

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 668 967**

51 Int. Cl.:

C07K 14/245 (2006.01)

C12N 15/31 (2006.01)

A61K 39/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.02.2009** **PCT/IB2009/000440**

87 Fecha y número de publicación internacional: **27.08.2009** **WO09104092**

96 Fecha de presentación y número de la solicitud europea: **23.02.2009** **E 09712707 (0)**

97 Fecha y número de publicación de la concesión europea: **21.03.2018** **EP 2254903**

54 Título: **Inmunógenos de Escherichia coli con solubilidad mejorada**

30 Prioridad:

22.02.2008 US 30902 P
09.07.2008 IT MI20081249

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.05.2018

73 Titular/es:

GLAXOSMITHKLINE BIOLOGICALS SA (100.0%)
Rue de l'Institut 89
1330 Rixensart, BE

72 Inventor/es:

SERINO, LAURA;
PIZZA, MARIAGRAZIA;
GOMES MORIEL, DANILO y
FONTANA, MARIA, RITA

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 668 967 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Inmunógenos de *Escherichia coli* con solubilidad mejorada

Campo de la invención

Esta invención se refiere a la inmunización contra cepas patógenas de *Escherichia coli*.

5 **Antecedentes de la técnica**

Las cepas de *E. coli* se han clasificado tradicionalmente como comensales o patógenas, y las cepas patógenas se subclasifican como cepas intestinales o extraintestinales. Las cepas patógenas de *E. coli* se discuten con más detalle en la referencia 1, y se incluyen en una serie de patotipos diferentes, es decir, un grupo de cepas de *E. coli* que causan una enfermedad común usando un conjunto común de factores de virulencia. La patotipificación de cepas es una técnica de rutina que se puede realizar genotípicamente o fenotípicamente. Un procedimiento de patotipificación reciente basado en el genotipo [2] utiliza un microchip de ADN.

Entre las cepas intestinales se conocen al menos seis patotipos bien descritos: enteropatógena (EPEC), enterohemorrágica (EHEC), enteroagregativa (EAEC), enteroinvasiva (EIEC), enterotoxigénica (ETEC) y adherente de forma difusa (DAEC).

15 Las cepas patógenas extraintestinales (o cepas 'ExPEC' [3,4]) de *E. coli* incluyen cepas uropatógenas (UPEC), cepas de meningitis neonatal (NMEC) y cepas asociadas a la septicemia (SEPEC). Las cepas ExPEC son la causa más común de infecciones del tracto urinario y una de las principales causas de meningitis neonatal y sepsis neonatal en los seres humanos, que puede conducir a complicaciones graves y la muerte. Otros tipos de infecciones extraintestinales incluyen osteomielitis, infecciones pulmonares, intraabdominales, de tejidos blandos y asociadas a dispositivos intravasculares. Otro patotipo de ExPEC fuera de los humanos es el patógeno aviar (APEC), que causa infecciones extraintestinales en las aves de corral.

La mayoría de las vacunas contra las ExPEC anteriores se han basado en lisados celulares o en estructuras celulares. SOLCOUROVAC™ incluye diez bacterias muertas por calor diferentes, incluidas seis cepas ExPEC. URO-VAXOM™ es una vacuna en forma de comprimido oral que contiene lisados bacterianos liofilizados de 18 cepas seleccionadas de *E. coli*. Baxter Vaccines desarrolló una vacuna UTI basada en pili de 6 a 10 cepas diferentes. MedImmune desarrolló un producto llamado MEDI 516 basado en el complejo de adhesina FimH. Por el contrario, las referencias 5 y 6 desvelan inmunógenos específicos de cepas ExPEC que pueden usarse como base de vacunas definidas contra cepas NMEC y UPEC. Schneider y col. (Infection and Immunity, 2004) describe la caracterización del determinante de la cápsula K15 de la cepa 536 de *E. coli*. Es un objeto de la invención proporcionar antígenos adicionales y mejores para su uso en la inmunización contra cepas patógenas de *E. coli* y más particularmente contra patotipos intestinales (p.ej., cepas EAEC, EIEC, EPEC y ETEC), así como patotipos ExPEC.

Divulgación de la invención

35 Uno de los muchos antígenos desvelados en la referencia 5 está descrito como el precursor del factor de colonización accesorio D (AcfD) (SEQ ID NOs: 7051 y 7052 en dicha referencia, SEQ ID NOs: 1 y 2 en el presente documento). La referencia 5 desvela la secuencia de la cepa NMEC IHE3034, y la presente invención se basa en variantes del 'precursor del AcfD' de la ExPEC que se han identificado en otros patotipos, que incluyen cepas APEC, UPEC, EAEC, EIEC, EPEC y ETEC. A diferencia de la divulgación de la referencia 5, estas variantes pueden ser particularmente útiles para tratar patotipos intestinales. Por lo tanto, la invención proporciona tales variantes, junto con su uso en la inmunización de pacientes contra las infecciones por *E. coli*. Además, esta divulgación incluye fragmentos de la proteína AcfD de todos los patotipos de *E. coli* en los que el fragmento ha aumentado la solubilidad en comparación con la longitud total a la vez que aumenta una respuesta inmunitaria sustancialmente similar en un sujeto como la provocada por la proteína de longitud completa. La presente invención está definida por las reivindicaciones adjuntas. Cualquier objeto en la solicitud que exceda el ámbito de la invención como se representa por las reivindicaciones no forma parte de la invención reivindicada, sino que solo sirve como información básica para comprender mejor la invención.

Polipéptidos utilizados con la invención

La invención proporciona un polipéptido que comprende una secuencia de aminoácidos que:

- 50 (a) es idéntica a una cualquiera de las SEQ ID NOs: 5 a 16;
 (b) tiene de 1 a 10 alteraciones de un solo aminoácido en comparación con (a);
 (c) tiene al menos un 95 % de identidad de secuencia con una cualquiera de las SEQ ID NOs: 5 a 16; y/o
 (d) es un fragmento de al menos 250 aminoácidos consecutivos de cualquiera de las SEQ ID NOs: 5 a 16; o
 (e) tiene al menos un 95 % de identidad con cualquiera de las SEQ ID NOs: 5 a 16 y comprende un fragmento de al menos 10 aminoácidos consecutivos de cualquiera de las SEQ ID NOs: 5 a 16;

con la condición de que dicha secuencia de aminoácidos no sea la SEQ ID NO: 2.

La divulgación también proporciona un polipéptido que comprende una secuencia de aminoácidos que:

- (a) es idéntica (es decir, 100 % idéntica) a una cualquiera de las SEQ ID NOs 3 a 16;
- (b) tiene al menos un % de identidad de secuencia con una o más de las SEQ ID NOs 3 a 16;
- 5 (c) es un fragmento de al menos b aminoácidos consecutivos de una o más de las SEQ ID NOs 3 a 16;
- (d) tiene 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 (o más) alteraciones de un solo aminoácido (deleciones, inserciones, sustituciones), que pueden estar en ubicaciones separadas o pueden ser contiguas, en comparación con las secuencias de (a) o (b); y/o
- 10 (e) cuando se alinea con una cualquiera de las SEQ ID NOs 3 a 16 usando un algoritmo de alineación por pares, cada ventana móvil de x aminoácidos desde el extremo N al extremo C (de modo que para una alineación que se extiende hasta p aminoácidos, donde $p > x$, existen $p-x+1$ de tales ventanas) tiene al menos $x \cdot y$ aminoácidos alineados idénticos, donde: x se selecciona de 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200; y se selecciona de 0,50, 0,60, 0,70, 0,75, 0,80, 0,85, 0,90, 0,91, 0,92, 0,93, 0,94, 0,95, 0,96, 0,97, 0,98, 0,99; y si $x \cdot y$ no es un número entero, entonces se redondea al número entero más cercano. El algoritmo preferido de alineación por pares es el algoritmo de alineación global de Needleman-Wunsch [7], que utiliza parámetros por defecto (p.ej., con una penalización por apertura de hueco = 10,0, y con una penalización por extensión de hueco = 0,5, usando la matriz de puntuación EBLOSUM62). Este algoritmo se implementa convenientemente en la herramienta *needle* en el paquete EMBOSS [8].
- 15
- 20 Estos polipéptidos incluyen variantes de las SEQ ID NOs 3 a 16, que incluyen variantes alélicas, formas polimorfas, homólogos, ortólogos, parálogos, mutantes, etc.
- El valor de a se puede seleccionar entre 50 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 87,5 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,5 % o más.
- El valor de b puede seleccionarse entre 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250 o más. Los fragmentos preferidos comprenden un epítipo o fragmento inmunógeno de las SEQ ID NOs 3 a 16. Otros fragmentos preferidos carecen de uno o más aminoácidos (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C y/o uno o más aminoácidos (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N de las SEQ ID NOs 3 a 16, preferiblemente mientras se retiene al menos un epítipo o fragmento inmunógeno de las SEQ ID NOs 3 a 16. Otros fragmentos omiten uno o más dominios de proteína p.ej., omisión de un péptido señal, de un dominio citoplásmico, de un dominio transmembrana, de un dominio extracelular, etc. Una deleción en el extremo N, hasta e incluyendo la secuencia GGGSG (es decir en la SEQ ID NO: 3, deleción de aminoácidos de los aminoácidos 1 a 30), proporciona un fragmento útil (véase, p.ej., SEQ ID NOs 99-113. Como se demuestra en el presente documento, tales deleciones aumentan la solubilidad del polipéptido AcfD a la vez que se conserva sustancialmente la misma inmunogenicidad.
- 25
- 30 Otro fragmento útil de la invención está formado por la escisión de una de las SEQ ID NOs: 3 a 16 alrededor de su región rica en Arg (p.ej., restos 770-775 de la SEQ ID NO: 3). Por ejemplo, uno de dichos fragmentos tiene, dentro o en los 20 aminoácidos desde su extremo C, una secuencia que tiene al menos un % de identidad con los aminoácidos 760-769 de la SEQ ID NO: 3; otro fragmento de este tipo tiene, dentro o en los 20 aminoácidos desde su extremo N, una secuencia que tiene al menos un % de identidad con los aminoácidos 776-785 de la SEQ ID NO: 3. Los fragmentos en dirección 3' de la región rica en Arg son particularmente útiles.
- 35
- 40 La invención también proporciona un polipéptido inmunógeno que comprende un fragmento de una proteína AcfD de *E. coli* en la que el fragmento contiene una deleción respecto a la proteína AcfD de *E. coli* que aumenta la solubilidad del fragmento en comparación con la proteína de longitud completa y en la que el fragmento provoca una respuesta inmunitaria sustancialmente similar en un sujeto a la de la proteína AcfD de *E. coli*.
- La invención también proporciona un polipéptido inmunógeno que comprende un fragmento de una proteína AcfD de *E. coli* en la que el fragmento contiene una deleción respecto a la proteína AcfD de *E. coli* que aumenta la solubilidad del fragmento en comparación con la proteína de longitud completa y en la que el fragmento provoca una respuesta inmunitaria sustancialmente similar en un sujeto a la de la proteína AcfD de *E. coli*, en el que
- 45
- (a) la proteína AcfD de *E. coli* tiene una secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NOs: 2-16; y/o
- 50 (b) la deleción es la eliminación de sustancialmente todos los aminoácidos N terminales hasta el enlazador gly-ser, la eliminación de la totalidad o una parte de la repetición rica en prolina N terminal, o ambas; y/o
- el fragmento de polipéptido inmunógeno comprende una secuencia de aminoácidos que comprende:
- (c) la secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOs 99 a 128;
- (d) de 1 a 10 alteraciones de un solo aminoácido en comparación con las SEQ ID NOs: 99 a 128; y/o

(e) al menos un 90 % de identidad de secuencia con una cualquiera de las SEQ ID NOs: 99 a 128.

La divulgación también proporciona un polipéptido que comprende una secuencia de aminoácidos que:

(a) es idéntica (es decir, 100 % idéntica) a una cualquiera de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128;

(b) tiene al menos un % de identidad de secuencia con una o más de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128;

(c) es un fragmento de al menos b aminoácidos consecutivos de una o más de las SEQ ID NOs 3 a 16; SEQ ID NOs 99-113, o SEQ ID NOs 114-128;

(d) tiene 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 (o más) alteraciones de un solo aminoácido (deleciones, inserciones, sustituciones), que pueden estar en ubicaciones separadas o pueden ser contiguas, en comparación con las secuencias de (a) o (b); y/o

(e) cuando se alinea con una cualquiera de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128 usando un algoritmo de alineación por pares, cada ventana móvil de x aminoácidos desde el extremo N al extremo C (de modo que para una alineación que se extiende hasta p aminoácidos, donde $p > x$, existen $p-x+1$ de tales ventanas) tiene al menos $x \cdot y$ aminoácidos alineados idénticos, donde: x se selecciona de 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200; y se selecciona de 0,50, 0,60, 0,70, 0,75, 0,80, 0,85, 0,90, 0,91, 0,92, 0,93, 0,94, 0,95, 0,96, 0,97, 0,98, 0,99; y si $x \cdot y$ no es un número entero, entonces se redondea al número entero más cercano. El algoritmo preferido de alineación por pares es el algoritmo de alineación global de Needleman-Wunsch [9], que utiliza parámetros por defecto (p.ej., con una penalización por apertura de hueco = 10,0, y con una penalización por extensión de hueco = 0,5, usando la matriz de puntuación EBLOSUM62). Este algoritmo se implementa convenientemente en la herramienta *needle* en el paquete EMBOSS [10].

Estos polipéptidos incluyen variantes de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, y SEQ ID NOs 114-128, que incluyen variantes alélicas, formas polimorfas, homólogos, ortólogos, parálogos, mutantes, etc. El valor de a se puede seleccionar entre 50 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 87,5 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, 99,5 % o más.

El valor de b puede seleccionarse entre 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250 o más. Los fragmentos preferidos comprenden un epítipo o fragmento inmunógeno de SEQ ID NOs 3 a 16, SEQ ID NOs 99-113 y SEQ ID NOs 114-128. Otros fragmentos preferidos carecen de uno o más aminoácidos (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C y/o uno o más aminoácidos (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128, preferiblemente a la vez que se retienen al menos un epítipo o fragmento inmunógeno de SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128. Otros fragmentos omiten uno o más dominios de proteína, p.ej., omisión de un péptido señal, de un dominio citoplásmico, de un dominio transmembrana, de un dominio extracelular, etc. Una deleción en el extremo N, hasta e incluyendo la secuencia GGGSG (es decir, en la SEQ ID NO: 3, deleción de aminoácidos de los aminoácidos 1 a 30), proporciona un fragmento útil (véase, p.ej., SEQ ID NOs 99-113. La eliminación de aproximadamente 90 aminoácidos en el extremo N, para eliminar la región rica en prolina, también es útil (véase, p.ej., SEQ ID NOs 114-128). Como se demuestra en el presente documento, tales deleciones aumentan la solubilidad del polipéptido AcfD a la vez que conservan sustancialmente la misma inmunogenicidad.

Un epítipo dentro de un fragmento puede ser un epítipo de linfocito B y/o un epítipo de linfocito T. Dichos epítipos se pueden identificar empíricamente (p.ej., utilizando PEPSCAN [11, 12] o procedimientos similares), o pueden predecirse (p.ej., utilizando el índice antigénico de Jameson-Wolf [13], enfoques basados en matriz [14], MAPITOPE [15], TEPITOPE [16,17], redes neuronales [18], OptiMer y EpiMer [19, 20], ADEPT [21], Tsites [22], hidrofiliidad [23], índice antigénico [24] o los procedimientos desvelados en las referencias 25-29, etc.) Los epítipos son las partes de un antígeno que se reconocen por y se unen a los sitios de unión al antígeno de los anticuerpos o receptores de linfocitos T y se unen a ellos, y también pueden denominarse "determinantes antigénicos".

Los fragmentos inmunógenos de las SEQ ID NOs 3 a 16 discutidos anteriormente incluyen, sin limitación, fragmentos inmunógenos que, cuando se administran a un sujeto en una composición adecuada que puede incluir un adyuvante (incluyendo sin limitación cualquiera de los coadyuvantes enumerados o discutidos en la sección "Composiciones inmunógenas y medicamentos" a continuación), o un vehículo adecuado acoplado al polipéptido, induce una respuesta inmunitaria mediada por anticuerpos o linfocitos T que reconoce el polipéptido de longitud completa aislado de las SEQ ID NOs 3 a 16, respectivamente, del cual se deriva el fragmento inmunógeno.

Los fragmentos particularmente útiles incluyen, pero no se limitan a, SEQ ID NOs: 17 a 95.

Los fragmentos inmunógenos de SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128 discutidos anteriormente incluyen, sin limitación, fragmentos inmunógenos que, cuando se administran a un sujeto en una composición adecuada que puede incluir un adyuvante (incluyendo sin limitación cualquiera de los adyuvantes enumerados o discutidos en la sección "Composiciones inmunógenas y medicamentos" a continuación), o un vehículo adecuado acoplado al polipéptido, induce una respuesta inmunitaria mediada por anticuerpos o linfocitos T que reconoce el polipéptido aislado de longitud completa de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128, respectivamente, de las cuales se deriva el fragmento inmunógeno.

Los fragmentos inmunógenos de SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-131, SEQ ID NOs 133-135, o SEQ ID NOs 137-139 incluyen, sin limitación, fragmentos inmunógenos que, cuando se administran a un sujeto en una composición adecuada que puede incluir un adyuvante (que incluye sin limitación cualquiera de los adyuvantes enumerados o discutidos en la sección "Composiciones inmunógenas y medicamentos" a continuación), o un vehículo adecuado acoplado al polipéptido, induce una respuesta inmunitaria mediada por anticuerpos o linfocitos T que reconoce el polipéptido aislado de longitud completa de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-131, SEQ ID NOs 133-135, o SEQ ID NOs 137-139, respectivamente, de las cuales se deriva el fragmento inmunógeno.

Un polipéptido de la invención puede, en comparación con uno cualquiera de las SEQ ID NOs 3 a 16, incluye uno o más (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.) sustituciones de aminoácidos, tales como sustituciones conservativas (es decir, sustituciones de un aminoácido por otro que tiene una cadena lateral relacionada). Los aminoácidos genéticamente codificados generalmente se dividen en cuatro familias: (1) ácidos, es decir aspartato, glutamato; (2) básicos, es decir, lisina, arginina, histidina; (3) no polares, es decir, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano; y (4) polares sin carga, es decir glicina, asparagina, glutamina, cisteína, serina, treonina, tirosina. La fenilalanina, el triptófano y la tirosina a veces se clasifican conjuntamente como aminoácidos aromáticos. En general, la sustitución de aminoácidos individuales dentro de estas familias no tiene un efecto importante sobre la actividad biológica.

Un polipéptido de la invención también puede, en comparación con uno cualquiera de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128, SEQ ID NOs 129-131, SEQ ID NOs 133-135, o SEQ ID NOs 137-139 incluye uno o más (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.) sustituciones de aminoácidos, tales como sustituciones conservativas (es decir, sustituciones de un aminoácido por otro que tiene una cadena lateral relacionada). Los aminoácidos genéticamente codificados generalmente se dividen en cuatro familias: (1) ácidos, es decir aspartato, glutamato; (2) básicos, es decir, lisina, arginina, histidina; (3) no polares, es decir, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano; y (4) polares sin carga, es decir glicina, asparagina, glutamina, cisteína, serina, treonina, tirosina. La fenilalanina, el triptófano y la tirosina a veces se clasifican conjuntamente como aminoácidos aromáticos. En general, la sustitución de aminoácidos individuales dentro de estas familias no tiene un efecto importante sobre la actividad biológica.

Un polipéptido puede incluir una o más (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.) deleciones de un solo aminoácido con respecto a una cualquiera de las SEQ ID NOs 3 a 16. De manera similar, un polipéptido puede incluir una o más (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.) inserciones (p.ej., cada una de 1, 2, 3, 4 o 5 aminoácidos) con respecto a una cualquiera de las SEQ ID NOs 3 a 16.

Un polipéptido también puede incluir una o más (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.) deleciones de un solo aminoácido con respecto a una cualquiera de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128. De manera similar, un polipéptido puede incluir una o más (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.) inserciones (p.ej., cada una de 1, 2, 3, 4 o 5 aminoácidos) con respecto a una cualquiera de las SEQ ID NOs 3 a 16.

Un polipéptido puede incluir adicionalmente una o más (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.) deleciones de un solo aminoácido con respecto a una cualquiera de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128, SEQ ID NOs 129-131, SEQ ID NOs 133-135 o SEQ ID NOs 137-139. De manera similar, un polipéptido puede incluir una o más (p.ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, etc.) inserciones (p.ej., cada una de 1, 2, 3, 4 o 5 aminoácidos) con respecto a una cualquiera de las SEQ ID NOs 3 a 16, 129, 133 o 137.

Dentro del grupo (c) de cualquiera de los anteriores, las deleciones o sustituciones pueden estar en el extremo N y/o el extremo C, o pueden estar entre los dos extremos. Por lo tanto, un truncamiento es un ejemplo de una deleción. Los truncamientos pueden implicar la deleción de hasta 40 (o más) aminoácidos en el extremo N y/o el extremo C. Como se ha mencionado anteriormente, por ejemplo, se puede usar el truncamiento para eliminar el extremo N hasta la secuencia GGGSG.

En general, cuando un polipéptido de la invención comprende una secuencia que no es idéntica a una secuencia completa de las SEQ ID NOs 3 a 16 (p.ej., cuando comprende una secuencia con <100 % de identidad de secuencia con la misma, o cuando comprende un fragmento de la misma), se prefiere que el polipéptido pueda provocar un anticuerpo que reconozca un polipéptido que consiste en la secuencia SEQ ID completa, es decir el anticuerpo se une a una o más de dichas SEQ ID NOs 3 a 16. Dicho anticuerpo se puede unir específicamente a las SEQ ID NOs 3 a 16, respectivamente, aunque no se une a proteínas que no con AcfD con una afinidad significativamente mayor que la afinidad no específica del anticuerpo a la albúmina sérica humana como patrón de referencia de unión no específica.

De manera similar, cuando un polipéptido de la invención comprende una secuencia que no es idéntica a una secuencia completa de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128 (p.ej., cuando comprende una secuencia con <100 % de identidad de secuencia con la misma, o cuando comprende un fragmento de la misma), se prefiere que el polipéptido pueda provocar un anticuerpo que reconozca un polipéptido que consiste en la secuencia SEQ ID completa es decir el anticuerpo se une a una o más de dichas SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID NOs 114-128. Tal anticuerpo se puede unir específicamente a las SEQ ID NOs 3 a 16, SEQ

ID NOs 99-113, o SEQ ID NOs 114-128, respectivamente, aunque no se une a proteínas que no son AcfD con una afinidad significativamente mayor que la afinidad no específica del anticuerpo a la albúmina sérica humana como patrón de referencia de unión no específica.

- 5 Adicionalmente, cuando un polipéptido de la invención comprende una secuencia que no es idéntica a una secuencia completa de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, SEQ ID NOs 114-128, SEQ ID NOs 129-131, SEQ ID NOs 133-135, o SEQ ID NOs 137-139 (p.ej., cuando comprende una secuencia con <100 % de identidad de secuencia con la misma, o cuando comprende un fragmento de la misma), se prefiere que el polipéptido pueda provocar un anticuerpo que reconozca un polipéptido que consiste en la secuencia SEQ ID completa es decir el anticuerpo se une a una o más de dichas SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, SEQ ID NOs 114-128, SEQ ID NOs 129-131, SEQ ID NOs 133-135 o SEQ ID NOs 137- 139. Tal anticuerpo se puede unir específicamente a las
- 10 SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, SEQ ID NOs 114-128, SEQ ID NOs 129-131, SEQ ID NOs 133-135, o SEQ ID NOs 137-139, aunque no se une a proteínas que no son AcfD con una afinidad significativamente mayor que la afinidad no específica del anticuerpo a la albúmina sérica humana como patrón de referencia de unión no específica.
- 15 En una realización, la divulgación proporciona un polipéptido que comprende una secuencia de aminoácidos: (a) que tiene al menos un % de identidad con una cualquiera de las SEQ ID NOs 3 a 16; y (b) que comprende un fragmento de al menos *b* aminoácidos consecutivos de dicha SEQ ID.
- En otra realización, la divulgación proporciona un polipéptido que comprende una secuencia de aminoácidos: (a) que tiene al menos un % de identidad con una cualquiera de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, o SEQ ID
- 20 NOs 114-128; y (b) que comprende un fragmento de al menos *b* aminoácidos consecutivos de dicha SEQ ID.
- En otra realización más, la divulgación proporciona un polipéptido que comprende una secuencia de aminoácidos: (a) que tiene al menos un % de identidad con una cualquiera de las SEQ ID NOs 3 a 16, SEQ ID NOs 99-113, SEQ ID NOs 114-128, SEQ ID NOs 129-131, SEQ ID NOs 133-135, o SEQ ID NOs 137-139; y (b) que comprende un fragmento de al menos *b* aminoácidos consecutivos de dicha SEQ ID.
- 25 Un polipéptido de la invención puede incluir un ion metálico p.ej., un ion metálico que está coordinado por uno o más aminoácidos en la cadena polipeptídica. Por ejemplo, el polipéptido puede incluir un catión metálico monovalente, divalente o trivalente. Los cationes divalentes son típicos, tales como Mn^{2+} , Fe^{2+} , Co^{2+} , Ni^{2+} , Cu^{2+} , etc. El catión divalente es preferiblemente Zn^{2+} . El ion puede coordinarse mediante una secuencia de aminoácidos HEAGH o HEVGH.
- 30 Los polipéptidos usados con la invención pueden adoptar diversas formas (p.ej., nativa, fusiones, glicosilada, no glicosilada, lipídada, no lipídada, fosforilada, no fosforilada, miristoilada, no miristoilada, monomérica, multimérica, particulada, desnaturalizada, etc.). Por ejemplo, un polipéptido de la invención puede tener una cisteína N terminal lipídada (p.ej., Cys-24 de las SEQ ID NOs: 3 a 16).
- Los polipéptidos usados con la invención pueden prepararse por diversos medios (p.ej., expresión recombinante, purificación del cultivo celular, síntesis química, etc.). Se prefieren las proteínas expresadas recombinantemente.
- 35 Los polipéptidos usados con la invención se proporcionan preferiblemente en forma purificada o sustancialmente purificada es decir sustancialmente exenta de otros polipéptidos (p.ej., exenta de polipéptidos naturales), particularmente de otros polipéptidos de *E. coli* o de células hospedadoras, y son generalmente al menos aproximadamente 50 % puros (en peso), y habitualmente al menos aproximadamente 90 % puros, es decir, menos de aproximadamente 50 %, y más preferiblemente menos de aproximadamente 10 % (p.ej., 5 %) de una composición está compuesta por otros polipéptidos expresados. Por lo tanto, los antígenos en las composiciones se separan del organismo completo con el que se expresa la molécula.
- Los polipéptidos usados con la invención son preferiblemente polipéptidos de *E. coli*. Dichos polipéptidos pueden seleccionarse adicionalmente entre polipéptidos de *E. coli* NMEC, APEC, UPEC, EAEC, EIEC, EPEC y ETEC.
- 45 El término "polipéptido" se refiere a polímeros de aminoácidos de cualquier longitud. El polímero puede ser lineal o ramificado, puede comprender aminoácidos modificados y puede estar interrumpido por no aminoácidos. Los términos también abarcan un polímero de aminoácido que se ha modificado naturalmente o por intervención; por ejemplo, formación de enlaces disulfuro, glicosilación, lipidación, acetilación, fosforilación o cualquier otra manipulación o modificación, tal como la conjugación con un componente de marcaje. También se incluyen, por
- 50 ejemplo, polipéptidos que contienen uno o más análogos de un aminoácido (que incluyen, por ejemplo, aminoácidos no naturales, etc.), así como otras modificaciones conocidas en la técnica. Los polipéptidos pueden ocurrir como cadenas simples o cadenas asociadas.
- La divulgación proporciona polipéptidos que comprenden una secuencia -P-Q- o -Q-P-, en la que: -P- es una secuencia de aminoácidos como se definió anteriormente y -Q- no es una secuencia como se definió anteriormente es decir la divulgación proporciona proteínas de fusión. Cuando el codón N terminal de -P- no es ATG, pero este codón no está presente en el extremo N de un polipéptido, se traducirá como el aminoácido estándar para ese codón en lugar de como un Met. Sin embargo, cuando este codón está en el extremo N de un polipéptido, se
- 55

traducirá como Met. Los ejemplos de restos Q incluyen, pero no se limitan a, etiquetas de histidina (es decir, His_n, donde $n = 3, 4, 5, 6, 7, 8, 9, 10$ o más), una proteína de unión a maltosa o glutatión-S-transferasa (GST).

La divulgación también proporciona una proteína oligomérica que comprende un polipéptido de la invención. El oligómero puede ser un dímero, un trímero, un tetramero, etc. El oligómero puede ser un homo-oligómero o un hetero-oligómero. Los polipéptidos en el oligómero pueden estar asociados de manera covalente o no covalente.

La invención también proporciona polipéptidos de *E. coli* que son fragmentos de AcfD de longitud completa (de los cuales las SEQ ID NOs 2-16 son ejemplos representativos) que tienen solubilidad incrementada respecto a la proteína de longitud completa a la vez que provocan una respuesta inmunitaria sustancialmente similar en un sujeto que la provocada por la proteína de longitud completa. Los ejemplos de dichos fragmentos polipeptídicos inmunógenos incluyen una cualquiera de las SEQ ID NOs 99-128. El aumento de la solubilidad puede medirse por cualquier medio disponible para un experto en la materia. Un procedimiento simple implica la sobreexpresión del fragmento en bacterias y el análisis de muestras comparativas de lisado bacteriano total frente al sobrenadante de lisado bacteriano después de la centrifugación o muestras de sedimento de lisado bacteriano después de la centrifugación frente a muestras de sobrenadante de lisado bacteriano después de la centrifugación. Un experto en la materia podría cultivar y expresar dichos fragmentos polipeptídicos inmunógenos usando técnicas estándar (por ejemplo, transformar bacterias BL21(DE3) con un vector de expresión pET21 que exprese el fragmento, cultivar las bacterias a razón de 0,6 DO₆₀₀ en LB e inducir con IPTG 1 mM y cultivar durante 3 horas después de la inducción), tales muestras se pueden analizar en SDS PAGE (por ejemplo, MOPS 4-12 %) y cuantificarse aproximadamente escaneando el gel teñido resultante y midiendo el tamaño relativo de las bandas. El aumento de solubilidad como se usa en el presente documento se determina a 25 °C. Tal solubilidad aumentada puede ser un aumento del 10 % en polipéptido soluble, un aumento del 20 % en polipéptido soluble, un aumento del 30 % en polipéptido soluble, un aumento del 50 % en polipéptido soluble, un aumento del 75 % en polipéptido soluble, un aumento del 100 % (es decir, dos veces) en polipéptido soluble, un aumento de tres veces en polipéptido soluble, un aumento de cuatro veces en polipéptido soluble, un aumento de cinco veces en polipéptido soluble, un aumento de siete veces en polipéptido soluble, o un aumento de diez veces en polipéptido soluble.

La invención además proporciona polipéptidos de *E. coli* que son fragmentos de AcfD de un tamaño menor que la longitud completa (de los cuales las SEQ ID NOs 2-16 son ejemplos representativos). Estos fragmentos de longitud inferior a la de longitud completa tienen una solubilidad incrementada en comparación con la proteína de longitud completa a la vez que provocan una respuesta inmunitaria sustancialmente similar en un sujeto en comparación con la provocada por la proteína de longitud completa. Los ejemplos de dichos fragmentos polipeptídicos inmunógenos incluyen cualquiera de las SEQ ID NOs 99-128. El aumento de la solubilidad puede medirse por cualquier medio disponible para un experto en la materia. Un procedimiento simple implica la sobreexpresión del fragmento en bacterias y el análisis de muestras comparativas de lisado bacteriano total frente al sobrenadante de lisado bacteriano después de la centrifugación o muestras de sedimento de lisado bacteriano después de la centrifugación frente a muestras de sobrenadante de lisado bacteriano después de la centrifugación. Un experto en la materia podría cultivar y expresar dichos fragmentos polipeptídicos inmunógenos usando técnicas estándar (por ejemplo, transformar bacterias BL21 (DE3) con un vector de expresión pET21 que exprese el fragmento, cultivar las bacterias a razón de 0,6 DO₆₀₀ en LB e inducir con IPTG 1 mM, y cultivar durante 3 horas después de la inducción). Dichas muestras se pueden analizar en SDS PAGE (por ejemplo, MOPS 4-12 %) y se pueden cuantificar aproximadamente escaneando el gel teñido resultante y midiendo el tamaño relativo de las bandas. El aumento de solubilidad como se usa en el presente documento se determina a 25 °C. Tal solubilidad aumentada puede ser un aumento del 10 % en polipéptido soluble, un aumento del 20 % en polipéptido soluble, un aumento del 30 % en polipéptido soluble, un aumento del 50 % en polipéptido soluble, un aumento del 75 % en el polipéptido soluble, un aumento del 100 % (es decir, dos veces) en polipéptido soluble, un aumento de tres veces en polipéptido soluble, un aumento de cuatro veces en polipéptido soluble, un aumento de cinco veces en polipéptido soluble, un aumento de siete veces en polipéptido soluble, o un aumento de diez veces en polipéptido soluble en cada caso en comparación con el polipéptido de longitud inferior en comparación con la proteína de longitud completa correspondiente medida en las mismas condiciones.

La invención además proporciona polipéptidos de *E. coli* que son fragmentos de un tamaño menor que la longitud completa de la AcfD de longitud completa (de los cuales las SEQ ID NOs 2-16, 129, 133 y 137 son ejemplos representativos). Estos fragmentos de longitud inferior a la de longitud completa tienen una solubilidad incrementada en comparación con la proteína de longitud completa a la vez que provocan una respuesta inmunitaria sustancialmente similar en un sujeto en comparación con la provocada por la proteína de longitud completa. Los ejemplos de dichos fragmentos polipeptídicos inmunógenos incluyen una cualquiera de las SEQ ID NOs 99-128, SEQ ID NOs 130-131, SEQ ID NOs 134-135 y SEQ ID NOs 138-139. El aumento de la solubilidad puede medirse por cualquier medio disponible para un experto en la materia. Un procedimiento simple implica la sobreexpresión del fragmento en bacterias y el análisis de muestras comparativas de lisado bacteriano total frente al sobrenadante de lisado bacteriano después de la centrifugación o muestras de sedimento de lisado bacteriano después de la centrifugación frente a muestras de sobrenadante de lisado bacteriano después de la centrifugación. Un experto en la materia podría cultivar y expresar dichos fragmentos polipeptídicos inmunógenos usando técnicas estándar (por ejemplo, transformar bacterias BL21 (DE3) con un vector de expresión pET21 que exprese el fragmento, cultivar las bacterias a razón de 0,6 DO₆₀₀ en LB e inducir con IPTG 1 mM, y cultivar durante 3 horas después de la inducción). Dichas muestras se pueden analizar en SDS PAGE (por ejemplo, MOPS 4-12 %) y se pueden cuantificar

aproximadamente escaneando el gel teñido resultante y midiendo el tamaño relativo de las bandas. El aumento de solubilidad como se usa en el presente documento se determina a 25 °C. Tal solubilidad aumentada puede ser un aumento del 10 % en polipéptido soluble, un aumento del 20 % en polipéptido soluble, un aumento del 30 % en polipéptido soluble, un aumento del 50 % en polipéptido soluble, un aumento del 75 % en polipéptido soluble, un aumento del 100 % (es decir, dos veces) en polipéptido soluble, un aumento de tres veces en polipéptido soluble, un aumento de cuatro veces en polipéptido soluble, un aumento de cinco veces en polipéptido soluble, un aumento de siete veces en polipéptido soluble, o un aumento de diez veces en polipéptido soluble en cada caso en comparación con el polipéptido de longitud inferior en comparación con la proteína de longitud completa correspondiente medida en las mismas condiciones.

La comparación de la respuesta inmunitaria provocada en un sujeto por el polipéptido con la respuesta inmunitaria provocada por la proteína de longitud completa puede llevarse a cabo mediante cualquier medio disponible para un experto en la materia. Un procedimiento simple como se usa en los ejemplos a continuación implica la inmunización de un sujeto modelo como el ratón y después la estimulación con una dosis letal de *E. coli*. Para una comparación adecuada, un experto en la materia seleccionará naturalmente el mismo adyuvante, tal como el adyuvante completo de Freund. En dicha prueba, los fragmentos polipeptídicos inmunógenos de la presente invención provocarán una respuesta inmunitaria sustancialmente similar en un sujeto (es decir, proporcionarán sustancialmente la misma protección frente a la estimulación letal) si, por ejemplo, el polipéptido proporciona al menos el 70 % de la protección proporcionada por la proteína de longitud completa, al menos el 80 % de la protección proporcionada por la proteína de longitud completa, al menos el 85 % de la protección proporcionada por la proteína de longitud completa, al menos el 90 % de la protección proporcionada por la proteína de longitud completa, al menos el 95 % de la protección proporcionada por la proteína de longitud completa, al menos el 97 % de la protección proporcionada por la proteína de longitud completa, al menos el 98 % de la protección proporcionada por la proteína de longitud completa o al menos el 99 % de la protección proporcionada por la proteína de longitud completa.

Además, la comparación entre la respuesta inmunitaria provocada en un sujeto por el polipéptido de longitud inferior a la completa y la respuesta inmunitaria provocada por la proteína de longitud completa correspondiente se puede llevar a cabo por cualquier medio disponible para un experto en la materia. Un procedimiento simple, tal como se usa en los ejemplos a continuación, implica la inmunización de un sujeto modelo como el ratón y la posterior estimulación con una dosis letal de *E. coli*. Para una comparación adecuada, un experto en la materia seleccionará naturalmente el mismo adyuvante, tal como el adyuvante completo de Freund. En dicha prueba, los fragmentos polipeptídicos inmunógenos de longitud inferior a los de la presente invención provocarán una respuesta inmunitaria sustancialmente similar en un sujeto en comparación con la respuesta inmunitaria provocada por la proteína de longitud completa correspondiente (es decir, proporcionará sustancialmente la misma protección contra la estimulación letal) si, por ejemplo, el polipéptido proporciona al menos el 70 % de la protección proporcionada por la proteína de longitud completa, al menos el 80 % de la protección proporcionada por la proteína de longitud completa, al menos el 85 % de la protección proporcionada por la proteína de longitud completa, al menos el 90 % de la protección proporcionada por la proteína de longitud completa, al menos el 95 % de la protección proporcionada por la proteína de longitud completa, al menos el 97 % de la protección proporcionada por la proteína de longitud completa, al menos el 98 % de la protección proporcionada por la proteína de longitud completa, o al menos el 99 % de la protección proporcionada por la proteína de longitud completa.

La proteína de AcfD de longitud completa contra la cual se compararía el fragmento de polipéptido inmunógeno (tanto para la solubilidad como para la respuesta inmunitaria provocada) puede ser cualquier proteína AcfD de *E. coli* representativa que incluye, sin limitación, las SEQ ID NOs 2-16. En realizaciones preferidas, la proteína AcfD será la proteína de longitud completa correspondiente a partir de la cual se obtiene el fragmento de polipéptido inmunógeno.

En algunas realizaciones, el polipéptido inmunógeno contendrá una delección respecto a la proteína AcfD de *E. coli* que da como resultado una mayor solubilidad. La delección puede incluir la eliminación de sustancialmente todos los aminoácidos N terminales hasta el enlazador gly-ser o la región gly-ser, la eliminación de la totalidad o una parte de la repetición rica en prolina N terminal, o ambas. Un experto en la materia entendería que los aminoácidos N terminales hasta el enlazador gly-ser o la región gly-ser que se corresponden a la región de la proteína AcfD de *E. coli* de interés es la porción de la proteína que se alinea con la región de las proteínas AcfD de *E. coli* identificadas en el presente documento y designada con una "G" debajo de la alineación en la Figura 1. De manera similar, un experto en la materia entendería que la repetición rica en prolina N terminal que corresponde a la región de la proteína AcfD de *E. coli* de interés es la porción de la proteína que se alinea con la región de la proteína AcfD de *E. coli* identificadas en el presente documento que se indican con una "P" bajo la alineación en la Figura 1.

En ciertos aspectos, el fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 99 o 114 (incluyendo cualquier fragmento que abarca la SEQ ID NO: 114 con un extremo N entre el extremo N de las SEQ ID NOs 99 y 114) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 2. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 100 o 115 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 115 con un extremo N entre el extremo N de las SEQ ID NOs 100 y 115), siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de SEQ ID NO 3. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprenda las SEQ ID NOs 101 o 116 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 116 con un extremo N entre el extremo N de las SEQ ID NOs 101 y 116), siempre que el fragmento de polipéptido inmunógeno

no tenga la secuencia de aminoácidos de la SEQ ID NO 4. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 102 o 117 (incluyendo cualquier fragmento que comprende la SEQ ID NO: 117 con un extremo N entre el extremo N de las SEQ ID NOs 102 y 117) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 5. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 103 o 118 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 118 con un extremo N entre el extremo N de las SEQ ID NOs 103 y 118) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 6. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 104 o 119 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 119 con un extremo N entre el extremo N de las SEQ ID NOs 104 y 119) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 7. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 105 o 120 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 120 con un extremo N entre el extremo N de las SEQ ID NOs 105 y 120) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 8. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 106 o 121 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 121 con un extremo N entre el extremo N de las SEQ ID NOs 106 y 121) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 9. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 107 o 122 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 122 con un extremo N entre el extremo N de las SEQ ID NOs 107 y 122) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 10. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 108 o 123 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 123 con un extremo N entre el extremo N de las SEQ ID NOs 108 y 123) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 11. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 109 o 124 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 124 con un extremo N entre el extremo N de las SEQ ID NOs 109 y 124) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 12. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 110 o 125 (incluyendo cualquier fragmento que incluye la SEQ ID NO: 125 con un extremo N entre el extremo N de las SEQ ID NOs 110 y 125) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 13. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 111 o 126 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 126 con un extremo N entre el extremo N de las SEQ ID NOs 111 y 126) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 14. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 112 o 127 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 127 con un extremo N entre el extremo N de las SEQ ID NOs 112 y 127) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 15. El fragmento de polipéptido inmunógeno puede ser un fragmento que comprende las SEQ ID NOs 113 o 128 (incluyendo cualquier fragmento que abarque la SEQ ID NO: 128 con un extremo N entre el extremo N de las SEQ ID NOs 113 y 128) siempre que el fragmento de polipéptido inmunógeno no tenga la secuencia de aminoácidos de la SEQ ID NO 16. Cualquiera de los fragmentos de polipéptidos inmunógenos anteriores también pueden incluir variaciones siempre que las variaciones no den como resultado que el polipéptido inmunógeno tenga la secuencia de cualquier proteína AcfD de longitud completa, incluyendo, sin limitación, cualquiera de las variaciones enumeradas en esta sección "Polipéptidos usados con la invención". Los ejemplos incluyen: de 1 a 10 alteraciones de aminoácidos individuales en comparación con las SEQ ID NOs aplicables; al menos 85 % de identidad de secuencia con las SEQ ID NOs aplicables; un fragmento de al menos 10 aminoácidos consecutivos de las SEQ ID NOs aplicables; y cuando se alinea con las SEQ ID NOs aplicables utilizando un algoritmo de alineación por pares, cada ventana móvil de x aminoácidos desde el extremo N hasta el extremo C tiene al menos x·y aminoácidos alineados idénticos, donde x es 30 e y es 0,75.

La divulgación también proporciona un procedimiento para producir un polipéptido de la invención, que comprende la etapa de cultivar una célula hospedadora transformada con ácido nucleico de la invención en condiciones que inducen la expresión de polipéptidos. El polipéptido puede purificarse a continuación, por ejemplo, a partir de los sobrenadantes de cultivo.

La invención proporciona una célula de *E. coli* que contiene un plásmido que codifica un polipéptido de la invención. El cromosoma de la célula de *E. coli* puede incluir un homólogo de AcfD, o tal homólogo puede estar ausente, pero en ambos casos el polipéptido de la invención puede expresarse a partir del plásmido. El plásmido puede incluir un gen que codifica un marcador, etc. A continuación se dan estos y otros detalles de plásmidos adecuados.

Aunque la expresión de los polipéptidos de la invención puede tener lugar en una cepa de *E. coli*, la invención habitualmente usará un hospedador heterólogo para la expresión. El hospedador heterólogo puede ser procariota (p.ej., una bacteria) o eucariota. Los hospedadores adecuados incluyen, entre otros, *Bacillus subtilis*, *Vibrio cholerae*, *Salmonella typhi*, *Salmonella typhimurium*, *Neisseria lactamica*, *Neisseria cinerea*, *Mycobacteria* (p.ej., *M. tuberculosis*), levaduras, etc.

La divulgación proporciona un procedimiento para producir un polipéptido de la invención, que comprende la etapa de sintetizar al menos parte del polipéptido por medios químicos.

Cualquiera y todas las proteínas, polipéptidos, polipéptidos híbridos, epítomos y fragmentos inmunógenos anteriores pueden estar en cualquiera de una serie de formas que incluyen, sin limitación, recombinantes, aisladas o sustancialmente purificadas (a partir de materiales que coexisten con tales proteínas, polipéptidos, polipéptidos híbridos, epítomos y fragmentos inmunógenos en su estado natural).

5 Ácidos nucleicos

La invención también proporciona ácidos nucleicos que codifican polipéptidos de la invención. También proporciona ácido nucleico que comprende una secuencia de nucleótidos que codifica uno o más polipéptidos de la invención.

La invención también proporciona ácido nucleico que comprende secuencias de nucleótidos que tienen identidad de secuencia con tales secuencias de nucleótidos. La identidad entre las secuencias se determina preferiblemente mediante el algoritmo de búsqueda de homología de Smith-Waterman como se describió anteriormente. Dichos ácidos nucleicos incluyen aquellos que usan codones alternativos para codificar el mismo aminoácido.

La divulgación también proporciona ácido nucleico que puede hibridarse con estos ácidos nucleicos. Las reacciones de hibridación pueden realizarse en condiciones de diferente "rigurosidad". Condiciones que aumentan la rigurosidad de una reacción de hibridación ampliamente conocida y publicada en la técnica (p.ej., página 7.52 de la referencia 224). Entre los ejemplos de condiciones relevantes se incluyen (en orden creciente de rigurosidad): temperaturas de incubación de 25 °C, 37 °C, 50 °C, 55 °C y 68 °C; concentraciones de tampón de 10 x SSC, 6 x SSC, 1 x SSC, 0,1 x SSC (donde SSC es NaCl 0,15 M y tampón citrato 15 mM citrato) y sus equivalentes usando otros sistemas tampón; concentraciones de formamida de 0 %, 25 %, 50 % y 75 %; tiempos de incubación de 5 minutos a 24 horas; 1, 2 o más etapas de lavado; tiempos de incubación de 1, 2 o 15 minutos; y soluciones de lavado de 6 x SSC, 1 x SSC, 0,1 x SSC o agua desionizada. Las técnicas de hibridación y su optimización son bien conocidas en la técnica (p.ej., ver referencias 30, 31, 224, 226, etc.).

En algunas realizaciones, el ácido nucleico de la invención se hibrida con una diana en condiciones de baja rigurosidad; en otras realizaciones hibrida en condiciones de rigurosidad intermedia; en realizaciones preferidas, hibrida en condiciones de alta rigurosidad. Un ejemplo de conjunto de condiciones de hibridación de baja rigurosidad es 50 °C y 10 x SSC. Un ejemplo de conjunto de condiciones de hibridación de rigurosidad intermedia es 55 °C y 1 x SSC. Un ejemplo de conjunto de condiciones de hibridación de alta rigurosidad es 68 °C y 0,1 x SSC.

La invención incluye ácido nucleico que comprende secuencias complementarias a estas secuencias (p.ej., para antisentido o sondas, o para usar como cebadores).

Los ácidos nucleicos de la invención se pueden usar en reacciones de hibridación (p.ej., transferencias Northern o Southern, o en microchips de ácidos nucleicos o 'chips génicos') y reacciones de amplificación (p.ej., PCR, SDA, SSSR, LCR, TMA, NASBA, etc.) y otras técnicas de ácidos nucleicos.

El ácido nucleico de acuerdo con la invención puede adoptar diversas formas (p.ej., monocatenario, bicatenario, vectores, cebadores, sondas, marcado etc.). Los ácidos nucleicos de la invención pueden ser circulares o ramificados, pero generalmente serán lineales. A menos que se especifique o se requiera lo contrario, cualquier realización de la invención que utilice un ácido nucleico puede utilizar tanto la forma bicatenaria como cada una de las dos formas complementarias monocatenarias que constituyen la forma bicatenaria. Los cebadores y las sondas generalmente son monocatenarios, al igual que los ácidos nucleicos antisentido.

Los ácidos nucleicos de la invención se proporcionan preferiblemente en forma purificada o sustancialmente purificada es decir sustancialmente exenta de otros ácidos nucleicos (p.ej., exenta de ácidos nucleicos naturales), particularmente de otros ácidos nucleicos de *E. coli* o de células hospedadoras, siendo generalmente al menos aproximadamente 50 % puros (en peso), y habitualmente al menos aproximadamente 90 % puros. Los ácidos nucleicos de la invención son preferiblemente ácidos nucleicos de *E. coli*.

Los ácidos nucleicos de la invención se pueden preparar de muchas maneras p.ej., por síntesis química (p.ej., síntesis de fosforamida de ADN) en su totalidad o en parte, mediante la digestión de ácidos nucleicos más largos utilizando nucleasas (p.ej., enzimas de restricción), uniendo ácidos nucleicos más cortos o nucleótidos (p.ej., usando ligasas o polimerasas), de bibliotecas genómicas o de ADNc, etc.

El ácido nucleico de la invención se puede unir a un soporte sólido (p.ej., una cuenta, placa, filtro, película, porta, soporte de microchip, resina, etc.). El ácido nucleico de la invención puede marcarse p.ej., con un marcador radioactivo o fluorescente, o un marcador de biotina. Esto es particularmente útil cuando el ácido nucleico se va a usar en técnicas de detección p.ej., cuando el ácido nucleico es un cebador o una sonda.

La expresión "ácido nucleico" incluye en general una forma polimérica de nucleótidos de cualquier longitud, que contienen desoxirribonucleótidos, ribonucleótidos y/o sus análogos. Incluye ADN, ARN, híbridos de ADN/ARN. También incluye análogos de ADN o ARN, como los que contienen cadenas principales modificadas (p.ej., ácidos nucleicos peptídicos (PNA) o fosforotioatos) o bases modificadas. Por lo tanto, la invención incluye ARNm, ARNt, ARNr, ribozimas, ADN, ADNc, ácidos nucleicos recombinantes, ácidos nucleicos ramificados, plásmidos, vectores, sondas, cebadores, etc. Cuando el ácido nucleico de la invención adopta la forma de ARN, puede tener o no un

grupo protector en 5'.

Los ácidos nucleicos de la invención pueden ser parte de un vector es decir parte de una construcción de ácido nucleico diseñada para la transducción/transfección de uno o más tipos de células. Los vectores pueden ser, por ejemplo, "vectores de clonación" diseñados para el aislamiento, la propagación y la replicación de nucleótidos insertados, "vectores de expresión" diseñados para la expresión de una secuencia de nucleótidos en una célula hospedadora, "vectores virales" diseñados para dar como resultado la producción de un virus recombinante o partícula similar a un virus, o "vectores lanzadera", que comprenden los atributos de más de un tipo de vector. Los vectores preferidos son plásmidos, como se menciona anteriormente. Una "célula hospedadora" incluye una célula individual o cultivo celular que puede ser o ha sido receptor de ácido nucleico exógeno. Las células hospedadoras incluyen la progenie de una sola célula hospedadora, y la progenie puede no ser necesariamente completamente idéntica (en morfología o en el complemento de ADN total) a la célula parental original debido a una mutación y/o cambio natural, accidental o deliberado. Las células hospedadoras incluyen células transfectadas o infectadas *in vivo* o *in vitro* con ácido nucleico de la invención.

