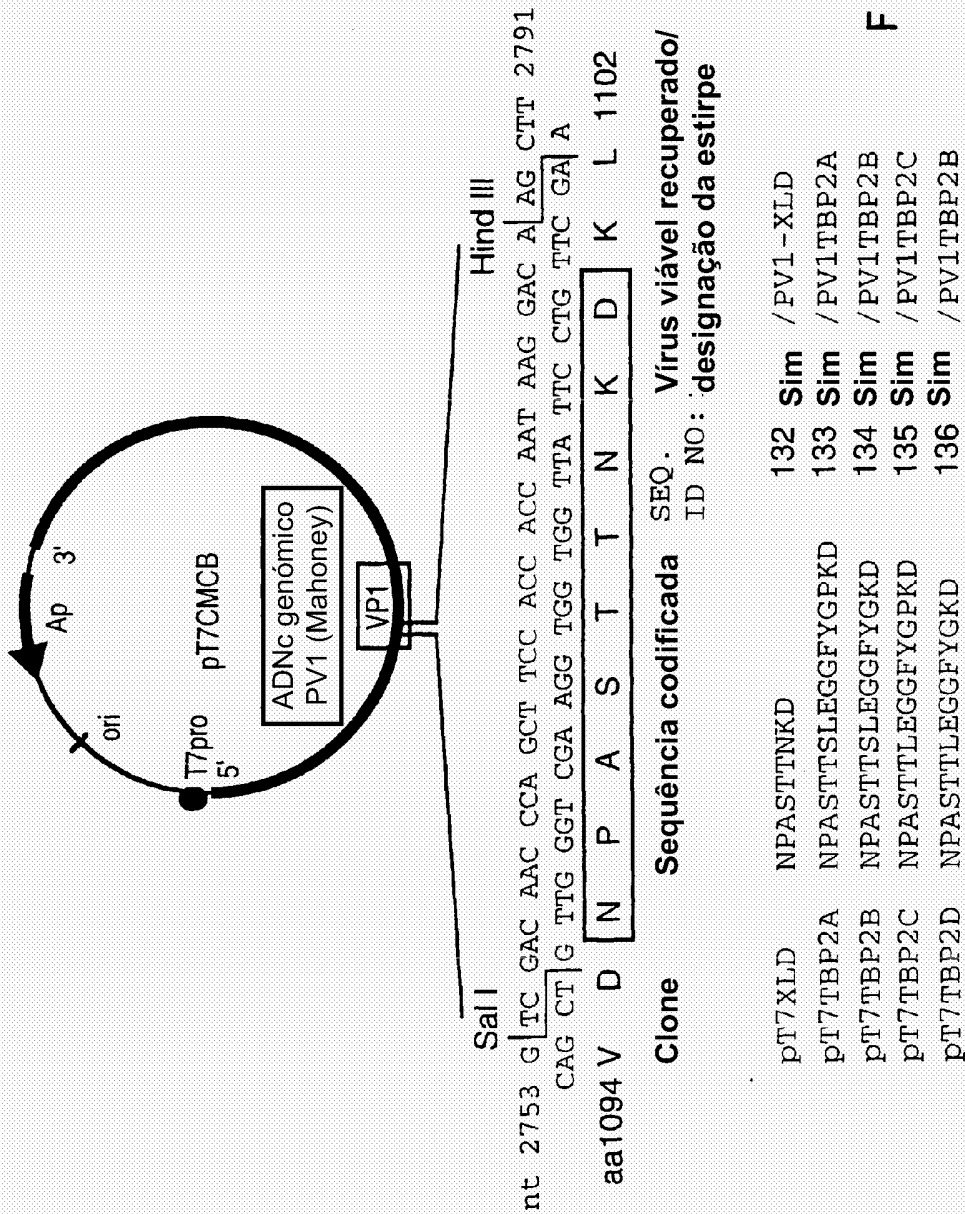


FIG.29.

FIG.30.