Cuando un ácido nucleico es ADN, se apreciará que "U" en una secuencia de ARN se reemplazará por "T" en el ADN. De forma similar, cuando un ácido nucleico es ARN, se apreciará que "T" en una secuencia de ADN se reemplazará por "U" en el ARN.

El término "complemento" o "complementario" cuando se usa en relación con ácidos nucleicos se refiere al emparejamiento de bases de Watson-Crick. Así, el complemento de C es G, el complemento de G es C, el complemento de A es T (o U), y el complemento de T (o U) es A. También es posible usar bases como I (la purina inosina) por ejemplo para complementar las pirimidinas (C o T).

Los ácidos nucleicos de la invención se pueden usar, por ejemplo: para producir polipéptidos; como sondas de hibridación para la detección de ácido nucleico en muestras biológicas; para generar copias adicionales de los ácidos nucleicos; para generar ribozimas u oligonucleótidos antisentido; como cebadores o sondas de ADN monocatenarios; o como oligonucleótidos que forman triple cadena.

La divulgación proporciona un procedimiento para producir ácido nucleico de la invención, en el que el ácido nucleico se sintetiza en parte o en su totalidad usando medios químicos.

La divulgación proporciona vectores que comprenden secuencias de nucleótidos de la invención (p.ej., vectores de clonación o expresión) y células hospedadoras transformadas con tales vectores.

La amplificación de ácidos nucleicos de acuerdo con la invención puede ser cuantitativa y/o en tiempo real.

Para ciertas realizaciones de la invención, los ácidos nucleicos son preferiblemente de al menos 7 nucleótidos de longitud (p.ej., 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 275, 300 nucleótidos o más largos).

Para ciertas realizaciones de la invención, los ácidos nucleicos son preferiblemente de 500 nucleótidos de longitud como máximo (p.ej., 450, 400, 350, 300, 250, 200, 150, 140, 130, 120, 110, 100, 90, 80, 75, 70, 65, 60, 55, 50, 45, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15 nucleótidos o más cortos).

Los cebadores y las sondas de la invención, y otros ácidos nucleicos usados para la hibridación, tienen preferiblemente una longitud de entre 10 y 30 nucleótidos (p.ej., 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, o 30 nucleótidos).

Composiciones inmunógenas y medicamentos

Los polipéptidos de la invención son útiles como principios activos (inmunógenos) en composiciones inmunógenas, y tales composiciones pueden ser útiles como vacunas. Las vacunas de acuerdo con la invención pueden ser profilácticas (es decir para prevenir la infección) o terapéuticas (es decir para tratar la infección), pero generalmente serán profilácticas.

Las composiciones inmunógenas serán farmacéuticamente aceptables. Por lo general, incluirán componentes además de los antígenos, p.ej., generalmente incluyen uno o más vehículos, excipientes y/o adyuvantes farmacéuticos. En la ref. 221 se proporciona una relación detallada de los vehículos y excipientes. En las referencias 32 y 33 se proporciona una relación detallada de los adyuvantes de vacunas.

Las composiciones generalmente se administrarán a un mamífero en forma acuosa. Sin embargo, antes de la administración, la composición puede haber estado en una forma no acuosa. Por ejemplo, aunque algunas vacunas se fabrican en forma acuosa, luego se llenan y distribuyen y se administran también en forma acuosa, otras vacunas se liofilizan durante la fabricación y se reconstituyen en una forma acuosa en el momento del uso. Por lo tanto, una composición de la invención puede secarse, tal como una formulación liofilizada.

La composición puede incluir conservantes tales como tiomersal o 2-fenoxietanol. Sin embargo, se prefiere que la vacuna esté sustancialmente exenta de (es decir menos de 5 µg/ml) de material mercurial, p.ej., exento de tiomersal. Las vacunas que no contienen mercurio son más preferidas. Las vacunas sin conservantes son particularmente preferidas.

- 5 Para mejorar la estabilidad térmica, una composición puede incluir un agente protector de la temperatura.

Para controlar la tonicidad, se prefiere incluir una sal fisiológica, tal como una sal de sodio. Se prefiere cloruro sódico (NaCl), que puede estar presente a entre 1 y 20 mg/ml, p.ej., aproximadamente 10 ± 2 mg/ml de NaCl. Otras sales que pueden estar presentes incluyen cloruro de potasio, dihidrógenofosfato de potasio, fosfato de disodio deshidratado, cloruro de magnesio, cloruro de calcio, etc.

- 10 Las composiciones generalmente tendrán una osmolalidad de entre 200 mOsm/kg y 400 mOsm/kg, preferiblemente entre 240-360 mOsm/kg, y estarán más preferiblemente dentro del intervalo de 290-310 mOsm/kg.

Las composiciones pueden incluir uno o más tampones. Los tampones típicos incluyen: un tampón de fosfato; un tampón Tris; un tampón de borato; un tampón de succinato; un tampón de histidina (particularmente con un adyuvante de hidróxido de aluminio); o un tampón de citrato. Los tampones generalmente se incluirán en el intervalo de 5-20 mM.

- 15 El pH de una composición generalmente estará entre 5,0 y 8,1, y más generalmente entre 6,0 y 8,0 p.ej., 6,5 y 7,5, o entre 7,0 y 7,8.

La composición es preferiblemente estéril. La composición es preferiblemente no pirógena, p.ej., contiene <1 UE (unidad de endotoxina, una medida estándar) por dosis, y preferiblemente <0,1 UE por dosis. La composición está preferiblemente exenta de gluten.

- 20 La composición puede incluir material para una sola inmunización, o puede incluir material para múltiples inmunizaciones (es decir un kit 'multidosis'). La inclusión de un conservante se prefiere en las formas multidosis. Como una alternativa (o además) a la inclusión de un conservante en composiciones multidosis, las composiciones pueden estar contenidas en un recipiente que tiene un adaptador aséptico para la extracción de material.

- 25 Las vacunas humanas se administran generalmente en un volumen de dosificación de aproximadamente 0,5 ml, aunque puede administrarse a niños media dosis (es decir aproximadamente 0,25 ml).

Las composiciones inmunógenas de la invención también pueden comprender uno o más agentes inmunorreguladores. Preferiblemente, uno o más de los agentes inmunorreguladores incluyen uno o más adyuvantes. Los adyuvantes pueden incluir un adyuvante TH1 y/o un adyuvante TH2, que se describen adicionalmente a continuación.

- 30

Los adyuvantes que pueden usarse en composiciones de la invención incluyen, pero no están limitados a:

A. Composiciones que contienen minerales

Las composiciones que contienen minerales adecuadas para usar como adyuvantes en la invención incluyen sales minerales, tales como sales de aluminio y sales de calcio (o mezclas de las mismas). Las sales de calcio incluyen fosfato de calcio (p.ej., las partículas "CAP" desveladas en la ref. 34). Las sales de aluminio incluyen hidróxidos, fosfatos, sulfatos, etc., adoptando las sales cualquier forma adecuada (p.ej., gel, cristalina, amorfa, etc.). Se prefiere la adsorción a estas sales. Las composiciones que contienen minerales también se pueden formular como una partícula de sal metálica [35].

- 35

Se pueden usar los adyuvantes conocidos como hidróxido de aluminio y fosfato de aluminio. Estos nombres son convencionales, pero se usan solo por conveniencia, ya que ninguna de las dos es una descripción precisa del compuesto químico real que está presente (p.ej., véase el capítulo 9 de la referencia 32). La invención puede usar cualquiera de los adyuvantes "hidróxido" o "fosfato" que son de uso general como adyuvantes. Los adyuvantes conocidos como "hidróxido de aluminio" son generalmente sales de oxihidróxido de aluminio, las cuales habitualmente son al menos parcialmente cristalinas. Los adyuvantes conocidos como "fosfato de aluminio" son generalmente hidroxifosfatos de aluminio, conteniendo a menudo también una pequeña cantidad de sulfato (es decir sulfato de hidroxifosfato de aluminio). Se pueden obtener por precipitación, y las condiciones de reacción y las concentraciones durante la precipitación influyen en el grado de sustitución del fosfato por hidroxilo en la sal.

- 40
- 45

Una morfología fibrosa (p.ej., como se ve en las micrografías electrónicas de transmisión) es típica de los adyuvantes de hidróxido de aluminio. El pl de los adyuvantes de hidróxido de aluminio es generalmente de aproximadamente 11 es decir el propio adyuvante tiene una carga superficial positiva a pH fisiológico. Se han descrito capacidades de adsorción de entre 1,8-2,6 mg de proteína por mg de Al^{+++} a pH 7,4 para los adyuvantes de hidróxido de aluminio.

- 50

Los adyuvantes de fosfato de aluminio generalmente tienen una relación molar PO_4/Al entre 0,3 y 1,2, preferiblemente entre 0,8 y 1,2, y más preferiblemente $0,95 \pm 0,1$. El fosfato de aluminio generalmente será amorfo,

particularmente para las sales de hidroxifosfato. Un adyuvante típico es el hidroxifosfato de aluminio amorfo con una relación molar PO_4/Al entre 0,84 y 0,92, incluida a razón de 0,6 mg de Al^{3+}/ml . El fosfato de aluminio generalmente será particulado (p.ej., morfología en forma de placa como se ve en las micrografías electrónicas de transmisión). Los diámetros típicos de las partículas están en el intervalo de 0,5-20 μm (p.ej., aproximadamente 5-10 μm) después de cualquier adsorción de antígeno. Se han descrito capacidades de adsorción de entre 0,7-1,5 mg de proteína por mg Al^{+++} a pH 7,4 para los adyuvantes de fosfato de aluminio.

El punto de carga cero (PZC) del fosfato de aluminio está inversamente relacionado con el grado de sustitución del fosfato por hidroxilo, y este grado de sustitución puede variar dependiendo de las condiciones de reacción y la concentración de los reactantes usados para preparar la sal por precipitación. El PZC también se altera cambiando la concentración de iones fosfato libres en solución (más fosfato = PZC más ácido) o añadiendo un tampón como un tampón de histidina (hace que el PZC sea más básico). Los fosfatos de aluminio usados de acuerdo con la invención generalmente tendrán un PZC de entre 4,0 y 7,0, más preferiblemente entre 5,0 y 6,5 p.ej., aproximadamente de 5,7.

Las suspensiones de sales de aluminio usadas para preparar las composiciones de la invención pueden contener un tampón (p.ej., un tampón de fosfato o una histidina o un tampón Tris), pero esto no siempre es necesario. Las suspensiones son preferiblemente estériles y están exentas de pirógenos. Una suspensión puede incluir iones de fosfato acuosos libres p.ej., presentes a una concentración entre 1,0 y 20 mM, preferiblemente entre 5 y 15 mM, y más preferiblemente aproximadamente de 10 mM. Las suspensiones también pueden comprender cloruro de sodio.

La invención puede usar una mezcla de un hidróxido de aluminio y un fosfato de aluminio. En este caso, puede haber más fosfato de aluminio que hidróxido p.ej., una relación en peso de al menos 2:1 p.ej., $\geq 5:1$, $\geq 6:1$, $\geq 7:1$, $\geq 8:1$, $\geq 9:1$, etc.

La concentración de Al^{+++} en una composición para administración a un paciente es preferiblemente menos de 10 mg/ml p.ej., ≤ 5 mg/ml, ≤ 4 mg/ml, ≤ 3 mg/ml, ≤ 2 mg/ml, ≤ 1 mg/ml, etc. Un intervalo preferido está entre 0,3 y 1 mg/ml. Se prefiere un máximo de 0,85 mg/dosis.

B. Emulsiones de aceite

Las composiciones de emulsiones oleosas adecuadas para uso como adyuvantes en la invención incluyen emulsiones de escualeno-agua, tales como MF59 [Capítulo 10 de la ref. 32; véase también la ref. 36] (5 % de escualeno, 0,5 % de Tween 80 y 0,5 % de Span 85, formulado en partículas submicrométricas usando un microfluidizador). También se puede usar el adyuvante de Freund completo (CFA) y el adyuvante de Freund incompleto (IFA).

Se conocen diversos adyuvantes de emulsión de aceite en agua, y estos generalmente incluyen al menos un aceite y al menos un tensioactivo, siendo el aceite o aceites y el tensioactivo o tensioactivos biodegradables (metabolizables) y biocompatibles. Las gotitas de aceite en la emulsión son generalmente de menos de 5 μm de diámetro, e idealmente tienen un diámetro submicrónico, alcanzándose estos tamaños pequeños con un microfluidizador para proporcionar emulsiones estables. Se prefieren las gotitas con un tamaño inferior a 220 nm ya que pueden someterse a esterilización por filtración.

La emulsión puede comprender aceites tales como los de una fuente animal (tal como pescado) o vegetal. Las fuentes de aceites vegetales incluyen frutos secos, semillas y granos. El aceite de cacahuete, el aceite de soja, el aceite de coco y el aceite de oliva, los más comúnmente disponibles, son ejemplos de aceites de frutos secos. Se puede utilizar aceite de jojoba p.ej., el obtenido de la semilla de jojoba. Los aceites de semillas incluyen aceite de cártamo, aceite de semilla de algodón, aceite de semilla de girasol, aceite de semilla de sésamo y similares. En el grupo de los granos de cereales, el aceite de maíz es el más fácilmente obtenible, pero también se puede usar el aceite de otros granos de cereal tales como trigo, avena, centeno, arroz, teff, triticale y similares. Los ésteres de ácidos grasos de 6-10 carbonos de glicerol y 1,2-propanodiol, aunque no se encuentran naturalmente en aceites de semillas, se pueden preparar por hidrólisis, separación y esterificación de los materiales apropiados a partir de los aceites de frutos secos y semillas. Las grasas y los aceites de la leche de mamífero son metabolizables y, por lo tanto, pueden usarse en la práctica de esta invención. Los procedimientos para la separación, purificación, saponificación y otros medios necesarios para obtener aceites puros a partir de fuentes animales son bien conocidos en la técnica. La mayoría de los peces contienen aceites metabolizables que pueden recuperarse fácilmente. Por ejemplo, el aceite de hígado de bacalao, los aceites de hígado de tiburón y el aceite de ballena, como el espermaceti, ejemplifican varios de los aceites de pescado que se pueden usar en el presente documento. Varios aceites de cadena ramificada se sintetizan bioquímicamente en unidades de isopreno de 5 carbonos y generalmente se denominan terpenoides. El aceite de hígado de tiburón contiene un terpenoide ramificado, insaturado, conocido como escualeno, 2,6,10,15,19,23-hexametil-2,6,10,14,12,22-tetracosahexaeno, que es particularmente preferido en el presente documento. El escualano, el análogo saturado del escualeno, también es un aceite preferido. Los aceites de pescado, incluyendo escualeno y escualano, se obtienen fácilmente de fuentes comerciales o se pueden obtener por procedimientos conocidos en la técnica. Otros aceites preferidos son los tocoferoles (ver a continuación). Se pueden usar mezclas de aceites.

Los tensioactivos se pueden clasificar por su "HLB" (equilibrio hidrófilo/lipófilo). Los tensioactivos preferidos de la

invención tienen un HLB de al menos 10, preferiblemente al menos 15, y más preferiblemente al menos 16. La invención se puede usar con tensioactivos que incluyen, pero no se limitan a: los tensioactivos de ésteres de polioxietilensorbitán (comúnmente denominados los Tweens), especialmente polisorbato 20 y polisorbato 80; copolímeros de óxido de etileno (EO), óxido de propileno (PO) y/u óxido de butileno (BO), vendidos bajo el nombre comercial DOWFAX™, tales como copolímeros de bloques EO/PO lineales; octoxinóles, que pueden variar en el número de grupos etoxi(oxi-1,2-etanodiilo) repetidos, siendo de particular interés el octoxinol-9 (Triton X-100 o t-octilfenoxipolietoxietanol); (octilfenoxi)polietoxietanol (IGEPAL CA-630/NP-40); fosfolípidos tales como fosfatidilcolina (lecitina); nonilfenol etoxilados, tales como la serie Tergitol™ NP; éteres grasos de polioxietileno derivados de alcoholes laurílico, cetílico, estearílico y oleílico (conocidos como tensioactivos Brij), tales como trietilenglicol monolauril éter (Brij 30); y ésteres de sorbitán (comúnmente conocidos como los SPAN), tales como el trioleato de sorbitán (Span 85) y monolaurato de sorbitán. Se prefieren tensioactivos no iónicos. Los tensioactivos preferidos para incluir en la emulsión son Tween 80 (monooleato de polioxietileno sorbitán), Span 85 (trioleato de sorbitán), lecitina y Triton X-100.

Se pueden usar mezclas de tensioactivo p.ej., mezclas de Tween 80/Span 85. También es adecuada una combinación de un éster de polioxietileno sorbitán tal como monooleato de polioxietileno sorbitán (Tween 80) y un octoxinol tal como t-octilfenoxipolietoxietanol (Triton X-100). Otra combinación útil comprende laureth 9 más un éster de polioxietileno sorbitán y/o un octoxinol.

Las cantidades preferidas de tensioactivos (% en peso) son: ésteres de polioxietileno sorbitán (tales como Tween 80) de 0,01 a 1 %, en particular aproximadamente 0,1 %; octil o nonilfenoxi polioxietanoles (tales como Triton X-100, u otros detergentes de la serie Triton) de 0,001 a 0,1 %, en particular de 0,005 a 0,02 %; éteres de polioxietileno (tales como laureth 9) de 0,1 a 20 %, preferiblemente de 0,1 a 10 % y en particular de 0,1 a 1 % o de aproximadamente 0,5 %.

Los coadyuvantes de emulsión preferidos tienen un tamaño de gotita promedio de <1 µm p.ej., ≤750 nm, ≤500 nm, ≤400 nm, ≤300 nm, ≤250 nm, ≤220 nm, ≤200 nm o menor. Estos tamaños de gotita se pueden lograr convenientemente mediante técnicas tales como la microfluidización.

Los adyuvantes de emulsión de aceite en agua específicos útiles con la invención incluyen, pero no están limitados a:

- Una emulsión submicrométrica de escualeno, Tween 80 y Span 85. La composición de la emulsión en volumen puede ser de aproximadamente 5 % de escualeno, aproximadamente 0,5 % de polisorbato 80 y aproximadamente 0,5 % de Span 85. En términos de peso, estas relaciones se convierten en 4,3 % de escualeno, 0,5 % polisorbato 80 y 0,48 % Span 85. Este adyuvante se conoce como 'MF59' [37-39], como se describe con más detalle en el Capítulo 10 de la ref. 40 y en el capítulo 12 de la ref. 41. La emulsión MF59 incluye ventajosamente iones citrato, p.ej., tampón de citrato sódico 10 mM.

- Una emulsión de escualeno, un tocoferol y Tween 80. La emulsión puede incluir solución salina tamponada con fosfato. También puede incluir Span 85 (p.ej., al 1 %) y/o lecitina. Estas emulsiones pueden tener de 2 a 10 % de escualeno, de 2 a 10 % de tocoferol y de 0,3 a 3 % de Tween 80, y la relación en peso de escualeno:tocoferol es preferiblemente ≤1 ya que esto proporciona una emulsión más estable. El escualeno y el Tween 80 pueden estar presentes en una relación de volumen de aproximadamente 5:2. Una de tales emulsiones puede prepararse disolviendo Tween 80 en PBS para dar una solución al 2 %, a continuación mezclando 90 ml de esta solución con una mezcla de (5 g de DL-α-tocoferol y 5 ml de escualeno), y después microfluidificando la mezcla. La emulsión resultante puede tener gotitas de aceite submicrométricas, p. con un diámetro promedio de entre 100 y 250 nm, preferiblemente de aproximadamente 180 nm.

- Una emulsión de escualeno, un tocoferol y un detergente Triton (p.ej., Triton X-100). La emulsión también puede incluir un 3d-MPL (ver a continuación). La emulsión puede contener un tampón de fosfato.

- Una emulsión que comprende un polisorbato (p.ej., polisorbato 80), un detergente Triton (p.ej., Triton X-100) y un tocoferol (p.ej., un succinato de α-tocoferol). La emulsión puede incluir estos tres componentes en una relación másica de aproximadamente 75:11:10 (p.ej., 750 µg/ml de polisorbato 80, 110 µg/ml de Triton X-100 y 100 µg/ml de succinato de α-tocoferol), y estas concentraciones deberían incluir cualquier contribución de estos componentes de los antígenos. La emulsión también puede incluir escualeno. La emulsión también puede incluir un 3d-MPL (ver a continuación). La fase acuosa puede contener un tampón de fosfato.

- Una emulsión de escualano, polisorbato 80 y poloxámero 401 ("Pluronic™ L121"). La emulsión se puede formular en solución salina tamponada con fosfato, pH 7,4. Esta emulsión es un vehículo de administración útil para péptidos de muramilo, y se ha utilizado con treonil-MDP en el adyuvante "SAF-1" [42] (0,05-1 % de Thr-MDP, 5 % de escualano, 2,5 % de Pluronic L121 y 0,2 % polisorbato 80). También se puede usar sin Thr-MDP, como en el adyuvante "AF" [43] (5 % de escualano, 1,25 % de Pluronic L121 y 0,2 % de polisorbato 80). Se prefiere la microfluidización.

- Una emulsión que comprende escualeno, un disolvente acuoso, un tensioactivo no iónico hidrófilo de

polioxietileno alquil éter (p.ej., polioxietileno (12) cetostearyl éter) y un tensioactivo no iónico hidrófobo (p.ej., un éster de sorbitán o éster de mannuro, tal como monoleato de sorbitán o "Span 80"). La emulsión es preferiblemente termorreversible y/o al menos el 90 % de las gotitas de aceite (en volumen) tiene un tamaño inferior a 200 nm [44]. La emulsión también puede incluir uno o más de: alditol; un agente crioprotector (p.ej., un azúcar, tal como dodecilmaltósido y/o sacarosa); y/o un alquilpoliglicósido. Tales emulsiones pueden ser liofilizadas.

- Una emulsión de escualeno, poloxámero 105 y Abil-Care [45]. La concentración final (peso) de estos componentes en las vacunas con adyuvante es 5 % de escualeno, 4 % de poloxámero 105 (poliol plurónico) y 2 % de Abil-Care 85 (Bis-PEG/PPG-16/16 PEG/PPG-16/16 dimeticona; triglicérido caprílico/cáprico).

- Una emulsión que tiene de 0,5 a 50 % de un aceite, 0,1-10 % de un fosfolípido y de 0,05 a 5 % de un tensioactivo no iónico. Como se describe en la referencia 46, los componentes fosfolípidos preferidos son fosfatidilcolina, fosfatidiletanolamina, fosfatidilserina, fosfatidilinositol, fosfatidilglicerol, ácido fosfatídico, esfingomielina y cardiolipina. Los tamaños de gotitas submicrométricas son ventajosos.

- Una emulsión submicrométrica de aceite en agua de un aceite no metabolizable (como el aceite mineral ligero) y al menos un tensioactivo (como lecitina, Tween 80 o Span 80). Se pueden incluir aditivos, tales como saponina QuilA, colesterol, un conjugado de saponina-lipófilo (como GPI-0100, descrito en la referencia 47, producido mediante la adición de amina alifática a desacilsaponina a través del grupo carboxilo del ácido glucurónico), bromuro de dimetildioctadecilamonio y/o N,N-dioctadecil-N,N-bis(2-hidroxietil)propandiamina.

- Una emulsión en la cual una saponina (p.ej., QuilA o QS21) y un esteroide (p.ej., un colesterol) se asocian como micelas helicoidales [48].

- Una emulsión que comprende un aceite mineral, un alcohol graso etoxilado lipófilo no iónico y un tensioactivo hidrófilo no iónico (p.ej., un alcohol graso etoxilado y/o un copolímero de bloques de polioxietileno-polioxipropileno) [49].

- Una emulsión que comprende un aceite mineral, un alcohol graso etoxilado hidrófilo no iónico y un tensioactivo lipófilo no iónico (p.ej., un alcohol graso etoxilado y/o un copolímero de bloques de polioxietileno-polioxipropileno) [49].

En algunas realizaciones, una emulsión puede mezclarse con el antígeno extemporáneamente, en el momento de la administración, y así el adyuvante y el antígeno se pueden mantener por separado en una vacuna empaquetada o distribuida, lista para la formulación final en el momento del uso. En otras realizaciones, una emulsión se mezcla con antígeno durante la fabricación, y de este modo la composición se envasa en una forma líquida adyuvada. El antígeno generalmente estará en forma acuosa, de manera que la vacuna finalmente se prepara mezclando dos líquidos. La relación de volumen de los dos líquidos para mezclar puede variar (p.ej., entre 5:1 y 1:5), pero generalmente es aproximadamente 1:1. Cuando se dan concentraciones de componentes en las descripciones anteriores de emulsiones específicas, estas concentraciones son generalmente para una composición no diluida, y la concentración después de mezclarse con una solución de antígeno disminuirá.

Cuando una composición incluye un tocoferol, se puede usar cualquiera de los tocoferoles α , β , γ , δ , ϵ o ξ , pero se prefieren los α -tocóferoles. El tocoferol puede adoptar varias formas, p.ej., diferentes sales y/o isómeros. Las sales incluyen sales orgánicas, como succinato, acetato, nicotinato, etc. Se pueden usar D- α -tocóferol y DL- α -tocóferol. Los tocoferoles se incluyen ventajosamente en vacunas para uso en pacientes de edad avanzada (p.ej., de 60 años o más) porque se ha informado que la vitamina E tiene un efecto positivo sobre la respuesta inmunitaria en este grupo de pacientes [50]. También tienen propiedades antioxidantes que pueden ayudar a estabilizar las emulsiones [51]. Un α -tocóferol preferido es el DL- α -tocóferol, y la sal preferida de este tocoferol es el succinato. Se ha observado que la sal succinato coopera con ligandos relacionados con el TNF *in vivo*.

C. Formulaciones de saponina [capítulo 22 de la ref. 32]

Las formulaciones de saponina también pueden usarse como adyuvantes en la invención. Las saponinas son un grupo heterogéneo de glucósidos de esteroide y glucósidos triterpenoides que se encuentran en la corteza, hojas, tallos, raíces e incluso flores de una amplia gama de especies de plantas. La saponina de la corteza del árbol *Quillaia saponaria* Molina ha sido ampliamente estudiada como adyuvante. La saponina también puede obtenerse comercialmente de *Smilax ornata* (zarzaparrilla), *Gypsophilla paniculata* (velo de novias) y *Saponaria officinalis* (raíz de jabón). Las formulaciones de adyuvantes de saponina incluyen formulaciones purificadas, tales como QS21, así como formulaciones de lípidos, tales como los ISCOM. QS21 se comercializa como Stimulon™.

Las composiciones de saponina se han purificado usando HPLC y RP-HPLC. Se han identificado fracciones específicas purificadas usando estas técnicas, que incluyen QS7, QS17, QS18, QS21, QH-A, QH-B y QH-C. Preferiblemente, la saponina es QS21. Un procedimiento de producción de QS21 se desvela en la ref. 52. Las formulaciones de saponina también pueden comprender un esteroide, como el colesterol [53].

Las combinaciones de saponinas y colesterol pueden usarse para formar partículas únicas llamadas complejos

inmunoestimulantes (ISCOM) [capítulo 23 de la ref. 32]. Los ISCOM generalmente también incluyen un fosfolípido tal como fosfatidiletanolamina o fosfatidilcolina. Cualquier saponina conocida puede usarse en los ISCOM. Preferiblemente, el ISCOM incluye uno o más de QuilA, QHA y QHC. Los ISCOM se describen con más detalle en las referencias. 53-55. Opcionalmente, los ISCOMS pueden estar desprovistos de detergente adicional [56].

- 5 Una revisión del desarrollo de adyuvantes basados en saponina se puede encontrar en las referencias 57 y 58.

D. Virosomas y partículas similares a virus

Los virosomas y las partículas similares a virus (VLP) también pueden usarse como adyuvantes en la invención. Estas estructuras generalmente contienen una o más proteínas de un virus opcionalmente combinado o formulado con un fosfolípido. Generalmente no son patógenos, no se replican y generalmente no contienen nada del genoma viral nativo. Las proteínas virales pueden producirse de forma recombinante o aislarse de virus completos. Estas proteínas virales adecuadas para su uso en virosomas o VLP incluyen proteínas derivadas del virus de la gripe (como HA o NA), virus de la hepatitis B (como proteínas del núcleo o de la cápside), virus de la hepatitis E, virus del sarampión, virus Sindbis, rotavirus, virus de la enfermedad de la boca y el pie, retrovirus, virus de Norwalk, virus del papiloma humano, VIH, fagos de ARN, fago Q β (como proteínas de la cubierta), fago GA, fago fr, fago AP205 y Ty (como la proteína p1 del retrotransposón de Ty). Las VLP se describen más detalladamente en las referencias 59-64. Los virosomas se describen más detalladamente en por ejemplo, la referencia 65.

E. Derivados bacterianos o microbianos

Los adyuvantes adecuados para su uso en la invención incluyen derivados bacterianos o microbianos tales como derivados no tóxicos de lipopolisacáridos (LPS) enterobacterianos, derivados de lípido A, oligonucleótidos inmunoestimuladores y toxinas de ribosilación del ADP y derivados destoxificados de los mismos.

Los derivados no tóxicos de LPS incluyen monofosforil lípido A (MPL) y MPL 3-O-desacilado (3dMPL). 3dMPL es una mezcla de monofosforil lípido A 3 des-O-acilado con 4, 5 o 6 cadenas aciladas. Una forma preferida de "partícula pequeña" de monofosforil lípido 3 des-O-acilado se desvela en la ref. 66. Tales "partículas pequeñas" de 3dMPL son lo suficientemente pequeñas como para ser filtradas de forma estéril a través de una membrana de 0,22 μ m [66]. Otros derivados de LPS no tóxicos incluyen imitadores de monofosforil lípido A, tales como derivados de aminoalquil glucosaminida fosfato p.ej., RC-529 [67,68].

Los derivados de lípido A incluyen derivados de lípido A de *Escherichia coli* como OM-174. OM-174 se describe, por ejemplo, en las ref. 69 y 70.

Los oligonucleótidos inmunoestimuladores adecuados para uso como adyuvantes en la invención incluyen secuencias de nucleótidos que contienen un motivo CpG (una secuencia de dinucleótidos que contiene una citosina no metilada unida por un enlace de fosfato a una guanósina). También se ha demostrado que los ARN y oligonucleótidos bicatenarios que contienen secuencias palindrómicas o poli (dG) son inmunoestimulantes.

Los CpG pueden incluir modificaciones/análogos de nucleótidos tales como modificaciones de fosforotioato y pueden ser de doble cadena o monocatenarios. Las referencias 71, 72 y 73 desvelan posibles sustituciones de análogos p.ej., reemplazo de guanósina por 2'-desoxi-7-desazaguanósina. El efecto adyuvante de los oligonucleótidos CpG se analiza adicionalmente en las referencias 74-79.

La secuencia CpG puede dirigirse a TLR9, como el motivo GTCGTT o TTCGTT [80]. La secuencia CpG puede ser específica para inducir una respuesta inmunitaria Th1, tal como un ODN de CpG-A, o puede ser más específica para inducir una respuesta de linfocitos B, como un ODN de CpG-B. Los ODN de CpG-A y CpG-B se discuten en las referencias 81-83. Preferiblemente, el CpG es un ODN de CpG-A.

Preferiblemente, el oligonucleótido CpG está construido de modo que el extremo 5' es accesible para el reconocimiento del receptor. Opcionalmente, se pueden unir dos secuencias de oligonucleótidos CpG en sus extremos 3' para formar "inmunómeros". Véase, por ejemplo, las referencias 80 y 84-86.

Un adyuvante CpG útil es CpG7909, también conocido como ProMune™ (Coley Pharmaceutical Group, Inc.). Otro es CpG 1826. Como alternativa, o además, para usar secuencias CpG, se pueden usar secuencias TpG [87], y estos oligonucleótidos pueden estar libres de motivos CpG no metilados. El oligonucleótido inmunoestimulador puede ser rico en pirimidina. Por ejemplo, puede comprender más de un nucleótido de timidina consecutivo (p.ej., TTTT, como se desvela en la ref. 87), y/o puede tener una composición de nucleótidos con > 25 % de timidina (p.ej., >35 %, >40 %, >50 %, >60 %, >80 %, etc.). Por ejemplo, puede comprender más de un nucleótido de citosina consecutivo (p.ej., CCCC, como se desvela en la ref. 87), y/o puede tener una composición de nucleótidos con > 25 % de citosina (p.ej., >35 %, >40 %, >50 %, >60 %, >80 %, etc.). Estos oligonucleótidos pueden estar exentos de motivos CpG no metilados. Los oligonucleótidos inmunoestimuladores generalmente comprenderán al menos 20 nucleótidos. Pueden comprender menos de 100 nucleótidos.

Un adyuvante particularmente útil basado en oligonucleótidos inmunoestimuladores se conoce como IC-31™ [88]. Por lo tanto, un adyuvante usado con la invención puede comprender una mezcla de (i) un oligonucleótido (p.ej.,

entre 15-40 nucleótidos) que incluye al menos un (y preferiblemente múltiples) motivos Cpl (es decir una citosina unida a una inosina para formar un dinucleótido), y (ii) un polímero policationico, como un oligopéptido (p.ej., entre 5-20 aminoácidos) que incluyen al menos una (y preferiblemente múltiples) secuencia(s) de tripéptido Lys-Arg-Lys. El oligonucleótido puede ser un desoxinucleótido que comprende una secuencia de 26 unidades de la secuencia 5'-(IC)₁₃-3' (SEQ ID NO: 96). El polímero policationico puede ser un péptido que comprende 11 unidades de la secuencia de aminoácidos KLKLLLLLKLK (SEQ ID NO: 97).

Las toxinas bacterianas de ribosilación del ADP y los derivados destoxificados de las mismas se pueden usar como adyuvantes en la invención. Preferiblemente, la proteína deriva de *E. coli* (enterotoxina termolábil de *E. coli* "LT"), del cólera ("CT") o de la tosferina ("PT"). El uso de toxinas de ribosilación del ADP destoxificadas como adyuvantes mucosos se describe en la ref. 89 y como adyuvantes parenterales en la ref. 90. La toxina o toxoide está preferiblemente en forma de una holotoxina, que comprende ambas subunidades A y B. Preferiblemente, la subunidad A contiene una mutación destoxificante; preferiblemente la subunidad B no está mutada. Preferiblemente, el adyuvante es un mutante LT destoxificado tal como LT-K63, LT-R72 y LT-G192. El uso de toxinas de ribosilación del ADP y sus derivados destoxificados, particularmente LT-K63 y LT-R72, como adyuvantes se puede encontrar en las referencias 91-98. Un mutante de CT útil es o CT-E29H [99]. La referencia numérica para las sustituciones de aminoácidos se basa preferiblemente en las alineaciones de las subunidades A y B de las toxinas de ribosilación del ADP expuestas en la ref. 100, únicamente con el propósito de la alineación y la numeración de aminoácidos en la misma.

F. Inmunomoduladores humanos

Los inmunomoduladores humanos adecuados para uso como adyuvantes en la invención incluyen citocinas, tales como interleucinas (p.ej., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12 [101], etc.) [102], interferones (p.ej., interferón-γ), factor estimulante de colonias de macrófagos y factor de necrosis tumoral. Un inmunomodulador preferido es IL-12.

G. Bioadhesivos y Mucoadhesivos

Los bioadhesivos y mucoadhesivos también pueden usarse como adyuvantes en la invención. Los bioadhesivos adecuados incluyen microesferas de ácido hialurónico esterificadas [103] o mucoadhesivos tales como derivados reticulados de poli(ácido acrílico), poli(alcohol vinílico), polivinilpirrolidona, polisacáridos y carboximetilcelulosa. El quitosano y sus derivados también pueden usarse como adyuvantes en la invención [104].

H. Micropartículas

Las micropartículas también pueden usarse como adyuvantes en la invención. Micropartículas (es decir una partícula de ~100 nm a ~150 μm de diámetro, más preferiblemente de ~200 nm a ~30 μm de diámetro, y lo más preferiblemente de ~500 nm a ~10 μm de diámetro) formadas a partir de materiales que son biodegradables y no tóxicos (p.ej., un poli(α-hidroxiácido), un ácido polihidroxibutírico, un poliortoéster, un polianhídrido, una policaprolactona, etc.), con poli(láctido-co-glicólido) son preferidos, opcionalmente tratados para tener una superficie cargada negativamente (p.ej., con SDS) o una superficie cargada positivamente (p.ej., con un detergente catiónico, como CTAB).

I. Liposomas (Capítulos 13 y 14 de la ref. 32)

Ejemplos de formulaciones de liposomas adecuadas para uso como adyuvantes se describen en las referencias 105-107.

J. Formulaciones de polioxietilen éter y polioxietilen éster

Los adyuvantes adecuados para usar en la invención incluyen éteres de polioxietileno y ésteres de polioxietileno [108]. Dichas formulaciones incluyen además tensioactivos de éster de polioxietileno sorbitan en combinación con un octoxinol [109] así como polioxietileno alquil éteres o tensioactivos éster en combinación con al menos un tensioactivo no iónico adicional tal como un octoxinol [110]. Los éteres de polioxietileno preferidos se seleccionan del siguiente grupo: polioxietilen-9-lauril éter (laureth 9), polioxietilen-9-estearil éter, polioxietilen-8-estearil éter, polioxietilen-4-lauril éter, polioxietilen-35-lauril éter y polioxietilen-23-lauril éter.

K. Fosfacenos

Se puede usar un fosfaceno, tal como poli(di(carboxilatofenoxi)fosfaceno) ("PCPP") como se describe, por ejemplo, en las referencias 111 y 112.

L. Péptidos de muramilo

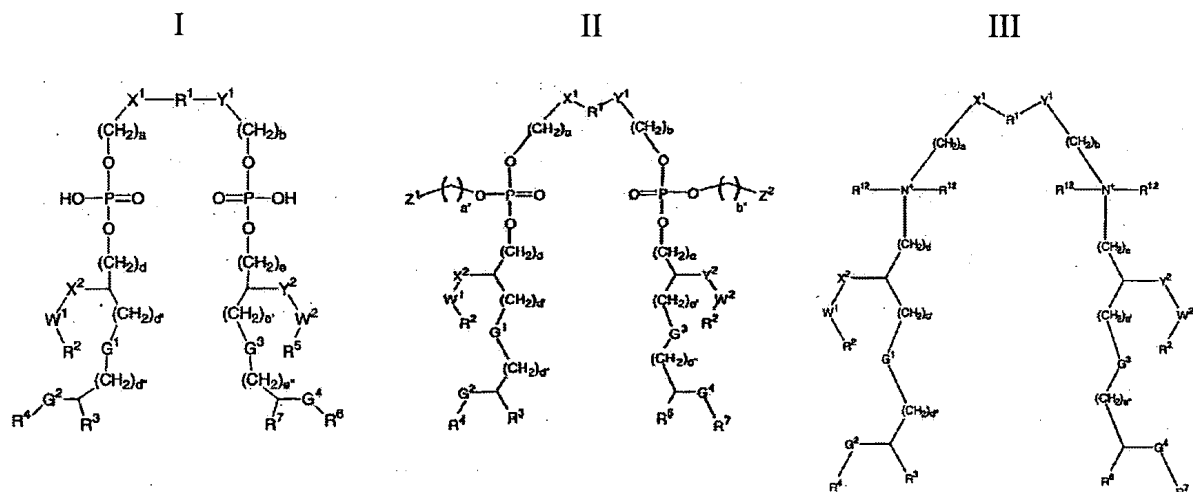
Los ejemplos de péptidos de muramilo adecuados para uso como adyuvantes en la invención incluyen N-acetil-muramil-L-treonil-D-isoglutamina (thr-MDP), N-acetil-normuramil-L-alanil-D-isoglutamina (nor-MDP) y N-acetilmuramil-L-alanil-D-isoglutaminil-L-alanina-2-(1'-2'-dipalmitoil-*sn*-glicero-3-hidroxifosforiloxi)-etilamina MTP-PE).

M. Compuestos de imidazoquinolona

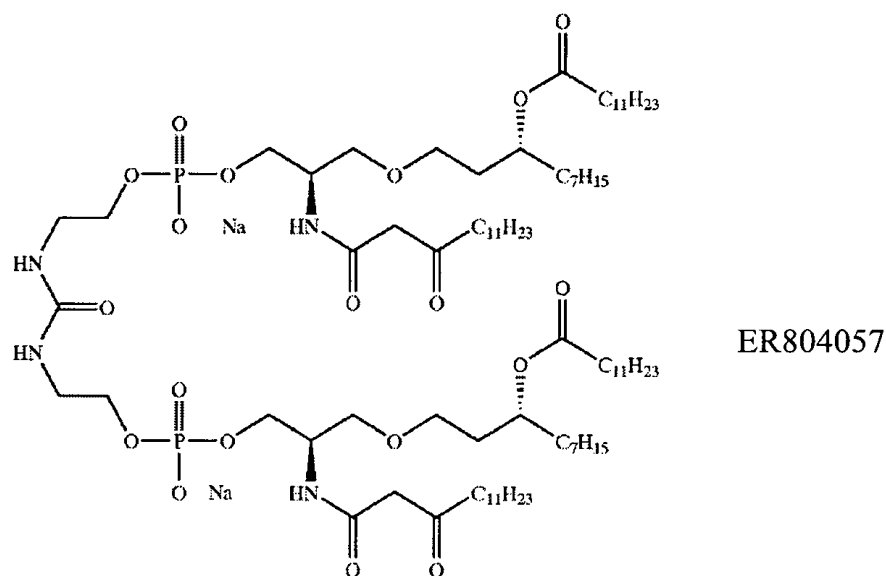
Los ejemplos de compuestos de imidazoquinolona adecuados para el uso de adyuvantes en la invención incluyen Imiquimod ("R-837") [113,114], Resiquimod ("R-848") [115] y sus análogos; y sus sales (p.ej., las sales clorhidrato). Se pueden encontrar más detalles acerca de las imidazoquinolinas inmunoestimuladoras en las referencias 116 a 120.

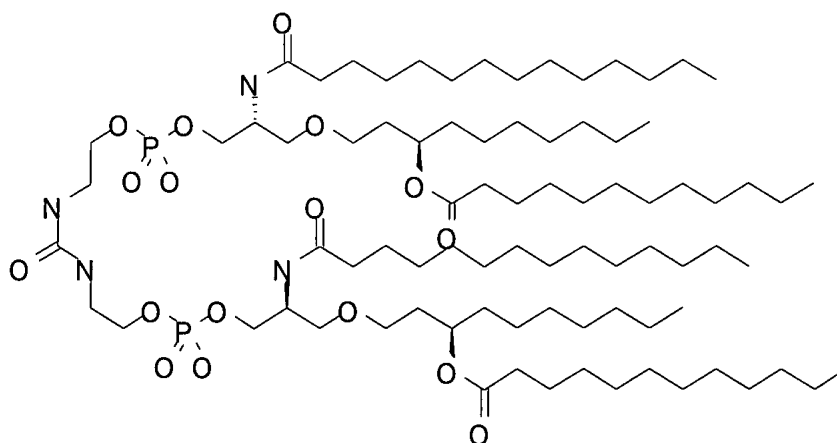
5 N. Ureas sustituidas

Las ureas sustituidas útiles como adyuvantes incluyen compuestos de fórmula I, II o III, o sus sales:



como se define en la referencia 121, como 'ER 803058', 'ER 803732', 'ER 804053', 'ER 804058', 'ER 804059', 'ER 804442', 'ER 804680', 'ER 804764', 'ER 803022' o 'ER 804057' p.ej.:



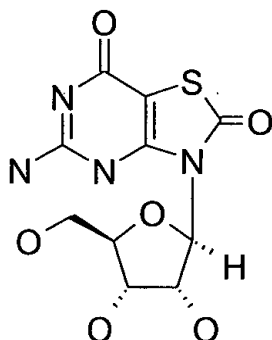


ER-803022:

O. Otros adyuvantes

Otros adyuvantes que pueden usarse con la invención incluyen:

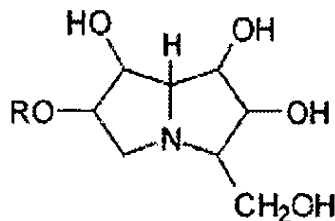
- 5 • Un derivado de aminoalquil glucosaminida fosfato, como RC-529 [122,123].
- Un compuesto de tiosemicarbazona, tal como los desvelados en la referencia 124. Los procedimientos de formulación, fabricación y selección de compuestos activos también se describen en la referencia 124. Las tiosemicarbazonas son particularmente eficaces en la estimulación de células mononucleares de sangre periférica humana para la producción de citocinas, como TNF- α .
- 10 • Un compuesto de triptantrina, tal como los desvelados en la referencia 125. Los procedimientos de formulación, fabricación y selección de compuestos activos también se describen en la referencia 125. Las tiosemicarbazonas son particularmente eficaces en la estimulación de células mononucleares de sangre periférica humana para la producción de citocinas, como TNF- α .
- Un análogo de nucleósido, tal como: (a) Isatorabina (ANA-245; 7-tia-8-oxoguanosina):



- 15 y profármacos del mismo; (b) ANA975; (c) ANA-025-1; (d) ANA380; (e) los compuestos desvelados en las referencias 126 a 128 Loxoribina (7-alil-8-oxoguanosina) [129].

- 20 • Los compuestos desvelados en la referencia 130, que incluyen: compuestos acilpiperazina, compuestos indoldiona, compuestos tetrahidraisoquinolina (THIQ), compuestos benzociclodiona, compuestos aminoazavinilo, compuestos aminobenzimidazol quinolinona (ABIQ) [131,132], compuestos hidraftalamida, compuestos benzofenona, compuestos isoxazol, compuestos estero, compuestos quinazolinona, compuestos pirrol [133], compuestos antraquinona, compuestos quinoxalina, compuestos triazina, compuestos pirazolpirimidina y compuestos benzazol [134].
- 25 • Compuestos que contienen lípidos unidos a un esqueleto acíclico que contiene fosfato, como el antagonista de TLR4 E5564 [135,136]:
- Un polímero de polioxidonio [137,138] u otro derivado de polietileno-piperazina N-oxidado.
- Metil inosina 5'-monofosfato ("MIMP") [139].

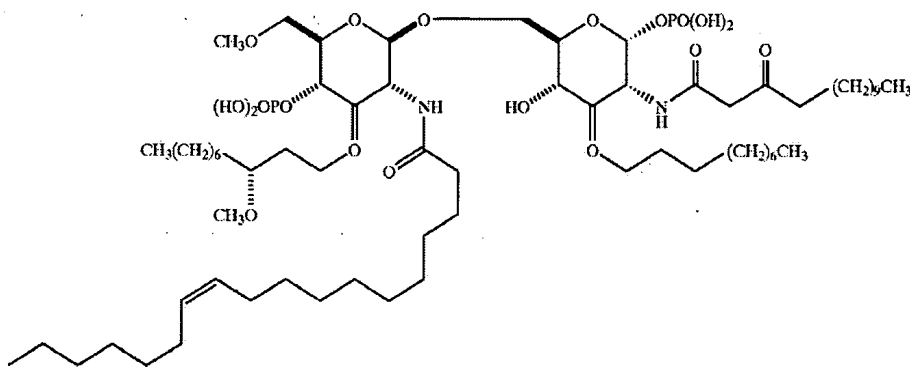
- Un compuesto de pirrolizidina polihidroxilada [140], como el que tiene la fórmula:



en la que R se selecciona del grupo que comprende hidrógeno, acilo, alquilo (p.ej. cicloalquilo), alqueno, alquino lineal o ramificado, no sustituido o sustituido, saturado o insaturado y grupos arilo o una sal o derivado farmacéuticamente aceptable de los mismos. Los ejemplos incluyen, pero no están limitados a: casuarina, casuarina-6- α -D-glucopiranos, 3-*epi*-casuarina, 7-*epi*-casuarina, 3,7-diepi-casuarina, etc.

- Un ligando CD1d, tal como una α -glicosilceramida [141-148] (p.ej., α -galactosilceramida), α -glicosilceramidas que contienen fitoesfingosina, OCH, KRN7000 [(2S,3S,4R)-1-O-(α -D-galactopiranosil)-2-(N-hexacosanoilamino)-1,3,4-octadecanotriol], CRONY-101, 3''-O-sulfo-galactosilceramida, etc.

- Una gamma inulina [149] o un derivado de la misma, como la algamulina.



Combinaciones adyuvantes

La invención también puede comprender combinaciones de aspectos de uno o más de los adyuvantes identificados anteriormente. Por ejemplo, las siguientes composiciones adyuvantes pueden usarse en la invención: (1) una saponina y una emulsión de aceite en agua [150]; (2) una saponina (p.ej., QS21) + un derivado de LPS no tóxico (p.ej., 3dMPL) [151]; (3) una saponina (p.ej., QS21) + un derivado de LPS no tóxico (p.ej., 3dMPL) + colesterol; (4) una saponina (p.ej., QS21) + 3dMPL + IL-12 (opcionalmente + un esteroide) [152]; (5) combinaciones de 3dMPL con, por ejemplo, QS21 y/o emulsiones de aceite en agua [153]; (6) SAF, que contiene 10 % de escualeno, 0,4 % de Tween 80™, 5 % de polímero de bloque plurónico L121, y thr-MDP, microfluidizado en una emulsión submicrométrica o agitado vorticialmente para generar una emulsión de mayor tamaño de partícula. (7) Sistema adyuvante Ribi™ (RAS), (Ribi Immunochem) que contiene 2 % de escualeno, 0,2 % de Tween 80 y uno o más componentes de la pared celular bacteriana del grupo que consiste en monofosforil lípido A (MPL), trehalosa dimicolato (TDM), y esqueleto de pared celular (CWS), preferiblemente MPL + CWS (Detox™); y (8) una o más sales minerales (como una sal de aluminio) + un derivado no tóxico de LPS (como 3dMPL).

Otras sustancias que actúan como agentes inmunoestimulantes se describen en el capítulo 7 de la ref. 32.

El uso de un adyuvante de hidróxido de aluminio y/o fosfato de aluminio es particularmente preferido, y los antígenos generalmente se adsorben a estas sales. El fosfato de calcio es otro adyuvante preferido. Otras combinaciones de adyuvantes preferidas incluyen combinaciones de adyuvantes Th1 y Th2 tales como CpG y alumbre o resiquimod y alumbre. Se puede usar una combinación de fosfato de aluminio y 3dMPL.

Las composiciones de la invención pueden provocar tanto una respuesta inmunitaria mediada por células como una respuesta inmunitaria humoral. Esta respuesta inmunitaria preferiblemente inducirá anticuerpos de larga duración (p.ej., neutralizantes) y una inmunidad mediada por células que pueden responder rápidamente después de la exposición a *Pneumococcus*.

En general, se considera que dos tipos de linfocitos T, linfocitos CD4 y CD8 son necesarios para iniciar y/o mejorar la inmunidad mediada por células y la inmunidad humoral. Los linfocitos T CD8 pueden expresar un correceptor CD8 y se conocen comúnmente como linfocitos T citotóxicos (CTL). Los linfocitos T CD8 son capaces de reconocer o

interactuar con antígenos exhibidos en las moléculas MHC de clase I.

Los linfocitos T CD4 pueden expresar un correceptor CD4 y comúnmente se les denomina linfocitos T cooperadores. Los linfocitos T CD4 son capaces de reconocer péptidos antigénicos unidos a moléculas MHC de clase II. Tras la interacción con una molécula MHC de clase II, los linfocitos CD4 pueden secretar factores tales como citocinas. Estas citocinas secretadas pueden activar los linfocitos B, los linfocitos T citotóxicos, los macrófagos y otras células que participan en una respuesta inmunitaria. Los linfocitos T colaboradores o los linfocitos CD4+ se pueden dividir además en dos subconjuntos funcionalmente distintos: el fenotipo TH1 y los fenotipos TH2 que difieren en su función citocina y efectora.

Los linfocitos TH1 activados potencian la inmunidad celular (incluido un aumento en la producción de CTL específicos de antígeno) y, por lo tanto, tienen un valor particular para responder a las infecciones intracelulares. Los linfocitos TH1 activados pueden secretar una o más de IL-2, IFN- γ y TNF- β . Una respuesta inmunitaria TH1 puede provocar reacciones inflamatorias locales al activar macrófagos, linfocitos citolíticos naturales (NK) y linfocitos T citotóxicos (CTL) CD8. Una respuesta inmunitaria TH1 también puede actuar para ampliar la respuesta inmunitaria estimulando el crecimiento de linfocitos B y T con IL-12. Los linfocitos B estimulados con TH1 pueden secretar IgG2a.

Los linfocitos TH2 activados a la producción de anticuerpos y, por lo tanto, son valiosos para responder a las infecciones extracelulares. Los linfocitos TH2 activados pueden secretar una o más de IL-4, IL-5, IL-6 e IL-10. Una respuesta inmunitaria TH2 puede dar como resultado la producción de IgG1, IgE, IgA y linfocitos B de memoria para la protección futura.

Una respuesta inmunitaria potenciada puede incluir una o más de una respuesta inmunitaria de TH1 y una respuesta inmunitaria TH2 potenciadas.

Una respuesta inmunitaria de TH1 puede incluir uno o más de un aumento en los CTL, un aumento en una o más de las citocinas asociadas con una respuesta inmunitaria TH1 (como IL-2, IFN- γ y TNF- β), un aumento en macrófagos activados, un aumento en la actividad de los linfocitos NK o un aumento en la producción de IgG2a. Preferiblemente, la respuesta inmunitaria mejorada de TH1 incluirá un aumento en la producción de IgG2a.

Se puede provocar una respuesta inmunitaria de TH1 usando un adyuvante de TH1. Un adyuvante de TH1 generalmente provocará niveles aumentados de producción de IgG2a en relación con la inmunización del antígeno sin adyuvante. Los adyuvantes TH1 adecuados para usar en la invención pueden incluir, por ejemplo, formulaciones de saponina, virosomas y partículas similares a virus, derivados no tóxicos de lipopolisacárido enterobacteriano (LPS), oligonucleótidos inmunoestimuladores. Los oligonucleótidos inmunoestimuladores, tales como oligonucleótidos que contienen un motivo CpG, son adyuvantes de TH1 preferidos para uso en la invención.

Una respuesta inmunitaria de TH2 puede incluir uno o más de un aumento en una o más de las citocinas asociadas con una respuesta inmunitaria de TH2 (como IL-4, IL-5, IL-6 e IL-10), o un aumento en la producción de IgG1, IgE, IgA y linfocitos B de memoria. Preferiblemente, la respuesta inmunitaria de TH2 potenciada incluirá un aumento en la producción de IgG1.

Se puede provocar una respuesta inmunitaria de TH2 usando un adyuvante de TH2. Un adyuvante de TH2 generalmente provocará niveles aumentados de producción de IgG1 en relación con la inmunización del antígeno sin adyuvante. Los adyuvantes de TH2 adecuados para su uso en la invención incluyen, por ejemplo, composiciones que contienen minerales, emulsiones de aceite y toxinas de ribosilación de ADP y derivados detoxificados de las mismas. Las composiciones que contienen minerales, tales como sales de aluminio son adyuvantes de TH2 preferidos para uso en la invención.

Preferiblemente, la invención incluye una composición que comprende una combinación de un adyuvante de TH1 y un adyuvante de TH2. Preferiblemente, tal composición provoca una respuesta potenciada de TH1 y TH2, es decir, un aumento en la producción tanto de la producción de IgG1 como de IgG2a con respecto a la inmunización sin un adyuvante. Todavía más preferiblemente, la composición que comprende una combinación de un adyuvante de TH1 y TH2 provoca un aumento de la respuesta inmunitaria de TH1 y/o un aumento de la respuesta inmunitaria de TH2 en relación con la inmunización con un único adyuvante (es decir, con respecto a la inmunización con un adyuvante de TH1 solo o inmunización con un adyuvante de TH2 solo).

La respuesta inmunitaria puede ser una o ambas de una respuesta inmunitaria TH1 y una respuesta TH2. Preferiblemente, la respuesta inmunitaria proporciona una o ambas de una respuesta de TH1 potenciada y una respuesta de TH2 potenciada.

La respuesta inmunitaria potenciada puede ser una o ambas de una respuesta inmunitaria sistémica y mucosal. Preferiblemente, la respuesta inmunitaria proporciona una o ambas de una respuesta inmunitaria sistémica potenciada y una respuesta inmunitaria mucosal mejorada. Preferiblemente, la respuesta inmunitaria mucosal es una respuesta inmunitaria TH2. Preferiblemente, la respuesta inmunitaria mucosal incluye un aumento en la producción de IgA.

E. coli puede provocar enfermedad en varias localizaciones anatómicas [4] y así las composiciones de la invención se pueden preparar de diversas formas. Por ejemplo, las composiciones se pueden preparar como inyectables, como soluciones líquidas o suspensiones. También se pueden preparar formas sólidas adecuadas para solución o suspensión en vehículos líquidos antes de la inyección (p.ej., una composición liofilizada o una composición sometida a criodesecación por pulverización). La composición puede prepararse para administración tópica p.ej., como una pomada, crema o polvo. La composición puede prepararse para administración oral p.ej., como un comprimido o cápsula, como un aerosol, o como un jarabe (opcionalmente con sabor). La composición puede prepararse para administración pulmonar p.ej., como inhalador, usando un polvo fino o un aerosol. La composición puede prepararse como un supositorio o pesario. La composición puede prepararse para administración nasal, aural u ocular p.ej., como gotas. La composición puede adoptar la forma de kit, diseñada de modo que una composición combinada se reconstituya justo antes de la administración a un paciente. Dichos kits pueden comprender uno o más antígenos en forma líquida y uno o más antígenos liofilizados.

Cuando una composición debe prepararse extemporáneamente antes de su uso (p.ej., cuando un componente se presenta en forma liofilizada) y se presenta como un kit, el kit puede comprender dos viales, o puede comprender una jeringa preparada y un vial, utilizándose el contenido de la jeringa para reactivar el contenido del vial antes de la inyección.

Las composiciones inmunógenas usadas como vacunas comprenden una cantidad inmunológicamente eficaz de antígeno(s), así como cualquier otro componente, según sea necesario. Por "cantidad inmunológicamente eficaz", se entiende que la administración de esa cantidad a un individuo, ya sea en una sola dosis o como parte de una serie, es eficaz para el tratamiento o la prevención. Esta cantidad varía según la salud y el estado físico del individuo a tratar, la edad, el grupo taxonómico del individuo que se tratará (p.ej., primate no humano, primate etc.), la capacidad del sistema inmunitario del individuo para sintetizar anticuerpos, el grado de protección deseado, la formulación de la vacuna, la evaluación del médico responsable del tratamiento de la enfermedad y otros factores relevantes. Se espera que la cantidad entre dentro de un rango relativamente amplio que se puede determinar a través de ensayos de rutina.

Procedimientos de tratamiento y administración de la vacuna

La divulgación también proporciona un procedimiento para provocar una respuesta inmunitaria en un mamífero que comprende la etapa de administrar una cantidad efectiva de una composición de la invención. La respuesta inmunitaria es preferiblemente protectora y preferiblemente implica anticuerpos y/o inmunidad mediada por células. El procedimiento puede provocar una respuesta de refuerzo.

La invención también proporciona un polipéptido de la invención para uso como medicamento, p.ej., para usar para provocar una respuesta inmunitaria en un mamífero.

La invención también proporciona el uso de un polipéptido de la invención en la fabricación de un medicamento para provocar una respuesta inmunitaria en un mamífero.

La divulgación también proporciona un dispositivo de administración prellenado con una composición inmunogénica de la invención.

Al provocar una respuesta inmunitaria en el mamífero por estos usos y procedimientos, el mamífero puede protegerse contra la infección por *E. coli*, incluidas las cepas ExPEC y no ExPEC. La invención es particularmente útil para proporcionar una amplia protección contra *E. coli* patógena, incluidos los patotipos intestinales, como los patotipos EPEC, EAEC, EIEC, ETEC y DAEC. Así, el mamífero puede estar protegido contra enfermedades que incluyen, entre otras, peritonitis, pielonefritis, cistitis, endocarditis, prostatitis, infecciones del tracto urinario (ITU), meningitis (particularmente meningitis neonatal), sepsis (o SIRS), deshidratación, neumonía, diarrea (infantil, diarrea del viajero, aguda, persistente, etc.), disentería bacilar, síndrome urémico hemolítico (SHU), pericarditis, bacteriuria, etc.

La SEQ ID NO: 3 y 12 y sus variantes son particularmente útiles para inmunizar contra el patotipo EAEC, y por lo tanto para prevenir la diarrea (tanto aguda como crónica).

La SEQ ID NO: 3, 12, 199, 109, 115 y 124 y sus variantes son particularmente útiles para inmunizar contra el patotipo CEEA, y por lo tanto para prevenir la diarrea (tanto aguda como crónica).

La SEQ ID NO: 3, 12, 199, 109, 115, 124, 129, 130 y 131 y sus variantes son particularmente útiles para inmunizar contra el patotipo CEEA, y por lo tanto para prevenir la diarrea (tanto aguda como crónica).

La SEQ ID NO: 4 y 10 y sus variantes son particularmente útiles para inmunizar contra el patotipo UPEC y, por lo tanto, para prevenir las ITU, que incluyen, entre otras, pielonefritis, cistitis (tanto aguda como recurrente), peritonitis, ITU asociadas a catéter, prostatitis y bacteriuria (incluida la bacteriuria asintomática).

La SEQ ID NO: 4, 10, 101, 107, 116 y 126 y sus variantes son particularmente útiles para inmunizar contra el patotipo UPEC y, por lo tanto, para prevenir las ITU, que incluyen, entre otras, pielonefritis, cistitis (tanto aguda como

recurrente), peritonitis, ITU asociadas a catéter, prostatitis y bacteriuria (incluida la bacteriuria asintomática).

La SEQ ID NO: 4, 10, 101, 107, 116, 126, 133, 134, 135, 137, 138 y 139 y sus variantes son particularmente útiles para inmunizar contra el patotipo UPEC y, por lo tanto, para prevenir las ITU, que incluyen, entre otras, pielonefritis, cistitis (tanto aguda como recurrente), peritonitis, ITU asociadas a catéter, prostatitis y bacteriuria (incluida la bacteriuria asintomática).

La SEQ ID NO: 5 y sus variantes son particularmente útiles para inmunizar contra el patotipo EIEC, y por lo tanto para prevenir la disentería (en particular la disentería bacilar) y el SUH (p.ej., en niños).

La SEQ ID NO: 5, 102 y 117 y sus variantes son particularmente útiles para inmunizar contra el patotipo EIEC, y por lo tanto para prevenir la disentería (en particular la disentería bacilar) y el SUH (p.ej., en niños).

La SEQ ID NO: 6, 9 y 11 y sus variantes son particularmente útiles para inmunizar contra el patotipo ETEC, y por lo tanto para prevenir la diarrea (incluyendo la diarrea del viajero y la diarrea infantil).

La SEQ ID NO: 6, 9, 11, 103, 106, 108, 118, 121 y 123 y sus variantes son también particularmente útiles para inmunizar contra el patotipo ETEC, y por lo tanto para prevenir la diarrea (incluyendo la diarrea del viajero y la diarrea infantil).

La SEQ ID NO: 7, 8 y 16 y sus variantes son particularmente útiles para inmunizar contra el patotipo de EPEC y, por lo tanto, para prevenir la diarrea (incluida la diarrea infantil).

La SEQ ID NO: 7, 8, 16, 104, 105, 113, 119, 120 y 128 y sus variantes también son particularmente útiles para inmunizar contra el patotipo EPEC, y por lo tanto para prevenir la diarrea (incluida la diarrea infantil).

El mamífero es preferiblemente un ser humano, pero puede ser p.ej., una vaca, un cerdo, un pollo, un gato o un perro, ya que la enfermedad por *E. coli* también es problemática en estas especies [4]. Cuando la vacuna es para uso profiláctico, el ser humano es preferiblemente un niño (p.ej., un niño pequeño o un bebé) o un adolescente; cuando la vacuna es para uso terapéutico, el ser humano es preferiblemente un adolescente o un adulto. Una vacuna para niños también se puede administrar a adultos p.ej., para evaluar la seguridad, la dosificación, la inmunogenicidad, etc.

Una forma de verificar la eficacia del tratamiento terapéutico consiste en controlar la infección por *E. coli* después de la administración de las composiciones de la invención. Una forma de controlar la eficacia del tratamiento profiláctico consiste en monitorizar respuestas inmunitarias, sistémicamente (como controlar el nivel de producción de IgG1 e IgG2a) y/o en la mucosa (como controlar el nivel de producción de IgA), contra los antígenos en las composiciones de la invención después de la administración de la composición. Generalmente, las respuestas de anticuerpos en suero específicas de antígeno se determinan después de la inmunización pero antes de la estimulación, mientras que las respuestas de anticuerpos de mucosa específicas de antígeno se determinan después de la inmunización y después de la exposición.

Otra forma de evaluar la inmunogenicidad de las composiciones de la presente invención es expresar las proteínas de forma recombinante para seleccionar sueros de pacientes o secreciones de mucosa mediante inmunotransferencia y/o microchips. Una reacción positiva entre la proteína y la muestra del paciente indica que el paciente ha creado una respuesta inmunitaria a la proteína en cuestión. Este procedimiento también puede usarse para identificar antígenos inmunodominantes y/o epítomos dentro de antígenos.

La eficacia de las composiciones de vacuna también puede determinarse *in vivo* provocando la infección por *E. coli* en modelos animales, p.ej., cobayas o ratones, con las composiciones de vacuna. Se describe un modelo murino de ExPEC y sepsis letal en la referencia 154. Un modelo de rata del algodón se describe en la ref. 155

Las composiciones de la invención generalmente se administrarán directamente a un paciente. La administración directa puede lograrse mediante inyección parenteral (p.ej., por vía subcutánea, intraperitoneal, intravenosa, intramuscular o en el espacio intersticial de un tejido), o por vía mucosal, como por vía rectal, oral (p.ej., comprimido, pulverización), vaginal, tópica, transdérmica o transcutánea, intranasal, ocular, aural, pulmonar u otra administración mucosal.

La invención puede usarse para provocar inmunidad sistémica y/o mucosal, preferiblemente para provocar una inmunidad sistémica y/o mucosal mejorada.

Preferiblemente, la inmunidad sistémica y/o mucosal potenciada se refleja en una respuesta inmunitaria TH1 y/o TH2 potenciada. Preferiblemente, la respuesta inmunitaria potenciada incluye un aumento en la producción de IgG1 y/o IgG2a y/o IgA.

La administración puede realizarse mediante un esquema de dosis única o un esquema de dosis múltiples. Se pueden usar dosis múltiples en un esquema de inmunización primaria y/o en un esquema de inmunización de refuerzo. En un programa de dosis múltiples, las diversas dosis pueden administrarse por la misma ruta o por rutas diferentes, p.ej., una estimulación parenteral y refuerzo a través de la mucosa, una estimulación a través de la

mucosa y refuerzo parenteral, etc. Generalmente, se administrarán dosis múltiples con al menos 1 semana de diferencia (p.ej., aproximadamente 2 semanas, aproximadamente 3 semanas, aproximadamente 4 semanas, aproximadamente 6 semanas, aproximadamente 8 semanas, aproximadamente 10 semanas, aproximadamente 12 semanas, aproximadamente 16 semanas, etc.).

- 5 Las vacunas de la invención pueden usarse para tratar tanto a niños como a adultos. Por lo tanto, un paciente humano puede tener menos de 1 año, 1-5 años, 5-15 años, 15-55 años o al menos 55 años. Los pacientes preferidos para recibir las vacunas son los ancianos (p.ej., ≥ 50 años, ≥ 60 años, y preferiblemente ≥ 65 años), los jóvenes (p.ej., ≤ 5 años), pacientes hospitalizados, trabajadores de la salud, personal armado y militar, mujeres embarazadas, pacientes con enfermedades crónicas o inmunodeficientes. Sin embargo, las vacunas no son adecuadas únicamente para estos grupos y pueden usarse en general en una población.

Las vacunas de la invención son particularmente útiles para pacientes que están a la espera de una operación quirúrgica u otros pacientes hospitalizados. También son útiles en pacientes que serán cateterizados. También son útiles en mujeres adolescentes (p.ej., entre 11 y 18 años) y en pacientes con infecciones crónicas del tracto urinario.

- 15 Las vacunas de la invención pueden administrarse a pacientes sustancialmente al mismo tiempo que (p.ej., durante la misma consulta médica o visita a un profesional de la salud o un centro de vacunación) otras vacunas, p.ej., sustancialmente al mismo tiempo que una vacuna contra el sarampión, una vacuna contra la parotiditis, una vacuna contra la rubéola, una vacuna contra el sarampión, paperas y rubéola, una vacuna contra la varicela, una vacuna contra el sarampión, paperas, rubéola y varicela, una vacuna contra la difteria, una vacuna antitetánica, una vacuna antitosferina, una vacuna DTP, una vacuna contra *H. influenzae* de tipo b, una vacuna inactivada contra la polio, una
20 vacuna contra el virus de la hepatitis B, una vacuna conjugada meningocócica (como la vacuna tetravalente A-C-W135-Y), una vacuna contra el virus sincitial respiratorio, etc.

Inmunización con ácido nucleico

- Las composiciones inmunógenas descritas anteriormente incluyen antígenos polipeptídicos. Sin embargo, en todos los casos, los antígenos polipeptídicos pueden reemplazarse por ácidos nucleicos (generalmente ADN) que
25 codifican esos polipéptidos, para proporcionar composiciones, procedimientos y usos basados en la inmunización con ácido nucleico. La inmunización con ácido nucleico es ahora un campo desarrollado (p.ej., ver referencias 156 a 163 etc.).

- El ácido nucleico que codifica el inmunógeno se expresa *in vivo* después del parto a un paciente y el inmunógeno expresado estimula a continuación el sistema inmunitario. El principio activo adoptará generalmente la forma de un vector de ácido nucleico que comprende: (i) un promotor; (ii) una secuencia que codifica el inmunógeno, unida operativamente al promotor; y opcionalmente (iii) un marcador seleccionable. Los vectores preferidos pueden comprender además (iv) un origen de replicación; y (v) un terminador de la transcripción en dirección 3' y operativamente asociado a (ii). En general, (i) y (v) serán eucariotas y (iii) y (iv) serán procariotas.

- Los promotores preferidos son promotores virales p.ej., de citomegalovirus (CMV). El vector también puede incluir secuencias reguladoras de la transcripción (p.ej., potenciadores) además del promotor y que interactúan funcionalmente con el promotor. Los vectores preferidos incluyen el potenciador/promotor de CMV inmediato temprano y los vectores más preferidos también incluyen el intrón A del CMV. El promotor está ligado operativamente a una secuencia en dirección 3' que codifica un inmunógeno, de modo que la expresión de la secuencia que codifica el inmunógeno está bajo el control del promotor.

- 40 Cuando se usa un marcador, preferiblemente funciona en un hospedador microbiano (p.ej., en un procariota, en una bacteria, en una levadura). El marcador es preferiblemente un marcador seleccionable procariota (p.ej., transcrito bajo el control de un promotor procariota). Por conveniencia, los marcadores típicos son genes de resistencia a antibióticos.

- El vector de la invención es preferiblemente un vector episomal o extracromosómico que se replica autónomamente, tal como un plásmido.

El vector de la invención preferiblemente comprende un origen de replicación. Se prefiere que el origen de la replicación sea activo en procariotas pero no en eucariotas.

- Los vectores preferidos incluyen así un marcador procariota para la selección del vector, un origen de replicación procariota, pero un promotor eucariota para dirigir la transcripción de la secuencia que codifica el inmunógeno. Por lo tanto, los vectores se (a) amplificarán y seleccionarán en hospedadores procariotas sin expresión de polipéptido, pero (b) se expresarán en hospedadores eucariotas sin amplificarse. Esta disposición es ideal para vectores de
50 inmunización con ácido nucleico.

- El vector de la invención puede comprender una secuencia terminadora de la transcripción eucariota en dirección 3' de la secuencia codificante. Esto puede mejorar los niveles de transcripción. Cuando la secuencia codificante no tiene la suya, el vector de la invención preferiblemente comprende una secuencia de poliadenilación. Una secuencia de poliadenilación preferida es la de la hormona de crecimiento bovina.

El vector de la invención puede comprender un sitio de clonación múltiple

Además de las secuencias que codifican el inmunógeno y un marcador, el vector puede comprender una segunda secuencia codificante eucariota. El vector también puede comprender un IRES (sitio de entrada al ribosoma) en dirección 5' de dicha segunda secuencia para permitir la traducción de un segundo polipéptido eucariota del mismo transcrito que el inmunógeno. En otra alternativa, la secuencia de codificación del inmunógeno puede estar en dirección 3' de un IRES.

El vector de la invención puede comprender motivos CpG no metilados, p.ej., secuencias de ADN no metiladas que tienen en común una citosina que precede a una guanosina, flanqueada por dos purinas 5' y dos pirimidinas 3'. En su forma no metilada, estos motivos de ADN han demostrado ser potentes estimuladores de varios tipos de células inmunitarias.

Los vectores se pueden administrar de una manera específica. Las técnicas de administración de ADN mediadas por receptor se describen, por ejemplo, en las referencias 164 a 169. Las composiciones terapéuticas que contienen un ácido nucleico se administran en un intervalo de aproximadamente 100 ng a aproximadamente 200 mg de ADN para administración local en un protocolo de terapia génica. Los intervalos de concentración de aproximadamente 500 ng a aproximadamente 50 mg, aproximadamente 1 µg a aproximadamente 2 mg, aproximadamente 5 µg a aproximadamente 500 µg, y aproximadamente 20 µg a aproximadamente 100 µg de ADN también pueden usarse durante un protocolo de terapia génica. Factores como el procedimiento de acción (p.ej., para potenciar o inhibir los niveles del producto génico codificado) y la eficacia de la transformación y la expresión son consideraciones que afectarán a la dosificación requerida para la eficacia final. Cuando se desea una mayor expresión en un área más grande de tejido, pueden requerirse mayores cantidades de vector o las mismas cantidades re-administradas en un protocolo sucesivo de administraciones, o pueden requerirse varias administraciones a diferentes porciones de tejido adyacentes o cercanas para conseguir un resultado terapéutico positivo. En todos los casos, la experimentación de rutina en ensayos clínicos determinará los intervalos específicos para lograr un efecto terapéutico óptimo.

Los vectores se pueden administrar utilizando vehículos de administración de genes. El vehículo de administración de genes puede ser de origen viral o no viral (véanse, en general, las referencias 170 a 173).

Los vectores basados en virus para la administración de un ácido nucleico deseado y la expresión en una célula deseada son bien conocidos en la técnica. Ejemplos de vehículos basados en virus incluyen, pero no se limitan a, retrovirus recombinantes (p.ej., referencias 174 a 184), vectores basados en alfavirus (p.ej., vectores del virus Sindbis, virus del bosque Semliki (ATCC VR-67, ATCC VR-1247), virus del río Ross (ATCC VR-373, ATCC VR-1246) y virus de la encefalitis equina venezolana (ATCC VR-923; ATCC VR-1250; ATCC VR 1249; ATCC VR-532); también se pueden usar los híbridos o quimeras de estos virus), vectores poxvirus (p.ej., vaccinia, viruela aviar, viruela del canario, vaccinia modificada de Ankara, etc.), vectores de adenovirus y vectores de virus adenoasociados (AAV) (p.ej., véase las referencias 185 a 190). La administración de ADN unido a adenovirus muerto [191] también puede emplearse.

También se pueden emplear vehículos y procedimientos de administración no virales, que incluyen, pero no se limitan a, ADN condensado policatiónico unido o no unido a adenovirus muerto solo [p.ej., 191], ADN ligado a ligandos [192], células vehículos de administración a células eucariotas [p.ej., referencias 193 a 197] y la neutralización o fusión de la carga nucleica con las membranas celulares. El ADN desnudo también se puede emplear. Los ejemplos de procedimientos de introducción de ADN desnudo se describen en las referencias 198 y 199. Los liposomas (p.ej., inmunoliposomas) que pueden actuar como vehículos de administración de genes se describen en las referencias 200 a 204. Se describen enfoques adicionales en las referencias 205 y 206.

La administración no viral adicional adecuada para su uso incluye sistemas de administración mecánica tales como el enfoque descrito en la ref. 206. Además, la secuencia codificante y el producto de la expresión de los mismos pueden administrarse mediante la deposición de materiales de hidrogel fotopolimerizados o el uso de radiación ionizante [p.ej., referencias 207 y 208]. Otros procedimientos convencionales para la administración de genes que pueden usarse para la administración de la secuencia codificante incluyen, por ejemplo, el uso de una pistola de partículas de transferencia génica manual [209] o el uso de radiación ionizante para activar genes transferidos [207 y 208].

La administración de ADN que usa micropartículas de PLG {poli(lactida-co-glicólido)} es un procedimiento particularmente preferido, p.ej., por adsorción a las micropartículas, que se tratan opcionalmente para tener una superficie cargada negativamente (p.ej., tratada con SDS) o una superficie con carga positiva (p.ej., tratada con un detergente catiónico, como CTAB).

Referencia 5 y exenciones

En algunas realizaciones, la invención puede no abarcar el uso de un polipéptido codificado por la SEQ ID NO: 1, p.ej., no abarca el uso de un polipéptido que comprende la secuencia de aminoácidos SEQ ID NO: 2, o no abarca el uso de un polipéptido que tiene la secuencia N-terminal SEQ ID NO: 98. Dichos polipéptidos y sus secuencias codificantes se desvelan en la referencia 5 para uso en la inmunización contra infecciones por NMEC.

En otras realizaciones, sin embargo, los polipéptidos de la referencia 5 están abarcados, pero p.ej., para nuevos propósitos médicos. Como se desvela en el presente documento, la estrecha homología entre diferentes patotipos de *E. coli* significa que una respuesta inmunitaria contra un polipéptido derivado de NMEC puede proporcionar protección cruzada contra cepas no NMEC. Por lo tanto, cuando la invención se refiere al tratamiento o la profilaxis de enfermedades causadas por cepas de *E. coli* que no están incluidas en el patotipo NMEC (p.ej., contra las cepas ExPEC APEC, UPEC o SEPEC; o contra los patotipos intestinales de *E. coli*, tales como las cepas EPEC, EAEC, EIEC, ETEC o DAEC), entonces se pueden abarcar los polipéptidos de referencia 5.

Anticuerpos

Los anticuerpos contra los antígenos de *E. coli* pueden usarse para la inmunización pasiva [210]. Por lo tanto, la divulgación proporciona un anticuerpo que se une a al menos 2 (p.ej., a 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 o los 13) de las 13 proteínas que consisten en cada una de las SEQ ID NOS: 3 a 16. Por lo tanto, la divulgación proporciona un anticuerpo que se une a al menos 2 (p.ej., a 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 o los 13) de las 13 proteínas que consisten cada una en las SEQ ID NOS: 3 a 16, SEQ ID NOS 99-113, y SEQ ID NOS 114-128. Los anticuerpos que se unen a solo uno de dicho grupo de 13 proteínas no están abarcados por la presente invención. Por lo tanto, la divulgación proporciona un anticuerpo que se une a al menos 2 (p.ej., a 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, o las 17) de las 17 proteínas que consisten cada una en la SEQ ID NOS: 3 a 16, 129, 133, 137 y 141, SEQ ID NOS 99-113, 130, 134, 138 y 142, y SEQ ID NOS 114-128, 131, 135, 139 y 143. En ciertas realizaciones, al menos una de las al menos 2 o más proteínas a las que se une el anticuerpo deben seleccionarse de una de las cuatro proteínas que consisten cada una en las SEQ ID NOS: 129, 133, 137 y 141, SEQ ID NOS 130, 134, 138 y 142, y SEQ ID NOS 131, 135, 139 y 143. Los anticuerpos que se unen a solo una de dicho grupo de 17 proteínas no están abarcados por la presente invención.

La divulgación también proporciona el uso de dichos anticuerpos en terapia. La divulgación también proporciona el uso de dichos anticuerpos en la fabricación de un medicamento. La divulgación también proporciona un procedimiento para tratar un mamífero que comprende la etapa de administrar una cantidad eficaz de un anticuerpo de la invención. Como se describió anteriormente para las composiciones inmunógenas, estos procedimientos y usos permiten que un mamífero esté protegido contra la infección por *E. coli*.

El término "anticuerpo" incluye moléculas de inmunoglobulina intactas, así como fragmentos de las mismas que son capaces de unirse a un antígeno. Estos incluyen moléculas de anticuerpos híbridas (quiméricas) [211, 212]; fragmentos F(ab')₂ y F(ab) y moléculas Fv; heterodímeros no covalentes [213, 214]; moléculas Fv monocatenarias (sFv) [215]; construcciones de fragmentos de anticuerpos diméricos y triméricos; minicuerpos [216, 217]; moléculas de anticuerpo humanizadas [218 - 220]; y cualquier fragmento funcional obtenido a partir de tales moléculas, así como los anticuerpos obtenidos a través de procedimientos no convencionales tales como presentación en fagos. Preferiblemente, los anticuerpos son anticuerpos monoclonales. Los procedimientos para obtener anticuerpos monoclonales son bien conocidos en la técnica. Se prefieren anticuerpos humanizados o completamente humanos.

Aspectos generales

La práctica de la presente invención empleará, a menos que se indique lo contrario, procedimientos convencionales de química, bioquímica, biología molecular, inmunología y farmacología, dentro de los conocimientos de la técnica. Tales técnicas se explican detalladamente en la bibliografía. Véase, p.ej., referencias 221-228, etc.

La expresión "que comprende" abarca "que incluyen" así como "que consiste" p.ej., una composición que comprende "X" puede consistir exclusivamente en X o puede incluir algo adicional, p. X + Y.

El término "aproximadamente" en relación con un valor numérico X significa, por ejemplo, $x \pm 10\%$.

La numeración "GI" se usa en el presente documento. Un número GI, o "GenInfo Identifier", es una serie de dígitos asignados consecutivamente a cada registro de secuencia procesado por NCBI cuando las secuencias se añaden a sus bases de datos. El número GI no tiene ninguna relación con el número de acceso del registro de secuencia. Cuando se actualiza una secuencia (p.ej., para la corrección, o para añadir más anotación o información), entonces recibe un nuevo número de GI. Por lo tanto, la secuencia asociada con un número de GI determinado nunca cambia.

Las referencias a una identidad de secuencia porcentual entre dos secuencias de aminoácidos significa que, cuando se alinean, ese porcentaje de aminoácidos es el mismo al comparar las dos secuencias. Esta alineación y el porcentaje de homología o identidad de secuencia se puede determinar usando programas de software conocidos en la técnica, por ejemplo los descritos en la sección 7.7.18 de la ref. 229. Se determina una alineación preferida mediante el algoritmo de búsqueda de homología Smith-Waterman usando una búsqueda de hueco afín con una penalización por apertura de hueco de 12 y una penalización por extensión de hueco de 2, matriz BLOSUM de 62. El algoritmo de búsqueda de homología Smith-Waterman se describe en la ref. 230.

Un experto en la materia entendería que "aislado" significa alterado "por la mano del hombre" respecto a su estado natural, es decir, si ocurre en la naturaleza, ha sido cambiado o eliminado de su entorno original, o ambos. Por ejemplo, un polinucleótido o un polipéptido presente de forma natural en un organismo vivo no está "aislado" cuando está en dicho organismo vivo, pero el mismo polinucleótido o polipéptido separado de los materiales coexistentes de

su estado natural está “aislado”, como se usa el término en esta divulgación. Además, un polinucleótido o polipéptido que se introduce en un organismo por transformación, manipulación genética o por cualquier otro procedimiento recombinante se entendería que está “aislado” incluso si todavía está presente en dicho organismo, que puede ser vivo o no vivo, excepto cuando dicha transformación, manipulación genética u otro procedimiento recombinante produce un organismo que de otro modo es indistinguible del organismo natural.

Breve descripción de los dibujos

La Figura 1 muestra una alineación de CLUSTALW de SEQ ID NOs: 2 a 16. Las regiones N terminales que pueden eliminarse para aumentar la solubilidad a la vez que se mantiene sustancialmente la misma inmunogenicidad se muestran en la parte inferior de la alineación. La región N terminal hasta el enlazador gly-ser o la región gly-ser se denota con “G” y la región rica en prolina se denota con “P.”

La Figura 2 muestra la identidad de aminoácidos entre pares de secuencias.

La Figura 3 muestra el análisis en gel de la proteína purificada, con bandas de alto PM visibles en ausencia de DTT.

La Figura 4 muestra la transferencia Western de cepas de *E. coli* patógenas y no patógenas usando un suero anti-AcfD. El panel (A) es una transferencia Western del lisado celular total. El panel (B) es una transferencia Western del sobrenadante del cultivo. Los carriles en cada uno de los paneles (A) y (B) de izquierda a derecha son los siguientes: Carril M - proteínas marcadoras con el peso molecular en kDa de cada proteína marcadora mostrada a lo largo del lado izquierdo del panel (A); 1 - IHE3034; 2 - CFT073; 3 - 536; 4 - BL21; 5 - MG 1655; 6 - W3110; 7 - NISSLE1917; 8 - IHE3034ΔActD. Como se observa a partir del análisis, las cepas patógenas (IHE3034, carril 1; 536, carril 3) expresan y secretan AcfD mientras que las cepas no patógenas (MG1655, carril 5; W3110, carril 6; Nissle 1917, carril 7) expresan la proteína pero su secreción es defectuosa. Las cepas CFT073 (carril 2) e IHE3034ΔAcfD (carril 8) se utilizan como control negativo, ya que no albergan el gen acfD. La cepa BL21 (carril 4) es una cepa de laboratorio utilizada como control positivo, ya que expresa y secreta AcfD.

La Figura 5 muestra una comparación de la solubilidad de la proteína AcfD y diversos fragmentos de la proteína. El panel (A) es un gel de gradiente SDS-PAGE (tampón MOPS 4-12 %) de muestras a 37 °C que compara el sedimento (carril izquierdo para cada proteína o fragmento) con el sobrenadante (carril derecho para cada proteína o fragmento). Los carriles de izquierda a derecha son: marcadores de peso molecular (bandas de 191 kDa, 97 kDa y 64 kDa), bacterias de control transformadas con el vector de expresión pET sin inserto, expresión bacteriana de 3526 (etiqueta his + péptido líder eliminado), bacterias expresión de L3526 (etiqueta his + longitud completa), expresión bacteriana de L3526-2stop (longitud completa), expresión bacteriana de 3526-DG (etiqueta his + eliminación del extremo N de AcfD hasta el enlazador de glicina-serina flexible), y expresión bacteriana de 3526-DP (etiqueta his + eliminación del extremo N de AcfD a través de la región rica en prolina). El panel (B) es un gel de gradiente SDS-PAGE (tampón MOPS 4-12 %) de muestras a 25 °C siguiendo el mismo orden para los carriles que en el panel (A).

La Figura 6 muestra una comparación de la expresión y la purificación de la proteína AcfD y diversos fragmentos de la proteína. El panel (A) es un gel de SDS-PAGE (tampón MOPS 12 %) de muestras de bacterias que expresan 3526 (etiqueta his + péptido líder eliminado), cultivadas a 25 °C comparando fracciones de varias etapas de la purificación. Los carriles de izquierda a derecha son: M: marcadores de peso molecular (bandas de 191 kDa, 97 kDa y 64 kDa), TOT: lisado bacteriano total, INS: fracción insoluble del lisado bacteriano, SM: fracción soluble del lisado bacteriano, FT: flujo a través de la columna de níquel; E1, E2 y E3 tres eluciones con tampón de imidazol 500 mM. El panel (B) es un gel SDS-PAGE (tampón MOPS al 12 %) de muestras de bacterias que expresan ΔG3526 (etiqueta his + eliminación del extremo N de AcfD al conector flexible de glicina-serina) cultivadas a 25 °C comparando fracciones de varias etapas de la purificación. Los carriles de izquierda a derecha son: M: marcadores de peso molecular (bandas de 191 kDa, 97 kDa y 64 kDa), TOT: lisado bacteriano total, INS: fracción insoluble del lisado bacteriano, SM: fracción soluble del lisado bacteriano, FT: flujo a través de la columna de níquel; E1, E2 y E3 tres eluciones con tampón de imidazol 500 mM. El panel (C) es un gel de SDS-PAGE (tampón MOPS 12 %) de muestras de bacterias que expresan ΔP3526 (etiqueta his + eliminación del extremo N de AcfD a través de la región rica en prolina), cultivadas a 25 °C comparando fracciones de varias etapas de la purificación. Los carriles de izquierda a derecha son: M: marcadores de peso molecular (bandas de 191 kDa, 97 kDa y 64 kDa), TOT: lisado bacteriano total, INS: fracción insoluble del lisado bacteriano, SM: fracción soluble del lisado bacteriano, FT: flujo a través de la columna de níquel; E1, E2 y E3 tres eluciones con tampón de imidazol 500 mM.

La Figura 7 muestra la identidad de aminoácidos entre pares adicionales de secuencias. En 16 *E. coli* enterohemorrágicas (ECEH) no se encontraron genes de AcfD (no se muestran). Las secuencias (donde se representan) de izquierda a derecha o de arriba a abajo son las siguientes: 10 cepas no patógenas o comensales (1: una cepa de *E. coli* comensal, 2: cepa DH10B, 3: cepa MG1655, 4: cepa W3110 (SEQ ID NOs: 14); 5: cepa HS (SEQ ID NO: 13); 9: otra cepa de *E. coli* comensal; y 10: otra cepa de *E. coli* comensal); tres cepas NMEC (1: cepa NMEC RS218; 2: cepa NMEC IHE3034 (SEQ ID NO: 2) y 3: cepa NMEC S88 (SEQ ID NO: 141)); una cepa APEC (1: cepa APEC O1); seis cepas UPEC (2: cepa UPEC 536 (SEQ ID NO: 4); 3: UTI89; 4: cepa UPEC F11

- 5 (SEQ ID NO: 10); 5: cepa UPEC IAI39 (SEQ ID NO: 133); y 6: cepa UPEC UMN026 (SEQ ID NO: 137)); tres cepas EAEC (1: cepa EAEC 101-1 (SEQ ID NO: 3); 2: cepa EAEC 042 (SEQ ID NO: 12); y 3: cepa EAEC 55989 (SEQ ID NO: 129)), una cepa EIEC (1: cepa EIEC 53638 (SEQ ID NO: 5)); cuatro cepas EPEC (2: cepa EPEC E22 (SEQ ID NO: 8)); 3: cepa EPEC E2348/69 (SEQ ID NO: 16); y 4: cepa EPEC E110019 (SEQ ID NO: 7)); tres cepas ETEC (1: cepa ETEC B7A (SEQ ID NO: 6); 2: cepa ETEC E24377A (SEQ ID NO: 9) y 3: cepa ETEC H10407 (SEQ ID NO: 11)); y una cepa resistente a antibióticos (1: cepa SECEC resistente a antibióticos (SEQ ID NO: 15)).

BREVE DESCRIPCIÓN DE LA LISTA DE SECUENCIAS

SEQ ID	Descripción
1	Secuencia codificadora de la cepa NMEC IHE3034
2	Secuencia de la cepa NMEC IHE3034
3	Secuencia de la cepa EAEC 101-1 (GI: 83587587)
4	Secuencia de la cepa UPEC 536 (GI: 110643204)
5	Secuencia de la cepa EIEC 53638 (GI: 75515237)
6	Secuencia de la cepa ETEC B7A (GI: 75227618)
7	Secuencia de la cepa EPEC E110019 (GI: 75239450)
8	Secuencia de la cepa EPEC E22 (GI: 75259912)
9	Secuencia de la cepa ETEC E24377A (GI: 157156747)
10	Secuencia de la cepa UPEC F11 (GI: 75241179)
11	Secuencia de la cepa ETEC H10407
12	Secuencia de la cepa EAEC 042
13	Secuencia de la cepa comensal HS (GI: 157162442)
14	Secuencia de la cepa comensal W3110 (GI: 89109748)
15	Secuencia de la cepa resistente a antibióticos SECEC
16	Secuencia de la cepa EPEC E2348/69
17-95	Fragmentos comunes a las SEQ ID NOs: 2 a 15
96	Nucleótido IC31
97	Péptido IC31
98	Secuencia del extremo N opcionalmente no contemplada
99-113	Deleciones representativas del extremo N de AcfD hasta el enlazador gly-ser o la región gly-ser
114-128	Delección representativa del extremo N de AcfD hasta la región rica en prolina
129	Secuencia de aminoácidos de la cepa EAEC 55989
130	Delección representativa del extremo N de AcfD de la cepa EAEC 55989 hasta el enlazador gly-ser o la región gly-ser
131	Delección representativa del extremo N de AcfD de la cepa EAEC 55989 hasta la región rica en prolina
132	Secuencia de ácido nucleico de la cepa EAEC 55989
133	Secuencia de aminoácidos de la cepa UPEC IAI39
134	Delección representativa del extremo N de AcfD de la cepa UPEC IAI39 hasta el enlazador gly-ser o la región gly-ser
135	Delección representativa del extremo N de AcfD de la cepa UPEC IAI39 hasta la región rica en prolina
136	Secuencia de ácido nucleico de la cepa UPEC IAI39
137	Secuencia de aminoácidos de la cepa UPEC UMN026
138	Delección representativa del extremo N de AcfD de la cepa UPEC UMN026 hasta el enlazador gly-ser o la región gly-ser
139	Delección representativa del extremo N de AcfD de la cepa UPEC UMN026 hasta la región rica en prolina
140	Secuencia de ácido nucleico de la cepa UPEC UMN026

(continuación)

SEQ ID	Descripción
141	Secuencia de aminoácidos de la cepa NMEC S88
142	Deleción representativa del extremo N de AcfD de la cepa NMEC S88 hasta el enlazador gly-ser o la región gly-ser
143	Deleción representativa del extremo N de AcfD de la cepa NMEC S88 hasta la región rica en prolina
144	Secuencia de ácido nucleico de la cepa NMEC S88

Escherichia coli 55989 (aislado asociado a la diarrea, sin plásmido - SEQ ID NOs: 129, 130 (Δ G), 131 (Δ P) y 132): Escherichia coli 55989 es un aislado clínico enteroagregativo. Las cepas de E. coli enteroagregativas se adhieren a las células de la mucosa y son una causa de gastroenteritis.

Escherichia coli IAI39 (aislado de infección del tracto urinario, sin plásmido SEQ ID NOs: 133, 134 (Δ G), 135 (Δ P) y 136): Escherichia coli IAI39 es una cepa de serotipo O7:K1 de una infección del tracto urinario. Escherichia coli UMN026 (aislado de infección del tracto urinario, 1 plásmido - SEQ ID NOs: 137, 138 (Δ G), 139 (Δ P) y 140): Escherichia coli UMN026 es un aislado clínico de serotipo O7:K1. Esta cepa es resistente a los medicamentos.

Escherichia coli S88 (aislado de meningitis, sin plásmido - SEQ ID NOs: 141, 142 (Δ G), 143 (Δ P) y 144): Escherichia coli S88 es una cepa de serotipo O45:K1 aislada de un caso de meningitis neonatal.

MODOS PARA REALIZAR LA INVENCION

Uno de los antígenos desvelados en la referencia 5 está designado como precursor del factor de colonización accesorio D (AcfD) (SEQ ID NO: 2 de aminoácidos del presente documento) de la cepa NMEC IHE3034. Esta proteína se ha expresado y purificado, y confiere protección contra las cepas ExPEC en un modelo animal de sepsis.

Se obtuvieron secuencias para los ortólogos en otras cepas de *E. coli*. La secuencia de aminoácidos observada en IHE3034 también se observó en las cepas APECO1 y UTI89, pero se encontraron 14 secuencias extra (SEQ ID NOs: 3 a 16). La Figura 1 muestra un alineamiento de las SEQ ID NOs: 2 a 16. Hay varios tramos de conservación a lo largo de las secuencias (SEQ ID NOs: 17 a 95). Los 30 aminoácidos N terminales están 100 % conservados, y estos incluyen el péptido señal (aa 1-23) y la cisteína N terminal de la lipoproteína nativa.

Algunas cepas tenían una mutación en el marco de lectura en el gen AcfD, lo que daba como resultado la falta de expresión del polipéptido. El gen *acfD* estaba totalmente ausente de las cepas CFT073, EDL933, Sakai y B171.

La Figura 2 muestra el % de identidad entre las secuencias de aminoácidos. Los marcadores son SEQ ID NOs, a excepción de MG1655, RS218, DH10B, APECO1 y UTI89 donde se usa el nombre de la cepa. El nivel más bajo de identidad (en recuadro en la Figura 2) fue del 85,9 %, entre las SEQ ID NOs: 2 y 4 (ambas cepas ExPEC).

La secuencia de AcfD de la cepa IHE3034 se clonó y luego se expresó a partir de un plásmido como una proteína recombinante marcada con His sin un péptido líder, en un hospedador de *E. coli*. La proteína fue purificada y analizada. La filtración en gel mostró un peso molecular mucho más alto que el previsto basándose únicamente en la secuencia de aminoácidos. El análisis en gel en ausencia de DTT, pero no en su presencia, muestra formas de proteína de mayor peso molecular (Figura 3). Por lo tanto, es probable que la proteína forme oligómeros.

Los sueros obtenidos contra AcfD se usaron en transferencias Western frente a lisados celulares totales (Figura 4 (A)) o sobrenadantes de cultivo precipitados con 60 % de TCA (Figura 4 (B)). Los sueros reconocieron una proteína de ~150 kDa en lisados tanto de cepas patógenas como comensales. No reaccionaron con esta banda en lisados de CFT073 o de un mutante inactivado para AcfD de IHE3034. La reactividad con proteínas en los sobrenadantes indica que la proteína puede ser secretada.

Se inmunizaron ratones CD1 (5 semanas de edad) por vía subcutánea usando 20 μ g del antígeno más adyuvante de Freund (u otro adyuvante como se indica a continuación). Los ratones se inocularon a los 0, 21 y 35 días. Catorce días después de la tercera inoculación, los ratones fueron expuestos a una dosis letal (DL80) de una cepa patógena de *E. coli*. Se recogió sangre de la cola 24 horas después de la estimulación para evaluar la bacteriemia. La mortalidad se controló durante cuatro días después de la exposición. La tasa de protección puede calcularse como (% de muertes en el grupo de control (sin inmunización) - % de muertes en el grupo experimental (inmunizado)) / % de muertes en el grupo de control x 100.

En experimentos adicionales usando un adyuvante de sal de aluminio, al que la proteína se adsorbió por completo, el 75 % de los ratones estaban protegidos frente al 0 % en los grupos de control. Las muertes en ambos grupos ocurrieron 1 día después de la estimulación letal.

En experimentos adicionales, el hospedador recombinante se cultivó en dos condiciones de pO₂ diferentes durante la expresión. En ambas condiciones, la proteína era visible en dos grupos con una carga diferente. No se observaron diferencias significativas en la eficacia protectora de los cuatro grupos diferentes de proteínas.

Aumento de la solubilidad

- 5 Inesperadamente, la proteína AcfD mostró baja solubilidad a pesar de que la proteína es una proteína secretada. Como se muestra en la Figura 5(B), la eliminación del extremo N terminal de AcfD hasta el enlazador gly-ser o la región gly-ser aumentó significativamente la solubilidad cuando se expresaba a 25 °C (Véase pK1-3526-DG Figura 5 (B)). De forma similar, la eliminación del extremo N de AcfD hasta la región rica en prolina aumentó significativamente la solubilidad cuando se expresaba a 25 °C (Véase pK1-3526-DP Figura 5(B)).
- 10 Para confirmar que ambos fragmentos tenían sustancialmente la misma inmunogenicidad que la AcfD de longitud completa, los fragmentos se purificaron. Los fragmentos purificados se usaron en tres experimentos de inmunización en ratones, adyuvados con adyuvante completo de Freund. Los ratones inmunizados fueron expuestos a una dosis letal de *E. coli*. La inmunización con AcfD con el extremo N hasta el enlazador gly-ser o la región gly-ser eliminada protegió al 100 % de los ratones de la muerte, mientras que la muerte ocurrió en el 90 % de los animales en el grupo control no inmunizado. La inmunización con AcfD con el extremo N terminal hasta la región rica en prolina eliminada protegió al 90 % de los ratones de la muerte, mientras que la muerte ocurrió en el 90 % de los animales en el grupo control no inmunizado.
- 15

Expresión y purificación

- 20 Las bacterias con una de las tres construcciones que expresan variantes etiquetadas con His de AcfD se cultivaron en 30 ml de medio y se indujeron para expresar la variante de AcfD a 25 °C (AcfD sin el péptido líder (3526), AcfD con el extremo N eliminado hasta el enlazador gly-ser o la región gly-ser (Δ G3526), y AcfD con el extremo N eliminado hasta la región rica en prolina (Δ G3526)). La bacteria fue cosechada y lisada por sonicación. Las fracciones solubles se aislaron y cargaron en una columna IMAC. La columna se lavó tres veces con tampón de imidazol 20 mM. Las variantes de AcfD se eluyeron a continuación con tres lavados de tampón de imidazol 500 mM.
- 25 Como se muestra en la Figura 6, la eliminación del extremo N de AcfD hasta el enlazador gly-ser o la región gly-ser aumentó significativamente la solubilidad y el rendimiento de la proteína purificada. El rendimiento obtenido estimado mediante el ensayo de Bradford fue el siguiente: 0,18 mg de 3526 y 2,34 mg de Δ G3526.

REFERENCIAS

- 30 [1] Kaper et al. (2004) Nat Rev Microbiol. 2(2):123-40.
 [2] Anjum et al. (2007) Appl Environ Microbiol 73 :5692-7.
 [3] Russo & Johnson (2000) J Infect Dis 181:1753-1754.
 [4] Smith et al. (2007) Foodborne Pathogens And Disease 4:134-63.
 [5] WO2006/089264.
 [6] WO2006/091517.
- 35 [7] Needleman & Wunsch (1970) J. Mol. Biol. 48, 443-453.
 [8] Rice et al. (2000) Trends Genet 16:276-277.
 [9] Needleman & Wunsch (1970) J. Mol. Biol. 48, 443-453.
 [10] Rice et al. (2000) Trends Genet 16:276-277.
 [11] Geysen et al. (1984) PNAS USA 81:3998-4002.
- 40 [12] Carter (1994) Methods Mol Biol 36:207-23.
 [13] Jameson, BA et al. 1988, CABIOS 4(1):181-186.
 [14] Raddizzani & Hammer (2000) Brief Bioinform 1(2):179-89.
 [15] Bublii et al. (2007) Proteins 68(1):294-304.
 [16] De Lalla et al. (1999) J. Immunol. 163:1725-29.
- 45 [17] Kwok et al. (2001) Trends Immunol 22:583-88.
 [18] Brusica et al. (1998) Bioinformatics 14(2):121-30
 [19] Meister et al. (1995) Vaccine 13(6):581-91.
 [20] Roberts et al. (1996) AIDS Res Hum Retroviruses 12(7):593-610.
 [21] Maksyutov & Zagrebelnaya (1993) Comput Appl Biosci 9(3):291-7.
- 50 [22] Feller & de la Cruz (1991) Nature 349(6311):720-1.
 [23] Hopp (1993) Peptide Research 6:183-190.
 [24] Welling et al. (1985) FEBS Lett. 188:215-218.
 [25] Davenport et al. (1995) Immunogenetics 42:392-297.
 [26] Tsurui & Takahashi (2007) J Pharmacol Sci. 105(4):299-316.
- 55 [27] Tong et al. (2007) Brief Bioinform. 8(2):96-108.
 [28] Schirle et al. (2001) J Immunol Methods. 257(1-2):1-16.
 [29] Chen et al. (2007) Amino Acids 33(3):423-8.
 [30] Patente US-5.707.829
- 60 [31] Current Protocols in Molecular Biology (F.M. Ausubel et al. eds., 1987) Supplement 30.
 [32] Vaccine Design: The Subunit and Adjuvant Approach (eds. Powell & Newman) Plenum Press 1995 (ISBN 0-

- 306-44867-X).
- [33] Vaccine Adjuvants: Preparation Methods and Research Protocols (Volume 42 of Methods in Molecular Medicine series). ISBN: 1-59259-083-7. Ed. O'Hagan.
- [34] Patente US-6355271.
- 5 [35] WO00/23105.
- [36] WO90/14837.
- [37] WO90/14837.
- [38] Podda & Del Giudice (2003) Expert Rev Vaccines 2:197-203.
- [39] Podda (2001) Vaccine 19: 2673-2680.
- 10 [40] Vaccine Design: The Subunit and Adjuvant Approach (eds. Powell & Newman) Plenum Press 1995 (ISBN 0-306-44867-X).
- [41] Vaccine Adjuvants: Preparation Methods and Research Protocols (Volume 42 of Methods in Molecular Medicine series). ISBN: 1-59259-083-7. Ed. O'Hagan.
- [42] Allison & Byars (1992) Res Immunol 143:519-25.
- 15 [43] Hariharan et al. (1995) Cancer Res 55:3486-9.
- [44] US-2007/014805.
- [45] Suli et al. (2004) Vaccine 22(25-26):3464-9.
- [46] WO95/11700.
- [47] Patente US-6.080.725.
- 20 [48] WO2005/097181.
- [49] WO2006/113373.
- [50] Han et al. (2005) Impact of Vitamin E on Immune Function and Infectious Diseases in the Aged at Nutrition, Immune functions and Health EuroConference, Paris, 9-10 June 2005.
- [51] Patente US-6630161.
- 25 [52] Patente US-5-057.540.
- [53] WO96/33739.
- [54] EP-A-0109942.
- [55] WO96/11711.
- [56] WO00/07621.
- 30 [57] Barr et al. (1998) Advanced Drug Delivery Reviews 32:247-271.
- [58] Sjolander et al. (1998) Advanced Drug Delivery Reviews 32:321-338.
- [59] Niiikura et al. (2002) Virology 293:273-280.
- [60] Lenz et al. (2001) J Immunol 166:5346-5355.
- [61] Pinto et al. (2003) J Infect Dis 188:327-338.
- 35 [62] Gerber et al. (2001) J Virol 75:4752-4760.
- [63] WO03/024480.
- [64] WO03/024481.
- [65] Gluck et al. (2002) Vaccine 20:B10-B16.
- [66] EP-A-0689454.
- 40 [67] Johnson et al. (1999) Bioorg Med Chem Lett 9:2273-2278.
- [68] Evans et al. (2003) Expert Rev Vaccines 2:219-229.
- [69] Meraldi et al. (2003) Vaccine 21:2485-2491.
- [70] Pajak et al. (2003) Vaccine 21:836-842.
- [71] Kandimalla et al. (2003) Nucleic Acids Research 31:2393-2400.
- 45 [72] WO02/26757.
- [73] WO99/62923.
- [74] Krieg (2003) Nature Medicine 9:831-835.
- [75] McCluskie et al. (2002) FEMS Immunology and Medical Microbiology 32:179-185.
- [76] WO98/40100.
- 50 [77] Patente US-6.207.646.
- [78] Patente US-6.239.116.
- [79] Patente US-6.429.199.
- [80] Kandimalla et al. (2003) Biochemical Society Transactions 31 (part 3):654-658.
- [81] Blackwell et al. (2003) J Immunol 170:4061-4068.
- 55 [82] Krieg (2002) Trends Immunol 23:64-65.
- [83] WO01/9593.
- [84] Kandimalla et al. (2003) BBRC 306:948-953.
- [85] Bhagat et al. (2003) BBRC 300:853-861.
- [86] WO03/035836.
- 60 [87] WO01/22972.
- [88] Schellack et al. (2006) Vaccine 24:5461-72.
- [89] WO95/17211.
- [90] WO98/42375.
- [91] Beignon et al. (2002) Infect Immun 70:3012-3019.
- 65 [92] Pizza et al. (2001) Vaccine 19:2534-2541.
- [93] Pizza et al. (2000) Int J Med Microbiol 290:455-461.

- [94] Scharton-Kersten et al. (2000) *Infect Immun* 68:5306-5313.
- [95] Ryan et al. (1999) *Infect Immun* 67:6270-6280.
- [96] Partidos et al. (1999) *Immunol Lett* 67:209-216.
- [97] Peppoloni et al. (2003) *Expert Rev Vaccines* 2:285-293.
- 5 [98] Pine et al. (2002) *J Control Release* 85:263-270.
- [99] Tebbey et al. (2000) *Vaccine* 18:2723-34.
- [100] Domenighini et al. (1995) *Mol Microbiol* 15:1165-1167.
- [101] WO99/40936.
- [102] WO99/44636.
- 10 [103] Singh et al] (2001) *J Cont Release* 70:267-276.
- [104] WO99/27960.
- [105] Patente US-6.090.406.
- [106] Patente US-5.916.588.
- [107] EP-A-0626169.
- 15 [108] WO99/52549.
- [109] WO01/21207.
- [110] WO01/21152.
- [111] Andrianov et al. (1998) *Biomaterials* 19:109-115.
- [112] Payne et al. (1998) *Adv Drug Delivery Review* 31:185-196.
- 20 [113] Patente US-4.680.338.
- [114] Patente US-4.988.815.
- [115] WO92/15582.
- [116] Stanley (2002) *Clin Exp Dermatol* 27:571-577.
- [117] Wu et al. (2004) *Antiviral Res.* 64(2):79-83.
- 25 [118] Vasilakos et al. (2000) *Cell Immunol.* 204(1):64-74.
- [119] Patentes US-4689338, US-4929624, US-5238944, US-5266575, US-5268376, US-5346905, US-5352784, US-5389640, US-5395937, US-5482936, US-5494916, US-5525612, US-6083505, US-6440992, US-6627640, US-6656938, US-6660735, US-6660747, US-6664260, US-6664264, US-6664265, US-6667312, US-6670372, US-6677347, US-6677348, US-6677349, US-6683088, US-6703402, US-6743920, US-6800624, US-6809203,
- 30 US-6888000 y US-6924293.
- [120] Jones (2003) *Curr Opin Investig Drugs* 4:214-218.
- [121] WO03/011223.
- [122] Johnson et al. (1999) *Bioorg Med Chem Lett* 9:2273-2278.
- [123] Evans et al. (2003) *Expert Rev Vaccines* 2:219-229.
- 35 [124] WO2004/060308.
- [125] WO2004/064759.
- [126] Patente US-6.924.271.
- [127] US2005/0070556.
- [128] Patente US-5.658.731.
- 40 [129] Patente US-5.011.828.
- [130] WO2004/87153.
- [131] Patente US-6.605.617.
- [132] WO02/18383.
- [133] WO2004/01845.
- 45 [134] WO03/082272.
- [135] Wong et al. (2003) *J Clin Pharmacol* 43(7):735-42.
- [136] US2005/0215517.
- [137] Dyakonova et al. (2004) *Int Immunopharmacol* 4(13):1615-23.
- [138] FR-2859633.
- 50 [139] Signorelli & Hadden (2003) *Int Immunopharmacol* 3(8):1177-86.
- [140] WO2004/064715.
- [141] De Libero et al, *Nature Reviews Immunology*, 2005, 5: 485-496
- [142] Patente US-5.936.076.
- [143] Oki et al, *J. Clin. Investig.*, 113: 1631-1640
- 55 [144] US2005/0192248
- [145] Yang et al, *Angew. Chem. Int. Ed.*, 2004, 43: 3818-3822
- [146] WO2005/102049
- [147] Goff et al, *J. Am. Chem., Soc.*, 2004, 126: 13602-13603
- [148] WO03/105769
- 60 [149] Cooper (1995) *Pharm Biotechnol* 6:559-80.
- [150] WO99/11241.
- [151] WO94/00153.
- [152] WO98/57659.
- [153] European patent applications 0835318, 0735898 and 0761231.
- 65 [154] Durant et al. (2007) *Infect Immun* 75:1916-25.
- [155] WO02/081653.

- [156] Donnelly et al. (1997) *Annu Rev Immunol* 15:617-648.
- [157] Strugnell et al. (1997) *Immunol Cell Biol* 75(4):364-369.
- [158] Cui (2005) *Adv Genet* 54:257-89.
- [159] Robinson & Torres (1997) *Seminars in Immunol* 9:271-283.
- 5 [160] Brunham et al. (2000) *J Infect Dis* 181 Suppl 3:S538-43.
- [161] Svanholm et al. (2000) *Scand J Immunol* 51(4):345-53.
- [162] DNA Vaccination - Genetic Vaccination (1998) eds. Koprowski et al. (ISBN 3540633928).
- [163] Gene Vaccination: Theory and Practice (1998) ed. Raz (ISBN 3540644288).
- [164] Findeis et al., *Trends Biotechnol.* (1993) 11:202
- 10 [165] Chiou et al. (1994) *Gene Therapeutics: Methods And Applications Of Direct Gene Transfer.* ed. Wolff
- [166] Wu et al., *J. Biol. Chem.* (1988) 263:621
- [167] Wu et al., *J. Biol. Chem.* (1994) 269:542
- [168] Zenke et al., *Proc. Natl. Acad. Sci. (USA)* (1990) 87:3655
- [169] Wu et al., *J. Biol. Chem.* (1991) 266:338
- 15 [170] Jolly, *Cancer Gene Therapy* (1994) 1:51
- [171] Kimura, *Human Gene Therapy* (1994) 5:845
- [172] Connelly, *Human Gene Therapy* (1995) 1:185
- [173] Kaplitt, *Nature Genetics* (1994) 6:148
- [174] WO 90/07936.
- 20 [175] WO 94/03622.
- [176] WO 93/25698.
- [177] WO 93/25234.
- [178] Patente US-5.219.740.
- [179] WO 93/11230.
- 25 [180] WO 93/10218.
- [181] Patente US-4.777.127.
- [182] Patente GB-2.200.651.
- [183] EP-A-0345242.
- [184] WO 91/02805.
- 30 [185] WO 94/12649.
- [186] WO 93/03769.
- [187] WO 93/19191.
- [188] WO 94/28938.
- [189] WO 95/11984.
- 35 [190] WO 95/00655.
- [191] Curiel, *Hum. Gene Ther.* (1992) 3:147
- [192] Wu, *J. Biol. Chem.* (1989) 264:16985
- [193] Patente US-5.814.482.
- [194] WO 95/07994.
- 40 [195] WO 96/17072.
- [196] WO 95/30763.
- [197] WO 97/42338.
- [198] WO 90/11092.
- [199] Patente US-5.580.859
- 45 [200] Patente US-5.422.120
- [201] WO 95/13796.
- [202] WO 94/23697.
- [203] WO 91/14445.
- [204] EP-0524968.
- 50 [205] Philip, *Mol. Cell Biol.* (1994) 14:2411
- [206] Woffendin, *Proc. Natl. Acad. Sci.* (1994) 91:11581
- [207] Patente US-5.206.152.
- [208] WO 92/11033.
- [209] Patente US-5.149.655.
- 55 [210] Brandt et al. (2006) *J Antimicrob Chemother.* 58(6):1291-4. Epub 2006 Oct 26
- [211] Winter et al., (1991) *Nature* 349:293-99
- [212] Patente US-4.816.567.
- [213] Inbar et al., (1972) *Proc. Natl. Acad. Sci. U.S.A.* 69:2659-62.
- [214] Ehrlich et al., (1980) *Biochem* 19:4091-96.
- 60 [215] Huston et al., (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85:5897-83.
- [216] Packet al., (1992) *Biochem* 31, 1579-84.
- [217] Cumber et al., (1992) *J. Immunology* 149B, 120-26.
- [218] Riechmann et al., (1988) *Nature* 332, 323-27.
- [219] Verhoeyan et al., (1988) *Science* 239, 1534-36.
- 65 [220] Patente GB-2.276.169.
- [221] Gennaro (2000) *Remington: The Science and Practice of Pharmacy.* 20th edition, ISBN: 0683306472.

- [222] Methods In Enzymology (S. Colowick and N. Kaplan, eds., Academic Press, Inc.)
 [223] Handbook of Experimental Immunology, Vols. I-IV (D.M. Weir and C.C. Blackwell, eds, 1986, Blackwell Scientific Publications)
 [224] Sambrook et al. (2001) Molecular Cloning: A Laboratory Manual, 3rd edition (Cold Spring Harbor Laboratory Press).
 [225] Handbook of Surface and Colloidal Chemistry (Birdi, K.S. ed., CRC Press, 1997)
 [226] Ausubel et al. (eds) (2002) Short protocols in molecular biology, 5th edition (Current Protocols).
 [227] Molecular Biology Techniques: An Intensive Laboratory Course, (Ream et al., eds., 1998, Academic Press)
 [228] PCR (Introduction to Biotechniques Series), 2nd ed. (Newton & Graham eds., 1997, Springer Verlag)
 [229] Current Protocols in Molecular Biology (F.M. Ausubel et al., eds., 1987) Supplement 30
 [230] Smith & Waterman (1981) Adv. Appl. Math. 2: 482-489.

LISTADO DE SECUENCIAS

SEQ ID NO: 1

TTGTCACTTGCCTTATTAATGAATAAGAAATTTAAATATAAGAAATCGCTTTTAGCGGCTATTTTAAGCGCAACCCCT
 GTTAGCCGGTGTGTATGGTGGTGGTTCAGGATCGTCTCCGATACGCCGTGTGTAGATTCTGGATCAGGGACTTTGTC
 CGGAAGTGAAACCCGATCCAACACCAACCCCGGAGCCGACACCTGAGCCGACGCCGACCCAGAACCTACGCCGGAT
 CCAACACCTGATCCTGAGCCGACACCCAGAACCCGAGCCAGAACCTGTTCTACGAAAACGGGTTATCTGACCCCTGGG
 CGGAAGCCAGCGGGTAACTGGTGTACCTGTAATGGTGAATCCAGCGATGGCTTTACCTTTACGCCAGGCAATACCG
 TGAGTTGTGTGGTGGGCAGTACGACCATTGCAACATTCAACACCCAGTCAGAAGCTGCGCGTAGCCTGCGTGGCGGTT
 GACAAAGTGTCTTTAGCCTGGAGGACGCGCAGGAGCTGGCGAATTTCTGAAAATAAGAAAACCAACGCCATCTCTCT
 GGTGACGTCAGCGACAGTTGCCCGCAGATGCAAGACAGCTTTGTCTTACTTTCTCGTCAGTGGTTGATCGCGCGC
 GATTTGAAAAACTGTATAAGCAAAATGATCTGGCAACAGACAATTCAGCAAGCTGGTCAATGAAGAGCTGGAAC
 AATGCTGCGACTGATAAAGCGCCGTCCACCCATACCTCAACGGTAGTGCCAGTCACGACAGAGGGAAACAAACCCGGA
 TCTGAACGCGTCTCTGTCGCGTAAACGCGGAACAGTTTATCAGTATCAACCCACTGAAATCATTTCTTTCCGAAG
 GCCAAGCTGGTGGATAGCCTGGGGAACGGTGTGCTGGCGTTGACTACTACCAATTCAGGCGGTGGCGTAAGTACG
 GAAAACGTAATTTTCTTTAGCTGGGGCGAAACCATCTCTTTTGGTATCGATACCTTTGAACTGGGCTCAGTACG
 TGGCAATAAGTCGACCATTGCGCTGACTGAATTGGGTGATGAAGTTCGCGGGGCAATATCGATCAGCTCAATTCATC
 GTTATTCGACGACTGGTCAAATAATACTCGTGTGTGTCGGACGATGTACGCAAGGTCTTTGCGCAATATCCCAAC
 GTGATCAACGAGATAATCAATCTTTCTGTTATCCACGGTGGCAGCTGGATGAAGCGGATCAAAACGTTGTGCTGCC
 TAACGAATTTATCGAGCAGTTTAAGACGGGTGAGGCCAAAGAGATCGATACCGCGAATTTGTCGGAAGACCGCGGTT
 GTAACGAGGCTCGCTGGTTCTCGCTGACACGCGCAATGTTAATGACGGCCAGATTGAGGCGTTTATTAACAGCTG
 TGGGCGGTGGATACGAATATCAGTCTGTGAGCAAGTTCACGCTCTTCCATGACTCTACCAACTTCTATGGCAGCAC
 CGGTAAACGCGCGGTGACGCGGTGGTAAATATCTCCAATCGGCATTCCCGATTCTGATGGCGCGTAATGATAAAA
 ACTATGGCTGGCGTTTGGCGAAAACGCGCCTGGGATATAAAATGAGCTGGCGTACATTACGGAAGCGCCTTCCATT
 GTGCAGCCAGAGAACGTTACGCGGATACCTGCGACTTTCAACCTGCGGTTTATTTGCTGGGGCAAGTCGGTGAAGG
 CAAACTGATGGTTATCGGTAACCCGCACTACAACAGCATCTGCGTTGCCCGAACGTTACAGTTGGGGCGGTGGTG
 TTAATAGTAAAGGTGAGTGACGTCAGCGGTGATTCTGATGACATGAAGCACTTTATGCAGAACGTAAGTGGCGTAC
 TTGTCAAATGACATCTGGCAGCCAAATACCAAGGATCATGACTGTGCGCACCAACCTGGAGAACGCTTTATTTCAA
 AAAAGCGGGCCAGGTATTGGGAAATAGTGACCATTTGCTTTCCATGAGGATTTCACTGGTATCACGGTTAAACAGT
 TGACCAGCTATGGCGATCTGAATCCGGAAGAGATTCCGTTGCTGATCCTCAACGGCTTTGAAATATGTGACTCAGTGG
 TCTGGCGATCCCTATGCTGTGCTCTGCGTGCAGATACGACAAACCGAAGCTGACTCAGCAGGATGTGACCGATCT
 GATCGCTTATCTGAACAAAGGTGGCTCGGTGCTGATCATGAAAACGTTGATGAGCAATCTTAAGGAAGAGCGCGT
 CCAGTTTGTGCGTCTGCTGGATGCCGCGGCTCTGTCAATGGCTCTGAACAAATCGGTGGTGAACACGATCCGCA
 GGGTATCCGGATCGCGTTCTGTCAGCGTCCGCGGACTGGCATTTGGGTTTATGAACGTTATCTGCTGCAGACGGCGC
 GCAACCCCGGTACACCATCGACCCAAATACAGGGGAAGTGACCTGAAAATACAGCAAGACAACAAGCCTGATGACA
 AGCCGAACTGGAAGTTGCGAGCTGGCAGGAGGAAGTTGAGGGCAAACAGGTAACGCGTTATGCGCTTTATGATGAA
 GCGGAATACACAACAGAAGATCTCTGGAAGCGGCAAGGCAAAATCTTTGAGAGCTTTCTGGGTTACAGGAGTG
 TAAGGACTCGACTTACCATTACGAGATTAAGTGTGAGAGCGCGCCAGGCACGGATGTTCCGGTAAACAGTGGCA
 TGTATGTTCCGCGTATACGCAACTGAATCTTGACGCCGACACCGCAAGCGATGGTGCAGGCGCGGATTTAGGC
 ACCAATACCTGACCGCTGATCAGCATGAGCTTTATTTCCGTACCAAGGCGATGAAGGTGAGCGTCTGAACAGTGT
 TGATCTGGAACGTTCTGTACAGAACATGTGCTGTGGCTGTGGAACGATACGAAATATCGTTACGAAGAGGGCAAGG
 AAGATGAGCTGGGCTTTAAACGTTTCAACGAGTTCTGAACTGTACGCCAATGATGCTATGCAGCGCGCAACAG
 TGCTCCGAGATCTGAAAAATCGCTGGTGCATAACAATGATCTACGGTGACGGTAGCAGCAAGCGGGCATGAT
 GAACCAAGCTATCCGCTCAACTATATGAAAAACCGCTGACCGCTGTGATGCTGGGCGGTTCTCTGGTGGGATCTGA
 ACATTAAGGTTGATGTGGAGAAGTACCCAGGATCCGTATCGGCAAGGGTGAGAGCGTTACGGAACATCAGCCTG
 TACTCGAATCCGACCAATGGTTTGGCGGTAACATGCAGTCAACCGGCTGTGGGCACCGGCCAGCAGGACGTAC
 CATTAAGCTTTCGGCTCAGTCCAGTGACTGTTACCGTGGCGCTGGCTGACGACCTGACTGGACGTGAGAAGCATG
 AAGTTGCGCTGACCGCTCCGCAAGAGTGACTAAACGTTATCTGGAAGGCTAACCGGTGAAGTGAAGCGGTAAGCAGGATGTTAC
 CTATAAAACTTGACGGGGCACAAGCATCGTTTACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGGTT
 ATCCGGTAAATGAACAGCAGCTTCTCGACGAACAGCACACGCTGCCGACGACGCCGTGAACGACTGGCTGATTTGG
 CACGAAGTCGGTCATAACGCTGCAGAAACACCGCTGAACGTACCGGGTGCAACTGAAGTGGCGAACAACGTGCTGGC
 GCTGTACATGCAGGATCGTATCTCGGTAAGATGAACGCTGTGCTGACGACATTACCGTCGCGCGGAATATCTGG
 ACGAGAGCAACGGTCAGGCTGGGCGCGCGGGTGGCGGTGACCGTCTGCTGATGTACGCACAGTTGAAGGAGTGG
 GCAGAGGAAAACCTTGATATCAAACAGTGGTATCCAGATGGTGAGCTGCCAAGTTCTACAGCGATCGTAAAGGGAT
 GAAGGGCTGGAACCTGTTCCAGTTGATGCACCGTAAAGCGCGCGCGATGATGTTGGTAAACGACCTTTGGTGGCA
 AGAATTACTGTGCTGAATCAATGGTAACGCTGCCGACACGCTGATGCTGTGTCATCCTGGGTCGCTCAGGCGGAT
 CTTTCGGAATTTCTTAAGAAATGGAATCCGGGTGCAAGTGCTTACCAAGTTGCCGGGAGCAACGGAGATGAGTTTCCA

GGGCGGTGTGAGCTCTTCGGCTTACAGCAGCTGGCGTCACTCAAGCTGCCGAAACCGGAAAAAGGGCCCGAAACCA
TTACAAGGTTACCGAGCATAAGATGTCTGCCGAGTAA

SEQ ID NO: 2

MSLALLMNNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPD
PTPDPEPTPEPEPEPVPTKTGYLTLGGSQRTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAARSLRAV
DKVSFSLDAQELANSENKKTNAISLVTSSDSCPADAELCLTFSSVVDRAFEKLYKQIDLATDNFSKLVNEEVEN
NAATDKAPSTHTSTVVPVTTTEGTPDLNASFVSANAEQFYQYQPTTEIILSEGQLVDSLNGVAGVDYNTNSGRGVT
ENGKFSFSWGETISFGIDTFELGSGVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPN
VINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTRNVNDGQIQGVINKL
WGVDTNYQSVSKFHVFDSTNFGYSTGNARGQAVVNISNAPFILMARNDKNYWLAFGEKRAWDKNELAYITEAPSI
VQPENVTRDTATFNLFPFISLGQVGEKLMVIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLR
LSNDIQWPNTKSIIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDNLPEEIPLLILNGFEYVTQW
SGDPYAVPLRADTSKPPLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLDAGLSMALNKSVMNDPQ
GYPDVRQRRTATGIWVYERYPAADGAQPPYITDPNTGEVTKWYQDQNKPDKPKLEVASWQEEVEGKQVTRYAFIDE
AEYTTESLEAAKAKIFEKFPGLKECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLADTAKAMVQAADLG
TNIQRLYQHLYFRTNGRKGERLSSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKTFTEFLNXYANDAYAGGTK
CSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSWSWDLNKKVDVEKYPGVSVAKGESVTENISL
YSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPVTKTYTEANGEVTFKV
PYGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGWKNLDSAPLGELESASFVYTTPKKNLEASNFTGGVAEFAK
DLDTFASSMNDFYGRNDEGDKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTPLNDWLIW
HEVGHNAAETPLNVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEDSENGQAWARGGAGDRLLMYAQLKEW
AEENFDIKQWYPDGELPKFYSDRKGKMGWNLFLQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMCASWVAQAD
LSEFFKKWNPAGASAYQLPGATEMSFQGGVSSSAYSSTLASLKLKPKPEKGPETINKVTEHKMSAE

SEQ ID NO: 3

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
PTPEPEPEPVPTKTGYLTLGGSQRTGATCNGESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVSF
SLEDAQELAASDDKKSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATD
KAPSTHTSPVVPVTTTTPGTPDLNASFVSANAEQFYQYQPTTEIILSEGLVDSMGNGVGVNYYTSSGRGVTGENGKF
NFSWGETISFGIDTFELGSGVRGNKSTIALTELGDVEVRGANIDQLIHRYSQAGKNDEREVPDVVRKVFAEYPNVINEI
INLSLSNGEALSEGQDTFERTNEFLEQFESGQAKEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVNDK
DYKSVTKFHVFDSTNFGYSTGNARGQAVVNISNAPFILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPEN
VTRDTATFNLFPFISLGQVGEKLMVIGNPHYNSILRCPNGYSWEGGVNDKNGQCTRNSDSNDMKHFMQNVLRYSDDK
WTPDAKASMTVGTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLIILNGFEYVTQVGNOPY
AIPLRADTSKPPLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDAGLSMALNKSVMNDPQGYPNR
VRQQRATGIWVYERYPAVDGALPYTIDSKTGEVKWKYQVENKPDKPKLEVASWLEDVDGKQETRYAFIDEADHTE
DSLKAAKAKIFEKFPGLKECKDPTYHYEVNCLYRPGTGVPVTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRL
YQHLYFRTNGRKGERLSSVDLERLYQNMSVWLWNLKIEYRYENDKDELGFKTFTEFLNXYANDAYTGGTQCSDELK
KSLVDNNMIYGEKSVNKGMMNPSYPLNYMEKPLTRLMLGRSWSWDLNKKVDVEKYPGAVSAEGEKVTETISLYSNPT
KWFAGNMQSTGLWAPAQKEVTIESSASVPVTVTVALADDLTGREKHEVALNRPPKVTKYTELKANGEVKTVPYGG
IYIKGNSPQNESAEFTFTGVVKAPFYKDGAWKNALNSAPLGELESDAFVYTTPKKNLEASNFTGGVAEFAKDLDTF
ASSMNDFYGRNDEGDKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTPLNDWLIWHEVGH
NAAETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNF
DIKQWYPEGDLPKFYSDREGMKWNLFLQLMHRKARGDEVGKTKFGERNYCAESNGNAADKMLCASWVAQTDLSEFF
KKWNPAGANAYQLPGASEMNFEGGVSQSAYETLAALNLPKPKQGPETINQVTEHKMSAE

SEQ ID NO: 4

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
PTPEPEPEPVPTKTGYLTLGGSQRTGATCNGESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVSF
SLEDAQELAASDDKKSNAVSLVTSSNSCPADTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATD
KAPSTHTSPVVPVTTTTPGTPDLNASFVSANAEQFYQYQPTTEIILSEGLVDSMGNGVGVNYYTSSGRGVTGENGKF
NFSWGETISFGIDTFELGSGVRGNKSTIALTELGDVEVRGANIDQLIHRYSQAGKNDEREVPDVVRKVFAEYPNVINEI
INLSLSNGEALSEGQDTFERTNEFLEQFESGQAKEIDTAICDSLGGCNSQRWFSLTARNVNDGQIQGVINKLWGVNDK
DYKSVTKFHVFDSTNFGYSTGNARGQAVVNISNAPFILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPEN
VTRDTATFNLFPFISLGQVGEKLMVIGNPHYNSILRCPNGYSWEGGVNDKNGQCTRNSDSNDMKHFMQNVLRYSDDK
WTPDAKASMTVGTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLIILNGFEYVTQVGNOPY
AIPLRADTSKPPLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDAGLSMALNKSVMNDPQGYPNR
VRQQRATGIWVYERYPAVDGALPYTIDSKTGEVKWKYQVENKPDKPKLEVASWLEDVDGKQETRYAFIDEADHTE
DSLKAAKAKIFEKFPGLKECKDPTYHYEVNCLYRPGTGVPVTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRL
YQHLYFRTNGRKGERLSSVDLERLYQNMSVWLWNLKIEYRYENDKDELGFKTFTEFLNXYANDAYTGGTQCSDELK
KSLVDNNMIYGEKSVNKGMMNPSYPLNYMEKPLTRLMLGRSWSWDLNKKVDVEKYPGAVSAEGEKVTETISLYSNPT

KWFAGNMQSTGLWAPAQKEVTIESTASVAVTVTVALADDLTGREKHEVALNRPPKVTKTYELKANGEVKTVPYGGI
IYIKGNSPQNESAEFTFTGVVKAPFYKDGAWKNALNSPAPLGELESDAFVYTTPKKNLEASNYKGGQEQFAEELDTF
ASSMNDIFYGRNDEGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEAGH
NAAETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKNF
DITKWYPEGNLPKFYSEREGMKGWNLFQLMHRKARGDEVGKTKFGERNYCAESNGNAADTLMLCASWVAQTDLSAFF
KKWNPANAYQLPGASEMNFEGGVSQSAYETLAALNLPKPQQGPETINKVTEYSMPAE

SEQ ID NO: 5

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
PTPEPEPEPVPTKTGYLTLGGGSRVGTGATCNGESSDGFTEKPGEDVTCVAGNTTIATFNTQSEAARSLRAVEKVSFS
LEDAQELAGSDDKKSNASLVTSNSNCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFFKKLVNEEVENNAATDK
APSTHTSPVVPVTTPGTKPDLNASFVSANAEQFYQYQPTFIIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFS
FSWGETISFGIDTFELGSVRGKNSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEI
NLSLSNGATLGEGEQVVNLPNEFIEQFNTGQAKEIDTAICAKTDGDCNEARWFSLTTRNVNDGQIQGVINKLWGVDTN
YKSVSKFHVFDSTNIFYGSTGNARGQAVVNIISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPEN
TRDTATFNLFFISLGQVGEGLMVGIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMEVNLRYLSDDKW
KPDASMTVGTNLDTVYFKRHGQVTGNSAAFDHFDFAGISVEHLSSYGDLDPEMPLLIILNGFEYVTQVGNDPYA
IPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSVMNNDPQGYPNRV
RQQRATGIWVYERYPAVDGALPYTIDSKTGEVKWKYQVENKPDDEKPKLEVASWLEDVVGKQETRYAFIDEADHKTED
SLKAAKEKIFAAFPGLKECTNPAYHYEVNCLYRPGTGVPVTTGGMYVPQYTLQSLNADTAKAMVQAADLGTNIQRLY
QHLYFRTNGRKGERLSSVDLERLYQNMSVWLWNDTSYRYEEGKNDELGFKTFTFEFLNCYANDAYAGGTKCSADLKK
SLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLYSNPTKW
FAGNMQSTGLWAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGIY
IKGNSSTNESASFTFTGVVKAPFYKDGAWKNLNSPAPLGELESDAFVYTTPKKNLNASNYTGGLEQFANDLDTFAS
SMNDFHGRDSEDGKHRMFTYKNLPGHKHRFTNDVQISIGDAHSGYPVMNSSFSPNSTTLPTTPLNDWLIWHEVGHNA
AETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKNFI
KKWYPDGTPLPEFYSEREGMKGWNLFQLMHRKARGDEVSNDFGKKNYCAESNGNAADTLMLCASWVAQTDLSEFFK
KWNPANAYQLPGASEMSFEGGVSQSAYNTLASLDLPKEPQGPETINQVTEHKMSAE

SEQ ID NO: 6

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
PTPEPEPEPVPTKTGYLTLGGSLRVGTGDTCNDESSDGFTEFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVS
FSLEDAQELAGSDNKKSNALSLVTSNMNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFFKKLVNEEVENNAAT
DKAPSTHTSPVVPVTTPGTKPDLNASFVSANAEQFYQYQPTFIIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGE
FSFSWGETISFGIDTFELGSVRGKNSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINE
IINLSLSNGATLDEGEQVVNLPNEFIEQFNTGQAKEIDTAICAKTDGDCNEARWFSLTTRNVNDGQIQGVINKLWGVDTN
YNYKSVSKFHVFDSTNIFYGSTGNARGQAVVNIISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPE
NVTRDTATFNLFFISLGQVGEGLMVGIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMEVNLRYLSDD
KWTPDAKASMTVGTNLDTVYFKRHGQVTGNSAAFDHFDFAGISVEHLSSYGDLDPEMPLLIILNGFEYVTQVGNDP
YAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSVMNNDPQGYPN
RVRQQRATGIWVYERYPAVDGALPYTIDSKTGEVKWKYQVENKPDDEKPKLEVASWLEDVVGKQETRYAFIDEADHKT
EDSLKAAKEKIFAAFPGLKECTNPAYHYEVNCLYRPGTGVPVTTGGMYVPQYTLQSLNADTAKAMVQAADLGTNIQRL
LYQHLYFRTNGRKGERLSCVDLERLYQNMSVWLWNDTSYRYEEGKNDELGFKTFTFEFLNCYANDAYAGGTKCSADL
KKSLLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLYSNPT
KWFAGNMQSTGLWAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGI
IYIKGNSSTNESASFTFTGVVKAPFYKDGAWKNLNSPAPLGELESDAFVYTTPKKNLNASNYTGGLEQFANDLDTFAS
ASSMNDIFYGRDSEDGKHRMFTYKNLPGHKHRFANDVQISIGDAHSGYPVMNSSFSPNSTTLPTTPLNDWLIWHEVGH
NAAETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKNF
DIKKWYPDGTPLPEFYSEREGMKGWNLFQLMHRKARGDEVSNDFGKKNYCAESNGNAADTLMLCASWVAQTDLSEFF
FKWNPANAYQLPGASEMSFEGGVSQSAYNTLASLKLKPEPQGPETINKVTEHKMSVE

SEQ ID NO: 7

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
PTPEPEPEPVPTKTGYLTLGGSLRVGTGDTCNDESSDGFTEFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVS
FSLEDAQELAGSDNKKSNALSLVTSNMNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFFKKLVNEEVENNAAT
DKAPSTHTSPVVPVTTPGTKPDLNASFVSANAEQFYQYQPTFIIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGE
FSFSWGETISFGIDTFELGSVRGKNSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINE
IINLSLSNGATLGEGEQVVNLPNEFIEQFNTGQAKEIDTAICAKTDGDCNEARWFSLTTRNVNDGQIQGVINKLWGVDTN
YNYKSVSKFHVFDSTNIFYGSTGNARGQAVVNIISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSVIRPE
NVTRETATFNLFFISLGQVGDGKLMVGIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMEVNLRYLSND
RWLPDAKSSMTVGTNLDTVYFKKHGQVLGNSAPFAFHKDFGTGIVKPMTSYGNLNPDEVPLLIILNGFEYVTQWGS
YSIPLRADTSKPKLTQQDVTDLIAYMKNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSVMNNDPQGYPD

RVRQRRSTPIWVYERYPAVDGKPPYTIIDDTTKEVIWKYQOENKPDCKPKLEVASWQEEVEGKQVTFQAFIDEADHKT
 PESLAAAKQRIIDAFPGLEVCKDSYHYEVNCLERYRPGSGVPVVTGGMYVPQYTQLDLGADTAKAMLOAADLGTNIQR
 LYQHLYFRTNGRQGERLNSVDLERLYQNMSVWLWNETKYRYEEGKEDELGFKTFTEFLNCTNNAYVGTQCSAELK
 KSLIDNKMIYGEESKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLYSNPTK
 WFAGNMQSTGLWAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGI
 YIKGNSSTNESASFTFTGVVKAPFYKDGAWKNDLNSPAPLGELESDAFVYTPPKNLNASNYTGGLEQFANDLDTFA
 SSMNDFYGRDSESGKHRMFTYKNLTGHKHRFANDVQISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGH
 AAETPLTVPGATEVANNVLALYMQDRYLGMNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFD
 IKKWYPEGELPKFFSDREGMKGWNLFQLMHRKARGDDVGNKTFGGKNYCAESNGNAADSLMLCASWVAQTDLSAFEK
 KWNPGANAYQLPGATEMSFEGGVSQSAYSTLASLKLKPKPEQGPETINKVTEHKMSLE

SEQ ID NO: 8

MNKKFKYKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
 PTPEPEPEPVPTKTGYLTLGGSLRVTDITCNDESSDGFFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVS
 FSLEDAQELAGSDNKKSNALSLVTSMNSCPANTEQVCLEFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAAT
 DKAPSTHTSPVVPVTPGTPKPDNLASFVSANAEQFYQYQPSIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGE
 FFSWGETISFGIDTFELGSRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEYPNVINE
 IINLSLSNGATLDEGEQVVNLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGV
 TNYKSVSKFHVFDSTNFGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPE
 NVTRDTATFNLPTISLGQVGECKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFENVRLYLSDD
 KWTDAKASMTVGTNLDTVYFKRHGQVTGNSAAFDHDPDFAGISVEHLSSYGDLPQEMPLLIILNGFEYVTQVGN
 YAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDAAGLSMALNKSVMNDPQGYPN
 RVRQRRSTPIWVYERYPAVDGALPYTIDSKTGEVKWKYQVENKPDCKPKLEVASWLEDVVGKQETRYAFIDEADHKT
 EDLSKAAKEKIFAAFPGLKECTNPAYHYEVNCLERYRPGTGVPVVTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQR
 LYQHLYFRTNGRQGERLSSVDLERLYQNMSVWLWNETSYRYEEGKNDELGFKTFTEFLNCTANDAYAGGTKCSADL
 KKSVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLYSNPT
 KWFAGNMQSTGLWAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGI
 IYIKGNSSTNESASFTFTGVVKAPFYKDGAWKNDLNSPAPLGELESDAFVYTPPKNLNASNYTGGLEQFANDLDTFA
 ASSMNDFYGRDETSGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGH
 NAAETPLTVPGATEVANNVLALYMQDRYLGMNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNF
 DIKKWYPEGELPKFFSDREGMKGWNLFQLMHRKARGDDVGDKTGGKNYCAESNGNAADTLMLCASWVAQTDLSAFEK
 KWNPGANAYQLPGATEMSFEGGVSQSAYSTLASLKLKPKPEQGPETINKVTEHKMSLE

SEQ ID NO: 9

MNKKFKYKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
 PTPEPEPEPVPTKTGYLTLGGSLRVTDITCNDESSDGFFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVS
 FSLEDAQELAGSDNKKSNALSLVTSMNSCPANTEQVCLEFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAAT
 DKAPSTHTSPVVPATTPGTPKPDNLASFVSANAEQFYQYQPTIILSEGRLVDSQGDGVGVNYTNSGRGVTGENGE
 FFSWGETISFGIDTFELGSRGNKSTIALTELGDVEVRGANIDQLIHRYSKAGQNHTRVVPDEVKVFAYEYPNVINE
 IINLSLSNGATLGEQVVNLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGV
 TNYKSVSKFHVFDSTNFGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPE
 NVTRDTATFNLPTISLGQVGDGKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFENVRLYLSDD
 KWTDAKASMTVGTNLDTVYFKRHGQVTGNSAAFDHDPDFAGISVEHLSSYGDLPQEMPLLIILNGFEYVTQVGN
 YAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDAAGLSMALNKSVMNDPQGYPN
 RVRQRRSTPIWVYERYPAVDGKPPYTIIDDTTKEVIWKYQOENKPDCKPKLEVASWQEEVEGKQVTFQAFIDEADHKT
 PESLAAAKQRIIDAFPGLEVCKDSYHYEVNCLERYRPGTDVPVVTGGMYVPQYTQLDLSADTAKAMLOAADLGTNIQR
 LYQHLYFRTNGRQGERLNSVDLERLYQNMSVWLWNETKYRYEEGKEDELGFKTFTEFLNCTNNAYVGTQCSAELK
 KSLIDNKMIYGEESKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGVVNTNGETVTQININLYSAPTK
 WFAGNMQSTGLWAPAQQEVSIKATVPVTVTVALADDLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGI
 YIKGDSKEVQSADFTFTGVVKAPFYKDGKWHDLNSPAPLGELESASFVYTPPKNLNASNYTGGLEQFANDLDTFA
 SSMNDFYGRDSESGKHRMFTYKNLPGHKHRFANDVQISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGH
 AAETPLTVPGATEVANNVLALYMQDRYLGMNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFD
 IKKWYPDGTPLEFYSEREGMKGWNLFQLMHRKARGDEVNDKFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFF
 KWNPGANAYQLPGASEMSFEGGVSQSAYNTLASLKLKPKPEQGPETINKVTEHKMSVE

SEQ ID NO: 10

MNKKFKYKSLLAAILSATLLAGCDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPT
 PEPTPEPEPEPVPTKTGYLTLGGSQRVGTGATCNDESSDGFFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVDKVS
 FSLEDAQELANSNKTNAISLVTSDDSCPADAEQLCLTFSSVVDRAFEKLYKQIDLATDNFSKLVNEEVENNAAT
 DKAPSTHTSTVVPVTTGTPKPDNLASFVSANAEQFYQYQPTIILSEGRLVDSLGNVAGVDYYTNSGRGVTDENGK
 FFSWGETISFGIDTFELGSRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEYPNVINE
 IINLSLSNGATLDEGDQNVNLPNEFIEQFKTGQAKEIDTAICAKTNGCNEARWFSLTTRNVNDGQIQGVINKLWGV

TNYQSVSKFHVFDSTNFYGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSI VQPE
 NVTRDTATFNL PFTSLGQVGDGKLMVIGNPHYNSILRCPNGYSWGGVNSKGECTLSGSDDMKHFQNVRLRYLSND
 IWQPNTKSIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDNLNPEEIPLLILNGFEYVTQWSGDP
 YAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLDDAAGLSMALNKS VVNNDPQGYPD
 RVRQQREKGIWVYERYPFVDGKPPYTI DETTKEVIWKYQQDNKPDDKPKLEVASWLEDVDGKQVKRYAFIDEAEHET
 NESLEAAKAKI IKAFFGLEECKDPTYHYEVNCL EYRPGTNVPVTGGMYVPRYTQLNLSADTAKAMVQAADLGTNIQR
 LYQHLYFRTNGRKGERLSSVDLERLYQNMSVWLWNEIEYSYDSSKEDELGFKTFTEFLNCYANDAYTKGTLCSAEL
 KQSLIDNKMIYEGESKAGMMNPSYPLNYMEKPLTRLMLGRS WWDLNIKVDVEKYPGAVSVGGEEVTETISLYSNPTK
 WFAGNMQSTGLWAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPRVTKTYSLDASGTVKFKVPYGGLI
 YIKGNSSTNESASFTFTGVVKAPFYKDGAWKNDLNSPAPLGELESASFVYTTPKKNLNASNYTGGLDQFAKDLDTFA
 SSMNDFYGRNDEGDKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTTPLNDWLIWHEVGHN
 AAETPLNVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLDENSGQAWARGGAGDRLLMYAQLKEWAEKNFD
 ITKWYPDGKLPAFYSEREGMKGNLFLQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFK
 KWNPGANAYQLPGAEMSFEAGVSSSAYSTLASLNLPKPEKGPETINKVTEHKMSAE

SEQ ID NO: 11

MNKKFKYKKSLLAAILSATLLAGCDGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
 PTPEPEPEPVPTKTGYLTGGSQRVTGATCNGESSDGTFTKPGEDVTCVAGNTTIATFNTQSEAARSLRAVEKVSFS
 LEDAQELAGSDDKKSNAVSLVTSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAP EEFKKL VNEEVENNAATDK
 APSTHTSPVVPVPTPGTKPDLNASFVSANAEQFYQYQPT EII LSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFS
 FSWGEATISFGIDTFELGSVRGNKSTIALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINELI
 NLSLSNGATLGEQEVVNLPNEFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTN
 YKSVSKFHVFDSTNFYGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSI VRPEN
 TRETASFNL PFTSLGQVGDGKLMVIGNPHYNSILRCPNGYSWGGVNSKGECTLSGSDDMKHFQNVRLRYLSNDRW
 LPDAKSSMTVGTNLETVYFKKHGQVLGNSAPFAFHEDFTGITVKPMTSYGNLNPDEVPLLI LNGFEYVTQWSGDPYS
 IPLRADTSKPKLTQQDVTDLIAYMKNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKS VVNNDPQGYPD
 RQRRSTPIWVYERYPAVDGKPPYTI DDTTKEVIWKYQQENKPKDDKPKLEVASWQEEVEGKQVTQFAFIDEADHKTPE
 SLAAAKQRI L DAFPGLEVCKDSYHYEVNCL EYRPGTDVVPVTGGMYVPQYTQLDLSADTAKAMLQAADLGTNIQR
 LYQHLYFRTNGRQGERLNSVDLERLYQNMSVWLWNEIEYSYDSSKEDELGFKTFTEFLNCYTNNAYVGTQCSAELKKS
 LIDNKMIYGEESKAGMMNPSYPLNYMEKPLTRLMLGRS WWDLNIKVDVEKYPGVVNTGETVTQININLYSAPT KWF
 AGNMQSTGLWAPAQKEVSI ESSTVPVTVTVALADDLTGREKHEVSLNRPRVTKTYDLKANDKVTFKVPYGGLIYI
 KGDSKEVQSADFTFTGVVKAPFYKDGKQHDLNSPAPLGELESASFVYTTPKKNLNASNYTGGLEQFANDLDTFASS
 MNDFYGRDSEDKHRMFTYKNLPGHKHRFANDVQISIGDAHSGYPVMNSSFSPNSTLPTTPLNDWLIWHEVGHNA
 ETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDI
 KWYPDGTPLPEFYSEREGMKGNLFLQLMHRKARGDEVSNDFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFK
 WNPANAYQLPGASEMSFEAGVSQSAYNTLASLDLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 12

MNKKFKYKKSLLAAILSATLLAGCDGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
 PTPEPEPEPVPTKTGYLTGGSQRITGATCNGESSDGTFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVSFS
 SLEDAQELAASDDKKSNAVSLVTSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAP EEFKKL VNEEVENNAATD
 KAPSTHTSPVVPVPTPGTKPDLNASFVSANAEQFYQYQPT EII LSEGRLVDSMGNGVGVNYTSSGRGVTGENGKF
 NFSWGETISFGIDTFELGSVRGNKSTIALTELGDEVRGANIDQLIHRYSQAGKNDEREVPDVVRKVFAYEPNVINEI
 INLSLSNGEALSEGDTFERTNEFLEQFESGQAKEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDK
 DYKSVTKFHVFDSTNFYGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPEN
 VTRDTATFNL PFTSLGQVGEKLMVIGNPHYNSILRCPNGYSWGGVNSKGECTLSGSDDMKHFQNVRLRYLSNDK
 WTPDAKASMTVGTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLPQKMPLLI LNGFEYVTQVGGDPY
 AVPLRADTSKPKLSQDDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKS VVNNDPQGYPD
 VRQRRATGIWVYERYPVVEGELPYTIDSKTGKVTWKYQIDNKPKDKPKLEVASWQEEVDGKQVTQFAFIDEADHKT
 ESLDAAKKILEKFKGLEECKDSTYHYEINCL EYRPGTNVPATGGMYVPRYTQLNLSADTAKAMVQAADLGTNIQR
 LYQHLYFRTNGRKGERLSSVDLERLYQNMSVWLWNEIEYSYDSSKEDELGFKTFTEFLNCYANDAYTGGTQCSDELK
 KSLVDNMIYGEKSVNKAAGMMNPSYPLNYMEKPLTRLMLGRS WWDLNIKVDVEKYPGAVSEEGQEVTESISLYSNPT
 KWFAGNMQSTGLWAPAQKEVTIKSNADVPTVTVALADDLTGREKHEVALNRPPKVTKYELKANGEVKTVPYGGI
 IYIKGNSKENNKASFTFTGVVKAPFYKNGAWKNALNSPAPLGELESDAFVYTTPKKNLEASNFTGGVAEFAKDLDT
 FASSMNDFYGRNDEGDKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTTPLNDWLIWHEVG
 HNAAETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKN
 FDIKKWYPDGTPLPEFYSEREGMKGNLFLQLMHRKARGDEVSNDFGGKNYCAESNGNADTLMLCASWVAQTDLSE
 FFKKWNPGANAYQLPGATEMSFEAGVSQSAYNTLASLDLPKPKQGPETINKVTEYSMPAE

SEQ ID NO: 13

MNKKFKYKKSLLAAILSATLLAGCDGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
 PTPEPEPEPVPTKTGYLTGGSQRVTGATCNGESSDGTFTKPGEDVTCVAGNTTIATFNTQSEAARSLRAVEKVSFS

I,EDAQELAGSDDKKSNAVSLVTSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDK
 APSTHTSPVVPVTPGTPKPDNLASFVSANAEQFYQYQPTETIILSEGRVDSQGYGVAGVNYTNSGRGVTGENGEFS
 FSWGETISFGIDTFELGSGVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDVRKVFAYEPNVINEII
 NLSLSNGATLGEQEVVNLNPNFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTRNVNDGQIQGVINKLWGVDTN
 YKSVSKFHVFDSTNFGSTGNARGQAVVNI SNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVPEENV
 TRETATFNLFPFISLGQVGDGKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFMEVNLRYLSNDRW
 LPDAKSNMTVGTNLDTVYFKKHGQVTGNSAAFGEHDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNDPYA
 IPLRADTSKPKLTQQDVTDLIAYMNKGGSVLIMENVMSNLKEESASGFVRLLOAAGLSMALNKSVMNNDPQGYPDRV
 RQRRSTPIWVYERYPAVDGKPPYTIDDTTKEVIWKYQQENKPDCKPKLEVASWQEEVEGQVTFQFAFIDEADHKTP
 SLAAAKKRILDAFPGLEECKDSYHYEVNCLYRPGTGPVPTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLY
 QHELYFRTNGRKGERLSSVDLERLYQNMSVWLWNKIEYRYENDKDDELGFKTFTFEFLNCRYANNAYDGGTQCSAELKQ
 SLIDNKMIYEGESKAGMMNPSYPLNYMEKPLTRMLGRSOWDLNPKVDVEKYPGAVSAEGEEVTETINLYSNPTKWF
 AGNMQSTGLWAPAQQEVSISNAKVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYI
 KSDSKEEKSANFTFTGVVKAPFYKDGKWKNDLKSAPLGELESASFVYTTPKKNLEASNYKGLKQFAEDLDTFASS
 MNDFYGRDGSSEKHRMFTYEALTGHKHRFTNDVQISIGDAHSGYPVMNSSFSNPTTLPTPLNDWLIWHEVGHNA
 ETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIK
 QWYPEGSLPAFYSEREGMKGNLFLQMLHRKARGDDVGNDFGNRNYCAESNGNAADTLMLCASWVAQTDLSAFEKKW
 NPGANAYQLPGATEMSFEGGVSQSAYNTLASLDLPKPKQGPETINKVTEYSMPAE

SEQ ID NO: 14

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPE
 PTPEPEPEPVPTTKTYLTLGGSQRVGTGATCNGESSDGTFFKPGEDVTCVAGNTTIATFNTQSEAARSLRAVEKVSFS
 LEDAQELAGSDDKKSNAVSLVTSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDK
 APSTHTSPVVPVTPGTPKPDNLASFVSANAEQFYQYQPTETIILSEGRVDSQGYGVAGVNYTNSGRGVTGENGEFS
 FSWGETISFGIDTFELGSGVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDVRKVFAYEPNVINEII
 NLSLSNGATLGEQEVVNLNPNFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTRNVNDGQIQGVINKLWGVDTN
 YKSVSKFHVFDSTNFGSTGNARGQAVVNI SNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVPEENV
 TRDTATFNLFPFISLGQVGEGLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFMEVNLRYLSDDKW
 KPDAKASMTVGTNLDTVYFKRHGQVTGNSAAFDFHDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNDPYA
 IPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLOAAGLSMALNKSVMNNDPQGYPNRV
 RQQRATGIWVYERYPAVDGALPYTIDSKTGEVWKYQVENKPDCKPKLEVASWLEDVVGQETRYAFIDEADHKTED
 SLKAAKEKIFAAPGLKECTNPAYHYEVNCLYRPGTGPVPTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLY
 QHELYFRTNGRKGERLSSVDLERLYQNMSVWLWNTSYRYEEGKNDELGFKTFTFEFLNCRYANDAYAGTKCSADLKK
 SLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRMLGRSOWDLNPKVDVEKYPGAVSEEGQNVTTETISLYSNPTKW
 FAGNMQSTGLWAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYI
 IKGNSSTNESASFFTGVVKAPFYKDGAWKNDLNSPAPLGELESADFVYTTPKKNLNASNYTGGLEQFANDLDTFAS
 SMNDFYGRDSEDGKHRMFTYKNLPGHKHRFTNDVQISIGDAHSGYPVMNSSFSNPTTLPTPLNDWLIWHEVGHNA
 AETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKNFDI
 KKWYPDGTPLPEFYSEREGMKGNLFLQMLHRKARGDEVNDKFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFK
 KWNPGANAYQLPGASEMSFEGGVSQSAYNTLASLDLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 15

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPE
 PTPEPEPEPVPTTKTYLTLGGSQRVGTGATCNGESSDGTFFKPGEDVTCVAGNTTIATFNTQSEAARSLRAVEKVSFS
 LEDAQELAGSDDKKSNAVSLVTSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDK
 APSTHTSPVVPVTPGTPKPDNLASFVSANAEQFYQYQPTETIILSEGRVDSQGYGVAGVNYTNSGRGVTGENGEFS
 FSWGETISFGIDTFELGSGVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDVRKVFAYEPNVINEII
 NLSLSNGATLGEQEVVNLNPNFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTRNVNDGQIQGVINKLWGVDTN
 YKSVSKFHVFDSTNFGSTGNARGQAVVNI SNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVPEENV
 TRETATFNLFPFISLGQVGDGKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFMEVNLRYLSNDRW
 LPDAKSNMTVGTNLDTVYFKKHGQVTGNSAAFGEHDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNDPYA
 IPLRADTSKPKLTQQDVTDLIAYMNKGGSVLIMENVMSNLKEESASGFVRLLOAAGLSMALNKSVMNNDPQGYPDRV
 RQRRSTPIWVYERYPAVDGKPPYTIDDTTKEVIWKYQQENKPDCKPKLEVASWQEEVEGQVTFQFAFIDEADHKTP
 SLAAAKKRILDAFPGLEECKDSYHYEVNCLYRPGTGPVPTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLY
 QHELYFRTNGRKGERLSSVDLERLYQNMSVWLWNKIEYRYENDKDDELGFKTFTFEFLNCRYANNAYDGGTQCSAELKQ
 SLIDNKMIYEGESKAGMMNPSYPLNYMEKPLTRMLGRSOWDLNPKVDVEKYPGAVSAEGEEVTETINLYSNPTKWF
 AGNMQSTGLWAPAQQEVSISNAKVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYI
 KSDSKEEKSANFTFTGVVKAPFYKDGKWKNDLKSAPLGELESASFVYTTPKKNLEASNYKGLKQFAEDLDTFASS
 MNDFYGRDGSSEKHRMFTYEALTGHKHRFTNDVQISIGDAHSGYPVMNSSFSNPTTLPTPLNDWLIWHEVGHNA
 ETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIK
 QWYPEGSLPAFYSEREGMKGNLFLQMLHRKARGDDVGNDFGNRNYCAESNGNAADTLMLCASWVAQTDLSAFEKKW
 NPGANAYQLPGATEMSFEGGVSQSAYNTLASLDLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 16

MNKKFKYKKSLLAAILSATLLAGCDGGSGPSSDTPPVDSTGSLPEVKPDPTPNPEPTPEPTPDPEPTPDPTPDPE
PTPEPEPEPVPTTKGYLTLGGSQRITGATCNGESSDGTFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVSF
SLEDAQELAASDDKKSNAVSLVTSSNSCPADTEQVCLTFSSVIESKRFDLSYKQIDLAPEEFKKLNVNEEVENNAATD
KAPSTHTSPVVPVPTPGTKPDLNASFVSANAEQFYQYQPTTEIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEF
SFSWGETISFGIDTFELGSGVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEI
INLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVOK
DYKSVTKFHVFDSTNFYGSTGNARQAVVNISNAAFILMARNDKNYWLAFGEKRAWOKNELAYITEAPSIQVQEN
VTRDTATFNLFPFISLGQVGEKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLSGDSDDMKHFMQNVRLYLSDDK
WTPDAKASMTVGTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLIILNGFEYVTQVGNPDY
AIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKS SVNTDPQGYPNR
VRQREKGIWVYERYPAVDSAQPPYTIIDPDGKVTWKYQEEGKPDCKPKLEVASWQEDVDGKQVTRYAFIDEAEHST
EESLEAAKAKIFEKFPGLQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLDADTAKAMVQAADLGTNIQR
LYQHELYFRTNGRKGERLSSVDLERLYQNMSVWLWNKIEYRYENDKDELGFKTTFTEFLNCYANNAYSEGQTCSADL
KKS LVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSWDLNLIKVDVEKYPGAVSAEGEKTETISLYSNPT
KWFAGNMQSTGLWAPAQQEVTIESTASVPVTVTVALADDLTGREKHEVALNRPPKVTKYDLKANDKVTFKVPYGGI
IYIKGNSPKNESAFTFTGVVKAPFYKDGWKNALNSPAPLGELESDFS VYTAPKNNLNASNYSTDGVAEFKEL
DTFASMNDFYGRDGESGNHRMFTYKALTGHKHRFANDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHE
VGHNAEETPLNVPGATEVANNVLALYMQDRYLGMNVRADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEWAE
ENFDIKQWYPDGELPKFYSDRKGMMGNLFLQLMHRKARGDDVSNDKFGGRNYCAESNGNAADTLMLCASWVAQADLS
EFFKKWNPANAYQLPGASEMSFEGGVSQSAYNTLAAMHLSKPEKGPETINKVTEYSMPAE

SEQ ID NO: 17

LPEVKPDPTP

SEQ ID NO: 18

PEPTPEPTPDPEPTP

SEQ ID NO: 19

EPTPEPEPEPVPTTKGYLTLGGS

SEQ ID NO: 20

ESSDGTFT

SEQ ID NO: 21

TQSEAARSLRAV

SEQ ID NO: 22

KVSFSLEDAQEL

SEQ ID NO: 23

LYKQIDLA

SEQ ID NO: 24

KLVNEEVENNAATDKAPSTHTS

SEQ ID NO: 25

GTKPDLNASFVSANAEQFYQYQP

SEQ ID NO: 26

EIILSEG

SEQ ID NO: 27

ISFGIDTFELGSGVRGNKSTIALTELGDVEVRGANIDQLIHRYS

SEQ ID NO: 28

VRKVFA

SEQ ID NO: 29

YPNVINEIINLSLSNG

SEQ ID NO: 30

GQAKEIDTAIC

SEQ ID NO: 31
RWFSLT

SEQ ID NO: 32
IQGVINKLWGVD

5 **SEQ ID NO: 33**
KFHVFHDSTNFGSTGNARGQAVVNISN

SEQ ID NO: 34
AFPILMARNDKNYWLAFGEKRAWDKN

10 **SEQ ID NO: 35**
LAYITEAPS

SEQ ID NO: 36
PENVTR

SEQ ID NO: 37
FNLPFISLGQVG

15 **SEQ ID NO: 38**
GKLMVIGNPHYNSILRCPNGYSW

SEQ ID NO: 39
NVLRYLS

20 **SEQ ID NO: 40**
MTVGTNL

SEQ ID NO: 41
PLLILNGFEYVTQ

SEQ ID NO: 42
PLRADTSKPKL

25 **SEQ ID NO: 43**
QQDVTDLIAY

SEQ ID NO: 44
NKGGSVLIMENVMSNLKEESAS

30 **SEQ ID NO: 45**
FVRLLDAAGLSMALNKSVVNNDPQGY

SEQ ID NO: 46
IWVYERYP

SEQ ID NO: 47
KPKLEVAS

35 **SEQ ID NO: 48**
AFIDEA

SEQ ID NO: 49
YHYEVNCL

40 **SEQ ID NO: 50**
TGGMYVP

SEQ ID NO: 51
ADTAKAM

SEQ ID NO: 52
QAADLGTNIQRLYQHELYFRT

5 **SEQ ID NO: 53**
VDLERLYQNMSVWLWN

SEQ ID NO: 54
DELGKTFTEFLNCY

10 **SEQ ID NO: 55**
KAGMMNPSPYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPG

SEQ ID NO: 56
PTKWFAGNMQSTGLWAPAQ

SEQ ID NO: 57
VTVTVALADDLTGREKHEV

15 **SEQ ID NO: 58**
VPYGGLIYIK

SEQ ID NO: 59
FTFTGVVKAPFYK

20 **SEQ ID NO: 60**
SPAPLGELES

SEQ ID NO: 61
FVYTTPKKNL

SEQ ID NO: 62
LDTFASSMNDP

25 **SEQ ID NO: 63**
GKHRMFTY

SEQ ID NO: 64
GHKHRF

30 **SEQ ID NO: 65**
NDVQISIGDAHSGYPVMNSSFS

SEQ ID NO: 66
NSTTLPTTPLNDWLIWHE

SEQ ID NO: 67
GHNAAETPL

35 **SEQ ID NO: 68**
VPGATEVANNVLALYMQDRYLGMNVRVADDITVAPEYL

SEQ ID NO: 69
QAWARGGAGDRLLMYAQLKEWAE

40 **SEQ ID NO: 70**
GMKGWNLFQLMHRKARGD

SEQ ID NO: 71
NYCAESNGN

SEQ ID NO: 72
LMLCASWVAQ

5 **SEQ ID NO: 73**
FFKKWNPGA

SEQ ID NO: 74
AYQLPGA

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNFEPTPEPTPDPEPTPDPTPDP
MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNFEPTPEPTPDPEPTPDPTPDP
EPTPEPEPEPVPTKTGYLTLGGS
CDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNFEPTPEPTPDPEPTPDPTPDP
CDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNFEPTPEPTPDPEPTPDPTPDP EPTPEPEPEPVPTKTGYLTLGGS
MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTP (S/P) VDSG (S/T) G (S/T) LPEVKPDPT (N/T) PEPTPE
PTPDPEPT (D/E) P (T/I) PDP
EPTPEPEPEPVPTKTGYLTLGGS (Q/L) R (V/I) TG
TCN (G/D) ESSDGFTF (T/K) PG
(S/N) TTIATF (N/D) TQSEAARSLRAV (D/E) KVSFSLEDAQEIA
SLVTS (S/M) (D/N) SCPA (DN) (A/T) EQ (L/V) CL (T/E) FSSV
LYKQIDLA (T/P) (D/E) (N/E) F (S/K) KLVNEEVENNAATDKAPSTIITS (T/P) VVF (V/A) TT (E/P) GTKPDL
NASTVSANAEQFYQYQP (T/S) EIILSEG (Q/R) LVDS
SGRGVT (D/G) ENG (K/E) F (S/N) FSWGE (T/A) ISFGIDTFELGSVRGNKSTIALTELGDEVIRGANIDQLIHYS
VRKVFA (A/E) YPNVINEIINLSLSNG
RWFSLT (T/A) RNVN
KFHFVHDSSTNFYGSTGNARGQAVVNISN (A/S) AFPILMARNDKNYWLAFGEKRAWDKN (D/E) LAYITEAPSIV
FNLFPFISLGQVG (D/E) GKLMVIGNPHYNSILRCPNGYSW
MTVGTNL (D/E) NVYFK
PLRADTSKPKL (S/T) QQDVTDLIAY (L/M) NKGGSVLIMENVMSNLKEESAS (G/S) FVRLDAAGLSMALNKSVV
NNDPQGY
ADTAKAM (V/L) QAADLGNTIQRLYQHELYFRT
DLDTFASSMNDFYGR
FFKKWNPGANAYQLPGA

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTP

10 **SEQ ID NO: 96**
ICICICICICICICICICIC

SEQ ID NO: 97
KLKLLLLLKLK

15 **SEQ ID NO: 98**
MSLALLM

SEQ ID NO: 99

DTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGYLTGGSQRVTGATCNGE
SSDGFTFTPGNTVSCVVGSTTIATFNTQSEAAARSLRAVDKVSFSLEDAQELANSSENKKTNAISLVTSSDSCPADAEQ
LCLTFSSVVDRAFEKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTTEGTPDLNASFVSANAEQF
YQYQPTTEIILSEGQLVDSLGNVAGVDYTNSSGRGVTDENGKFSFSWGETISFGIDTFELGVSVRGNKSTIALTELGD
EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAK
EIDTAICAKTDGNCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHVFDSTNFGSTGNARGQAVVNISN
SAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIQVENVTRDTATFNLFFISLGQVGEGLMVGIGNPHYNSI
LRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVRLYLSNDIWPNTKSI MTVGTLNENVYFKKAGQVLGNSAPFA
FHEDFTGITVKQLTSYGDNLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIM
ENVMSNLKEESASSFVRLDLAAGLSMALNKS VVNNDPQGYPDVRVQRATGIWVYERYPAADGAQPPYITIDPNTGEV
TWKYQQDNKPDDPKLEVASWQEEVEGKQVTRYAFIDEAEYTTESLEAAKAKI FEKFPGLQECKDSTYHYEINCLE
RRPGTDVPVTTGGMYVPRYTQLNLADATKAMVQAADLGTNIQRLYQHELYFRTKGSKGERLNSVDLERLYQNMSVWL
WNDTKYRYEKGDELGFKTFTEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPL
TRLMLGRSWWDLNIKVDVEKYPGVSASAKGESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTV
ALADDLTGREKHEVALNRPRVTCTYLEANGEVTFKVPYGGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGWKN
LDSPAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEDEGKHRMFTYKNLTGHKHRFTN
DVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNAAEPLTVPGATEVANNVLALYMQDRYLGMNR
VADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGLPKFYSDRKGKMGWNLQLMHRKA
RGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKWNPASAYQLPGATEMSFQGGVSSSAYSTLAS
LKLKPEKGPETINKVTEHKMSAE

SEQ ID NO: 100

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPTPEPEPEPVPTKTGYLTGGSQRITGATCNGE
SSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAASDDKKSNAVSLVTSSNSCPANTE
QVCLTFSSVIESKRFDSLKYQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTPGTKPDLNASFVSANAEQ
FYQYQPTTEIILSEGRVDSMGNGVVGNYTSSGRGVTGENGKFNFSWGETISFGIDTFELGVSVRGNKSTIALTELG
DEVRGANIDQLIHRYSQAGKNDEREVDPVVRKVFAEYPNVINEIINLSLSNGEALSEGDTFERTNEFLEQFESGQA
KEIDTAICDSLGGCNSQRWFSLTARNVNDGQIQGVINKLWGVDDKYKSVTKFHVFDSTNFGSTGNARGQAVVNIS
NAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLFFISLGQVGEGLMVGIGNPHYNS
ILRCPNGYSWEGVDKNGQCTRNDSNDMKHFMQNVRLYLSDDKWT PDAKASMTVGTLNLT VYFKRHGQVTGNSAEF
GFHPDFAGISVEHLSSYGDLPQEMPLLIILNGFEYVTQVGNDPYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLI
MENVMNSNLKEESASGFVRLDLAAGLSMALNKS VVNNDPQGYPNRVRQQRATGIWVYERYPAVDGALPYITIDSKTGEV
KWYQVENKPPDDPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKAKI FEKFPGLQECKDPTYHYEVNCL
YRPGTGVPVTTGGMYVPPYTQLSLNADATKAMVQAADLGTNIQRLYQHELYFRTNGRKGERLSSVDLERLYQNMSVWL
WNKIEYRYENDKDELGFKTFTEFLNCYANDAYTGGTQCSDELKKSLVDNNMIYGEKSVNKAGMMNPSYPLNYMEK
LTRLMLGRSWWDLNIKVDVEKYPGAVSAEGEKVTETISLYSNPTKWFAGNMQSTGLWAPAQKEVTIESSASVPVTVT
VALADDLTGREKHEVALNRPPKVTCTYELKANGEVTFKVPYGGGLIYIKGNSPQNESAEFTFTGVVKAPFYKDGAWN
ALNSPAPLGELESADFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEDEGKHRMFTYKNLTGHKHRFT
NDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNAAEPLTVPGATEVANNVLALYMQDRYLGMNR
RVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKQWYPEGDLPKFYSDREGMKGWNLQLMHRK
ARGDEVGKTKFGERNYCAESNGNAADKLMLCASWVAQTDLSEFFKKWNPANAYQLPGASEMNFEGGVSQSAYETLA
ALNLKPKPQQGPETINQVTEHKMSAE

SEQ ID NO: 101

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPTPEPEPEPVPTKTGYLTGGSQRITGATCNGE
SSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAASDDKKSNAVSLVTSSNSCPANTE
QVCLTFSSVIESKRFDSLKYQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTPGTKPDLNASFVSANAEQ
FYQYQPTTEIILSEGRVDSMGNGVVGNYTSSGRGVTGENGKFNFSWGETISFGIDTFELGVSVRGNKSTIALTELG
DEVRGANIDQLIHRYSQAGKNDEREVDPVVRKVFAEYPNVINEIINLSLSNGEALSEGDTFERTNEFLEQFESGQA
KEIDTAICDSLGGCNSQRWFSLTARNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFGSTGNARGQAVVNIS
NAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLFFISLGQVGEGLMVGIGNPHYNS
ILRCPNGYSWEGVDKNGQCTRNDSNDMKHFMQNVRLYLSDDKWT PDAKASMTVGTLNLT VYFKRHGQVTGNSAEF
GFHPDFAGISVEHLSSYGDLPQEMPLLIILNGFEYVTQVGNDPYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLI
MENVMNSNLKEESASGFVRLDLAAGLSMALNKS VVNNDPQGYPNRVRQQRATGIWVYERYPAVDGALPYITIDSKTGEV
KWYQVENKPPDDPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKAKI FEKFPGLQECKDPTYHYEVNCL
YRPGTGVPVTTGGMYVPPYTQLSLNADATKAMVQAADLGTNIQRLYQHELYFRTNGRKGERLSSVDLERLYQNMSVWL
WNKIEYRYENDKDELGFKTFTEFLNCYANDAYTGGTQCSDELKKSLVDNNMIYGEKSVNKAGMMNPSYPLNYMEK
LTRLMLGRSWWDLNIKVDVEKYPGAVSAEGEKVTETISLYSNPTKWFAGNMQSTGLWAPAQKEVTIESTASVAVTVT
VALADDLTGREKHEVALNRPPKVTCTYELKANGEVTFKVPYGGGLIYIKGNSPQNESAEFTFTGVVKAPFYKDGAWN
ALNSPAPLGELESADFVYTTPKKNLEASNFKGGQFAEELDTFASSMNDFYGRNDEDEGKHRMFTYKNLTGHKHRFT
NDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEAGHNAAEPLTVPGATEVANNVLALYMQDRYLGMNR

RVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKNFDITKWPENLPKFYSEREGMKGWNLQLMHRK
ARGDEVGKTKFGERNYCAESNGNAADTLMLCASWVAQTDLSAFFKKWNPANAYQLPGASEMNFEGGVSQSAYETLA
ALNLKPKPQQGPETINKVTEYSMPAE

SEQ ID NO: 102

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPTPEPEPEPVPTKTGYLTLGGSQVRVTGATCNGE
SSDGFTFKPGEDVTCVAGNTTIATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSDDKKSNASVLTSSNSCPANTEQ
VCLTFSSVIESKRFDSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTPGTKPDLNASFVSANAEQF
YQYQPTTEILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSVRGNKSTIALTELGD
EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLGEGEQVNNLPNEFIEQFNTGQAK
EIDTAICAKTDGECNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFGSTGNARGQAVVNISN
AAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLFFISLGQVGEGLMVGIGNPHYNSI
LRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFEMENVLRYSDDKWKPDAKASMTVGTNLDTVYFKRHGQVTGNSAAFD
FHPDFAGISVEHLSSYGDLDPQEMPLLIINGFEYVTQVGNOPYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIM
ENVMNSNLKEESASGFVRLDDAAGLSMALNKSVVNNOPQGYPNRVRQQRATGIWVYERYPAVDGALPYTIDSKTGEVK
WKYQVENKPDCKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKEKIFAAFPGLKECTNPAYHYEVNCL
RPGTGVPVTTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQHLYFRTNGRKGERLSSVDLERLYQNMSVWLW
NDTSYRYEKGKDELGFKTFTFEFLNXYANDAYAGGTKCSADLKKSLVDNNMIYDGGSSKAGMMNPSYPLNYMEKPLT
RLMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLNSNPTKWFAGNMQSTGLWAPAQKEVTIKSNANVPVTVTV
LADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNESASFSTFTGVVKAPFYKDGAWKNDL
NSPAPLGELESDAFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFFHGRDSEDKHRMFTYKNLPGHKHRTND
VQISIGDAHSGYPVMNSSFSPNSTTLPPTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLAALYMQDRYLGMKNR
ADDITVAPEYLEESNQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWPYDGTPLPEFYSEREGMKGNLFLQMLHRKA
RGDEVSNDFGGKKNYCAESNGNAADTLMCASWVAQTDLSEFFKKWNPANAYQLPGASEMSFEGGVSSQAYNTLAS
LDLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 103

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPTPEPEPEPVPTKTGYLTLGGSRLVTGDITCND
ESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSVTSMNSCPANT
EQVCLEFSSVIESKRFDSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTPGTKPDLNASFVSANAE
QFYQYQPTSEILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSVRGNKSTIALTEL
GDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLGEGEQVNNLPNEFIEQFNTGQ
AKEIDTAICAKTDGECNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFGSTGNARGQAVVNI
SNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLFFISLGQVGEGLMVGIGNPHYN
SILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFEMENVLRYSDDKWTPDAKASMTVGTNLDTVYFKRHGQVTGNSAA
FDFHPDFAGISVEHLSSYGDLDPQEMPLLIINGFEYVTQVGNOPYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVL
IMENVMNSNLKEESASGFVRLDDAAGLSMALNKSVVNNOPQGYPNRVRQQRATGIWVYERYPAVDGALPYTIDSKTGE
VKWKYQVENKPDCKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKEKIFAAFPGLKECTNPAYHYEVNCL
EYRPGTGVPVTTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQHLYFRTNGRKGERLSCVDLERLYQNMSVW
LWNDTSYRYEKGKDELGFKTFTFEFLNXYANDAYAGGTKCSADLKKSLVDNNMIYDGGSSKAGMMNPSYPLNYMEK
LTRLMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLNSNPTKWFAGNMQSTGLWAPAQKEVTIKSNANVPVTVT
VALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNESASFSTFTGVVKAPFYKDGAWKN
DLNSPAPLGELESDAFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFFHGRDSEDKHRMFTYKNLPGHKHRTFA
NDVQISIGDAHSGYPVMNSSFSPNSTTLPPTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLAALYMQDRYLGMKN
RVADDITVAPEYLEESNQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWPYDGTPLPEFYSEREGMKGNLFLQMLHR
KARGDEVSNDFGGKKNYCAESNGNAADTLMCASWVAQTDLSEFFKKWNPANAYQLPGASEMSFEGGVSSQAYNTLAS
ASLKL2KPEQGPETINKVTEHKMSVE

SEQ ID NO: 104

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPTPEPEPEPVPTKTGYLTLGGSRLVTGDITCND
ESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSVTSMNSCPANT
EQVCLEFSSVIESKRFDSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTPGTKPDLNASFVSANAE
QFYQYQPTSEILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSVRGNKSTIALTEL
GDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLGEGEQVNNLPNEFIEQFNTGQ
AKEIDTAICAKTDGECNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFGSTGNARGQAVVNI
SNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLFFISLGQVGEGLMVGIGNPHYN
SILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFEMENVLRYSDDKWTPDAKASMTVGTNLDTVYFKRHGQVTGNSAA
FDFHPDFAGISVEHLSSYGDLDPQEMPLLIINGFEYVTQVGNOPYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVL
IMENVMNSNLKEESASGFVRLDDAAGLSMALNKSVVNNOPQGYPNRVRQQRATGIWVYERYPAVDGALPYTIDSKTGE
VKWKYQVENKPDCKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKEKIFAAFPGLKECTNPAYHYEVNCL
EYRPGTGVPVTTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQHLYFRTNGRKGERLSCVDLERLYQNMSVW
LWNDTSYRYEKGKDELGFKTFTFEFLNXYANDAYAGGTKCSADLKKSLVDNNMIYDGGSSKAGMMNPSYPLNYMEK
LTRLMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLNSNPTKWFAGNMQSTGLWAPAQKEVTIKSNANVPVTVT
VALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNESASFSTFTGVVKAPFYKDGAWKN
DLNSPAPLGELESDAFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFFHGRDSEDKHRMFTYKNLPGHKHRTFA
NDVQISIGDAHSGYPVMNSSFSPNSTTLPPTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLAALYMQDRYLGMKN
RVADDITVAPEYLEESNQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWPYDGTPLPEFYSEREGMKGNLFLQMLHR
KARGDEVSNDFGGKKNYCAESNGNAADTLMCASWVAQTDLSEFFKKWNPANAYQLPGASEMSFEGGVSSQAYNTLAS
ASLKL2KPEQGPETINKVTEHKMSVE

5

TRLMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLNSNPTKWFAGNMQSTGLWAPAQKEVTIKSNANVPVTVT
ALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNESASFSTFTGVVKAPFYKDGAWKND
LNSPAPLGELESDAFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFFHGRDSESGKHRMFTYKNLPGHKHRTFAN
DVQISIGDAHSGYPVMNSSFSPNSTTLPPTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLAALYMQDRYLGMKNR
VADDITVAPEYLEESNQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWPYEGELPKFFSDREGMKGNLFLQMLHRKA
RGDDVGNKTFGGKKNYCAESNGNAADTLMCASWVAQTDLSAFFKKWNPANAYQLPGATEMSFEGGVSSQAYSTLAS
LKLKPEQGPETINKVTEHKMSLE

SEQ ID NO: 105

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPTPEPEPEPVPTKTGYLTLGGSLRVTGDITCND
 ESSDGFFTFTPGDKVTCVAGNNTTATFDTQSEAAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSLVTSMNSCPANT
 EQVCLEFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVPTPGTKPDLNASFVSANAE
 QFYQYQFTEIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSRVGNKSTIALTEL
 GDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLDEGEQVVNLPNEFIEQFKTGQ
 AKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVHFDSTNFGYSTGNARGQAVVNI
 SNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLFFISLGQVGGKLMVIGNPHYN
 SILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFEMENVLRLYLSDDKWTPDAKASMTVGTNLDTVYFKRHGQVTVNSAA
 FDFHDPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNPDYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVL
 IMENVMSNLKEESASGFVRLDAAGLSMALNKSVMNNDPQGYPNRVRQQRATGIWVYERYPAVDGALPYTIDSKTGE
 VKWKYQVENKPDCKPKEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKEKIFAAPFGLKECTNPAYHYEVNCL
 EYRPGTGVPVTTGGMYVPQYTQLSLADTAKAMVQAADLGTNIQRLYQHELYFRTNGRKGERLSSVDLERLYQNMVSW
 LWNETSRYEEGKNDLGFKTFTEFLNCYANDAYAGTKCSADLKKSLVDNNMIYGDSSKAGMMNPSYPLNYMEKPL
 LTRMLGRSNNWDLNIKVDVEKYPGAVSEEGQNVTTETISLSNPTKWFAGNMQSTGLWAPAQKEVTIKSNANVPVTV
 VALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGNSSTNESASFTFTGVVKAPFYKDGAWKN
 DLNSPAPLGELESADFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFYGRDETSKGHRMFTYKNTLGHKHRFTN
 NDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNAAEPLTVPGATEVANNVLALYMQDRYLGMKNR
 VADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDITKKWYPGFTLPEFYSEREGMKGNLFLQLMHRK
 ARGDDVGDKTFGGKKNYCAESNGNAADTLMCLCASWVAQTDLSEFFKKWNPGANAYQLPGATEMSFEGGVSQSAYSTLA
 SLKLPKPEQGPETINKVTEHKMSLE

SEQ ID NO: 106

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPTPEPEPEPVPTKTGYLTLGGSLRVTGDITCND
 ESSDGFFTFTPGDKVTCVAGNNTTATFDTQSEAAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSLVTSMNSCPANT
 EQVCLEFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVPTPGTKPDLNASFVSANAE
 QFYQYQFTEIILSEGRLVDSQGDGVVGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSRVGNKSTIALTEL
 GDEVRGANIDQLIHRYSKAGQNHTRVVPDEVRKVFAEYPNVINEIINLSLSNGATLDEGEQVVNLPNEFIEQFKTGQ
 AKEIDTAICAKTDGCNEARWFSLTTRNVNDGKIQGVINKLWGVDTNYKSVSKFHVHFDSTNFGYSTGNARGQAVVNI
 SNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETATFNLFFISLGQVGGKLMVIGNPHYN
 SILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFEMENVLRLYLSDDKWTPDAKASMTVGTNLDTVYFKRHGQVTVNSAA
 FDFHDPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNPDYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVL
 IMENVMSNLKEESASGFVRLDAAGLSMALNKSVMNNDPQGYPNRVRQRRSTPIWVYERYPAVDGKPPYTIIDTTKE
 VIWKYQENKPDCKPKEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKQRIIDAFPGLEVCKDSYHYEVNCL
 EYRPGTDVPTGGMYVPQYTQLDLSADTAKAMVQAADLGTNIQRLYQHELYFRTNGRQGERLNSVDLERLYQNMVSW
 LWNETYRYEEGKEDELGFKTFTEFLNCYNNAYVGTQCSAELKKSLIDNKMIYGEESKAGMMNPSYPLNYMEKPL
 TRMLGRSNNWDLNIKVDVEKYPGVVNTGETVTQNTQNSAPTQKWFAGNMQSTGLWAPAQQEVESKATVPVTVTV
 ALADDLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGLIYIKGDSKEVQSADFTFTGVVKAPFYKDGKWOHD
 LNSPAPLGELESADFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFYGRDSEDKGHRMFTYKNTLPGHKHRFTN
 DVQISIGDAHSGYPVMNSSFSPNSTTLPTTPLNDWLIWHEVGHNAAEPLTVPGATEVANNVLALYMQDRYLGMKNR
 VADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDITKKWYPDGTPLPEFYSEREGMKGNLFLQLMHRK
 ARGDEVNDKTFGGKKNYCAESNGNAADTLMCLCASWVAQTDLSEFFKKWNPGANAYQLPGASEMSFEGGVSQSAYNTLA
 SLKLPKPEQGPETINKVTEHKMSVE

SEQ ID NO: 107

DTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGYLTLGGSQRVGTGATCN
 GESSDGFFTFTPGNTVSCVVGSTTATFNTQSEAAARSLRAVDKVSFSLEDAQELANSNKKTNALSLVTSDDSCPAD
 EQCLTFSSVDRARFEKLYKQIDLATDNFSLVNEEVENNAATDKAPSTHTSTVVPVTTGKTPDLNASFVSANAE
 QFYQYQFTEIILSEQLVDSLNGVAGVDYITNSGRGVTDENGKFSFSWGETISFGIDTFELGSRVGNKSTIALTEL
 GDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLDEGDQNVNLPNEFIEQFKTGQ
 AKEIDTAICAKTNGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHVHFDSTNFGYSTGNARGQAVVNI
 SNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIQPENVTTRDTATFNLFFISLGQVGGKLMVIGNPHYN
 SILRCPNGYSWGGGVNSKGECTLSGSDSDMKHFMQNVRLYLSNDIWPNTKSIIMTVGTNLENVYFKKAGQVLGNSAP
 FAFHEDEFTGITVKQLTSYGDLPNEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVL

5

IMENVMSNLKEESASSFVRLDAAGLSMALNKSVMNNDPQGYPNRVRQQRREKGIWVYERYPFVDGKPPYTIIDETTKE
 VIWKYQDNKPDCKPKEVASWLEDVDGKQVKRYAFIDEAEHETNESLEAAKAKIKAFPGLEECKDPTYHYEVNCL
 EYRPGTNVPTGGMYVPRYTQNLNSADTAKAMVQAADLGTNIQRLYQHELYFRTNGRKGERLSSVDLERLYQNMVSW
 LWNEIEYSYDSSKEDELGFKTFTEFLNCYANDAYKTGLCSAELKQSLIDNKMIYEGESKAGMMNPSYPLNYMEKPL
 TRMLGRSNNWDLNIKVDVEKYPGAVSVGGEEVTTETISLSNPTKWFAGNMQSTGLWAPAQKEVTIKSNANVPVTVTV
 ALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGNSSTNESASFTFTGVVKAPFYKDGAWKND
 LNSPAPLGELESADFVYTTPKKNLNASNYTGGLDQFAKDLDTFASSMNDFYGRNDEDKGHRMFTYKNTLGHKHRFTN
 DVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNAAEPLTVPGATEVANNVLALYMQDRYLGMKNR
 VADDITVAPEYLDSENGQAWARGGAGDRLLMYAQLKEWAEKNFDITKKWYPDGKLPFAYSEREGMKGNLFLQLMHRKA
 RGDDVGNSSTFGGKKNYCAESNGNAADTLMCLCASWVAQTDLSEFFKKWNPGANAYQLPGAEMSFEFGVSSSAYSTLAS
 LNLKPKPEKGPETINKVTEHKMSAE

SEQ ID NO: 108

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPEPTPEPEPEPVPTKTGYLTILGGSQRVTGATCNGE
SSDGFTEFKPGEDVTCVAGNTTITATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSNSCPANTEQ
VCLTFSSVIESKRFDLSLYKQIDLAPEEFKLVNEEVENNAATDKAPSTHTSPVVPVTTTPTGTPDLNASFVSANAEQF
YQYQPTTEIILSEGRVLDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGEAISFGIDTFELGSVRGNKSTIALTELGD
EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLGEGEQVNNLPNEFIEQFNTGQAK
EIDTAICAKTDGDCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFYGSTGNARGQAVVNISN
AAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETASFNLFFISLGQVGDGKLMVIGNPHYNSI
LRCPNGYSWNGGVNKDQCTLNSDPDDMKNFENVLRYLSNDRWLPDAKSSMTVGTNLETVYFKKHGQVLGNSAPFA
FHKDFGTITVKPMTSYGNLNPDEVPLLIILNGFEYVTQWGSDDPYISPLRADTSKPKLTQQDVTDLIAYMNKGGSVLIM
ENVMSNLKEESASGCFVRLDDAAGLSMALNKSVMNNDPQGYPDRVRQRRSTPIWVYERYPAVDGKPPYTTIDTTKEVI
WKYQQENKPDCKPLEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKQRILDAFPGLEVCKDSYHYEVNCLLEY
RPGTDVPVTGGMYVPQYTTQLDLSADTAKAMLQAADLTGNIQRLYQHLYFRTNGRQGERLNSVDLERLYQNMSVWLW
NETKYRYEKGEDLGFKTFTEFLNCTYNNAYVGTQCSAELKKSLLDNKMIYGEESKAGMMNPSYPLNYMEKPLTR
LMLGRSWWDLNIKVDVEKYPGVVNTNGETVTQNNINLYSAPTQWAGNMQSTGLWAPAQQEVSIESKSTVPVTVTVAL
ADDLTGREKHEVSLNRPPRVTKTYDLKANDKVTFFKVPYGGLIYIKGDSKEVQSADFTFTGVVKAPFYKDGKWQHDLN
SPAPLGELESASFVYTTPKKNLNASNYTGGLEQGFANLDLTFASSMNDFYGRDSEDKHRMFTYKNLPGHKHRFANDV
QISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLALYMQDRYLGMNRVA
DDITVAPEYLEESNQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWYPDGTPLPEFYSEREGMKGWNLFQLMHRKAR
GDEVSNDFGKKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPANAYQLPGASEMSFEGGVSQSAYNTLASL
DLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 109

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPEPTPEPEPEPVPTKTGYLTILGGSQRITGATCNGE
SSDGFTEFKPGDVTCVAGNNTTITATFDTQSEAAARSLRAVEKVSFSLEDAQELAASDDKKSSNAVSLVTSNSCPANTE
QVCLTFSSVIESKRFDLSLYKQIDLAPEEFKLVNEEVENNAATDKAPSTHTSPVVPVTTTPTGTPDLNASFVSANAEQ
FYQYQPTTEIILSEGRVLDSMGNGVGVNYTSSGRGVTGENGFNFWSGETISFGIDTFELGSVRGNKSTIALTELGD
DEVGANIDQLIHRYSQAGKNDEREVPDVRKVFAEYPNVINEIINLSLSNGEALSEGQDTFERTNEFLEQFESGQA
KEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDDKDYKSVTKFHVFDSTNFYGSTGNARGQAVVNIS
NAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTREDTATFNLFFISLGQVGEGLMVIGNPHYNS
ILRCPNGYSWEGGVKNGQCTRNDSNDMKHFMQNVRLYLSNDKWTQDAKASMTVGTNLTDTVYFKRHGQVTGNSAEF
GFHPDFAGISVEHLSSYGDLDLPQKMPLLIILNGFEYVTQVGGDPYAVPLRADTSKPKLSQQDVTDLIAYLNKGGSVLI
MENVMSNLKEESASGCFVRLDDAAGLSMALNKSVMNNDPQGYPDRVRQRRATGIWVYERYPVVEGELPYTTIDSKTGK
TWKYQIDNKPDCKPLEVASWQEEVDGKQVTQFAFIDEADHKTPESLDAKKKILEKFKGLEECKDSTYHYEINCLE
YRPGTNVPATGGMYVPRYTQNLNSADTAKAMVQAADLTGNIQRLYQHLYFRTNGRKGERLSSVDLERLYQNMSVWL
WNEIEYSYDSSKEDELGFKTFTEFLNCTYANDAYTGGTQCSDELKKSVDNNMIYGEKSVNKGMMNPSYPLNYMEK
LTRLMLGRSWWDLNIKVDVEKYPGAVSEEGQEVTEISLSYSNPTKWFAGNMQSTGLWAPAQKEVTIKSNADVPVTVT
VALADDLTGREKHEVALNRPPKVTKTYELKANGEVKFTVPYGGLIYIKGNSKENNKSASFFTFTGVVKAPFYKNGAWK
NALNSPAPLGELESADFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEGKHRMFTYKNLTGKHHRF
TNDVQISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLALYMQDRYLGM
NRVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWYPDGTPLPEFYSEREGMKGWNLFQLMH
RKARGDEVSNDFGGRNYCAESNGNTADTLMLCASWVAQTDLSEFFKKWNPANAYQLPGATEMSFEGGVSQSAYNT
LASLDLPKPKQGPETINKVTEYSMPAE

SEQ ID NO: 110

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPEPTPEPEPEPVPTKTGYLTILGGSQRVTGATCNGE
SSDGFTEFKPGEDVTCVAGNTTITATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSNSCPANTEQ
VCLTFSSVIESKRFDLSLYKQIDLAPEEFKLVNEEVENNAATDKAPSTHTSPVVPVTTTPTGTPDLNASFVSANAEQF
YQYQPTTEIILSEGRVLDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSVRGNKSTIALTELGD
EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLGEGEQVNNLPNEFIEQFNTGQAK

EIDTAICAKTDGDCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFYGSTGNARGQAVVNISN
AAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETATFNLFFISLGQVGDGKLMVIGNPHYNSI
LRCPNGYSWNGGVNKDQCTLNSDPDDMKNFENVLRYLSNDRWLPDAKSNMTVGTNLTDTVYFKKHGQVTGNSAAG
FHPDFAGISVEHLSSYGDLDPQEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPKLTQQDVTDLIAYMNKGGSVLIM
ENVMSNLKEESASGCFVRLDDAAGLSMALNKSVMNNDPQGYPDRVRQRRSTPIWVYERYPAVDGKPPYTTIDTTKEVI
WKYQQENKPDCKPLEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKQRILDAFPGLECKDSYHYEVNCLLEY
RPGTGVPVVTGGMYVPQYTTQLSLNADTAKAMVQAADLTGNTQRLYQHLYFRTNGRKGERLSSVDLERLYQNMSVWLW
NKIEYRYENDKDELGFKTFTEFLNCTYANNAYDGGTQCSAELKQSLIDNKMIYEGESKAGMMNPSYPLNYMEKPLTR
LMLGRSWWDLNIKVDVEKYPGAVSAEGEETETINLYSNPTKWFAGNMQSTGLWAPAQQEVSIESKNAKVPVTVTVAJ
ADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFVPYGGLIYIKSDSKEEKSANFTFTGVVKAPFYKDGKWNOLK
SPAPLGELESASFVYTTPKKNLEASNYKGLKQFAEDLDTFASSMNDFYGRDGESGKHRMFTYKALTGKHHRFTNDV
QISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLALYMQDRYLGMNRVA
DDITVAPEYLEESNQAWARGGAGDRLLMYAQLKEWAEKNFDIKQWYPEGSLPAFYSEREGMKGWNLFQLMHRKARG
DDVGNDFGGRNYCAESNGNAADTLMLCASWVAQTDLSAFFKKWNPANAYQLPGATEMSFEGGVSQSAYNTLASL
LPPKPKQGPETINKVTEYSMPAE

SEQ ID NO: 111

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPEPTPEPEPEPVPTKTGYLTLLGGSQRVTGATCNGE
SSDGFTEFKPGEDVTCVAGNTTIATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSNSCPANTEQ
VCLTFSSVIESKRFDSLYKQIDLAPEEFKKLNVNEEVENNAATDKAPSTHTSPVVPVTTPGTKPDLNASFVSANAEQF
YQYQPTTEILSEGRVLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSGVRGNKSTIALTELGD
EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLGEQEVVNLNPNEFIEQFNTGQAK
EIDTAICAKTDGNCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFGSTGNARGQAVVNISN
AAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLFFISLGQVGEGLMVGIGNPHYNSI
LRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFEMENVLRYLSDDKWKPDAKASMTVGTNLDTVYFKRHGQVTGNSAAFD
FHPDFAGISVEHLSSYGDLDPQEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIM
ENVMSNLKEESASGFVRLDDAAGLSMALNKSVMNDPQGYPNRVRQQRATGIWVYERYPAVDGALPYTIDSKTGEVK
WKYQVENKPDGKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKEKIFAAFPGLKECTNPAYHYEVNCLLEY
RPGTGVPVTTGGMYVPQYTLQSLNADTAKAMVQAADLTGNIQRLYQHELYFRTNGRKGERLSSVDLERLYQNMSVWLW
NDTSYRYEKGNDLGLFKTFTEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLT
RLMLGRSOWDLNIKVDVEKYPGAVSEEGQNVTTETISLYSNPTKWFAGNMQSTGLWAPAQKEVTIKSNANVPVTVTVA
LADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGNSSTNESASFTFTGVVKAPFYKDGAWKNDL
NSPAPLGELESDAFVYTTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFYGRDSEDKHRMFTYKLNPLGKHKRFTND
VQISIGDAHSGYPVMNSSFSPNSTTLPTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLALYMQDRYLGMKNRV
ADDITVAPEYLEESNNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWPYDGTPLPEFYSEREGMKGNLFLQLMHRKA
RGDEVSNDFKGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPANAYQLPGASEMSFEGGVSQSAYNTLAS
LDLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 112

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPEPTPEPEPEPVPTKTGYLTLLGGSQRVTGATCNGE
SSDGFTEFKPGEDVTCVAGNTTIATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSNSCPANTEQ
VCLTFSSVIESKRFDSLYKQIDLAPEEFKKLNVNEEVENNAATDKAPSTHTSPVVPVTTPGTKPDLNASFVSANAEQF
YQYQPTTEILSEGRVLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSGVRGNKSTIALTELGD
EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLGEQEVVNLNPNEFIEQFNTGQAK
EIDTAICAKTDGNCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFGSTGNARGQAVVNISN
AAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVRPENVTRETATFNLFFISLGQVGDGKLMVGIGNPHYNSI
LRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFEMENVLRYLSNDRWLPDAKSNMTVGTNLDTVYFKKHGQVTGNSAAFG
FHPDFAGISVEHLSSYGDLDPQEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPKLTQQDVTDLIAYMKNKGGSVLIM
ENVMSNLKEESASGFVRLDDAAGLSMALNKSVMNDPQGYPNRVRQRRSTPIWVYERYPAVDGKPPYTIIDDTTKEVI
WKYQENKPDGKPKLEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKRILDAFPLGEECKDSYHYEVNCLLEY
RPGTGVPVTTGGMYVPQYTLQSLNADTAKAMVQAADLTGNIQRLYQHELYFRTNGRKGERLSSVDLERLYQNMSVWLW
NKIEYRYENDKDELGLFKTFTEFLNCYANNAYDGGTQCSAELKQSLIDNKMIYGEKSKAGMMNPSYPLNYMEKPLTR
LMLGRSOWDLNIKVDVEKYPGAVSAEGEEVTETINLYSNPTKWFAGNMQSTGLWAPAQQEVSIKSNKVPVTVTVAL
ADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKSDSKEEKSANFTFTGVVKAPFYKDGKWKNDLK
SPAPLGELESASFVYTTTPKKNLEASNYKGLKQFAEDLDTFASSMNDFYGRDGESGKHRMFTYKALTGHKHRTNDV
QISIGDAHSGYPVMNSSFSPNSTTLPTPLNDWLIWHEVGHNAEETPLTVPGATEVANNVLALYMQDRYLGMKNRVA
DDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKQWYPEGSLPAFYSEREGMKGNLFLQLMHRKARG
DDVGNDKFGNRNYCAESNGNAADTLMLCASWVAQTDLSAFFKKWNPANAYQLPGATEMSFEGGVSQSAYNTLASLD
LPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 113

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGYLTLLGGSQRITGATCNGE
SSDGFTEFTPGDKVTCVAGNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAASDDKKSSNAVSLVTSNSCPADTE
QVCLTFSSVIESKRFDSLYKQIDLAPEEFKKLNVNEEVENNAATDKAPSTHTSPVVPVTTPGTKPDLNASFVSANAEQ
FYQYQPTTEILSEGRVLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSGVRGNKSTIALTELGD
DEVIRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQA
KEIDTAICAKTDGNCNEARWFSLTTRNVNDGQIQGVINKLWGVOKDYKSVTKFHVFDSTNFGSTGNARGQAVVNIS
NAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVQPENVTTRDTATFNLFFISLGQVGEGLMVGIGNPHYNS
ILRCPNGYSWNGGVNKDGQCTLSGSDDMKHFQNVLRYSDDKWTDAKASMTVGTNLDTVYFKRHGQVTGNSAEF
GFHPDFAGISVEHLSSYGDLDPQEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLI
MENVMSNLKEESASGFVRLDDAAGLSMALNKSVMNDPQGYPNRVRQQRREKGIWVYERYPAVDSAPPPYTIIDPDTGK
VTWKYQEEGKPDGKPKLEVASWQEDVDGKQVTRYAFIDEAEHSTESLEAAKAKIFEKFPGLQECKDSTYHYEINCL
ERRPGTDVPVTTGGMYVPYTLQNLADTAKAMVQAADLTGNIQRLYQHELYFRTNGRKGERLSSVDLERLYQNMSVW
LWNKIEYRYENDKDELGLFKTFTEFLNCYANNAYSEGTCQCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEK
LTRLMLGRSOWDLNIKVDVEKYPGAVSAEGEKVTETISLYSNPTKWFAGNMQSTGLWAPAQQEVTIESTASVPVTVT
VALADDLTGREKHEVALNRPPKVTCTYDLKANDKVTFKVPYGGLIYIKGNSPKNESAEFTFTGVVKAPFYKDGKWK
ALNSPAPLGELESDFSFTYAPKNNLNASNYSTYDGAFAKELDTFASSMNDFYGRDGESGNHRMFTYKALTGHK
RFANDVQISIGDAHSGYPVMNSSFSTNSTTLPTPLNDWLIWHEVGHNAEETPLNVPGATEVANNVLALYMQDRYL
KMNRVADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDRKGMKGNLFLQLM
HRKARGDDVSNDFGGRNYCAESNGNAADTLMLCASWVAQADLSEFFKKWNPANAYQLPGASEMSFEGGVSQSAYN
TLAAMHLSKPEKGPETINKVTEYSMPAE

SEQ ID NO: 114

KTGYLTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAAARSLRAVDKVSFSLEDAQELANSEN
KKTNAISLVTSSDSCPADAEQLCLTFSSVVDRAREFELYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPV
TTEGTPKPDNLASFVSANAEQFYQYQPTTEIILSEGQLVDSLNGVAGVDYITNSGRGVTDENGKFSFSWGETISFGID
TFELGSVRGNKSTIALTELGEVVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLDE
GDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHVFD
STNFGYSTGNARGQAVVNISNSAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIQPENVTTRDTATFNLPI
SLGQVGEGLMVGINPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYSNDIWQPNTKSIMTVGT
NLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDNLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKL
TQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLDDAAGLSMALNKSVVNNDPQGYPDVRVQRATGIWVYE
RYPADGAQPPYITDPNTGEVTKYQQDNKPDDKPKLEVASWQEEVEGKQVTRYAFIDEAEYTTESLEAAKAKIFE
KFPGQLQECKDSTYHYEINCLERRPGTDVPVTGGMYPVRYTQLNLADTAKAMVQAADLTGNIQRLYQHELYFRTKGS
KGERLNSVDLERLYQNMSVWLWNDTKYRYEKGDELGFKFTTEFLNICYANDAYAGGKCSADLKKSLVDNNMIYGD
GSSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGVSVAKGESVTENISLYSNPTKWFAGNMQSTGLW
APAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPRVTKYTLANGEVTFKVPYGGLIYIKGDSKDDVSA
NFTFTGVVKAPFYKDGWKNLDSAPPLGELESASFVYTPKKNLEASNFTGGVAEFAKDLDTFASSMNDIFYGRNDE
DGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTPLNDWLIWHEVGHNAETPLNVPGAT
EVANNVLALYMQDRYLGMKNRVADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPK
FYSDRKGMKGWNLFLQMLHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKWNPASAYQLP
GATEMSFQGGVSSSAYSTLASLKLKPKPEKGPETINKVTEHKMSAE

SEQ ID NO: 115

KTGYLTLGGSQRITGATCNGESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAASD
DKKSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVP
VTTPTGTPKPDNLASFVSANAEQFYQYQPTTEIILSEGRVDSMGNGVGVVNYTSSGRGVTGENGKFNFSWGETISFGI
DTFELGSVRGNKSTIALTELGEVVRGANIDQLIHRYSQAGKNDEREVDPDVVRKVFAEYPNVINEIINLSLSNGEALS
EGDQTFERTNEFLEQFESGQAKEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDDKYKSVTKFHVH
DSTNFGYSTGNARGQAVVNISNAAPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLPI
ISLGQVGEGLMVGINPHYNSILRCPNGYSWEGGVVDKNGQCTRNSDSNDMKHFMQNVLRYSDDKWTDPKASMTVG
TNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDLPQEMPLLIILNGFEYVTQVGNDPYAIPLRADTSKPK
LTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVVNNDPQGYPNRVQRATGIWVY
ERYPAVDGALPYTIDSKTGEVVKWYQVENKPDDKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKAKIFE
KFPGKKECKDPTYHYEVNCLERYRPGTGVPVTGGMYPVRYTQLSLNADTAKAMVQAADLTGNIQRLYQHELYFRTNGR
KGERLSSVDLERLYQNMSVWLWNLKIEYRYENDKDDELGFKFTTEFLNICYANDAYTGGTQCSDELKKSLVDNNMIYGE
KSVNAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGAVSAEGEKVTETISLYSNPTKWFAGNMQSTGL
WAPAQKEVTIASSASVPVTVTVALADDLTGREKHEVALNRPPKVTKYELKANGEVKFTVPYGGLIYIKGNPQNES
AEFTFTGVVKAPFYKDGAWKNALNSPAPLGELESDAFVYTPKKNLEASNFTGGVAEFAKDLDTFASSMNDIFYGRND
EDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTPLNDWLIWHEVGHNAETPLTVPGA
TEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKQWYPEGDLF

KFYSDREGMKGNLFLQMLHRKARGDEVGKTKFGERNYCAESNGNAADKMLCASWVAQTDLSEFFKKWNPANAYQL
PGASEMNFEGGVQSAYETLAALNLPKPQOGPETINQVTEHKMSAE

SEQ ID NO: 116

KTGYLTLGGSQRITGATCNGESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAASD
DKKSNAVSLVTSSNSCPADTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVP
VTTPTGTPKPDNLASFVSANAEQFYQYQPTTEIILSEGRVDSMGNGVGVVNYTSSGRGVTGENGKFNFSWGETISFGI
DTFELGSVRGNKSTIALTELGEVVRGANIDQLIHRYSQAGKNDEREVDPDVVRKVFAEYPNVINEIINLSLSNGEALS
EGDQTFERTNEFLEQFESGQAKEIDTAICDSLGGCNSQRWFSLTARNVNDGQIQGVINKLWGVDTNYKSVSKFHVH
DSTNFGYSTGNARGQAVVNISNAAPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNLPI
ISLGQVGEGLMVGINPHYNSILRCPNGYSWEGGVVDKNGQCTRNSDSNDMKHFMQNVLRYSDDKWTDPKASMTVG
TNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDLPQEMPLLIILNGFEYVTQVGNDPYAIPLRADTSKPK
LTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVVNNDPQGYPNRVQRATGIWVY
ERYPAVDGALPYTIDSKTGEVVKWYQVENKPDDKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKAKIFE
KFPGKKECKDPTYHYEVNCLERYRPGTGVPVTGGMYPVRYTQLSLNADTAKAMVQAADLTGNIQRLYQHELYFRTNGR
KGERLSSVDLERLYQNMSVWLWNLKIEYRYENDKDDELGFKFTTEFLNICYANDAYTGGTQCSDELKKSLVDNNMIYGE
KSVNAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGAVSAEGEKVTETISLYSNPTKWFAGNMQSTGL
WAPAQKEVTIESTASVAVTVTVALADDLTGREKHEVALNRPPKVTKYELKANGEVKFTVPYGGLIYIKGNPQNES
AEFTFTGVVKAPFYKDGAWKNALNSPAPLGELESDAFVYTPKKNLEASNFKGQEQFAEELDTFASSMNDIFYGRND
EDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTPLNDWLIWHEAGHNAETPLTVPGA
TEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKNFDITKWPYEGNLP
KFYSEREGMKGNLFLQMLHRKARGDEVGKTKFGERNYCAESNGNAADTLMLCASWVAQTDLSAFFKKWNPANAYQL
PGASEMNFEGGVQSAYETLAALNLPKPQOGPETINKVTEYSMPAE

SEQ ID NO: 117

KTGYLTGGSSQRTGATCNGESSDGFTEFKPGEDVTCVAGNTTITATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSD
 KKSNAVSLVTSNNSCPANTEQVCLTFSSVIESKRFDLSYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPV
 TTPGTPKPDNLASFVSANAEQFYQYQPTTEILSEGRVLDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGGETISFGID
 TFEGLSVRGNKSTIALTELGEDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLGE
 GEQVVNLNPNFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFD
 STNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTREDTATFNLPI
 SLGQVGEGLMVGIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMEVNLRYLSDDKWKPDAKASMTVGT
 NLDTVYFKRHGQVTGNSAADFDPDFAGISVEHLSSYGDLDPPQEMPLLILNGFEYVTVQVNDPYAIPLRADTSKPKL
 TQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVVNNDPQGYPNRVQRATGIWVYE
 RYPAVDGALPYTIDSKTGEVKWKYQVENKPDDKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKEKIFAA
 FPGLKECTNPAYHYEVNCLYRPGTGPVPTGGMYVPQYTQLSLNADTAKAMVQAADLTGNIQRLYQHELYFRTNGRK
 GERLSSVDLERLYQNMSVWLWNTSYRYEEGKNDELGFKTFTEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYGDG
 SSKAGMNPSPYPLNYMEKPLTRLMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLYSNPTKWFAGNMQSTGLWA
 PAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNESAS
 FTFTGVVKAPFYKDGAWKNDLNSPAPLGELESDAFYVYTPKKNLNASNYTGGLAQFANDLDTFASSMNDFFGRDSE
 GKHRMFTYKNLPGHKHRTNDVQISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGHNAETPLTVPGATE
 VANNVLALYMQDRYLKGMNRVADDITVAPEYLEESNQAWAREGGADRLLMYAQLKEWAEKNFDIKKWPYDGTPLPE
 FYSEREGMKGNLFLQMLHRKARGDEVSNDFGKKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPANAYQLP
 GASEMSFEGGVSQSAYNTLASLDLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 118

KTGYLTGGSLRVTDITCNDESSDGFTEFTPGDKVTCVAGNTTITATFDTQSEAAARSLRAVEKVSFSLEDAQELAGS
 DNKKSNAVSLVTSMNNSCPANTEQVCLTFSSVIESKRFDLSYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPV
 PVTTPGTPKPDNLASFVSANAEQFYQYQPTSEILSEGRVLDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGGETISFG
 IDTFELGSVRGNKSTIALTELGEDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATL
 DEGEQVVNLNPNFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFD
 HDSTNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTREDTATFNLPI
 FISLGQVGEGLMVGIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMEVNLRYLSDDKWTDPDAKASMTV
 GTNLDTVYFKRHGQVTGNSAADFDPDFAGISVEHLSSYGDLDPPQEMPLLILNGFEYVTVQVNDPYAIPLRADTSKP
 KLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVVNNDPQGYPNRVQRATGIWV
 YERYPAVDGALPYTIDSKTGEVKWKYQVENKPDDKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKEKIF
 AAFPGLKECTNPAYHYEVNCLYRPGTGPVPTGGMYVPQYTQLSLNADTAKAMVQAADLTGNIQRLYQHELYFRTNG
 RKGERLSCVDLERLYQNMSVWLWNTSYRYEEGKNDELGFKTFTEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYGDG
 DGSSKAGMNPSPYPLNYMEKPLTRLMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLYSNPTKWFAGNMQSTGL
 WAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNES
 ASFTFTGVVKAPFYKDGAWKNDLNSPAPLGELESDAFYVYTPKKNLNASNYTGGLAQFANDLDTFASSMNDFFGRDS
 EDGKHRMFTYKNLPGHKHRTNDVQISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGHNAETPLTVPGA

TEVANNVLALYMQDRYLKGMNRVADDITVAPEYLEESNGQAWARGGADRLLMYAQLKEWAEKNFDIKKWPYDGTPL
 PEFYSEREGMKGNLFLQMLHRKARGDEVSNDFGKKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPANAYQ
 LPGASEMSFEGGVSQSAYNTLASLKLKPEQGPETINKVTEHKMSVE

SEQ ID NO: 119

KTGYLTGGSLRVTDITCNDESSDGFTEFTPGDKVTCVAGNTTITATFDTQSEAAARSLRAVEKVSFSLEDAQELAGS
 DNKKSNAVSLVTSMNNSCPANTEQVCLTFSSVIESKRFDLSYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPV
 PVTTPGTPKPDNLASFVSANAEQFYQYQPTSEILSEGRVLDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGGETISFG
 IDTFELGSVRGNKSTIALTELGEDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATL
 GEQVVNLNPNFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFD
 HDSTNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTREDTATFNLPI
 FISLGQVGDGKLMVGIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMEVNLRYLSNDRWLPDAKSSMTV
 GTNLDTVYFKKHGQVLGNSAPFAFHKDFGTITVKPMTSYGNLNPDEVPLLIILNGFEYVTVQWSDPYIPLRADTSKP
 KLTQQDVTDLIAYMKNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVVNNDPQGYPNRVQRATGIWV
 YERYPAVDGKPPYTIDDTTKEVIWKYQVENKPDDKPKLEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKQRIIL
 DAFPGLVECKDSYHYEVNCLYRPGSGVPVPTGGMYVPQYTQLDLGADTAKAMLQAADLTGNIQRLYQHELYFRTNG
 RQGERLNSVDLERLYQNMSVWLWNETKYRYEEGKEDELGFKTFTEFLNCYTNNAYVGTQCSAELKKSLIDNKMIGE
 ESSKAGMNPSPYPLNYMEKPLTRLMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLYSNPTKWFAGNMQSTGLW
 APAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNES
 SFTFTGVVKAPFYKDGAWKNDLNSPAPLGELESDAFYVYTPKKNLNASNYTGGLAQFANDLDTFASSMNDFFGRDSE
 SGKHRMFTYKNLTGHKHRFANDVQISIGDAHSGYPVMNSSFSFNSTTLPTTPLNDWLIWHEVGHNAETPLTVPGAT
 EVANNVLALYMQDRYLKGMNRVADDITVAPEYLEESNGQAWARGGADRLLMYAQLKEWAEKNFDIKKWPYEGELPK
 FFSREGMKGNLFLQMLHRKARGDDVGNTFTGGKNYCAESNGNAADSLMLCASWVAQTDLSAFFKKWNPANAYQLP
 GATEMSFEGGVSQSAYSTLASLKLKPEQGPETINKVTEHKMSLE

SEQ ID NO: 120

KTGYLTLGGSLRVTDITCNDESSDGFFTFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVSFSLEDAQELAGS
DNKKSNAISLVTSMNSCPANTEQVCLEFSSVIESKRFDSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVV
PVTTPTGKPDNLASFVSANAEQFYQYQPSIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFG
IDTFELGSVRGNKSTIALTELGEVIRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEYPNVINEIINLSLSNGATL
DEGEQVNLNPNFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHV
HDSTNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTTRDTATFNL
FISLGQVVGEGKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFEMENVRLYLSDDKWTPDAKASMTV
GTNLDTVYFKRHGQVTGNSAADFDPFAGISVEHLSSYGDLDPQEMPLLLILNGFEYVTQVGNPDYAIPLRADTSKP
KLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVMNNDPQGYPNRVQRATGIWV
YERYPAVDGALPYTIDSKTGEVKWKYQVENKPDCKPKLEVASWLEDVVGKQETRYAFIDEADHKTEDSLKAAKEKIF
AAFPGLKECTNPAYHYEVNCLERYPGTGPVPTGGMYVPQYTQLSLNADTAKAMVQAADLTGNIQRLYQHELYFRITNG
RKGERLSSVDLERLYQNMSVWLWNTSYRYEEGKDELGFKTFTEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYG
DGSSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNLIKVDVEKYPGAVSEEGQNVTTETISLYSNPTKWFAGNMQSTGL
WAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNES
ASFTFTGVVKAPFYKDGAWKNDLNSPAPLGELESADFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFYGRDE
TSGKHRMFTYKNLTGKHHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTPLNDWLIWHEVGHNAETPLTVPGA
TEVANNVLAALYMQDRYLGMNVRADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWPYEGELP
KFFSDREGMKGNLFLQMLHRKARGDDVGDKTFGGKNYCAESNGNAADTLMLCASWVAQTDLSAFFKKWNPANAYQL
PGATEMSFEGGVSQSAYSTLASLKLKPKPEQGPETINKVTEHKMSLE

SEQ ID NO: 121

KTGYLTLGGSLRVTDITCNDESSDGFFTFTPGDKVTCVAGNNTTIATFDTQSEAARSLRAVEKVSFSLEDAQELAGS
DNKKSNAISLVTSMNSCPANTEQVCLEFSSVIESKRFDSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVV
PATTPGKPDNLASFVSANAEQFYQYQPTIILSEGRLVDSQGDGVVGVNYTNSGRGVTGENGEFSFSWGETISFG
IDTFELGSVRGNKSTIALTELGEVIRGANIDQLIHRYSKAGQNHTRVVPDEVKVFAYEYPNVINEIINLSLSNGATL
GEGEQVNLNPNFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHV
HDSTNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETATFNL
FISLGQVVGDKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFEMENVRLYLSDDKWTPDAKASMTV
GTNLDTVYFKRHGQVTGNSAADFDPFAGISVEHLSSYGDLDPQEMPLLLILNGFEYVTQVGNPDYAIPLRADTSKP
KLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVMNNDPQGYPNRVQRORSTPIWV
YERYPAVDGKPPYTIIDTTKEVIWKYQQENKPDCKPKLEVASWEEVEGKQVTQFAFIDEADHKTPESLAAAKQRIIL
DAFPGLEVCCKDSYHYEVNCLERYPGTDVPVPTGGMYVPQYTQLDLSADTAKAMLQAADLTGNIQRLYQHELYFRITNG
RQGERLNSVDLERLYQNMSVWLWNETKYRYEEGKDELGFKTFTEFLNCYTNNAYVGTQCSAELKKSLIDNKMIYGE
ESSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNLIKVDVEKYPGVVNTNGETVTQININLYSAPTKWFAGNMQSTGLW
APAQQEVSIESKATVPVTVTVALADDLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGGLIYIKGDSKEVQSA
DFTFTGVVKAPFYKDGKQWDLNSPAPLGELESASFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDFYGRDSE

DGKHRMFTYKNLPGHKHREFANDVQISIGDAHSGYPVMNSSFSPNSTTLPTPLNDWLIWHEVGHNAETPLTVPGAT
EVANNVLAALYMQDRYLGMNVRADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWPYDGTPLP
EFYSEREGMKGNLFLQMLHRKARGDEVSNDFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPANAYQL
PGASEMSFEGGVSQSAYNTLASLKLKPKPEQGPETINKVTEHKMSVE

SEQ ID NO: 122

KTGYLTLGGSQRVGTGATCNDESSDGFFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVDKVSFSLEDAQELANSEN
KKTNATISLVTSDDSCPADAEQLCLTFSSVVDRAFEKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPV
TTEGTPKPDNLASFVSANAEQFYQYQPTIILSEGQLVDSLGNVGAGVDYNTNSGRGVTDENGKFSFSWGETISFGID
TFELGSVRGNKSTIALTELGEVIRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEYPNVINEIINLSLSNGATLDE
GDQNVVLPNEFIEQFKTGQAKEIDTAICAKTNGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHVFDH
STNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIQVPENVTTRDTATFNL
PFIISLGQVVGDKLMVIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVRLYLSNDIWPNTKSI
MTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDNLNPEIPLLLILNGFEYVTQWSGDPYAVPLRADTSKP
KLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLDDAAGLSMALNKSVMNNDPQGYPNRVQRQREKGIWV
YERYPAVDGKPPYTIIDTTKEVIWKYQQDNKPDCKPKLEVASWLEDVVGKQVKRYAFIDEAEHETNESLEAAKAKI
KAFPGLEECKDPTYHYEVNCLERYPGTNVPTGGMYVPRYTQLNLSADTAKAMVQAADLTGNIQRLYQHELYFRITNG
RKGERLSSVDLERLYQNMSVWLWNEIEYSYDSSKEDELGFKTFTEFLNCYANDAYTKGTLCSAELKQSLIDNKMIY
GEGSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNLIKVDVEKYPGAVSVGGEEVTETISLYSNPTKWFAGNMQSTGL
WAPAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNES
ASFTFTGVVKAPFYKDGAWKNDLNSPAPLGELESASFVYTTPKKNLNASNYTGGLDQFAKDLDTFASSMNDFYGRNEDG
KHRMFTYKNLTGKHHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTPLNDWLIWHEVGHNAETPLNVPGATEV
ANNVLAALYMQDRYLGMNVRADDITVAPEYLDENSGQAWARGGAGDRLLMYAQLKEWAEKNFDITKWPYDGLPAFY
SEREGMKGNLFLQMLHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPANAYQLPGA
AEMSFEFGVSSSAYSTLASLNLKPKPEKGPETINKVTEHKMSAE

SEQ ID NO: 123

KTGYLTGGSQRTVGATCNGESSDGFTFKPGEDVTCVAGNTT IATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSDD
 KKSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPV
 TTPGTPKPDNLASFVSANAEQFYQYQPTTEILSEGRVLDSDQGYGVAGVNYTNSGRGVTGENGEFSFSWGEAISFGID
 TFEGLSVRGNKSTIALTELGEVIRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLGE
 GEQVVNLPNEFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVHFD
 STNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETASFNLPEFI
 SLGQVGDKLWVIGNPHYNSILRCPNGYSWNGGVNKDQCTLNSDPDDMKNFENVLRYSNDRWLPDAKSSMTVGT
 NLETVYFKKHGQVLGNSAPFAFHKDFGTGIVTKPMTSYGNLNPDEVPLLI LINGFEYVTQWGS DPYSIPLRADTSKPKL
 TQQDVTDLIAYMNGKGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKS VVNNDPQGYPDRVRQRRSTPIWVYE
 RYPAVDGKPPYTIIDTTKEVIWKYQQENKPDCKPKLEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKQRILDA
 FPGLEVCKDSYHYEVNCLERYRPGTDVPTGGMYVPQYTQLDLSADTAKAMLQAADLGTNIQRLYQHLYFRTNGRQ
 GERLNSVDLERLYQNMSVWLWNETKYRYEEGKEDELGFKFTFEFLNCYTNNAYVGTQCSAELKKS LIDNKMIYGEES
 SKAGMNPSPYPLNYMEKPLTRLMLGRS WWDNLNIKVDVEKYPGAVVNTNGETVTQININLYSAPT KWFAGNMQSTGLWAP
 AQQEVSIKSTVPVTVTVALADDLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGGLIYIKGDSKEVQSADF
 TFTGVVKAPFYKDGKWKQHDNLSPAPLGELESASFVYTTPKKNLNASNYTGGLEQFANOLDTFASSMNDFYGRDSEEG
 KHRMFTYKNLPGHKHFRFANDVQISIGDAHSGYPVMNSSFSNSTTLPTTPLNDWLIWHEVGHNAAEPTLTVPGATEV
 ANNVLALYMQDRYLKGKMN RVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWYPDGTPLPEF
 YSEREGMKGNL FQLMHRKARGDEVSNDFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPANAYQLPG
 ASEMSFEGGVSQSAYNTLASLDLPKPKQGPETINQVTEHKMSAE

SEQ ID NO: 124

KTGYLTGGSQRTVGATCNGESSDGFTFTPGDKVTCVAGNNTT IATFDTQSEAAARSLRAVEKVSFSLEDAQELAASD
 DKKSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVP
 VTTPTGTPKPDNLASFVSANAEQFYQYQPTTEILSEGRVLDSDMGNGVGVVNYTSSGRGVTGENGKFNFSWGETISFGI
 DTFELGSVRGNKSTIALTELGEVIRGANIDQLIHRYSQAGKNDEREVDPVVRKVFAAYPNVINEIINLSLSNGEALS
 EGDQTFERTNEFLEQFESGQAKEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDKDYKSVTKFHVHFD
 DSTNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVT RDTATFNLPEFI
 ISLGGQVGDKLWVIGNPHYNSILRCPNGYSWEGGVNDKNGQCTRNSDSDNDMKHFQNVLRYSNDKWTDAKASMTVG
 TNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDLPQKMPLLI LINGFEYVTQVGGDPYAVPLRADTSKPK
 LSQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKS VVNNDPQGYPDRVRQRRATGIWVY
 ERYPVVEGELPYYTIDSKTGKVTWKYQIDONKPDCKPKLEVASWQEEVDGKQVTQFAFIDEADHKTTESLDAAKKKILE
 KFKGLEECKDSTYHYEINCLEYPGTNPATGGMYVPRYTQLNLSADTAKAMVQAADLGTNIQRLYQHLYFRTNGR
 KGERLSSVDLERLYQNMSVWLWNEIEYSYDSSKEDELGFKFTFEFLNCYANDAYTGGTQCSDELKKS LVDNMMIYGE
 KSVNKAGMNPSPYPLNYMEKPLTRLMLGRS WWDNLNIKVDVEKYPGAVSEEGQEVTSISLNSNPTKWFAGNMQSTGL
 WAPAQKEVTIKSNADVPVTVTVALADDLTGREKHEVALNRPPKVTCTYELKANGEVKFTVPYGGGLIYIKGNSKENNK

SASFTFTGVVKAPFYKNGAWKNALNSPAPLGELESDAFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRN
 DEDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSNSTTLPTTPLNDWLIWHEVGHNAAEPTLTVPG
 ATEVANNVLALYMQDRYLKGKMN RVADDITVAPEYLEESNNQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWYPDGTPL
 LPEFYSEREGMKGNL FQLMHRKARGDEVSNDFGGKNYCAESNGNTADTLMLCASWVAQTDLSEFFKKWNPANAY
 QLPGATEMSFEGGVSQSAYNTLASLDLPKPKQGPETINKVTEYSMPAE

SEQ ID NO: 125

KTGYLTGGSQRTVGATCNGESSDGFTFKPGEDVTCVAGNTT IATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSDD
 KKSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPV
 TTPGTPKPDNLASFVSANAEQFYQYQPTTEILSEGRVLDSDQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGID
 TFEGLSVRGNKSTIALTELGEVIRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLGE
 GEQVVNLPNEFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVHFD
 STNFGYSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETATFNLPEFI
 SLGQVGDKLWVIGNPHYNSILRCPNGYSWNGGVNKDQCTLNSDPDDMKNFENVLRYSNDRWLPDAKSNMTVGT
 NLDTVYFKKHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDLPQEMPLLI LINGFEYVTQVGNPDYAIPLRADTSKPKL
 TQQDVTDLIAYMNGKGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKS VVNNDPQGYPDRVRQRRSTPIWVYE
 RYPAVDGKPPYTIIDTTKEVIWKYQQENKPDCKPKLEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKKRILDA
 FPGLEECKDSYHYEVNCLERYRPGTGVPVPTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQHLYFRTNGRK
 GERLSSVDLERLYQNMSVWLWNKIEYRYENDKDDELGFKFTFEFLNCYANNAYDGGTQCSAELKQSLIDNKMIYEGG
 SKAGMNPSPYPLNYMEKPLTRLMLGRS WWDNLNIKVDVEKYPGAVSAEGEEVTETINLYSNPTKWFAGNMQSTGLWAP
 AQQEVSIKSNAKVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKSDSKEEKSANF
 TFTGVVKAPFYKDGKWKNDLKSAPLGELESASFVYTTPKKNLEASNKGGGLKQFAEDLDTFASSMNDFYGRDGESG
 KHRMFTYALTGHKHRFTNDVQISIGDAHSGYPVMNSSFSNSTTLPTTPLNDWLIWHEVGHNAAEPTLTVPGATEV
 ANNVLALYMQDRYLKGKMN RVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWYPEGSLPAFY
 SEREGMKGNL FQLMHRKARGDDVGNDFGGRNYCAESNGNAADTLMLCASWVAQTDLSAFFKKWNPANAYQLPGA
 TEMSFEGGVSQSAYNTLASLDLPKPKQGPETINKVTEYSMPAE

SEQ ID NO: 126

KTGYLTLGGSQRVTGATCNGESSDGFTEFKPGEDVTCVAGNTTITATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSD
 KKSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLNVNEEVENNAATDKAPSTHTSPVVPV
 TTPGTPKPDNLNASFVSANAEQFYQYQPTTEILSEGRVLDVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGID
 TFEELGVSVRGNKSTIALTELGEVIRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLGE
 GEQVVNLNPNFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVHFD
 STNFYGSTGNARGQAVVNISNAAFPIILMARNDKKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTRODTATFNLFFI
 SLGQVGEGLKMLVIGNPHYSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFENVLRYSDDKWKPKDAKASMTVGT
 NLDTVYFKRHGQVTGNSAAFDFHPDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPKL
 TQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDAAAGLSMALNKSVVNNDPQGYPNRVROQRATGIWVYE
 RYPAVDGALPYTIDSKTGEVKWKYQVENKPDCKPKEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKEKIFAA
 FPGLKECTNPAYHYEVNCLERYRPGTGVPTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQHLYFRTNGRK
 GERLSSVDLERLYQNMSVWLWNTSYRYEEGKNDELGFKFTFEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYGDG
 SSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGAVSEEGQNVTTETISLYSNPTKWFAGNMQSTGLWA
 PAQKEVTIKSNANVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKGNSSTNESAS
 FTFTGVVKAPFYKDGAWKNDLNSPAPLGELESDFVYTPKKNLNASNYTGGLEQFANDLDTFASSMNDIFYGRDSED
 GKHRMFTYKKNLPGHKHRTNDVQISIGDAHSGYPVMNSSFSPNSTTLPTTPLNDWLIWHEVGHNAETPLTVPGATE
 VANNVLAALYMQDRYLGMKNRVADDITVAPEYLEESNQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWYPDGTPLPE
 FYSEREGMKGNLFLQLMHRKARGDEVNDKFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPANAYQLP
 GASEMSFEGGVSQSAYNTLASLDLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 127

KTGYLTLGGSQRVTGATCNGESSDGFTEFKPGEDVTCVAGNTTITATFNTQSEAAARSLRAVEKVSFSLEDAQELAGSD
 KKSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLNVNEEVENNAATDKAPSTHTSPVVPV
 TTPGTPKPDNLNASFVSANAEQFYQYQPTTEILSEGRVLDVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGID
 TFEELGVSVRGNKSTIALTELGEVIRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLGE
 GEQVVNLNPNFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVHFD
 STNFYGSTGNARGQAVVNISNAAFPIILMARNDKKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVTRETATFNLFFI
 SLGQVGDGKMLVIGNPHYSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFENVLRYSNDRWLPDAKSNMTVGT
 NLDTVYFKKHGQVTGNSAAFGEHDPDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPKL
 TQQDVTDLIAYMKNKGGSVLIMENVMSNLKEESASGFVRLDAAAGLSMALNKSVVNNDPQGYPNRVROQRSTPIWVYE
 RYPAVDGKPPYITDDTTKEVIWKYQENKPDCKPKEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAKKRILDA
 FPGLKECKDSYHYEVNCLERYRPGTGVPTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQHLYFRTNGRK
 GERLSSVDLERLYQNMSVWLWNKIEYRYENDKDELGFKFTFEFLNCYANNAYDGGTQCSAELKQSLIDNKMIYGE
 SKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGAVSAEGEEVTETINLYSNPTKWFAGNMQSTGLWAP

AQQEVSISNAKVPVTVTVALADDLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGGLIYIKSDSKEEKSANF
 TFTGVVKAPFYKDGKWKNDLKSAPLGELESASFVYTPKKNLEASNYKGLKQFAEDLDTFASSMNDIFYGRDGESG
 KHRMFTYEALTGHKHRTNDVQISIGDAHSGYPVMNSSFSPNSTTLPTTPLNDWLIWHEVGHNAETPLTVPGATEV
 ANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNQAWARGGAGDRLLMYAQLKEWAEKNFDIKQWYPEGSLPAFY
 SEREGMKGNLFLQLMHRKARGDDVGNDFGNRNKYCAESNGNAADTLMLCASWVAQTDLSAFFKKWNPANAYQLPGA
 TEMSFEGGVSQSAYNTLASLDLPKPEQGPETINQVTEHKMSAE

SEQ ID NO: 128

KTGYLTLGGSQRITGATCNGESSDGFTEFTPGDKVTCVAGNNTTITATFDTQSEAAARSLRAVEKVSFSLEDAQELAASD
 DKKSNAVSLVTSSNSCPADTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLNVNEEVENNAATDKAPSTHTSPVVP
 VTTPTGTPKPDNLNASFVSANAEQFYQYQPTTEILSEGRVLDVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGI
 DTFELGVSVRGNKSTIALTELGEVIRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLD
 EGDQNVVLPNPFIEQFNTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDKYKSVTKFHVHFD
 DSTNFYGSTGNARGQAVVNISNAAFPIILMARNDKKNYWLAFGEKRAWDKNELAYITEAPSLVQPENVTRODTATFNLFFI
 ISLGQVGEGLKMLVIGNPHYSILRCPNGYSWNGGVNKDGQCTLSGSDDMKHFQMNVLRYLSDDKWTPDAKASMTVG
 TNLDTVYFKRHGQVTGNSAEFGHDPDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPK
 LTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLDAAAGLSMALNKSVVNTDPOGYPNRVROQREKGIWVY
 ERYPAVDSAQPPYITDPDTGKVTWKYQEEGKPDCKPKEVASWQEDVDGKQVTRYAFIDEAEHSTESLEAAKAKIF
 EKFPGLQECKDSTYHYEINCLERRPGTGVPTGGMYVPQYTQLNLADTAKAMVQAADLGTNIQRLYQHLYFRTNG
 RKGERLSSVDLERLYQNMSVWLWNKIEYRYENDKDELGFKFTFEFLNCYANNAYSEGTCQCSADLKKSLVDNNMIY
 DGSSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGAVSAEGEKVTETISLYSNPTKWFAGNMQSTGL
 WAPAQQEVTIESTASVPVTVTVALADDLTGREKHEVALNRPPKVTCTYDLKANDKVTFKVPYGGGLIYIKNSPKNES
 AEFTFTGVVKAPFYKDGWKNALNSPAPLGELESDFVYTPKKNLNASNYSTYDGVAFKAFIDTFASSMNDIFYG
 RDGESGNHRMFTYKALTGHKHRTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNAETPLNV
 PGATEVANNVLAALYMQDRYLGMKNRVADDITVAPEYLDENQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDG
 ELPKFYSDRKGKGNLFLQLMHRKARGDDVNDKFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKWNPANAYQLP
 GASEMSFEGGVSQSAYNTLAAMHLSKPEKGPETINKVTEYSMPAE

SEQ ID NO: 129

MNKKFKYKKSLLAAILSATLLAGCDGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPE
 PTPEPEPEPVPTKTGYLTLLGSLRVTGDITCNDESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVS
 FSLEDAQELAGSDNKKSNALSLVTSMNSCPANTEQVCLEFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAAT
 DKAPSTHTSPVVPATTPGTPDLNASFVSANAEQFYQYQPTTEIILSEGRVDSQGDGVVGVNYYTNSGRGVTGENGE
 FSFSWGETISFGIDTFELGSLVRGNKSTIALTELGDVEVRGANIDQLIHRYSKAGQNHTRVVPDEVKVFAYEPNVINE
 IINLSLSNGATLGEQVQVNLNPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGV
 TNYKSVSKFHVFDSTNFGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPE
 NVTRETATFNLFFISLGQVGDGKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFENVLRYLSND
 RWLPDAKSSMTVGTNLDTVYFKKHGQVLGNSAPFAFHKDFGTGITVKPMTSYGNLNPDEVPLLLILNGFEYVTQWGS
 DYSIPLRADTSKPKLTQQDVTDLIAYMNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVMNNDPQGY
 PDVRQRRSTPIWVYERYPAVDGKPPYTIDDTTKEVIWKYQQENKPDCKPKLEVASWQEEVEGKQVTQFAFIDEADHKT
 PESLAAAKQRIIDAFPGLEVCKDSYHYEVNCLERYRPGTGVPVVTGGMYVPQYTQDLGADTAKAMLQAADLGTNIQR
 LYQHELYFRTNGLQGERLNSVDLERLYQNMSVWLWNETKYRYEEGKEDELGFKFTFEFLNCYTNNAYVGTQCSAELK
 KSLIDNKMIYGEESKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNIKVDVEKYPGVVNTNGETVTQINILYSAPTK
 WFAGNMQSTGLWAPAQQEVSIESKATVPVTVTVLADDLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGGLI
 YIKGDSKEVQSADFTFTGVVKAPFYKDGKQWHDLSNPAPLGELESASFVYTTPKKNLNASNYTGGLEQFANDLDTFA
 SSMNDFYGRDSEDKHRMFTYKNLPGHKHFRFANDVQISIGDAHSGYPVMNSSFSPNSTTLPTPLNDWLIWHEVGH
 NAAETPLTVPGATEVANNVLAALYMDRYLGKMNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNF
 DIKKWYPDGTPLPEFYSEREGMKGNLFLQMLHRKARGDEVSNDFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFF
 KKNWPGANAYQLPGATEMSFEGGVSQSAYNTLASLNLKPKKQGPETINKVTEYSMPAE

SEQ ID NO: 130

DTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPTPEPEPEPVPTKTGYLTLLGSLRVTGDITCN
 DESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSLVTSMNSCPANT
 EQVCLEFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPATTPGTPDLNASFVSANAE
 QFYQYQPTTEIILSEGRVDSQGDGVVGVNYYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSLVRGNKSTIALTEL
 GDEVRGANIDQLIHRYSKAGQNHTRVVPDEVKVFAYEPNVINEIINLSLSNGATLGEQVQVNLNPNEFIEQFKTGQ
 AKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHVFDSTNFGSTGNARGQAVVN
 ISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETATFNLFFISLGQVGDGKLMVIGNPHY
 NSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFENVLRYLSNDRWLPDAKSSMTVGTNLDTVYFKKHGQVLGNSAP
 FAFHKTGTGITVKPMTSYGNLNPDEVPLLLILNGFEYVTQWGSDPYSIPLRADTSKPKLTQQDVTDLIAYMNKGGSVL
 IMENVMSNLKEESASGFVRLDDAAGLSMALNKSVMNNDPQGYPDVRQRRSTPIWVYERYPAVDGKPPYTIDDTTKE
 VIWKYQQENKPDCKPKLEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKQRIIDAFPGLEVCKDSYHYEVNCL

EYRPGTGVPVVTGGMYVPQYTQDLGADTAKAMLQAADLGTNIQRLYQHELYFRTNGLQGERLNSVDLERLYQNMSVW
 LWNETKYRYEEGKEDELGFKFTFEFLNCYTNNAYVGTQCSAELKSLIDNKMIYGEESKAGMMNPSYPLNYMEKPL
 TRMLGRSWWDLNIKVDVEKYPGVVNTNGETVTQINILYSAPTKWFAGNMQSTGLWAPAQQEVSIESKATVPVTVTV
 ALADDLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGGLIYIKGDSKEVQSADFTFTGVVKAPFYKDGKQW
 HDLSNPAPLGELESASFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDYGRDSEDKHRMFTYKNLPGHKHFRFAN
 DVQISIGDAHSGYPVMNSSFSPNSTTLPTPLNDWLIWHEVGHNAETPLTVPGATEVANNVLAALYMDRYLGKMN
 RVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWYPDGTPLPEFYSEREGMKGNLFLQMLHR
 KARGDEVSNDFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKNWPGANAYQLPGATEMSFEGGVSQSAYNTLA
 SLNLKPKKQGPETINKVTEYSMPAE

SEQ ID NO: 131

KTGYLTLLGSLRVTGDITCNDESSDGFTFTPGDKVTCVAGNNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAGS
 DNKKSNALSLVTSMNSCPANTEQVCLEFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPV
 PATTPGTPDLNASFVSANAEQFYQYQPTTEIILSEGRVDSQGDGVVGVNYYTNSGRGVTGENGEFSFSWGETISFG
 IDTFELGSLVRGNKSTIALTELGDVEVRGANIDQLIHRYSKAGQNHTRVVPDEVKVFAYEPNVINEIINLSLSNGATL
 GEQVQVNLNPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYKSVSKFHV
 HDSTNFGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETATFNL
 FFISLGQVGDGKLMVIGNPHYNSILRCPNGYSWNGGVNKGQCTLNSDPDDMKNFENVLRYLSNDRWLPDAKSSMTV
 GTNLDTVYFKKHGQVLGNSAPFAFHKDFGTGITVKPMTSYGNLNPDEVPLLLILNGFEYVTQWGSDPYSIPLRADTSK
 PKLTQQDVTDLIAYMNKGGSVLIMENVMSNLKEESASGFVRLDDAAGLSMALNKSVMNNDPQGYPDVRQRRSTPIW
 VYERYPAVDGKPPYTIDDTTKEVIWKYQQENKPDCKPKLEVASWQEEVEGKQVTQFAFIDEADHKTPESLAAAKQRI
 IDAFPGLEVCKDSYHYEVNCLERYRPGTGVPVVTGGMYVPQYTQDLGADTAKAMLQAADLGTNIQRLYQHELYFRTN
 GLQGERLNSVDLERLYQNMSVWLWNETKYRYEEGKEDELGFKFTFEFLNCYTNNAYVGTQCSAELKSLIDNKMIYGE
 ESSKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNIKVDVEKYPGVVNTNGETVTQINILYSAPTKWFAGNMQSTGLW
 APAQQEVSIESKATVPVTVTVLADDLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGGLIYIKGDSKEVQSA
 DFTFTGVVKAPFYKDGKQWHDLSNPAPLGELESASFVYTTPKKNLNASNYTGGLEQFANDLDTFASSMNDYGRDSE
 DKHRMFTYKNLPGHKHFRFANDVQISIGDAHSGYPVMNSSFSPNSTTLPTPLNDWLIWHEVGHNAETPLTVPGAT
 EVANNVLAALYMDRYLGKMNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKKWYPDGTPLP
 EPHYSEREGMKGNLFLQMLHRKARGDEVSNDFGGKNYCAESNGNAADTLMLCASWVAQTDLSEFFKKNWPGANAYQL
 PGATEMSFEGGVSQSAYNTLASLNLKPKKQGPETINKVTEYSMPAE

SEQ ID NO: 132

ATGAATAAGAAATTTAAATATAAGAAATCGCTTTTAGCGGCTATTTTGTAGTGCAACCCTGTTAGCCGGTTGTGATGG
 CGGTGGGTCCGGATCTTCTCCGATACGCCGCTGTAGATTCTGGAACAGGGTCTTTGCCGGAAGTGAAACCTGATC
 CAACACCAAAACCCGGAGCCGACACCTGAGCCAACGCCGACCCAGAACCTACGCCGGAACCGACACCTGATCCTGAG
 CCAACACCAGAACCCGGAGCCAGAACCTGTTCCTACGAAAACGGGTATCTGACTCTGGGCGGAAGCCTGCGGGTAAC
 TGGTGATATCACCTGTAATGATGAATCCAGCGATGGCTTTACCTTTACACCAGGCGACAAAGTCACCTGTGTGGCAG
 GGAACAACACGACAATTGCTACCTTCGACACCCAGTCAGAAGCTGCGCGTAGCCTGCGTGCAGTTGAAAAAGTGTGCG
 TTTAGTCTTGAGGACGCGCAAGAACTGGCGGGTCCGACAACAAGAAAAGCAATGCGCTCTCGCTGGTCACCTCCAT
 GAACAGTTGCCCGCGAATACAGAACAGGTGTGCCTGGAGTTCTCCTCGGTGATCGAGAGTAAACGTTTCGACTCGC
 TGTATAAGCAAATCGATCTGGCACCCGGAAGAATCAAAAAGCTGGTCAATGAAGAGGTGGAAAACAATGCCGCGACC
 GATAAAGCGCCATCCACTCATACTTCACCGGTGCTGCCGCCACCCTCCGGGAACAAAACCGGATCTAAACGCTTC
 CTTCTGTCGGCTAACCGGGAACAGTTTATCAGTATCAACCCACTGAAATCATTCTCTCTGAAGGTGCGACTGGTCG
 ATAGTCAGGGGGGATGGTGTGTTGGTGTCAACTATTACACCAATTCGGCCGTTGGTGTAAACCGGAGAAAACGGGGAA
 TTTTCCTTTAGTTGGGGGGAACCATCTCCTTTGGCATCGACACTTTTGAGCTTGGTTCTGTGCGTGGTAACAAGTC
 GACTATTGCATTGACTGAAGTTCGGGTGATGAAGTTCGCGGGGCAATATCGATCAGTTGATTACCCGCTATTTCGAAGG
 CTGGACAAAATCACACGCGTGTAGTTCGGATGAAGTGCACAAGTTTGTGCTGAATATCCCAACGTGATTAACGAG
 ATTATCAATCTCTGTTATCCAATGGTGGCAGCTGGGGGAAGGTGAGCAAGTCGTTAATCTGCTTACGAATTTAT
 CGAGCAGTTTAAGACGGGTGAGGCCAAAGAGATCGATACCGCGATTGTGCGAAAACCGATGGTTGTAACGAGGCTC
 GCTGTTCTCGCTGACAACGCGCAATGTTAATGACGCCAGATTGAGGGCGTTATCAACAGCTGTGGGGCGTGGAT
 ACGAACTACAAATCTGTGATGAAGTTCATGTATTCCATGACTCTACCAACTTCTATGGCAGCACGGGTAATGCCGCG
 CGGTGAGCGGTGGTGAATATCTCCAACGCGGCTTCCGATTTCTGATGGCGCGTAATGATAAAACTACTGGCTGG
 CGTTTGGCGAAAAACGCGCCTGGGATAAAAATGAGCTGGCGTACATTACTGAAGCGCTTCCATTGTGCGACCCAGAG
 AACGTGACACGCGAAACCGCCACCTTCAACCTGCCGTTTATCTCGCTGGGGCAAGTGGGCGATGGCAAGCTGATGGT
 TATCGGTAACCCACACTACAACAGCATCTGCGTTGCCGGAACGGCTACAGCTGGAACGGGGCGTTAATAAAGACG
 GACAGTGTACGCTCAACAGCGACCCGGATGACATGAAGAATTCATGGAGAAGCTGCTGCGCTATCTGTCAAATGAT
 CGCTGGTTGCCGGATGCAAAATCCAGTATGACCGTGGGTACTAACCTGGACACGGTATATTTCAAAAAACATGGTCA
 GGTGCTGGGAAATAGCGCACCGTTTGGCTTCCACAAGGATTTCACTGGCATCACGGTCAAACCAATGACAGCTATG
 GCAATCTGAATCCAGATGAAGTTCCTCTGTTGATCTCAATGGCTTTGAATACGTACACAATGGGGTAGCGATCCT
 TACTCCATTCTCTGCGCGCAGATACCAGCAAACCGAAGCTGACCCAGCAGGATGTGACCGATTGATCGCCTATAT
 GAACAAAGGTGGATCGGTGCTGATCATGAAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGCATCTGGCTTTGTAC
 GTCTGCTTGATGCCGAGGTTTGTGATGGCGCTTAACAAGTCGGTAGTAAATAACGATCCGCAAGGCTACCCCTGAC

CGTGTTCTGTCACGACGTTCAACGCCAATTTGGGTCTATGAGCGTTATCCGGCTGTCGATGGTAAACCACCGTATAC
 CATTGATGACACCACGAAAGAACTTATCTGGAAATATCAGCAAGAAAAACAAACCTGATGACAAACCGAAGCTGGAG
 TTGCCAGCTGGCAGGAAGAAGTTGAGGGTAAACAGGTAAACCAATTCGCCCTTATTGATGAAGCCGACCAAAAAACG
 CCTGAGTCACGTGGCTGCGGCAAAACAGAGAATTTCTGGACGCTTCCAGGGCTGGAAGTGTGTAAGGATCTGACTA
 TCACTATGAGGTCAACTGTCTGGAATACCGCCAGGCGAGGTGTGCCGGTAACCGGTGGCATGTATGTTCCGCACT
 ATACGCGAGCTGGATCTTGGAGCTGACACTGCGAAACCGATGCTGCGAGGCTGCGGATTTAGGCACCAATATTGAGCGC
 CTGTATCAGCATGAGCTTTATTTCCGTACCAATGGCCTCCAGGGTGAGCGTCTCAACAGCGTTGATCTGGAACGTTT
 ATACCAAAACATGTCCGTCTGGCTGTGGAACGAGACGAAATATCGTTATGAAGAGGGTAAAGAAGACGAGCTGGGCT
 TTAAACGTTTCACTGAGTTTCTGAAGTCTACACCAACAAATGCAATACGTTGGCACGCGAGTGTCTGCTGAGCTGAAA
 AAATCGCTGATCGATAACAAGATGATTIATGGTGAAGAAAGCAGCAAAGCGGGCATGATGAACCCGAGCTACCCGCT
 CAACTATATGAAAAACCGCTGACACGCCGTGATGCTGGGCCCTTCCCTGCTGCGATCTGAACATCAAAGTTGATGTTG
 AGAAGTATCCGGGAGTGGTCAATACAAACCGCGAAACCGTCAACACAAAACATTAACCTGTACTCAGCTCCAACCAAA
 TGGTTTGCAGGTAACATGCAGTCAACTGGCCTGTGGGCACTGCCAGCAGGAAGTCAGCATGAGTCAAAAGCGGAC
 AGTTCCCTGTGACCGTGAAGTGTGCGCTGGCCGACGACCTGACAGGACGAGAGAAGCATGAAGTTAGCCTGAATCGTC
 CACCCAGAGTGACAAAAACCTATGACCTGAAAGCCAATGATAAGGTGACGTTCAAAGTCCCTTACGGTGGTCTGATT
 TACATCAAGGGCGACAGCAAAGAGGTGCAATCAGCTGACTTCACCTTTACCGGTGTAGTAAAACGCGCGTTCTATAAA
 AGACGGTAAGTGGCAACACGATCTGAATCCCTGCCCGCTGGGCGAACTGGAGTCTGCCTCGTCTATACCA
 CACCGAAGAAGAACCITGAATGCCAGCAATTAACCTGGCGGACTGGAGCAATTCGCTAACGATCTGGATACCTTTGCC
 AGCTCGATGAATGACTTCTACGGCCGTGATAGCGAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGCCGGE
 CCACAAACATCGTTTTCGCCAACGATGTGCAGATCTCCATCGGTGATGCGCATTCGGGTATCCGGTAATGAACAGCA
 GCTTCTCGCCGAAACAGCACACCGCTGCCGACGACCGCGCTGAACGACTGGCTGATCTGGCATGAAGTTCGCTCATAAC
 GCCGCAAAACCGCGTTGACTGTACCGGTGCAACTCAACTCGCTTAACACGCTGCTGGCGCTGTACATGCAGGATCG
 TTATCTCGGCAACATGAACCGTCTCGCTGACCATATTACCGTCGCACCGGAATACTCGGAGGAAGCAACGGTCAAG
 CATGGGCGCGCGCGGTGCGGGTGACCGTCTGCTGATGTAGCACAGCTGAAGGAATGGGCGAGAGAAAACTTTGAT
 ATCAAGAAATGGTATCCAGATGGCACTCCTCTGCCAGAGTTTACAGCGAGCGTGAAGGGATGAAGGGCTGGAACCT
 GTTCCAGTTGATGCATCGTAAAGCACGCGCGATGAGGTGAGCAATGACAAGTTTGGCGGCAAGAATTACTGTGCTG
 AATCAACCGTAAACGACGCGGACACGCTGATGCTGTGTGCTCCTGGGTGCGCCAGACGGATCTTTCGGAGTTCTTT
 AAGAAATGGAATCCGGGCGCAATGCTTACCAGTTGCCGGGAGCGACGAGATGAGCTTGAAGGCGGTGTGAGCCA
 GTCGCGGTACAACACGCTGGCGTCACTCAATCTGCCGAACCGAAGCAAGGCGCGAAACCAITTAACAAGGTTACCG
 AGTATTGATGCCTGCTGAA

SEQ ID NO: 133

MNKKFKYKKSLLAAILSATLLAGCDGGSGSSSDTPPVDSTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPEPEPEPVPTKTGYLTLLGGSQRVTGATCNGESSDGFFTFTPGEDVTCVAGNTTIATFDTQSEAAARSLRAVEKVSFSLE
 DAQELAASDDKKSNVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNNEEVENNAATDKAPSTHTSPVVPVTPGTCPDLNASFVSANAEQFYQYQPTTEIILSEGRVLDVSMGNGVVGVNYYTSSGRGVTGENGKFNFS
 WGETISFGIDTFELGSGVRGNKSTIALTELGEVRGANIDQLIHRYSQAGKNDEREVPDVVRKVFAEYPNVINEIINLSLSNGEALSEGQDQTFERTNEFLEQFESGQAKEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDKDYK
 SVTKFHFVHDSTNFGSTGNARGQAVVNI SNAAPFILMARNDKNYWLA FGEKRAWDKNELAYITEAPSLVEPENVT
 DTATFNLPIISLGQVGEGLMVIGNPHYNSILRCPNGYSWEGGVKNGQCTRNSDSNDMKHFMQNVRLRYLSDDKWTP
 DAKASMTVGTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNPDYAIPL
 RADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKS VVNNDPQGYPNRVRQ
 QRATGIWVYERYPAVDGALPYTIDSKTGEVKWKYQVENKPDCKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSL
 KAAKAKIFEKFPGLKECKDPTYHYEVNCLERYRPGTGVPVTTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQH
 ELYFRTNGRKGERLSSVDLERLYQNMSVWLWNKIEYRYENDKDELGFKTFTEFLNCYANDAYTGGTQCSDELKKS
 LDNNMIYGEKSVNKGMMNPSPYPLNYMEKPLTRLMLGRSWWDLNLIKVDVEKYPGAVSAEGEKVTETISLSYNSPTKWF
 AGNMQSTGLWAPAQKEVTIESSASVPVTVTVALADDLTGREKHEVALNRPPKVTKTYDLKANDKVTFKVPYGGGLIY
 KGNSPKNESAEFTFTGVVKAPFYKDGWKNALNSPAPLGELESDAFVYTTTPKKNLEASNFTGGVAEFAKDLDTFASS
 MNDFYGRNDEGDKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNA
 AETPLTVPGATEVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIK
 QWYPEGDLPKFYS DREGMKGNLFLQLMHRKARGDDVGKTKFGERNYCAESNGNAADKMLCASFVAQTDLSEFFKKW
 NPGANAYQLPGASEMNFEGGVSQSAYETLAALNLPKPQQGPETINQVTEHKMSAE

SEQ ID NO: 134

DTPPVDSTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPTPDPEPEPEPEPVPTKTGYLTLLGGSQRVTGATCNGESS
 DGFFTFTPGEDVTCVAGNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAASDDKKSNVSLVTSSNSCPANTEQVC
 LTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNNEEVENNAATDKAPSTHTSPVVPVTPGTCPDLNASFVSANAEQFYQ
 YQPTTEIILSEGRVLDVSMGNGVVGVNYYTSSGRGVTGENGKFNFSWGETISFGIDTFELGSGVRGNKSTIALTELGEV
 RGANIDQLIHRYSQAGKNDEREVPDVVRKVFAEYPNVINEIINLSLSNGEALSEGQDQTFERTNEFLEQFESGQAKEI
 DTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDKDYKSVTKFHFVHDSTNFGSTGNARGQAVVNI SNA
 APFILMARNDKNYWLA FGEKRAWDKNELAYITEAPSLVEPENVTROTATFNLPIISLGQVGEGLMVIGNPHYNSILR
 CPNGYSWEGGVKNGQCTRNSDSNDMKHFMQNVRLRYLSDDKWTPDAKASMTVGTNLDTVYFKRHGQVTGNSAEFGFH

PDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMEN
 VMSNLKEESASGFVRLLDAAGLSMALNKS VVNNDPQGYPNRVRQQRATGIWVYERYPAVDGALPYTIDSKTGEVKWK
 YQVENKPDCKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKAKIFEKFPGLKECKDPTYHYEVNCLERYR
 GTGVPVTTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQH ELYFRTNGRKGERLSSVDLERLYQNMSVWLWNK
 IEYRYENDKDELGFKTFTEFLNCYANDAYTGGTQCSDELKKS LDNNMIYGEKSVNKGMMNPSPYPLNYMEKPLTRL
 MLGRSWWDLNLIKVDVEKYPGAVSAEGEKVTETISLSYNSPTKWFAGNMQSTGLWAPAQKEVTIESSASVPVTVTVAL
 ADDLTGREKHEVALNRPPKVTKTYDLKANDKVTFKVPYGGGLIY KGNSPKNESAEFTFTGVVKAPFYKDGWKNALN
 SPAPLGELESDAFVYTTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEGDKHRMFTYKNLTGHKHRFTNDV
 QISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNA AETPLTVPGATEVANNVLALYMQDRYLGMKNRVA
 DDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKQWYPEGDLPKFYS DREGMKGNLFLQLMHRKARG
 DDVGKTKFGERNYCAESNGNAADKMLCASFVAQTDLSEFFKKWNPGANAYQLPGASEMNFEGGVSQSAYETLAALN
 LPKPQQGPETINQVTEHKMSAE

SEQ ID NO: 135

KTGYLTLLGGSQRVTGATCNGESSDGFFTFTPGEDVTCVAGNTTIATFDTQSEAAARSLRAVEKVSFSLEDAQELAASDD
 KKSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNNEEVENNAATDKAPSTHTSPVVPV
 TTPGTCPDLNASFVSANAEQFYQYQPTTEIILSEGRVLDVSMGNGVVGVNYYTSSGRGVTGENGKFNFSWGETISFGID
 TFELGSGVRGNKSTIALTELGEVRGANIDQLIHRYSQAGKNDEREVPDVVRKVFAEYPNVINEIINLSLSNGEALSE
 GDQTFERTNEFLEQFESGQAKEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDKDYKSVTKFHFVHD
 STNFGSTGNARGQAVVNI SNAAPFILMARNDKNYWLA FGEKRAWDKNELAYITEAPSLVEPENVTROTATFNLPII
 SLGQVGEGLMVIGNPHYNSILRCPNGYSWEGGVKNGQCTRNSDSNDMKHFMQNVRLRYLSDDKWTPDAKASMTVGT
 NLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPOEMPLLIILNGFEYVTQVGNPDYAIPLRADTSKPKL
 TQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKS VVNNDPQGYPNRVRQQRATGIWVY
 ERYPAVDGALPYTIDSKTGEVKWKYQVENKPDCKPKLEVASWLEDVDGKQETRYAFIDEADHKTEDSLKAAKAKIFEK
 FPGLKECKDPTYHYEVNCLERYRPGTGVPVTTGGMYVPQYTQLSLNADTAKAMVQAADLGTNIQRLYQH ELYFRTNGRK
 GERLSSVDLERLYQNMSVWLWNKIEYRYENDKDELGFKTFTEFLNCYANDAYTGGTQCSDELKKS LDNNMIYGEK
 SVNKGMMNPSPYPLNYMEKPLTRLMLGRSWWDLNLIKVDVEKYPGAVSAEGEKVTETISLSYNSPTKWFAGNMQSTGLW
 APAQKEVTIESSASVPVTVTVALADDLTGREKHEVALNRPPKVTKTYDLKANDKVTFKVPYGGGLIY KGNSPKNESAE
 FTFTGVVKAPFYKDGWKNALNSPAPLGELESDAFVYTTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDE
 GDKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNA AETPLTVPGAT
 EVANNVLALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKQWYPEGDLPK
 FYS DREGMKGNLFLQLMHRKARGDDVGKTKFGERNYCAESNGNAADKMLCASFVAQTDLSEFFKKWNPGANAYQLP
 GASEMNFEGGVSQSAYETLAALNLPKPQQGPETINQVTEHKMSAE

SEQ ID NO: 136

ATGAATAAGAAATTTAAATATAAGAAATCGCTTTTAGCGGCTATTTTGTAGTGAACCCCTGTTAGCCGGTTGTGATGG
 CGGTGGCTCCGGATCTTCTCCGATACGCGCCTGTAGATTCTGGAACAGGGTCTTTGCCGGAAGTGAAACCTGATC
 CAACACCAAACCCGGAGCCGACGCTGAGCCAACGCCGACCCAGAACCTACGCCGGAACCGACACCTGATCCTGAG
 CCAGAACCGGAGCCAGAACCTGTTCCTACGAAAACGGGTTATCTGACCCTGGCGGGAAGCCAGCGGGTAACTGGTGC
 TACTTGTATGGTGAATCCAGCGATGGCTTTACATTTACACCTGGTGAGGACGTTACTTGGCTGGCGGGTAACACGA
 CAATTGCTACCTTCGACACCCAGTCAGAAGCTGCGCGTAGCCTGCGTGCGGTTGAAAAAGTGTCTGTTAGTCTTGAG
 GACGCGCAAGAACTGGCAGCTTCCGATGACAAGAAAAGCAATGCGGTTTTCGCTGGTAACGTCCAGTAACAGCTGTCC
 GGCGAATACAGAACAGGTTTGCTGACGTCTCTCTCAGTGATCGAGAGTAAACGTTTCGACTCGCTGTATAAGCAAA
 TCGATCTGGCACCAGGAGTTCAAAAAGCTGGTCAATGAAGAGGTTGAAAAACAATGCCCGACCGATAAAGCGCCA
 TCCACTCATACTTACCCGTTGTGCCGTCACACGCCGGGAACAAAACCGGATCTGAACGCTTCTTTCGTGTCCGG
 TAACCGGAACAGTTTTATCAGTATCAACCCACTGAAATCATTTCTCCGAAGGCCGACTGGTAGATAGCATGGGCA
 ATGGTGTGGTTGGCGTAAATTACTACACCAGCTCAGGCCGTGGCGTAACTGGCGAAAACGGCAAATTCAACTTCAGC
 TGGGGCGAAACCATCTCTTTGGTATCGACACCTTTGAACTGGGCTCAGTGCGCGGCAATAAGTCGACCATTGCGTT
 GACTGAACTGGGTGACGAAGTTGCGCGCGCGAATATTGATCAGCTTATTCACTCGTTACTCCCAGGCCGAAAAAATG
 ATGAGCGTGAAGTGCCGACGCTAGTGCGCAAGGCTTTTGGCGAGTATCCCAACGTAATCAACGAGATTATCAATCTC
 TCGTTATCCCAATGGCGAGGCTTGAGCGAAGGCGATCAAACTTTGAGCGGACAAACGAATTTCTTGAGCAGTTTGA
 ATCCGGGCAGGCTAAAGAGATTGATACGGCGATTTGTGACTCCCTTGGGGGCTGCAACTCTCAGCGTTGGTTCTCGT
 TGACAGCACGCAATGTTAACGAAGGTCAGATTGAGGCGGTTATCAACAAGCTGTGGGGTGTGGATAAAGATTACAAA
 TCAGTGACGAAATCCACGCTTCTCATGACTCTACCAACTTCTATGGCAGTACCGGTAATGCGCGCGGTCAGGCAGT
 GGTGAATATCTCCAACGCGGCTTCCCGATTCTGATGGCGCGTAATGATAAAAACTACTGGCTGGCGTTTGGCGAAA
 AACCGCCTGGGATAAAAAAGAGCTGGCGTACATTACGGAAGCGCTTCTCTCGTTGAGCCGGAACGTTACGCGC
 GATACCGCCACCTTTAACTGCCGTTTATTTTCGCTGGGGCAAGTCGGTGAGGGCAAACTGATGGTTATCGGTAACCC
 GCACTACAACAGCATTTTGGCTTGTCCGAACGGTTACAGCTGGGAAGGCGGTTGTGATAAAAAAGGTCAGTGTACGC
 GTAACAGTGATTCTAATGATATGAAGCACTTTATGCAGAACGTTGTGCGCTATCTGTCCGACGATAAATGGACGCCG
 GACCGGAAAGCCAGCATGACCGTAGGCACCAACCTGGATACTGTCTATTTCAAACGTCATGGTCAGGTTACAGGAAA
 CAGCGCTGAGTTCCGCTTTTCATCCGATTTTGGCGGTATCTCTGTTGAGCATTTAAGTAGCTATGGCGATCTCGACC
 CGCAGGAAATGCCGCTGCTGATCCTCAACGGCTTTGAGTATGTGACTCAGGTTGGTAACGATCCTTATGCAATCCCG

CTGCGAGCAGATACCAGCAAACCGAAGCTGACTCAGCAGGATGTGACCGATCTGATCGCCTATCTGAACAAAGGTGG
 ATCGGTGCTGATCATGGAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGGCTGTGGCTTTGTACGTCTGTTGATG
 CCGCAGGTCTGTTCGATGGCACTGAACAAGTCGGTAGTAATAACGATCCGCAAGGGTATCCCAACCGGTTCTGTCAG
 CAGCGCGCAACCGGCATTTGGGTCTATGAACGTTATCTGCGCTAGATGGTGCGCTGCCGTACCCATCGATAGTAA
 GACAGGGGAAGTTAAGTGGAAATATCAGGTAGAAAACAACTGATGACAACCGAAGCTGGAAGTTGCCAGCTGGC
 TGGAAGATGTAGATGGCAAACAGGAAACGCGTTATGCCCTTTATGATGAGGCCGATCATAAAACAGAGGATTCTCTG
 AACGCTCCCAAGCCAAAAATCTTTGAGAAGTTTCTGGATTAAAGGAGTGTAAAGACCCAACCTTACCCTACGAGGT
 CAACGTCTCTGGAATATCTGCTGGCACGGGGGTTCCGGTTACTGGTGGCATGTATGTTCCACAGTATACGCAACTAA
 GCCTTAACGCCGACACGGCAAAAGCGATGGTGCAGGCTGCGGATTTAGGCACCAACATTCAGCGTCTGTATCAGCAT
 GAGCTCTACTTCCGGACCAATGGTTCGCAAAGGTGAGCGCTCTGAGCAGCGTCGATCTGGAACGTCGTACCGAAGCAT
 GTCGGTCTGGCTGTGGAATAAAATGAATATCGCTATGAAAACGACAAGGATGACGAGCTGGGCTTTAAACGTTTCA
 CGAGTTCTTGAACCTCTACGCCAACGATGCCATACTGGCGGCACGCAAGTGTCTGATGAGCTGAAAAAATCGCTG
 GTCGATAACAACATGATCTACGGCGAGAAGAGCGTTAATAAAGCGGGCATGATGAACCCGAGCTATCCGCTCAACTA
 TATGAAAAACCGCTGACGCGCCTGATGCTGGGTCTTCTGTTGGGATCTAAACATCAAAGTTGATGTGCGAGAAGT
 ATCCGGGAGCGGTATCGGCAGAGGTGAGAAGTTACTGAAACCATCAGCCTGTACTCCAATCCGACCAATGGTTT
 CCGCTTAACATGCAGTCTACTGGCCTGTGGGCTCCGGCTCAGAAAAGAGGTCACCATTGAGTCTTCTGATCAGTTCC
 TGTGACTGTCAACGTGGCGCTGGCCGACGATCTGACCGGACGCGAGAAGCATGAAGTCGCGCTGAACCGTCCGCCAA
 AAGTGACGAAAACCTATGATCTGAAAGCCAATGATAAGGTGACCTTCAAGGTGCCTTACGGTGGTCTGATTATATC
 AAGGGCAACAGCCCAAAGATGAGTCAGCCGAATTCACCTTTACTGGTGTGGTAAAAGCGCGCTTCTATAAAGACGG
 CGAATGGAAAAACGCTCTGAACCTCCCTGCGCGCTGGGCGAGCTGGAATCAGACGCTTTTGGTTTATACCACGCCGA
 AGAAGAACCTTGAGGCTAGCAATTTCACTGCTGCTTAGCAGAATTGCTAAAGATCTGGATACCTTTGCCAGCTCG
 ATGAATGACTTCTACGGTCGTAAATGATGAAGACGGTAAACACCGGATGTTTACCTATAAAAACCTTGACGGGGCACA
 GCATCGTTTACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGTTTATCCGGTAAATGAACAGCAGCTTCT
 GACGACACGACACGCTGCCGACGACGCGCTGAACGACTGTTGATCTGGCAGCAAGTAGGGCACAACGCGCGC
 GAAACGCGCTGACTGTTCTTGGCGCACTGAAGTAGCGAACACGTTCTGGCGCTGTACATGCAGGATCGCTATCT
 CGGCAAGATGAACCGTGTCTGCTGACGATATTACTGTGCGCGCGGAATATCTGACGAGACCAACCGCTCAGGCATGGG
 CGCGCGGCGGTGCGGGCGACCGTCTGCTGATGTACGCGCAGCTGAAGGAGTGGGCGAGAAAAAATTTGATATCAAA
 CAGTGGTATCCAGAAGGTGACCTGCCTAAGTTCTACAGCGATCGTGAAGGGATGAAAGGCTGGAACCTGTTCCAGTT
 GATGCACCGTAAAGCACGCGCGATGATGTTGGCAAAACCAAGTTTGGCGAAAGAAATTTACTGTGCCGAATCCAACG
 GTAACGCTCCCGACAAGCTGATGCTGTGTGCCCTCTGGGTGCTCAGACGGATCTTTCGGAGTTCTTTAAGAAATGG
 AATCCGGGCGCAATGCTTACCAGTTACCGGGAGCGAGCGAGATGAACCTCGAGGGCGGTGTGAGCCAGTCCGGTTTA
 CGAGACGCTGGCGCGCTTAATCTGCCGAACCGCAGCAAGGGCGGAAACCATTAATCAGGTTACCGAGCATAAGA
 TGTACGCTGAG

SEQ ID NO: 137

MNKKFKYKKSLLAAILSATLLAGCDGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPE
PTPEPEPEPVPTKTGYLTLGGSQRVTGATCNGESSDGTFTPGEDVTCVAGNTTIATFNTQSEAARSLRAVEKVSFS
LEDAQELAASDNKKSNVSLVTSNSCPADTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDK
APSTHTSPVVPVPTPGTKPDLNASFVSANAEQFYQYQPTTEIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFS
FSWGETISFGIDTFELGSRGNKSTIALTELGD EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEII
NLSLSNGATLDEGEQVVNLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDT
YKSVSKFHVHFDSTNFGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPEN
TRETATFNLFFISLGQVGKGLMVIIGNPHYNSILRCPNGYSWEGGVDKNGQCTRNSDSNDMKHFMQNVRLRYLSNDKW
TPDAKASMTVGTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSYGDLDLPQKMPLLIILNGFEYVTQVGGDPYA
VPLRADTSKPKLTQQDVTDLIAYMNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSVMNNDPQGYPDRV
RQRRATGIWVYERYPVVEGELPYTIDSKTGKVTWKYQIDNKPDKPKLEVASWQEEVDGKQVTQFAFIDEADHKTTE
SLDAAKKKILEKFKGLEECKDSTYHIEINCLEYRPGTNVPVTGGMYVPRYTQNLNSADTAKAMVQAADLGTNIQRLY
QHELYFRTNGRKGERLSSVDLERLYQNMSVWLWNEIEYSYDSSKEDELGFKTFTEFLNCYANDAYTGGTQCSDELKK
SLVDNNMIYGEKSVNKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGAVSAEGEKVTETISLYSNPTK
WFAGNMOSTGLWAPAQKEVTIESSASVPVTVTVLADDLTGREKHEVALNRPPKVTCTYDLKANDKVTFKVPYGGGLI
YIKGNSPKNESAEFTFTGVVKAPFYKDGWKNALNSPAPLGELESDAFVYTTPKKNLEASNFTGGVAEFAKDLDTFA
SSMNDIFYGRNDEGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTPLNDWLIWHEVGH
AAETPLTVPGATEVANNVLAALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAENFD
IKTWYPDGNLPAFYSEREGMKGWNLQLMHRKARGDEVGKTKFGERNYCAESNGNAADTLMLCASWVAQTDLSEFFK
KWNPGANAYQLPGASEMNFEGGVSQSAYETLAAALNLPKPQOGPETINQVTEHKMSAE

SEQ ID NO: 138

DTTPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTPEPIPDPEPTPEPEPEPVPTKTGYLTLGGSQRVTGATCNGE
SSDGTFTPGEDVTCVAGNTTIATFNTQSEAARSLRAVEKVSFSLEDAQELAASDNKKSNVSLVTSNSCPADTEQ
VCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVPTPGTKPDLNASFVSANAEQF
YQYQPTTEIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGIDTFELGSRGNKSTIALTELGD
EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIIINLSLSNGATLDEGEQVVNLPNEFIEQFKTGQAK

EIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTSYKSVSKFHVHFDSTNFGSTGNARGQAVVNISN
AAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETATFNLFFISLGQVGKGLMVIIGNPHYNSI
LRCPNGYSWEGGVDKNGQCTRNSDSNDMKHFMQNVRLRYLSNDKWTPDAKASMTVGTNLDTVYFKRHGQVTGNSAEFG
FHPDFAGISVEHLSYGDLDLPQKMPLLIILNGFEYVTQVGGDPYAVPLRADTSKPKLTQQDVTDLIAYMNKGGSVLIM
ENVMSNLKEESASGFVRLLDAAGLSMALNKSVMNNDPQGYPDRVRQRRATGIWVYERYPVVEGELPYTIDSKTGKVT
WKYQIDNKPDKPKLEVASWQEEVDGKQVTQFAFIDEADHKTTESLDAAKKKILEKFKGLEECKDSTYHIEINCLEY
RPGTNVPVTGGMYVPRYTQNLNSADTAKAMVQAADLGTNIQRLYQHELYFRTNGRKGERLSSVDLERLYQNMSVWLW
NEIEYSYDSSKEDELGFKTFTEFLNCYANDAYTGGTQCSDELKKSVDNNMIYGEKSVNKAGMMNPSYPLNYMEKPL
TRMLGRSWWDLNLIKVDVEKYPGAVSAEGEKVTETISLYSNPTKWFAGNMOSTGLWAPAQKEVTIESSASVPVTVTV
ALADDLTGREKHEVALNRPPKVTCTYDLKANDKVTFKVPYGGGLIYIKGNSPKNESAEFTFTGVVKAPFYKDGWKN
ALNSPAPLGELESDAFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDIFYGRNDEGKHRMFTYKNLTGHKHRFT
DVQISIGDAHSGYPVMNSSFSTNSTLPTPLNDWLIWHEVGHNAETPLTVPGATEVANNVLAALYMQDRYLGMKNR
VADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAENFDIKTWYPDGNLPAFYSEREGMKGWNLQLMHRKA
RGDEVGKTKFGERNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPGANAYQLPGASEMNFEGGVSQSAYETLAA
LNLKPQOGPETINQVTEHKMSAE

SEQ ID NO: 139

KTGYLTLGGSQRVTGATCNGESSDGTFTPGEDVTCVAGNTTIATFNTQSEAARSLRAVEKVSFSLEDAQELAASDN
KKSNAVSLVTSNSCPADTEQVCLTFSSVIESKRFDLSLYKQIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPV
TTPGTPKPDNLNASFVSANAEQFYQYQPTTEIILSEGRLVDSQGYGVAGVNYTNSGRGVTGENGEFSFSWGETISFGID
TFELGSRGNKSTIALTELGD EVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIIINLSLSNGATLDE
GEQVVNLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTSYKSVSKFHVHFD
STNFGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVRPENVTRETATFNLFFI
SLGQVGKGLMVIIGNPHYNSILRCPNGYSWEGGVDKNGQCTRNSDSNDMKHFMQNVRLRYLSNDKWTPDAKASMTVGT
NLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSYGDLDLPQKMPLLIILNGFEYVTQVGGDPYAVPLRADTSKPKL
TQQDVTDLIAYMNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSVMNNDPQGYPDRVRQRRATGIWVY
RYPVVEGELPYTIDSKTGKVTWKYQIDNKPDKPKLEVASWQEEVDGKQVTQFAFIDEADHKTTESLDAAKKKILEK
FKGLEECKDSTYHIEINCLEYRPGTNVPVTGGMYVPRYTQNLNSADTAKAMVQAADLGTNIQRLYQHELYFRTNGRK
GERLSSVDLERLYQNMSVWLWNEIEYSYDSSKEDELGFKTFTEFLNCYANDAYTGGTQCSDELKKSVDNNMIYGEK
SVNKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGAVSAEGEKVTETISLYSNPTKWFAGNMOSTGLW
APAQKEVTIESSASVPVTVTVLADDLTGREKHEVALNRPPKVTCTYDLKANDKVTFKVPYGGGLIYIKGNSPKNES
AEFTFTGVVKAPFYKDGWKNALNSPAPLGELESDAFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDIFYGRNDE
DGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTPLNDWLIWHEVGHNAETPLTVPGAT
EVANNVLAALYMQDRYLGMKNRVADDITVAPEYLEESNGQAWARGGAGDRLLMYAQLKEWAENFDIKTWYPDGNLPA
FYSEREGMKGWNLQLMHRKARGDEVGKTKFGERNYCAESNGNAADTLMLCASWVAQTDLSEFFKKWNPGANAYQLP
GASEMNFEGGVSQSAYETLAAALNLPKPQOGPETINQVTEHKMSAE

SEQ ID NO: 140

ATGAATAAGAAATTTAAATATAAGAAATCGCTTTTAGCGGCTATTTTGTAGCGCAACCCCTGTTAGCCGGTTGTGATGG
 CGGTGGTTCCGGATCTTCCCTCGATACGCCGCTGTAGATTCTGGAACAGGGTCTTTGCCGGAAGTGAAACCTGATC
 CAACACCAAACCCGGAGCCGACGCTGAGCCAACGCCGAGCCAGAACCTACGCCGGAACCGATACCTGATCCTGAG
 CCAACACCAGAACCCGAGCCAGAACCTGTTCTACGAAAACGGGTATCTGACCTGGGCGGAAGCCAGCGGGTAAC
 TGGTGTCTACCTGTAATGGTGAATCCAGCGATGGCTTACATTTACACCTGGTGAGGACGTTACTTGGCTGGCGGGTA
 ACACGACAATTGCCACCTTCAACACTCAGTCAGAAGCTGCGCGTAGCCTGCGTGCGGTTGAAAAAGTGTCTTTAGT
 CTTGAGGACGCGCAAGAACTGGCGGCTTCCGACAATAAGAAAAGCAATGCGGTTTTCGCTGGTAACGTCCAGTAACAG
 CTGTCCGGCGGATACAGAACAGGTTTGCTGACGTTCTCCTCTGTGATCGAGAGTAAACGTTTCGACTCGCTGTATA
 AGCAAATCGATCTGGCACCAGGAGTTCAAAAAGCTGGTCAATGAAGAGGTGGAAAAAATGCCCGGACCGGATAAAA
 GCGCCATCCACTCATCTTACCAGGTCGTGCTGTACCACGCCGGGAACAAAACCGGATCTGAACGCTCTTTCGT
 GTCGGCTAACCGGGAACAGTTTATCAGTATCAACCTACGGAATCATTCTCTCCGAAGCCGACTGGTAGATAGCC
 AGGGATATGGTGTGTGCTGGCGTCAACTACTACCAATTCAGGCCGTGGCGTGACAGGGGAAAATGGTGAATTTTCC
 TTTAGCTGGGGCGAAACCATCTCCTTTGGTATCGATACCTTTGAACTGGGTTCAGTGCAGCGCAATAAGTCGACCAT
 TGGCGTGAAGTGAAGTTTCGCGGGCGCAATATTGATCAGCTTATTCATCGCTATTCGACGACCGGGC
 AAAATAATACCCGTGTTGTTCCGGACGATGTACGCAAGGTCTTTGCCGAATATCCCAACGTGATCAACGAGATTATC
 AATCTCTCGTTATCCAACGGTGGCAGCGTGGATGAAGGTGAGCAAGTTGTTAATCTGCCAAACGAATTTATTGAGCA
 GTTTAAGACGGGTCAAGCCAAAGAGATCGATACCGGATTTGTGCGAAAACCGAGCGTTGTAACGAGGCTCGCTGGT
 TCTCGCTGACGACGCGCAATGTTAATGACGCGCAGATTACAGGCCGTATTAAACAAGCTGTGGGGTGTGGATACAAGT
 TACAAATCTGTGACGAAGTTCCATGTATCCATGACTCTACCAACTTCTATGCGAGCAGCGGTAATGCGCGCGGTCA
 GCGGTGGTGAATATCTCCAACGCGGCATTCCCGATTCTGATGGCGGTAATGATAAAAACTACTGGCTGGCCTTCG
 CGGAAAACGCGCTGGGATAAAAAAGCTGGCGTACATTACGGAAGCGCTTCCATTGTGCGCCCGGAAAACGTG
 ACTCGCGAATCTGCGAACCTTCAACCTGCGGTTTATCTCGCTGGGCAAGTCGCGCAAGTGAAGTGAAGTTATCGG
 TAACCCGCACTACAACAGCATTTCTGCGCTGTCCGAACGGTTACAGCTGGGAAGGCGGTGTGATAAAAAACGGTCAGT
 GTACGCGTAACAGTGATTCTAATGATATGAAGCACTTTATGCAGAACGTGTTGCGCTATCTGTCCAACGATAAATGG

ACGCCGACGCGAAAGCCAGCATGACCGTAGGCACCAACCTGGATACTGTCTATTTCAAACGTGATGGTCAGGTTAC
 AGGAAACAGTGCTGAGTTCCGCTTCCATCCGGATTTTTCGCGGTATCTCCGTTGAGCATTAAAGTAGCTATGGCGATC
 TCGATCCGAGAAAATGCCGCTGCTGATCCTTAACGGCTTTGAGTATGTGACTCAGGTGGGGGGCGATCCCTATGCA
 GTGCCCTCTGCTGCAGATACCAGCAAACCGAAGCTGACCCAGCAGGATGTGACCGATTGATCGCCTATATGAACAA
 AGGTGGCTCGGTGCTGATCATGGAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGCGCTTGGCTTGTGCGTCTGC
 TGGATGCCGCGAGGCTGTCAATGGCACTGAACAAGTCGGTGGTGAATAACGATCCGCGAGGTTATCCAGACCGCGTG
 CGCCAGCTCGAGCGACTGGCATTTGGGTTTATGAACGTATCCGTTGTGGAGGTTGAGCTGCCGTACACCATTGA
 TTCCAAAACAGGAAAGTTACCTGGAAATATCAGATTGATAACAAGCCTGATAAGAAACCGAAACGGAAGTAGCCA
 GCTGGCAAGAAGAGTTGATGGTAAACAGGTAACCTCAATTCGCCCTTATCGATGAAGCCGACACAAAACGACTGAA
 TCTCTEGATGCTGCGAAGAAGAAAATTTCTGAAAAAATCAAGGGATTAGAGGAGTGAAGGACTCGCATATCACTA
 CGAAATCAACTGCCGGAATACCGTCTTGGCAGCAACGTTCCGGTTACTGGTGGCATGTATGTTCCGCGCTATACGC
 AACTGAATCTCAGCGCCGACACTGCAAAAGCGATGGTGCAGGCGCGGATTTAGGCACCAACATTACGCGTCTGTAT
 CAGCATGAGCTCTACTTCCGGACCAATGGTCGCAAGGTGAGCGCTGAGCAGCGTCGATCTGGAACGTCTGTACCA
 GAACAIGTCGGTCTGGCTGTGGAATGAAATGAATACAGCTATGACAGCAGCAAGGAAGACGAGCTGGGCTTTAAAA
 CGTTACCCGAGTTCTTGAACCTGCTACGCCAACGATGCCCTATACCTGGCGGCACGCAAGTGTCTGATGAGCTGAAAAA
 TCGCTGGTCGATAACAACATGATCTACGCCGAGAAGAGCGTTAATAAAGCGGCGATGATGAACCCAGCTATCCGCT
 CAACTATATGAAAAACCGCTGACGCGCTGATGCTGGTCTGTTCTTGGTGGGATCTGAACATCAAAGTTGATGTCG
 AGAAGTATCCGGGAGCGGTATCGGCAGAAAGGTGAGAAGGTACTGAAACCATCAGCCTGACTCGAATCCGACCAAA
 TGGTTTTCGCGGTACATGCAGTCTACTGGCCTATGGGCTCCGGCTCAGAAAGAGGTCAACATTGAGTCTTCTGCATC
 AGTTCCCTGTGACTGTACCGTGGCGCTGGCCGACGATCTGACCGGACGCGAGAAGCATGAAGTCGCGCTGAACCGTC
 CGCCAAAAGTGACGAAAACCTATGATCTGAAAGCCAATGATAAGGTGACCTTCAAGGTGCCTTACGGTGGTCTGATT
 TATATCAACGCCAACAGCCCAAGAAATGAGTCAGCCGAATTCACCTTTACTGGTGTGGTAAAAGCGCGCTTCTATAA
 AGACGGCGAATGAAAAACGCTCTGAATCCCTGCGCCGCTGGGCGAGCTGGAATCAGACGCTTTCGTTTATACCA
 CGCCGAAGAAGAACCTTGAGGCCAGCAATTTCACTGGTGGTGTAGCAGAATTCGCTAAAGATCTGGATACCTTTGCC
 AGCTCGATGAATGACTTCTACGGTCGTAATGATGAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGACGGG
 GCACAAGCATCGTTTACCAACGATGTGCAGATCTCATCGGTGATGCGCACTCTGGTTATCCGGTAATGAACAGCA
 GCTTCTCGACGAACAGCACCAGCTGCCGACGACGCGCTGAACGACTGGCTGATCTGGCAGCAAGTTGGTCACAAC
 GCCGCCGAAACGCCGCTGACTGTTCTTGGCGCAACTGAAGTAGCGAACAACGTGCTGGCGCTGTACATGCAGGATCG
 CTATCTCGGCAAGATGAACCGTGTGCTGACGACATTACCGTCGCGCCGGAATATCTGGAGGAGAGCAACCGTCAAG
 CATGGGCGCGCGGCTGCGGGTGACCGTCTGCTGATGACGACAGCTGAAAGAGTGGGCGAGAGAAAACTTTGAT
 ATCAAGACGTGGTATCCAGATGGCAATCTGCCAGCGTTCTACAGCGAGCGTGAAGGGATGAAAGGCTGGAACCTGTT
 CCAGTTGATGCATCGTAAAGCACGCGCGGATGAGGTTGGCAAAACCAAGTTTGGCGAAAGAAATTAATGTGCCGAAT
 CCAACGTAACGCTGCCGACAGCTGATGCTGTGTGATCCTGGGTGCGCCAGACGGATCTTTTCGGAGTTCTTTAAG
 AAATGGAATCCGGGCGCAATGCTTACCAGTTGCCGGGAGCGAGCGAGATGAACTTCGAGGGCGGTGTGAGCCAGTC
 GGCTTACGAGACGCTGGCGGCTTAATCTGCCGAACCGCAGCAAGGGCCGGAACCATTAATCAGGTTACCGAGC
 ATAAGATGTCAGCTGAG

SEQ ID NO: 141

MNKKFKYKKSLLAAILSATLLAGCDGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGYLTLGGSQRTVGATCNGESSDGFFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRAEFKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTPDLNASFVSANAEQFYQYQPTTEIILSEGQLVDSLGNVAGVDYTTNSGRGVTDENGKFSFSWGCTISFGIDTFELGSVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHFVHDSTNFGYSTGNARGQAVVNISNSAFFILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIQPVENVTRDTATFNLFFISLGQVGEGLMVGIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGSDSDMKHFMQNVRLRYLSNDIWPNTKSIIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLDAAGLSMALNKSVMNDPQGYPDVRQRATGIWVYERYPAADGAQPPYITIDPNTGEVTWKYQQDNKPKDDPKLEVASWQEEVEGKQVTRYAFIDEAEYTTESLEAAKAKIFEKFPGLQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLADTAKAMVQAADLGTNIQRLYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKFTFEFLNCYANDAYAGGTKCSADLKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGSVSAGKESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGWKNLDSAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEGDKHRMFTYKNLTGKHHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTTPLNDWLIWHEVGHNAEETPLNVPGATEVANNVLALYMQDRYLKGMNRVADDITVAPEYLDENSGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKNWPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLKPKPEKGPETINKVTEHKMSAE

SEQ ID NO: 142

DTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGYLTLGGSQRTVGATCNGESSDGFFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRAEFKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTPDLNASFVSANAEQFYQYQPTTEIILSEGQLVDSLGNVAGVDYTTNSGRGVTDENGKFSFSWGCTISFGIDTFELGSVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHFVHDSTNFGYSTGNARGQAVVNISNSAFFILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIQPVENVTRDTATFNLFFISLGQVGEGLMVGIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGSDSDMKHFMQNVRLRYLSNDIWPNTKSIIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLDAAGLSMALNKSVMNDPQGYPDVRQRATGIWVYERYPAADGAQPPYITIDPNTGEVTWKYQQDNKPKDDPKLEVASWQEEVEGKQVTRYAFIDEAEYTTESLEAAKAKIFEKFPGLQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLADTAKAMVQAADLGTNIQRLYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKFTFEFLNCYANDAYAGGTKCSADLKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGSVSAGKESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGWKNLDSAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEGDKHRMFTYKNLTGKHHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTTPLNDWLIWHEVGHNAEETPLNVPGATEVANNVLALYMQDRYLKGMNRVADDITVAPEYLDENSGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKNWPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLKPKPEKGPETINKVTEHKMSAE

LCLTFSSVVDRAEFKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTPDLNASFVSANAEQFYQYQPTTEIILSEGQLVDSLGNVAGVDYTTNSGRGVTDENGKFSFSWGCTISFGIDTFELGSVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHFVHDSTNFGYSTGNARGQAVVNISNSAFFILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIQPVENVTRDTATFNLFFISLGQVGEGLMVGIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGSDSDMKHFMQNVRLRYLSNDIWPNTKSIIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLDAAGLSMALNKSVMNDPQGYPDVRQRATGIWVYERYPAADGAQPPYITIDPNTGEVTWKYQQDNKPKDDPKLEVASWQEEVEGKQVTRYAFIDEAEYTTESLEAAKAKIFEKFPGLQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLADTAKAMVQAADLGTNIQRLYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKFTFEFLNCYANDAYAGGTKCSADLKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGSVSAGKESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGWKNLDSAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEGDKHRMFTYKNLTGKHHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTTPLNDWLIWHEVGHNAEETPLNVPGATEVANNVLALYMQDRYLKGMNRVADDITVAPEYLDENSGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKNWPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLKPKPEKGPETINKVTEHKMSAE

SEQ ID NO: 143

KTGYLTLGGSQRTVGATCNGESSDGFFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRAEFKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTPDLNASFVSANAEQFYQYQPTTEIILSEGQLVDSLGNVAGVDYTTNSGRGVTDENGKFSFSWGCTISFGIDTFELGSVRGNKSTIALTELGDVEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAYEPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHFVHDSTNFGYSTGNARGQAVVNISNSAFFILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIQPVENVTRDTATFNLFFISLGQVGEGLMVGIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGSDSDMKHFMQNVRLRYLSNDIWPNTKSIIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLDAAGLSMALNKSVMNDPQGYPDVRQRATGIWVYERYPAADGAQPPYITIDPNTGEVTWKYQQDNKPKDDPKLEVASWQEEVEGKQVTRYAFIDEAEYTTESLEAAKAKIFEKFPGLQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLADTAKAMVQAADLGTNIQRLYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKFTFEFLNCYANDAYAGGTKCSADLKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRMLGRSWWDLNLIKVDVEKYPGSVSAGKESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGWKNLDSAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEGDKHRMFTYKNLTGKHHRFTNDVQISIGDAHSGYPVMNSSFSTNSTLPTTPLNDWLIWHEVGHNAEETPLNVPGATEVANNVLALYMQDRYLKGMNRVADDITVAPEYLDENSGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKNWPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLKPKPEKGPETINKVTEHKMSAE

SEQ ID NO: 144

ATGAATAAGAAATTTAAATATAAGAAATCGCTTTTAGCGGCTATTTTAAGCGCAACCCTGTTAGCCGGTTGTGATGG
 TGGTGGTTTCAGGATCGTCTCCGATACGCCGCTCTGTAGATTCTGGATCAGGGACTTTGCCGGAAGTGAAACCCGATC
 CAACACCAACCCCGGAGCCGACACCTGAGCCGACGCCGAGCCAGAACCTACGCCGGATCCAACACCTGATCCTGAG
 CCGACACCAGAACCCGAGCCAGAACCTGTTCTACGAAAACGGGTTATCTGACCCCTGGGCGGAAGCCAGCGGGTAAC
 TGGTGTCTACCTGTAATGGTGAATCCAGCGATGGCTTTACCTTTACGCCAGGCAATACCGTGAGTTGTGTGGTGGGCA
 GTACGACCATTTGCAACATTCAACACCCAGTCAGAAGCTGCGCGTAGCCTGCGTGCGGTTGACAAAGTGTCTGTTTACC
 CTGGAGGACGCGCAGGAGCTGGCGAATTCTGAAAATAAGAAAACCAACGCCATCTCTCTGGTGACGTCCAGCGACAG
 TTGCCCGCAGATGCAGAACAGCTTTGTCTTACTTTCTCGTCAGTGTTGATCGCGCGGATTTGAAAACTGTATA
 AGCAAATTGATCTGGCAACAGACAATTTTCAGCAAGCTGGTCAATGAAGAGGTGGAAAAACAATGCTGCGACTGATAAA
 GCGCCGTCACCCATACCTCAACGGTAGTGCCAGTCACGACAGAGGGAACAAAACCGGATCTGAACGCGTCTCTCGT
 GTCGGCTAACGCGGAACAGTTTATCAGTATCAACCCACTGAAATCATTCTTTCCGAAGGCCAAGCTGGTGATAGCC
 TGGGGAACGGTGTGTCTGGCGTTGACTACTACCCAATTCAGGCCGTGGCGTAAGTACGAGAAACGGTAAATTTTCC
 TTTAGCTGGGGCGAAACCATCTCTCTTTGGTATCGATACCTTTGAACTGGGCTCAGTACGTGGCAATAAGTCGACCAT
 TGCGCTGACTGAATTGGGTGATGAAGTTCGCGGGGCAAAATATCGATCAGCTCATTTCATGTTATTCGACGACTGGTC
 AAAATAATACTCGTGTGTGTTCCGGACGATGTACGCAAGGTCTTTGCCGAATATCCCAACGTGATCAACGAGATAATC
 AATCTTTTCGTTATCCAACGGTGGCAGCTGGATGAAGCGCATCAAAACGTTGTGCTGCCCTAACGAATTTATCGAGCA
 GTTTTAAGACGGGTACGGCCAAAGAGATCGATACCGCGATTTGTGCGAAAACCGACGGTTGTAACGAGGCTCGCTGGT
 TCTCGCTGACAACGCGCAATGTTAATGACGGCCAGATTTCAGGGCGTTATTAACAAGCTGTGGGGCGTGGATACGAAC
 TATCAGTCTGTGACGAAGTTCACAGTCTTCCATGACTCTACCAACTTCTATGGCAGTACCGGTAACGCGCGCGGTCA
 GCGGTGGTAAATATCTCCAACCTCGGCATTCCCGATTCTGATGGCGCGTAATGATAAAAACTACTGGCTGGCGTTTG
 CGAAAAACCGCCTGGGATAAAAAATGAGCTGGCGTACATTACGGAAGCGCCTTCCATTGTGCAGCCAGAGAATGTT

ACGCGCGATACTGCGACTTTCAACCTGCCGTTTATTTTCGCTGGGGCAAGTCGGTCAAGCCAAACTGATCCTTATCGG
 TAACCCCGCTACAAACAGCATCCTGCCGTTGCCGGAACGGTTACAGTTGGGGCGGTGGTGTAAATAGTAAAGGTGAGT
 GTACGCTCAGCGGTGATTCTGATGACATGAAGCACTTTATGCAGAACGTACTGCGCTACTTGTCAATGACATCTGG
 CAGCCAAATACCAAGAGCATCATGACTGTGCGCACCAACCTGGAGAACGTTTATTTCAAAAAAGCGGGCCAGGTATT
 GGGAAATAGTGCACCATTTGCTTTCCATGAGGATTTCACTGGTATCACGGTTAAACAGTTGACCAGCTATGGCGATC
 TGAATCCGGAAGAGATTCCGTTGCTGATCCTCAACGGCTTTGAATATGTGACTCAGTGGTCTGGCGATCCCTATGCT
 GTGCTCTGCGTGACATACAGCAAAACCGAAGCTGACTCAGCAGGATGTGACCGATCTGATCGCTTATCTGAACAA
 AGGTGGCTCGGTGCTGATCATGGAACGCTGATGAGCAATCTTAAGGAAGAGAGCGGTCAGTTTGTGCGTCTGCT
 TGGATGCCGCGGGTCTGTCAATGGCTCTGAACAAATCGGTGTTGAACAACGATCCGCAAGGGTATCCGGATCCGCTT
 CGTCAGCGTCGCGGACTGGCATTGGGTTTATGAACGTTATCCTGCTGCAGACGCGCGCAACCGCGTACACCAT
 CGACCCAAATACAGGGGAAGTGACCTGGAAATACCAGCAAGACAAACAGCTGATGACAAGCCGAAACTGGAAGTTG
 CGAGCTGGCAGGAGGAAGTTGAGGGCAACAGGTAACGCGTTATGCCCTTATGATGAAGCGGAATACACAACAGAA
 GAATCTCTGGGAAGCGCAAAAGGCAAAATCTTTGAGAAGTTTCTGGGTTACAGGAGTGTAAAGGACTCGACTTTACCA
 TTACGAGATTAAGTGTGAGGCGCGCCAGGACCGGATGTTCCGGTAACAGGTGGCATGTATGTTCCGCGCTATA
 CGCAACTGAATCTTGACGCGGACACCGCGAAAGCGATGGTGCAGGCGCGGATTTAGGCACCAACATTACCGCGCTG
 TATCAGCATGAGCTTTATTTCCGTACCAAGGCAGTAAGGTGAGCGTCTGAACAGTGTGATCTGGAACGCTGTGTA
 CCAGAACAATGTCGGTCTGGCTGTGGAACGATACGAAATATCGTTACGAAGAGGGCAAGGAAGATGAGCTGGGCTTTA
 AAACGTTTCCAGGATCTCTGAACTGCTACGCCAATGATGCTATGCAGGCGGCAACCAAGTGTCTCGGAGATCTGAAA
 AAATCGCTGGTCTGATAACAACATGATCTACGGTGACGGTAGCAGCAAGCGGGCATGATGAACCAAGCTATCCGCT
 CAATATATGGAACAAACCGCTGACGCGTCTGATGCTGGGCGTTCTGTTGGGATCTGAACATTAAGGTTGATGTGG
 AGAAGTACCAGGATCCGATACGGCAAGGGTGAGAGCGTTACGGAACCAATCAGCCTGTACTCGAATCCGACCAAA
 TGGTTTTCGGGTAACATGCAGTCAACCGGCTGTGGGCACCGGCCAGCAGGACGTCAACCATTAAGTCTTCCGCGCTC
 AGTCCCAGTGACTGTTACCGTGGCGCTGGCTGACGACCTGACTGGACGTGAGAAGCATGAAGTTGCGCTGAACCGTCTC
 CGCCAAGATGACTAAAACGTATCTCTGGAGGCTAACGGTGAAGTGACCTTCAAGGTGCCCTTATGCTGCTCTGATT
 TATATCAAGGGCGACAGTAAGGATGATGTTTCTGCTAACTTCACTTTACCGGTGTAGTAAAGCGCGCTTCTATAA
 AGACGCGAATGGAAAAACGATCTGGACTCACCGCGCGGCTTGGCCACCTGGAGTCTGCGTCTGCTCTATACCA
 CCGCGAAGAACCTTGAGGCCAGCAATTTCACTGGTGGTGTAGCAGAAATTCGCTAAAGATCTGGATACCTTTGCC
 AGCTCGATGAATGACTTCTACGGTCCATATGATCAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGACGGG
 GCACAAGCATCGTTTACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGTTATCCGGTAATGAACAGCA
 GCTTCTCGACGAACAGCACCGCTGCCGACGACCGGCTGAACGACTGGCTGATTGGCACGAAGTCCGTCATAAC
 GCTGCAGAAACACCGCTGAACGTACCGGGTGCAACTGAAGTGGCGAACAACGCTGCTGGCGCTGTACATGCAGGATCG
 CTATCTCGGTAAGATGAACCGTGTGCTGACGACATTACCGTCCGCGCGGAATATCTGGACGAGAGCAACGGTCAGG
 CCTGGGCGCGCGCGGTGCGGCTGACCGTCTGCTGATGTACGCACAGTTGAAGGAGTGGGCGAGAGGAAAACCTTTGAT
 ATCAACAGTGGTATCCAGATGGTGAAGTCTGCTAAGTCTACAGCGATCGTAAAGGGATGAAGGGCTGGAACCTGTT
 CCAGTTGATGCACCGTAAAGCGCGCGCGATGATGTTGGTAAACGACCTTTGGTGGCAAGAATTACTGTGCTGAAT
 CCAATGGTAACGCTGCCGACACGCTGATGCTGTGTGCATCCTGGGTCGCTCAGGCGGATCTTTCCGGAATCTTTAAG
 AAATGGAATCCGGGTGCAAGTGTACAGTTGCCGGGAGCAACGGAGATGAGTTTCCAGGGCGGTGTGAGCTCTTC
 GGCTTACAGCACGCTGGCGTCACTCAAGCTGCCGAAACCGGAAAAAGGGCCGGAACCATTAACAAGGTTACCGGAG
 ATAAGATGTCTGCCGAG

REIVINDICACIONES

1. Un polipéptido que comprende una secuencia de aminoácidos que:

- 5 (a) es idéntica a una cualquiera de las SEQ ID NOs: 5 a 16;
 (b) tiene de 1 a 10 alteraciones de un solo aminoácido en comparación con (a);
 (c) tiene al menos un 95 % de identidad de secuencia con una cualquiera de las SEQ ID NOs: 5 a 16; y/o
 (d) es un fragmento de al menos 250 aminoácidos consecutivos de cualquiera de las SEQ ID NOs: 5 a 16; o
 (e) tiene al menos un 95 % de identidad con cualquiera de las SEQ ID NOs: 5 a 16 y comprende un fragmento de al menos 10 aminoácidos consecutivos de cualquiera de las SEQ ID NOs: 5 a 16;

con la condición de que dicha secuencia de aminoácidos no sea la SEQ ID NO: 2.

10 2. Un polipéptido inmunógeno que comprende un fragmento de una proteína AcfD de *E. coli* en el que el fragmento contiene una delección respecto a la proteína AcfD de *E. coli* que aumenta la solubilidad del fragmento en comparación con la proteína de longitud completa y en el que el fragmento provoca una respuesta inmunitaria en un sujeto sustancialmente similar a la de la proteína AcfD de *E. coli*.

3. El polipéptido inmunógeno de la reivindicación 2, en el que

- 15 (a) la proteína AcfD de *E. coli* tiene una secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NOs: 2-16; y/o
 (b) la delección es la eliminación de sustancialmente todos los aminoácidos N terminales hasta el enlazador gly-ser, la eliminación de la totalidad o una parte de la repetición rica en prolina N terminal, o ambas; y/o

el fragmento de polipéptido inmunógeno comprende una secuencia de aminoácidos que comprende:

- 20 (c) la secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOs 99 a 128;
 (d) de 1 a 10 alteraciones de un solo aminoácido en comparación con las SEQ ID NOs: 99 a 128; y/o
 (e) al menos un 90 % de identidad de secuencia con una cualquiera de las SEQ ID NOs: 99 a 128.

4. Un polipéptido inmunógeno de acuerdo con la reivindicación 3, en el que

- 25 (a) dicha proteína AcfD de *E. coli* tiene una secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NOs: 129, 133 y 137, y la delección es opcionalmente la eliminación de todos los aminoácidos N terminales hasta el enlazador gly-ser, la eliminación de toda o una parte de la repetición rica en prolina N terminal, o ambas; y/o

el fragmento de polipéptido inmunógeno comprende una secuencia de aminoácidos que comprende:

- 30 (b) la secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NOs 130, 131, 134, 135, 138 y 139;
 (c) de 1 a 10 alteraciones de un solo aminoácido en comparación con las SEQ ID NOs: 130, 131, 134, 135, 138 y 139;
 (d) al menos un 92 % de identidad de secuencia con una cualquiera de las SEQ ID NOs: 130, 131, 134, 135, 138 y 139 y un fragmento de al menos 10 aminoácidos consecutivos de cualquiera de las SEQ ID NOs: 130, 131, 134, 135, 138 y 139.

5. El polipéptido de cualquier reivindicación anterior, que incluye un ion Zn^{2+} .

6. El polipéptido de cualquier reivindicación anterior, en el que dicho polipéptido es aislado, purificado o recombinante.

7. Un polinucleótido que codifica el polipéptido de cualquiera de las reivindicaciones 1 a 4.

40 8. Un kit que comprende uno o más polipéptidos o polinucleótidos de acuerdo con cualquiera de las reivindicaciones anteriores, opcionalmente en el que uno o más componentes están liofilizados.

9. Un vector que comprende el polinucleótido de la reivindicación 7, opcionalmente en el que el vector es un plásmido.

45 10. Una célula de *E. coli*, que contiene un plásmido que codifica el polipéptido de cualquiera de las reivindicaciones 1 a 4.

11. Un procedimiento de producción de un polipéptido de acuerdo con cualquiera de las reivindicaciones 1 a 4, que comprende cultivar una célula hospedadora transformada con un polinucleótido de acuerdo con la reivindicación 7 o el vector de acuerdo con la reivindicación 8, en condiciones que inducen la expresión del polipéptido.

50 12. Una composición inmunogénica que comprende un polipéptido de acuerdo con cualquiera de las reivindicaciones 1-6, que comprende opcionalmente uno o más vehículos, excipientes y/o adyuvantes farmacéuticos.

13. Un polipéptido que comprende una secuencia de aminoácidos que:
- (a) es idéntica a una cualquiera de las SEQ ID NOs: 3 a 16;
 - (b) tiene de 1 a 10 alteraciones de un solo aminoácido en comparación con (a);
 - (c) tiene al menos 85 % de identidad de secuencia con una cualquiera de las SEQ ID NOs: 3 a 16;
 - (d) es un fragmento de al menos 10 aminoácidos consecutivos de cualquiera de las SEQ ID NOs: 3 a 16;
 - y/o
 - (e) cuando se alinea con cualquiera de las SEQ ID NOs: 3 a 16 usando un algoritmo de alineación por pares, cada ventana móvil de x aminoácidos desde el extremo N hasta el extremo C tiene al menos $x \cdot y$ aminoácidos alineados idénticos, donde x es 30 e y es 0,75;
- 5 para su uso como un medicamento.
- 10 14. El polipéptido para su uso de acuerdo con la reivindicación 13 para su uso en la provocación de una respuesta inmunitaria para proteger a un mamífero contra la infección intestinal por *E. coli*.
- 15 15. El polipéptido para su uso de acuerdo con la reivindicación 13 o la reivindicación 14, que incluye un ion Zn^{2+} .
- 15 16. El polipéptido para su uso de acuerdo con una cualquiera de las reivindicaciones 13 a 15, en el que el polipéptido proporciona protección contra *E. coli* patógena, incluyendo patotipos intestinales como ETEC, EPEC, EAEC, EIEC y DAEC.
- 20 17. El polipéptido para su uso de acuerdo con una cualquiera de las reivindicaciones 13 a 16, en el que el polipéptido es para proteger contra una enfermedad seleccionada de la lista: peritonitis, diarrea, por ejemplo diarrea infantil, diarrea del viajero, aguda, persistente, disentería bacilar y síndrome urémico hemolítico (HUS).
- 20 18. El polipéptido para su uso de acuerdo con una cualquiera de las reivindicaciones 13 a 17 en el que el mamífero es un ser humano.
- 25 19. El uso de un polipéptido en la fabricación de un medicamento para provocar una respuesta inmunitaria para proteger a un mamífero contra la infección intestinal por *E. coli*, en el que el polipéptido comprende una secuencia de aminoácidos que:
- (a) es idéntica a una cualquiera de las SEQ ID NOs: 3 a 16;
 - (b) tiene de 1 a 10 alteraciones de un solo aminoácido en comparación con (a);
 - (c) tiene al menos 85 % de identidad de secuencia con una cualquiera de las SEQ ID NOs: 3 a 16;
 - (d) es un fragmento de al menos 10 aminoácidos consecutivos de cualquiera de las SEQ ID NOs: 3 a 16;
 - y/o
 - (e) cuando se alinea con cualquiera de las SEQ ID NOs: 3 a 16 usando un algoritmo de alineación por pares, cada ventana móvil de x aminoácidos desde el extremo N hasta el extremo C tiene al menos $x \cdot y$ aminoácidos alineados idénticos, donde x es 30 e y es 0,75.
- 30

FIGURA 1

	10	20	30	40	50	60	70
SEQID3	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID4	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID12	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID16	MNKKFKYKKSLLAAILSATLLAGCDGGGSGPSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID5	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID14	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID6	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID8	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID9	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID11	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID7	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID13	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID15	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
SEQID2	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPSVDSGSGLPEVKPDPTTPEPTPEPTPDPEPTP						
SEQID10	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPSVDSGSGLPEVKPDPTTPEPTPEPTPDPEPTP						

SEC. CONS.	MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPPVDSGTGSLPEVKPDPTPNPEPTPEPTPDPEPTP						
REG N-TERM	GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP						

```

      80          90          100          110          120          130          140
      |           |           |           |           |           |           |
SEQID3  EPTPDP--EPTPEPEPEPVPTKTGYLTLGGSQRITG-ATCNGESSDGFFTPGDKVTCVAGNNTTIATFD
SEQID4  EPTPDP--EPTPEPEPEPVPTKTGYLTLGGSQRITG-ATCNGESSDGFFTPGDKVTCVAGNNTTIATFD
SEQID12 EPTPDP--EPTPEPEPEPVPTKTGYLTLGGSQRITG-ATCNGESSDGFFTPGDKVTCVAGNNTTIATFD
SEQID16 DPTPDP--EPTPEPEPEPVPTKTGYLTLGGSQRITG-ATCNGESSDGFFTPGDKVTCVAGNNTTIATFD
SEQID5  EPIPD--EPTPEPEPEPVPTKTGYLTLGGSQRVTG-ATCNGESSDGFTEKPGEDVTCVAG-NTTIATFN
SEQID14 EPIPD--EPTPEPEPEPVPTKTGYLTLGGSQRVTG-ATCNGESSDGFTEKPGEDVTCVAG-NTTIATFN
SEQID6  EPTPDP--EPTPEPEPEPVPTKTGYLTLGGSLRVTGDI TCNDESSDGFFTPGDKVTCVAGNNTTIATFD
SEQID8  EPTPDP--EPTPEPEPEPVPTKTGYLTLGGSLRVTGDI TCNDESSDGFFTPGDKVTCVAGNNTTIATFD
SEQID9  EPTPDP--EPTPEPEPEPVPTKTGYLTLGGSLRVTGDI TCNDESSDGFFTPGDKVTCVAGNNTTIATFD
SEQID11 EPIPD--EPTPEPEPEPVPTKTGYLTLGGSQRVTG-ATCNGESSDGFTEKPGEDVTCVAG-NTTIATFN
SEQID7  EPTPDP--EPTPEPEPEPVPTKTGYLTLGGSLRVTGDI TCNDESSDGFFTPGDKVTCVAGNNTTIATFD
SEQID13 EPIPD--EPTPEPEPEPVPTKTGYLTLGGSQRVTG-ATCNGESSDGFTEKPGEDVTCVAG-NTTIATFN
SEQID15 EPIPD--EPTPEPEPEPVPTKTGYLTLGGSQRVTG-ATCNGESSDGFTEKPGEDVTCVAG-NTTIATFN
SEQID2  DPTPDP--EPTPEPEPEPVPTKTGYLTLGGSQRVTG-ATCNGESSDGFFTPGNTVSCVVG-STTIATFN
SEQID10 DPTPDPDEPTPEPEPEPVPTKTGYLTLGGSQRVTG-ATCNGESSDGFFTPGNTVSCVVG-STTIATFN
:* *** ***** *:*** *.*****.***: *:*.* .*****:
SEC. CONS. EPTPDPDEPTPEPEPEPVPTKTGYLTLGGSQRVTGDATCNGESSDGFFTPGDKVTCVAGNNTTIATFD
REG N-TERM PPPPPPPPPPPPPPPPPPPPP

```

	150	160	170	180	190	200	210
SEQID3	TQSEAARSLRAVEKVSFSLEDAQELAASDDKKSSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYK						
SEQID4	TQSEAARSLRAVEKVSFSLEDAQELAASDDKKSSNAVSLVTSSNSCPADTEQVCLTFSSVIESKRFDLSLYK						
SEQID12	TQSEAARSLRAVEKVSFSLEDAQELAASDDKKSSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYK						
SEQID16	TQSEAARSLRAVEKVSFSLEDAQELAASDDKKSSNAVSLVTSSNSCPADTEQVCLTFSSVIESKRFDLSLYK						
SEQID5	TQSEAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYK						
SEQID14	TQSEAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYK						
SEQID6	TQSEAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSLVTSMNSCPANTEQVCLEFSSVIESKRFDLSLYK						
SEQID8	TQSEAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSLVTSMNSCPANTEQVCLEFSSVIESKRFDLSLYK						
SEQID9	TQSEAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSLVTSMNSCPANTEQVCLEFSSVIESKRFDLSLYK						
SEQID11	TQSEAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYK						
SEQID7	TQSEAARSLRAVEKVSFSLEDAQELAGSDNKKSNALSLVTSMNSCPANTEQVCLEFSSVIESKRFDLSLYK						
SEQID13	TQSEAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYK						
SEQID15	TQSEAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYK						
SEQID2	TQSEAARSLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEKLKY						
SEQID10	TQSEAARSLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEKLKY						
	*****:***** *:*****:***** :*****:k:*x *****: ***:***						
SEC. CONS.	TQSEAARSLRAVEKVSFSLEDAQELAGSDDKKSSNAVSLVTSSNSCPANTEQVCLTFSSVIESKRFDLSLYK						

	220	230	240	250	260	270	280
SEQID3	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID4	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID12	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID16	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID5	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID14	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID6	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID8	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID9	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID11	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID7	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID13	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID15	QIDLAPEEFKKLVNEEVENNAATDKAPSTHTSPVVPVTTT	PGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID2	QIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTT	EGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEQID10	QIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTT	EGTKPDLNASFVSANAEQFYQYQPTTEIILSE					
SEC. CONS.	*****. :*.*****.*****.***.*** *****.*****.*****						

	290	300	310	320	330	340	350
SEQID3	GRLVDSMGNVGVVNYTSSGRGVTGENGKFNFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID4	GRLVDSMGNVGVVNYTSSGRGVTGENGKFNFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID12	GRLVDSMGNVGVVNYTSSGRGVTGENGKFNFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID16	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID5	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID14	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID6	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID8	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID9	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID11	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID7	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID13	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID15	GRLVDSQGYGVAGVNYTNSGRGVTGENGFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID2	GQLVDSLNGVAGVDYTTNSGRGVTDENGKFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEQID10	GQLVDSLNGVAGVDYTTNSGRGVTDENGKFSFSWGETISF	GIDTFELGSVRGNKSTIALTELGDVEVRGA					
SEC. CONS.	*:**** * *.*:***.*****.***:*.*****:*****.*****						

	360	370	380	390	400	410	420
SEQID3	NIDQLIHRYSQAGKNDEREVPDVVRKVFAYPNVINEIINL	SLNGEALSEGDTFERTNEFLEQFESGQ					
SEQID4	NIDQLIHRYSQAGKNDEREVPDVVRKVFAYPNVINEIINL	SLNGEALSEGDTFERTNEFLEQFESGQ					
SEQID12	NIDQLIHRYSQAGKNDEREVPDVVRKVFAYPNVINEIINL	SLNGEALSEGDTFERTNEFLEQFESGQ					
SEQID16	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGDQNVVLPNEFIEQFKTGQ					
SEQID5	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFNTGQ					
SEQID14	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFNTGQ					
SEQID6	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFKTGQ					
SEQID8	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFKTGQ					
SEQID9	NIDQLIHRYSKAGQNNHTRVVPDEVVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFKTGQ					
SEQID11	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFNTGQ					
SEQID7	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFKTGQ					
SEQID13	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFNTGQ					
SEQID15	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGEQVVNLPNEFIEQFNTGQ					
SEQID2	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGDQNVVLPNEFIEQFKTGQ					
SEQID10	NIDQLIHRYSTTGQNNTRVVPDDVRKVFAYPNVINEIINL	SLNGATLDEGDQNVVLPNEFIEQFKTGQ					
SEC. CONS.	***** :*:*. * **:***** ***** :*.***:*. .***:***:***						

	430	440	450	460	470	480	490
SEQID3	AKEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDDKDYKSVTKFHFVHDS	TN	FN	Y	GS	T	GNAR
SEQID4	AKEIDTAICDSLGGCNSQRWFSLTARNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID12	AKEIDTAICDSLGGCNSQRWFSLTARNVNEGQIQGVINKLWGVDDKDYKSVTKFHFVHDS	TN	FN	Y	GS	T	GNAR
SEQID16	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDDKDYKSVTKFHFVHDS	TN	FN	Y	GS	T	GNAR
SEQID5	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID14	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID6	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID8	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID9	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID11	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID7	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID13	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID15	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID2	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
SEQID10	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN
	*****	.	***.	*****	****:	*****	*****
SEC. CONS.	AKEIDTAICAKTDGCEARWFSLTTRNVNDGQIQGVINKLWGVDTN	YKS	VS	KF	HV	FD	STN

	500	510	520	530	540	550	560
SEQID3	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVT	RDTATFNL	PFI	SLGQV			
SEQID4	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVT	RDTATFNL	PFI	SLGQV			
SEQID12	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVT	RDTATFNL	PFI	SLGQV			
SEQID16	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSI	VQPENVT	RDTATFNL	PFI	SLGQV		
SEQID5	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVT	RDTATFNL	PFI	SLGQV			
SEQID14	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVT	RDTATFNL	PFI	SLGQV			
SEQID6	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVT	RDTATFNL	PFI	SLGQV			
SEQID8	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSLVEPENVT	RDTATFNL	PFI	SLGQV			
SEQID9	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSVIRP	PENVTR	ETATFNL	PFI	SLGQV		
SEQID11	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSVIRP	PENVTR	ETASFNL	PFI	SLGQV		
SEQID7	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNDLAYITEAPSVIRP	ENVTR	ETATFNL	PFI	SLGQV		
SEQID13	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSVIRP	ENVTR	ETATFNL	PFI	SLGQV		
SEQID15	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSVIRP	ENVTR	ETATFNL	PFI	SLGQV		
SEQID2	GQAVVNISNSA	FILMARNDKNYWLAFGEKRAWDKNELAYITEAPSI	VQPE	NVT	RDTATFNL	PFI	SLGQV
SEQID10	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSI	VQPE	NVT	RDTATFNL	PFI	SLGQV	
	*****	*****	*****	*****	*****	*****	*****
SEC. CONS.	GQAVVNISNAAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSI	VQPE	NVT	RDTATFNL	PFI	SLGQV	

```

          570          580          590          600          610          620          630
          |           |           |           |           |           |           |
SEQID3   GEGKLMVIGNPHYNSILRCPNGYSWEGGVDKNGQCTRNSDSNDMKHFMQNVRLRYSDDKWTDPDAKASMTV
SEQID4   GEGKLMVIGNPHYNSILRCPNGYSWEGGVDKNGQCTRNSDSNDMKHFMQNVRLRYSDDKWTDPDAKASMTV
SEQID12  GEGKLMVIGNPHYNSILRCPNGYSWEGGVDKNGQCTRNSDSNDMKHFMQNVRLRYSDDKWTDPDAKASMTV
SEQID16  GEGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSGSDDMKHFMQNVRLRYSDDKWTDPDAKASMTV
SEQID5   GEGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSDDKWKPDAKASMTV
SEQID14  GEGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSDDKWKPDAKASMTV
SEQID6   GEGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSDDKWTDPDAKASMTV
SEQID8   GEGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSDDKWTDPDAKASMTV
SEQID9   GDGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSDDKWTDPDAKASMTV
SEQID11  GDGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSNDRWLDPDAKSSMTV
SEQID7   GDGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSNDRWLDPDAKSSMTV
SEQID13  GDGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSNDRWLDPDAKSNMTV
SEQID15  GDGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSNDRWLDPDAKSNMTV
SEQID2   GEGKLMVIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGSDDMKHFMQNVRLRYSNDIWPNTKSIIMTV
SEQID10  GDGKLMVIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGSDDMKHFMQNVRLRYSNDIWPNTKSIIMTV
*****
SEC. CONS. GEGKLMVIGNPHYNSILRCPNGYSWNGGVNKDGQCTLNSDPDDMKNFMENVRLRYSDDKWTDPDAKASMTV

```

	640	650	660	670	680	690	700
SEQID3	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID4	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID12	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID16	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID5	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID14	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID6	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID8	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID9	GTNLDTVYFKRHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID11	GTNLETVYFKKHGQVLGNSAPFAFHKDFGTITVKPMTSYGNLNPDEVPLLILNGFEYVTQWGSDPYSIPL						
SEQID7	GTNLDTVYFKKHGQVLGNSAPFAFHKDFGTITVKPMTSYGNLNPDEVPLLILNGFEYVTQWGSDPYSIPL						
SEQID13	GTNLDTVYFKKHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID15	GTNLDTVYFKKHGQVTGNSAEFGFHPDFAGISVEHLSSYGDLDPQEMPLLILNGFEYVTQVGNDPYAIPL						
SEQID2	GTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLPNPEIPLILNGFEYVTQWGSDPYAVPL						
SEQID10	GTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLPNPEIPLILNGFEYVTQWGSDPYAVPL						
SEC. CONS.	*****:*****: *** **** * ** ***:***: :****:***:***** ***** ..****:***						

	710	720	730	740	750	760	770
SEQID3	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID4	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID12	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID16	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID5	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID14	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID6	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID8	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID9	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID11	RADTSKPCLTQQDVTDLIAYMKNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID7	RADTSKPCLTQQDVTDLIAYMKNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID13	RADTSKPCLTQQDVTDLIAYMKNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID15	RADTSKPCLTQQDVTDLIAYMKNKGGSVLIMENVMSNLKEESASGFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID2	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLLDAAGLSMALNKSUVVNDPQGY						
SEQID10	RADTSKPCLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLLDAAGLSMALNKSUVVNDPQGY						
SEC. CONS.	*****:*****:*****:*****:*****:*****:*****:*****:*****:*****						

	780	790	800	810	820	830	840
SEQID3	PNRVRRQRATGIWYERYPAVDGALP-YTIDSKTGEVKWKYQVENKPDDBPKPLEVASWLEDVDGKQVTRY						
SEQID4	PNRVRRQRATGIWYERYPAVDGALP-YTIDSKTGEVKWKYQVENKPDDBPKPLEVASWLEDVDGKQVTRY						
SEQID12	PDRVRQRRATGIWYERYPAVDGELP-YTIDSKTGKVTWKYQIDNKPDKPKLEVASWQEEVDGKQVTF						
SEQID16	PNRVRRQRREKGIWYERYPAVDSAPPPYTDIDPDGKVTWKYQEEGKPDDBPKPLEVASWQEDVDGKQVTRY						
SEQID5	PNRVRRQRATGIWYERYPAVDGALP-YTIDSKTGEVKWKYQVENKPDDBPKPLEVASWLEDVDGKQVTRY						
SEQID14	PNRVRRQRATGIWYERYPAVDGALP-YTIDSKTGEVKWKYQVENKPDDBPKPLEVASWLEDVDGKQVTRY						
SEQID6	PNRVRRQRATGIWYERYPAVDGALP-YTIDSKTGEVKWKYQVENKPDDBPKPLEVASWLEDVDGKQVTRY						
SEQID8	PNRVRRQRATGIWYERYPAVDGALP-YTIDSKTGEVKWKYQVENKPDDBPKPLEVASWLEDVDGKQVTRY						
SEQID9	PNRVRRRRSTPIWYERYPAVDGKPP-YTIDDTTKEVIWKYQENKPDDBPKPLEVASWQEEVEGKQVTF						
SEQID11	PDRVRQRRSTPIWYERYPAVDGKPP-YTIDDTTKEVIWKYQENKPDDBPKPLEVASWQEEVEGKQVTF						
SEQID7	PDRVRQRRSTPIWYERYPAVDGKPP-YTIDDTTKEVIWKYQENKPDDBPKPLEVASWQEEVEGKQVTF						
SEQID13	PDRVRQRRSTPIWYERYPAVDGKPP-YTIDDTTKEVIWKYQENKPDDBPKPLEVASWQEEVEGKQVTF						
SEQID15	PDRVRQRRSTPIWYERYPAVDGKPP-YTIDDTTKEVIWKYQENKPDDBPKPLEVASWQEEVEGKQVTF						
SEQID2	PDRVRQRRATGIWYERYPAADGAAPPPYTDIDPNTGEVTWKYQDNKPDDBPKPLEVASWQEEVEGKQVTRY						
SEQID10	PDRVRQRRATGIWYERYPAADGAAPPPYTDIDPNTGEVTWKYQDNKPDDBPKPLEVASWQEEVEGKQVTRY						
SEC. CONS.	*:****:* . ***** .:. * **** * : * **** :****.***** *:****:..						

	850	860	870	880	890	900	910		
SEQID3	AFIDEADHK	TEDSLKAAK	AKIFEKFPGL	KECKDPTYHY	EVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD	
SEQID4	AFIDEADHK	TEDSLKAAK	AKIFEKFPGL	KECKDPTYHY	EVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD	
SEQID12	AFIDEADHK	TTESLDA	AKKILEFKGLE	ECKDSTYHY	EVNCL	EYRPGTN	VPATGGMYVPRYT	QNLNSAD	
SEQID16	AFIDEAEH	STEESLE	AAKAKIFEKFPGL	QECKDSTYHY	EVNCL	ERRPGTD	VPVTTGGMYVPRYT	QNLNLDAD	
SEQID5	AFIDEADHK	TEDSLKAAK	EKIFA	AFPGLKECTN	PAYHYEVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD	
SEQID14	AFIDEADHK	TEDSLKAAK	EKIFA	AFPGLKECTN	PAYHYEVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD	
SEQID6	AFIDEADHK	TEDSLKAAK	EKIFA	AFPGLKECTN	PAYHYEVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD	
SEQID8	AFIDEADHK	TEDSLKAAK	EKIFA	AFPGLKECTN	PAYHYEVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD	
SEQID9	AFIDEADHK	TPESLAA	AKQRILDA	FPGLVECKD	SDYHYEVNCL	EYRPGTD	VPVTTGGMYVPQY	TQLDLSAD	
SEQID11	AFIDEADHK	TPESLAA	AKQRILDA	FPGLVECKD	SDYHYEVNCL	EYRPGTD	VPVTTGGMYVPQY	TQLDLSAD	
SEQID7	AFIDEADHK	TPESLAA	AKQRILDA	FPGLVECKD	SDYHYEVNCL	EYRPGSG	VPVTTGGMYVPQY	TQLDLGAD	
SEQID13	AFIDEADHK	TPESLAA	AKKRILDA	FPGLEECKD	SDYHYEVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD	
SEQID15	AFIDEADHK	TPESLAA	AKKRILDA	FPGLEECKD	SDYHYEVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD	
SEQID2	AFIDEAEY	TTEESLE	AAKAKIFEKFPGL	QECKDSTYHY	EVNCL	ERRPGTD	VPVTTGGMYVPRYT	QNLNLDAD	
SEQID10	AFIDEAEH	TNESLE	AAKAKI	AFPGLKECKD	PTYHYEVNCL	EYRPGTN	VPVTTGGMYVPRYT	QNLNSAD	
	*****	: * * *	* * * *	: * * *	: . .	* * * * *	* * * * *	* * * * * : * * *	
SEC. CONS.	AFIDEADHK	TEESLKA	AAKAKIF	2AFPGL	EECKD	STYHYEVNCL	EYRPGTG	VPVTTGGMYVPQY	TQLSLNAD

	920	930	940	950	960	970	980
SEQID3	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	WNKIEYRYEND	KDDDEL	GFK	
SEQID4	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	WNKIEYRYEND	KDDDEL	GFK	
SEQID12	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NEIEYSYDSS	KEDEL	GFK	
SEQID16	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	WNKIEYRYEND	KDDDEL	GFK	
SEQID5	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NDTSYRYEEGK	NDEL	GFK	
SEQID14	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NDTSYRYEEGK	NDEL	GFK	
SEQID6	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NDTSYRYEEGK	NDEL	GFK	
SEQID8	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NDTSYRYEEGK	NDEL	GFK	
SEQID9	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NETKYRYEEGK	EDDEL	GFK	
SEQID11	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NETKYRYEEGK	EDDEL	GFK	
SEQID7	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NETKYRYEEGK	EDDEL	GFK	
SEQID13	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	UNKIEYRYEND	KDDDEL	GFK	
SEQID15	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	UNKIEYRYEND	KDDDEL	GFK	
SEQID2	TAKAMVQAADLGTNIQRLYQHEL	YFRTKGSKGERLNSVD	LRLYQNMMSVULW	NDTKYRYEEGK	EDDEL	GFK	
SEQID10	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NEIEYSYDSS	KEDEL	GFK	
	*****	*****	*****	*****	*****	*****	*****
SEC. CONS.	TAKAMVQAADLGTNIQRLYQHEL	YFRTNGRKGERLSSVD	LRLYQNMMSVULW	NTIEYRYEEGK	EDDEL	GFK	

	990	1000	1010	1020	1030	1040	1050
SEQID3	TFTEFLNCYANDAYTGGTQCSDELKKS	LVDDNNMIYGEKS	VNKAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID4	TFTEFLNCYANDAYTGGTQCSDELKKS	LVDDNNMIYGEKS	VNKAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID12	TFTEFLNCYANDAYTGGTQCSDELKKS	LVDDNNMIYGEKS	VNKAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID16	TFTEFLNCYANNAYSEGTOCSADLKKS	LVDDNNMIYGDGSS	-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID5	TFTEFLNCYANDAYAGGTKCSADLKKS	LVDDNNMIYGDGSS	-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID14	TFTEFLNCYANDAYAGGTKCSADLKKS	LVDDNNMIYGDGSS	-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID6	TFTEFLNCYANDAYAGGTKCSADLKKS	LVDDNNMIYGDGSS	-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID8	TFTEFLNCYANDAYAGGTKCSADLKKS	LVDDNNMIYGDGSS	-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID9	TFTEFLNCYTNNAYVG-TQCSAELKKS	LIDNKM	YIGEES-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID11	TFTEFLNCYTNNAYVG-TQCSAELKKS	LIDNKM	YIGEES-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID7	TFTEFLNCYTNNAYVG-TQCSAELKKS	LIDNKM	YIGEES-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID13	TFTEFLNCYANNAYDGGTQCSAELKQSL	IDNKM	YIGE-GS-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID15	TFTEFLNCYANNAYDGGTQCSAELKQSL	IDNKM	YIGE-GS-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID2	TFTEFLNCYANDAYAGGTKCSADLKKS	LVDDNNMIYGDGSS	-KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
SEQID10	TFTEFLNCYANDAYTGGTQCSAELKQSL	IDNKM	YIGES--KAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		
	*****:***	* ** :****:*****	. *****				
SEC. CONS.	TFTEFLNCYANDAYAGGTQCSAELKKS	LVDDNNMIYGE	GSSNKAGMMNPSYPLNYMEKPL	TRMLGRS	WWD		

	1060	1070	1080	1090	1100	1110	1120
SEQID3	LN	IKVD	VEKY	PGAV	SAEG	EKVT	TETIS
SEQID4	LN	IKVD	VEKY	PGAV	SAEG	EKVT	TETIS
SEQID12	LN	IKVD	VEKY	PGAV	SEEG	EVTE	TETIS
SEQID16	LN	IKVD	VEKY	PGAV	SAEG	EKVT	TETIS
SEQID5	LN	TKVD	VEKY	PGAV	SEEG	ANVT	TETIS
SEQID14	LN	IKVD	VEKY	PGAV	SEEG	ANVT	TETIS
SEQID6	LN	IKVD	VEKY	PGAV	SEEG	ANVT	TETIS
SEQID8	LN	IKVD	VEKY	PGAV	SEEG	ANVT	TETIS
SEQID9	LN	IKVD	VEKY	PGVNT	NGET	VTQ	NINLY
SEQID11	LN	IKVD	VEKY	PGVNT	NGET	VTQ	NINLY
SEQID7	LN	IKVD	VEKY	PGAV	SEEG	ANVT	TETIS
SEQID13	LN	IKVD	VEKY	PGAV	SAEG	EEVT	TETIN
SEQID15	LN	IKVD	VEKY	PGAV	SAEG	EEVT	TETIN
SEQID2	LN	IKVD	VEKY	PGSV	SAKG	EEVT	TETIN
SEQID10	LN	IKVD	VEKY	PGAV	SVG	EEVT	TETIS
***** *	*	***:*.***	*****	*****	*****	*****	*****
SEC. CONS.	LN	IKVD	VEKY	PGAV	SEEG	ANVT	TETIS

	1130	1140	1150	1160	1170	1180	1190
SEQID3	DLTGREKHEVALNRPPKVTKTYELKANGEVKFTVPYGGLIYIKGN	SPQN-ESAEFTFTGVVKAPFYKDGA					
SEQID4	DLTGREKHEVALNRPPKVTKTYELKANGEVKFTVPYGGLIYIKGN	SPQN-ESAEFTFTGVVKAPFYKDGA					
SEQID12	DLTGREKHEVALNRPPKVTKTYELKANGEVKFTVPYGGLIYIKGN	SKENKNSASFSTFTGVVKAPFYKNGA					
SEQID16	DLTGREKHEVALNRPPKVTKTYDLKANDKVTFKVPYGGLIYIKGN	SPKN-ESAEFTFTGVVKAPFYKDGE					
SEQID5	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGN	SSTN-ESASFSTFTGVVKAPFYKDGA					
SEQID14	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGN	SSTN-ESASFSTFTGVVKAPFYKDGA					
SEQID6	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGN	SSTN-ESASFSTFTGVVKAPFYKDGA					
SEQID8	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGN	SSTN-ESASFSTFTGVVKAPFYKDGA					
SEQID9	DLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGLIYIKGD	SKEV-QSADFSTFTGVVKAPFYKDGG					
SEQID11	DLTGREKHEVSLNRPPRVTKTYDLKANDKVTFKVPYGGLIYIKGD	SKEV-QSADFSTFTGVVKAPFYKDGG					
SEQID7	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGN	SSTN-ESASFSTFTGVVKAPFYKDGA					
SEQID13	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKSD	SKEE-KSANFTFTGVVKAPFYKDGG					
SEQID15	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKSD	SKEE-KSANFTFTGVVKAPFYKDGG					
SEQID2	DLTGREKHEVALNRPPRVTKTYTLEANGVTFKVPYGGLIYIKGD	SKDD-VSANFTFTGVVKAPFYKDGE					
SEQID10	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGN	SSTN-ESASFSTFTGVVKAPFYKDGA					
	*****	*****	*****	*****	*****	*****	*****
SEC. CONS.	DLTGREKHEVALNRPPRVTKTYSLDASGTVKFKVPYGGLIYIKGN	S2TNNESASFSTFTGVVKAPFYKDGA					

```

1200      1210      1220      1230      1240      1250      1260
|          |          |          |          |          |          |
SEQID3    WKNALNSPAPLGELESDAFVYTTPKKNLEAS---NFTGGVAEFAKDLDTFASMMNDFYGRNDEDEGKHRMF
SEQID4    WKNALNSPAPLGELESDAFVYTTPKKNLEAS---NYKGGGEQFAEELDTFASMMNDFYGRNDEDEGKHRMF
SEQID12   WKNALNSPAPLGELESDAFVYTTPKKNLEAS---NFTGGVAEFAKDLDTFASMMNDFYGRNDEDEGKHRMF
SEQID16   WKNALNSPAPLGELESDFVYTPAKNNLNASNYSNYDTGVAEFAKELDTFASMMNDFYGRDGESGNHRMF
SEQID5    WKNDLNSPAPLGELESDAFVYTTPKKNLNAS---NYTGGLEQFANDLDTFASMMNDFYHGRDSEDEGKHRMF
SEQID14   WKNDLNSPAPLGELESDAFVYTTPKKNLNAS---NYTGGLEQFANDLDTFASMMNDFYGRDSEDEGKHRMF
SEQID6    WKNDLNSPAPLGELESDAFVYTTPKKNLNAS---NYTGGLEQFANDLDTFASMMNDFYGRDSEDEGKHRMF
SEQID8    WKNDLNSPAPLGELESDAFVYTTPKKNLNAS---NYTGGLEQFANDLDTFASMMNDFYGRDSEDEGKHRMF
SEQID9    WQHDLNSPAPLGELESASFVYTTPKKNLNAS---NYTGGLEQFANDLDTFASMMNDFYGRDSEDEGKHRMF
SEQID11   WQHDLNSPAPLGELESASFVYTTPKKNLNAS---NYTGGLEQFANDLDTFASMMNDFYGRDSEDEGKHRMF
SEQID7    WKNDLNSPAPLGELESDAFVYTTPKKNLNAS---NYTGGLEQFANDLDTFASMMNDFYGRDSEDEGKHRMF
SEQID13   WKNDLNSPAPLGELESASFVYTTPKKNLEAS---NYKGGGLKQFAEDLDTFASMMNDFYGRDGESGKHRMF
SEQID15   WKNDLNSPAPLGELESASFVYTTPKKNLEAS---NYKGGGLKQFAEDLDTFASMMNDFYGRDGESGKHRMF
SEQID2    WKNDLNSPAPLGELESASFVYTTPKKNLEAS---NFTGGVAEFAKDLDTFASMMNDFYGRNDEDEGKHRMF
SEQID10   WKNDLNSPAPLGELESASFVYTTPKKNLNAS---NYTGGLEQFAKDLDTFASMMNDFYGRNDEDEGKHRMF
*: *.*****: *****: *****: *****: *****: *****: *****: *.*****
SEC. CONS. WKNDLNSPAPLGELESDAFVYTTPKKNLNASNYSNYTGGLEQFANDLDTFASMMNDFYGRDSEDEGKHRMF

```

	1270	1280	1290	1300	1310	1320	1330
SEQID3	TYKNLTGHHKRF	TNDVQISIGDAHSGYPVMNSSFS	TNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID4	TYKNLTGHHKRF	TNDVQISIGDAHSGYPVMNSSFS	TNSTTLP	PTPLNDWL	IWHEAGHNAAETPL	TVPGA	
SEQID12	TYKNLTGHHKRF	TNDVQISIGDAHSGYPVMNSSFS	TNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID16	TYKALTGHHKRF	ANDVQISIGDAHSGYPVMNSSFS	TNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	NVPGAT	
SEQID5	TYKNLPGHKHRF	TNDVQISIGDAHSGYPVMNSSFS	PNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID14	TYKNLPGHKHRF	TNDVQISIGDAHSGYPVMNSSFS	PNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID6	TYKNLPGHKHRF	ANDVQISIGDAHSGYPVMNSSFS	PNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID8	TYKNLTGHHKRF	TNDVQISIGDAHSGYPVMNSSFS	TNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID9	TYKNLPGHKHRF	ANDVQISIGDAHSGYPVMNSSFS	PNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID11	TYKNLPGHKHRF	ANDVQISIGDAHSGYPVMNSSFS	PNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID7	TYKNLTGHHKRF	ANDVQISIGDAHSGYPVMNSSFS	PNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID13	TYEALTGHHKRF	TNDVQISIGDAHSGYPVMNSSFS	PNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID15	TYEALTGHHKRF	TNDVQISIGDAHSGYPVMNSSFS	PNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	TVPGAT	
SEQID2	TYKNLTGHHKRF	TNDVQISIGDAHSGYPVMNSSFS	TNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	NVPGAT	
SEQID10	TYKNLTGHHKRF	TNDVQISIGDAHSGYPVMNSSFS	TNSTTLP	PTPLNDWL	IWHEVGHNAAETPL	NVPGAT	
SEC. CONS.	***	*****	*****	*****	*****	*****	*****

	1340	1350	1360	1370	1380	1390	1400
SEQID3	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID4	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNNQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID12	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNNQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID16	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID5	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNNQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID14	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNNQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID6	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID8	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID9	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID11	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID7	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID13	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID15	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID2	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		
SEQID10	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		

SEC. CONS.	EVANNVLALYMQDRYL	GKMNRVADDITVAPEY	LEESNGQAWARGGAGDR	LLMYAQLKEWAEKNFD	IKQWY		

	1410	1420	1430	1440	1450	1460	1470
SEQID3	PEGD-LPKFYS	DREGMKGWNL	FQLMHRKARG	DEVGKTKFG	ERNYCAESNG	NAADKLMLCAS	WVAQTDLSE
SEQID4	PEGN-LPKFYS	EREGMKGWNL	FQLMHRKARG	DEVGKTKFG	ERNYCAESNG	NAADTLMLCAS	WVAQTDLSA
SEQID12	PDGTPLEFYS	EREGMKGWNL	FQLMHRKARG	DEVNDKFGG	KNYCAESNG	NAADTLMLCAS	WVAQTDLSE
SEQID16	PDGE-LPKFYS	DRKMGKWNLF	QLMHRKARGD	DVSNDKF	GGKNYCAESNG	NAADTLMLCAS	WVAQADLSE
SEQID5	PDGTPLEFYS	EREGMKGWNL	FQLMHRKARG	DEVNDKFGG	KNYCAESNG	NAADTLMLCAS	WVAQTDLSE
SEQID14	PDGTPLEFYS	EREGMKGWNL	FQLMHRKARG	DEVNDKFGG	KNYCAESNG	NAADTLMLCAS	WVAQTDLSE
SEQID6	PDGTPLEFYS	EREGMKGWNL	FQLMHRKARG	DEVNDKFGG	KNYCAESNG	NAADTLMLCAS	WVAQTDLSE
SEQID8	PEGE-LPKFFS	DREGMKGWNL	FQLMHRKARG	DDVGDKTF	GGKNYCAESNG	NAADTLMLCAS	WVAQTDLSA
SEQID9	PDGTPLEFYS	EREGMKGWNL	FQLMHRKARG	DEVNDKFGG	KNYCAESNG	NAADTLMLCAS	WVAQTDLSE
SEQID11	PDGTPLEFYS	EREGMKGWNL	FQLMHRKARG	DEVNDKFGG	KNYCAESNG	NAADTLMLCAS	WVAQTDLSE
SEQID7	PEGE-LPKFFS	DREGMKGWNL	FQLMHRKARG	DDVGDKTF	GGKNYCAESNG	NAADSLMLCAS	WVAQTDLSA
SEQID13	PEGS-LPAFYS	EREGMKGWNL	FQLMHRKARG	DDVGNDK	FGNRYCAESNG	NAADTLMLCAS	WVAQTDLSA
SEQID15	PEGS-LPAFYS	EREGMKGWNL	FQLMHRKARG	DDVGNDK	FGNRYCAESNG	NAADTLMLCAS	WVAQTDLSA
SEQID2	PDGE-LPKFYS	DRKMGKWNLF	QLMHRKARGD	DVGNSTF	GGKNYCAESNG	NAADTLMLCAS	WVAQADLSE
SEQID10	PDGK-LPAFYS	EREGMKGWNL	FQLMHRKARG	DDVGNSTF	GGKNYCAESNG	NAADTLMLCAS	WVAQADLSE
	:	**	*:*****	*****	*.*	*:*****	*****
SEC. CONS.	PDGTPLP2FYS	EREGMKGWNL	FQLMHRKARG	DEVGNDK	FGGKNYCAESNG	NAADTLMLCAS	WVAQTDLSE

	1480	1490	1500	1510	1520	1530
SEQID3	FFKKWNP	GANAYQ	LPGASE	MNFEGG	VQSAYE	TAAALN
SEQID4	FFKKWNP	GANAYQ	LPGASE	MNFEGG	VQSAYE	TAAALN
SEQID12	FFKKWNP	GANAYQ	LPGATE	MSFEGG	VQSAYN	TASLDL
SEQID16	FFKKWNP	GANAYQ	LPGASE	MSFEGG	VQSAYN	TAAAMH
SEQID5	FFKKWNP	GANAYQ	LPGASE	MSFEGG	VQSAYN	TASLDL
SEQID14	FFKKWNP	GANAYQ	LPGASE	MSFEGG	VQSAYN	TASLDL
SEQID6	FFKKWNP	GANAYQ	LPGASE	MSFEGG	VQSAYN	TASLKL
SEQID8	FFKKWNP	GANAYQ	LPGATE	MSFEGG	VQSAYS	TASLKL
SEQID9	FFKKWNP	GANAYQ	LPGASE	MSFEGG	VQSAYN	TASLKL
SEQID11	FFKKWNP	GANAYQ	LPGASE	MSFEGG	VQSAYN	TASLDL
SEQID7	FFKKWNP	GANAYQ	LPGATE	MSFEGG	VQSAYS	TASLKL
SEQID13	FFKKWNP	GANAYQ	LPGATE	MSFEGG	VQSAYN	TASLDL
SEQID15	FFKKWNP	GANAYQ	LPGATE	MSFEGG	VQSAYN	TASLDL
SEQID2	FFKKWNP	GANAYQ	LPGATE	MSFQGG	VSSSAY	STLASL
SEQID10	FFKKWNP	GANAYQ	LPGAAE	MSFEGG	VSSSAY	STLASL
SEC. CONS.	*****.*****:*.*:****.***.***:..*.*:*****:***!.*. *					

FIGURA 2

MG65	14	13	RS28	2	ARECI	4	U19	10	3	12	5	16	8	7	6	9	11	15	Seins
10	10	924	88	88	88	922	88	882	922	905	999	914	973	925	984	925	924	987	D10B
	10	924	88	88	88	922	88	882	922	905	999	914	973	925	984	925	924	987	MG65
		924	88	88	88	922	88	882	922	905	999	914	973	925	984	925	924	987	14
			876	876	876	902	876	889	897	896	984	908	923	936	922	982	939	936	13
				10	10	859	10	941	866	861	88	836	882	879	878	867	877	867	RS28
					10	859	10	941	866	862	88	836	882	879	878	867	878	877	2
						859	10	941	866	862	88	836	882	879	878	867	878	877	ARECI
							859	889	983	944	921	913	92	881	919	887	871	889	4
								941	866	862	88	836	882	879	878	867	878	877	U19
									867	869	881	881	88	884	888	871	882	881	10
										949	921	916	92	882	919	887	873	889	3
											905	90	895	879	902	888	877	882	12
												913	972	924	984	924	924	986	5
													913	889	915	889	888	906	16
														915	986	924	889	924	8
															987	913	950	987	7
																987	912	923	6
																	984	983	9
																		941	11

FIGURA 3

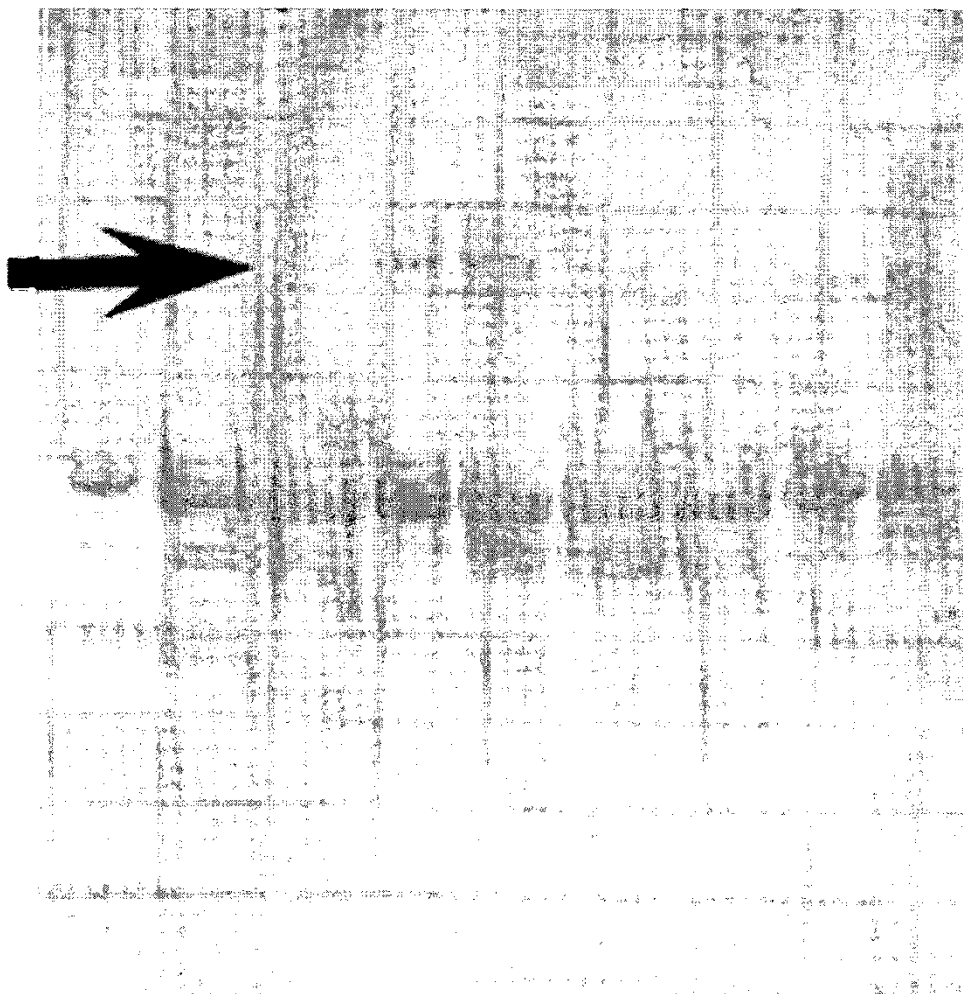
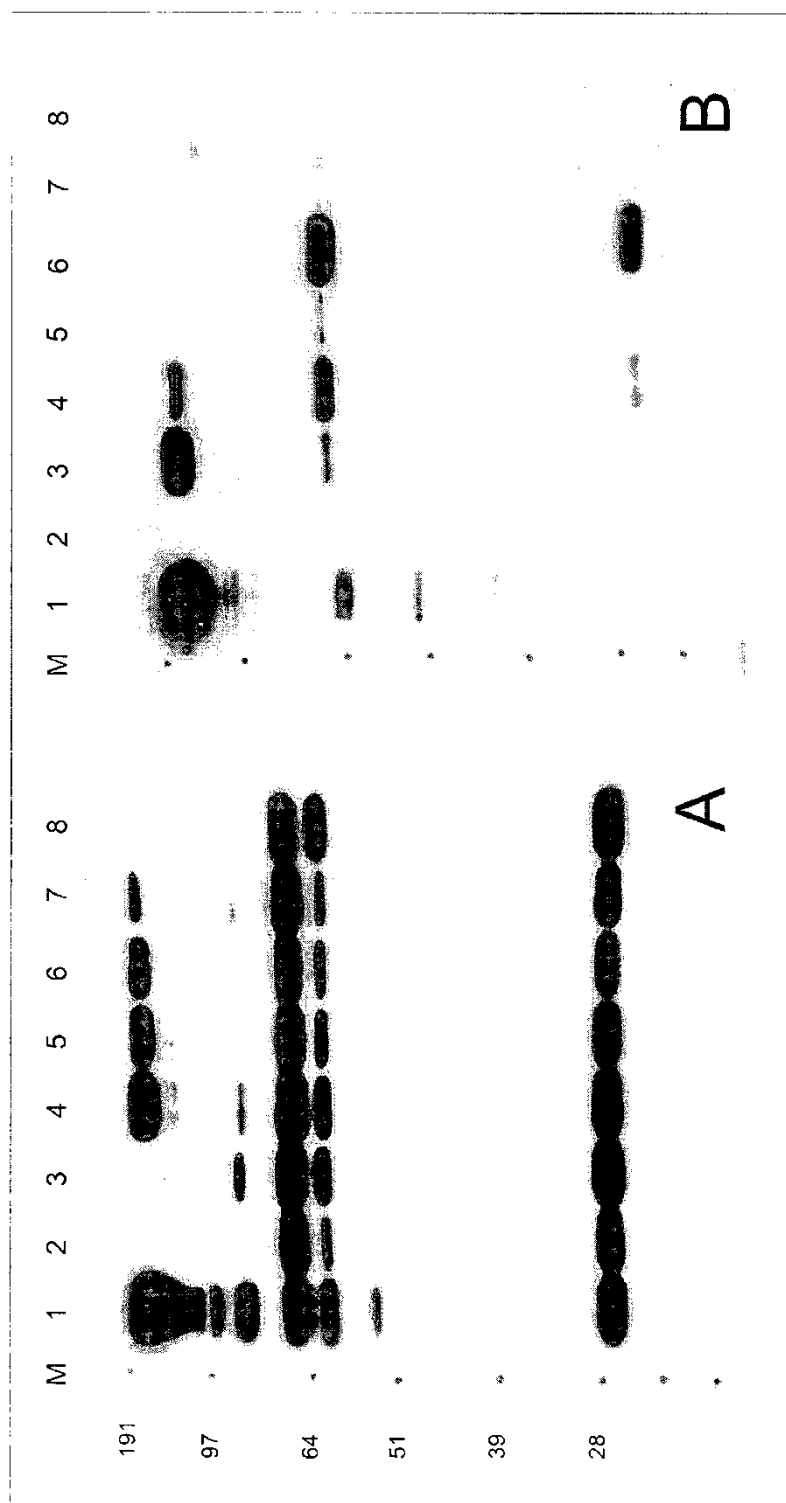


FIGURA 4



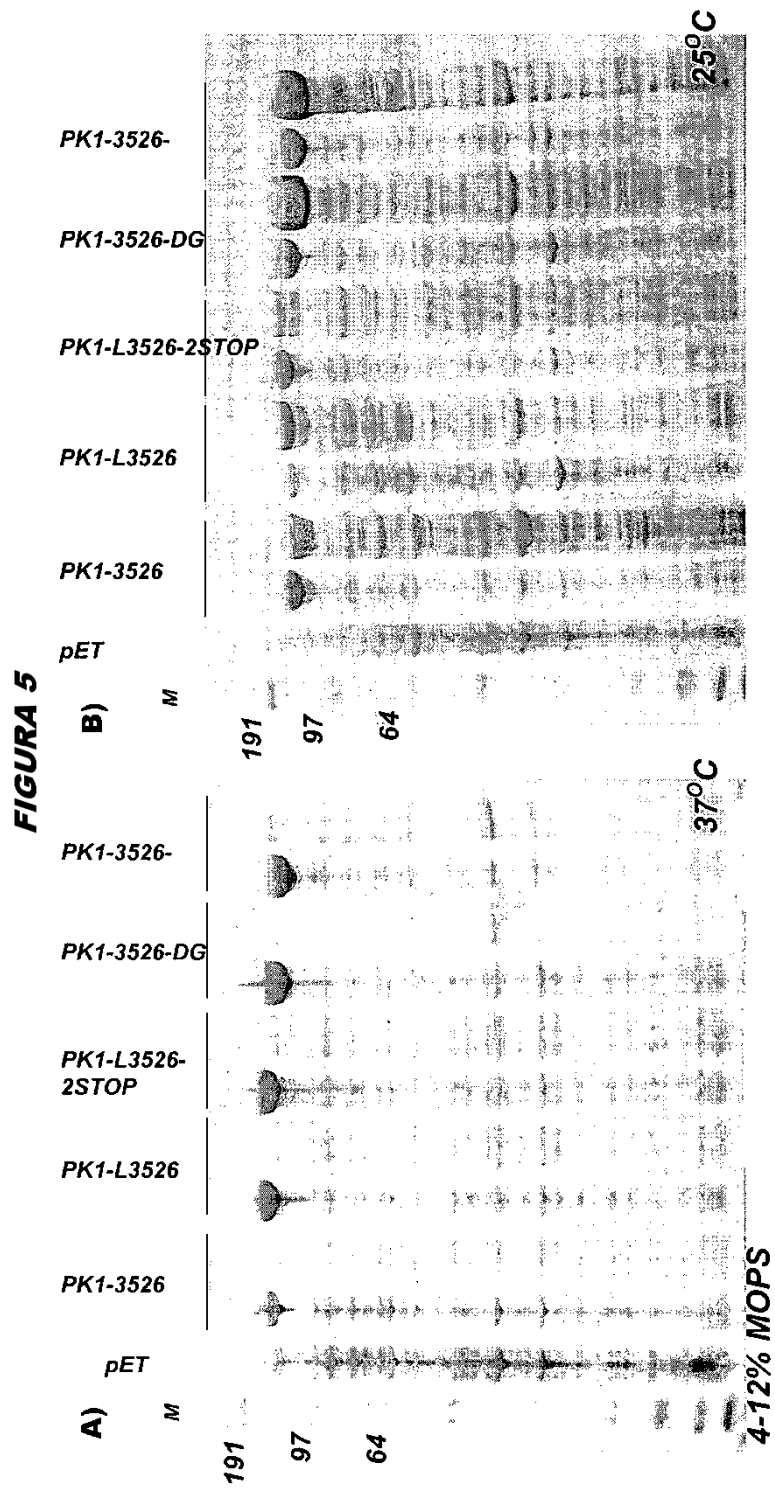


FIGURA 6

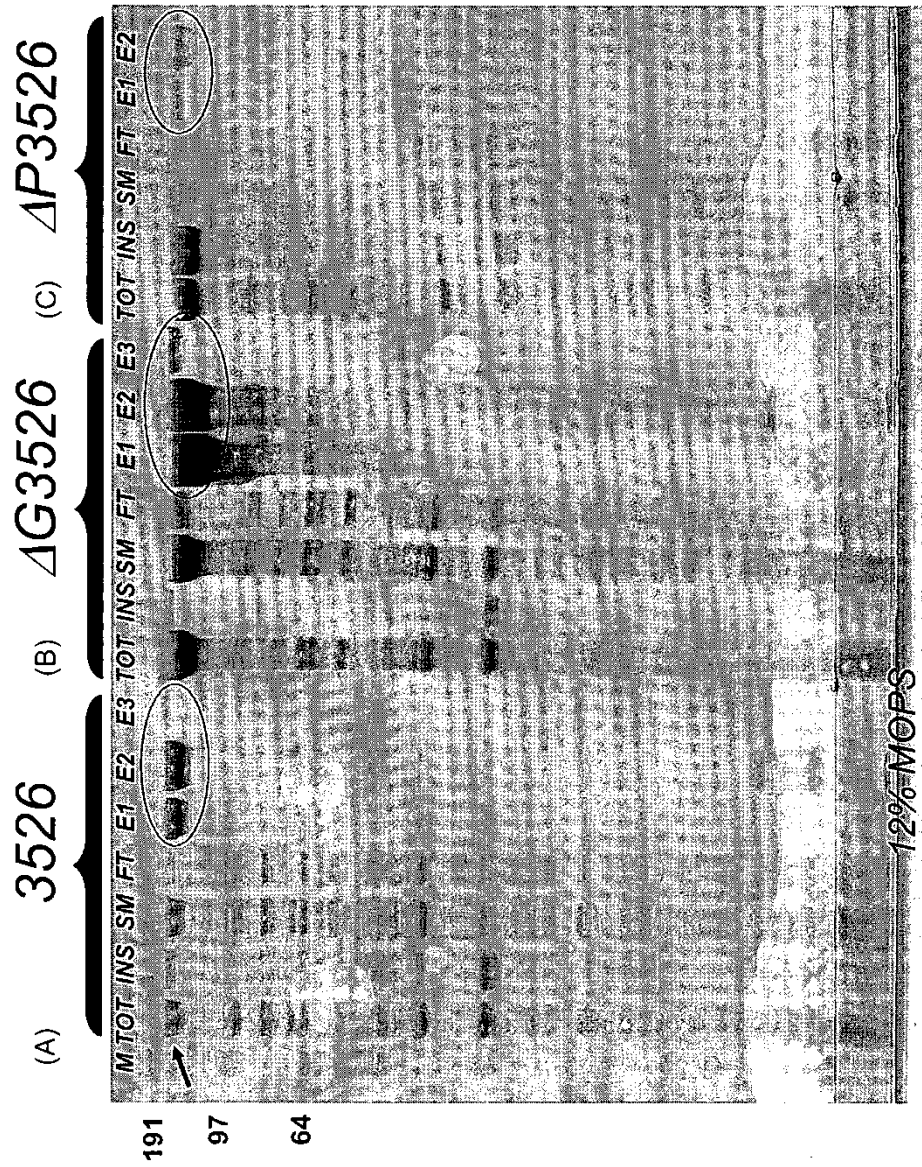


FIGURA 7

