

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 622 280**

51 Int. Cl.:

C07K 14/33 (2006.01)

C12N 15/09 (2006.01)

C12N 15/75 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.05.2011 PCT/US2011/038358**

87 Fecha y número de publicación internacional: **01.12.2011 WO11150342**

96 Fecha de presentación y número de la solicitud europea: **27.05.2011 E 11733719 (6)**

97 Fecha y número de publicación de la concesión europea: **22.03.2017 EP 2576604**

54 Título: **Difocinas y métodos de utilización de las mismas**

30 Prioridad:

27.05.2010 US 349145 P

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

06.07.2017

73 Titular/es:

**AVIDBIOTICS CORP. (100.0%)
100 Kimball Way
South San Francisco, CA 94080-6218, US**

72 Inventor/es:

**SCHOLL, DEAN M.;
GEBHART, DANA M.;
WILLIAMS, STEVEN R.;
GOVONI, GREGORY R. y
MARTIN, DAVID W.**

74 Agente/Representante:

UNGRÍA LÓPEZ, Javier

Observaciones :

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 622 280 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

Difocinas y métodos de utilización de las mismas5 **Campo de la invención**

Esta solicitud se refiere generalmente a la identificación y aislamiento de una agrupación de genes suficiente para producir una bacteriocina, y más específicamente, una bacteriocina de elevado peso molecular de tipo R que destruye *Clostridium difficile*, y a los métodos para alterar su especificidad bactericida, producirla y utilizarla.

10

Antecedentes de la invención

Clostridium difficile es una bacteria aerobia obligada, formadora de esporas, gram positiva que es un patógeno destacado para los seres humanos y otros mamíferos (Bartlett et al., 1977, Bartlett et al., 1979; Keel et al., 2007; Sunenshine & McDonald, 2006). La secuencia completa del genoma de *C. difficile* R20291 está disponible en la Base de Datos EMBL (número de acceso EBI EM_PRO:FN545816). A bajas densidades *C. difficile* puede residir inocuamente en el tracto gastrointestinal (GI) de los mamíferos, pero tras la expansión, frecuentemente como resultado de los antibióticos administrados que reducen las bacterias comensales, la bacteria *C. difficile* produce exotoxinas suficientes para causar un abanico de enfermedades que van desde una enfermedad diarreica leve a una colitis pseudo-membranosa característica, mortal, particularmente en seres humanos de edad avanzada y otros con co-morbilidades significativas (Bartlett, 2002).

Debido a que las esporas formadas por este patógeno se diseminan ampliamente y son difíciles de erradicar o inactivar en hospitales y centros de cuidados crónicos, la probabilidad de que los pacientes sean colonizados por *C. difficile* aumenta bruscamente al entrar en dicha instalación (Bartlett, 2007). De hecho, una cepa relativamente nueva de *C. difficile* que es una cepa bacteriana toxigénica hipervirulenta de *C. difficile*, BI/NAP1/027, que causa una enfermedad grave en situaciones de brote masivo, ha sido documentada recientemente (Spigaglia et al., 2002, Pépin et al., 2004, McDonald et al., 2005, Muto et al., 2005, Loo et al., 2005, Belmares et al., 2009). La incidencia de *C. difficile* (CDAD) en niños, anteriormente con bajo riesgo, también ha aumentado sustancialmente (Benson et al., 2007; Zilberberg et al., 2010).

La eliminación del patógeno profilácticamente en portadores asintomáticos o sujetos colonizados mediante la administración de antibióticos está fuertemente contraindicada debido al alto riesgo de inducir enfermedades asociadas con *C. difficile*.

Se han descrito bacteriocinas de tipo R elaboradas por bacterias gram negativas y han sido desplegadas por tales bacterias para eliminar otras cepas gram negativas competitivas, incluso en algunas circunstancias otras especies o géneros de bacterias gram negativas (Kageyama et al., 1964; Kageyama et al 1964a; Kingsbury, D, 1966; Blackwell y Law, 1981; Blackwell et al., 1982; Campagnari et al., 1994; Strauch et al., 2001; Jabrane et al., 2002). Se ha descrito la fusión de regiones de unión a la placa basal (BPAR) de piocinas de tipo R a dominios de unión a receptores heterólogos (RBD), dando como resultado la creación de nuevas piocinas de tipo R con nuevas especificidades bactericidas para bacterias gram negativas (Williams et al., 2008; Scholl et al., 2009).

Otras bacteriocinas de elevado peso molecular o bacteriocinas de tipo R se han descrito en bacterias gram positivas (Coetsee et al., 1968; Thompson y Pattee, 1981; Zink et al., 1995). Sin embargo, se sabe mucho menos sobre las estructuras de la bacteriocina de elevado peso molecular de tipo R producidas por bacterias gram positivas. Pero si bien éstas han sido descritas, ninguna se ha caracterizado a un nivel genético o se ha manipulado de una manera propicia o necesaria para desarrollar un agente útil. Se han descrito bacteriocinas de elevado peso molecular para 2 especies de *Clostridium*, *botulinum* y *perfringens* (Ellison y Kautter, 1970; Anastasio et al., 1971; Nieves et al., 1981). No se ha descrito ninguna que sea producida por *C. difficile* o que elimine *C. difficile*.

Compendio de la invención

Esta invención se basa en el aislamiento de todo el locus genético o de una agrupación genética que codifica las bacteriocinas de tipo R específicas de *C. difficile* (denominadas en la presente memoria "difocinas") que son bactericidas frente a otras cepas de *C. difficile*; la expresión de la agrupación de genes de difocina y la producción de difocinas en bacterias aerobias; y el descubrimiento de que el marco de lectura abierto (ORF) 1374 de la agrupación de genes de difocina determina el espectro bactericida de esa difocina frente a cepas de *C. difficile*. Esta invención proporciona un medio práctico para alterar la especificidad de las difocinas por medio ingeniería genética para producir nuevas difocinas y para fabricar y administrar directa o indirectamente difocinas para eliminar *C. difficile* del tracto gastrointestinal (GI) de animales colonizados, incluyendo seres humanos. La administración de difocinas puede tratar o prevenir el desarrollo de la infección por *C. difficile* y de enfermedades asociadas sin dañar, como hacen los antibióticos tradicionales, las bacterias GI comensales tan necesarias para una buena salud.

60

De acuerdo con la presente invención, se proporciona una molécula de ácido nucleico aislada que codifica la bacteriocina de elevado peso molecular de tipo R (epm) descrita en la presente memoria, en donde la molécula de ácido nucleico codifica una bacteriocina de epm de tipo R en la que el polipéptido del dominio de unión al receptor nativo (RBD) se ha sustituido por un polipéptido RBD codificado por una secuencia de polinucleótidos heteróloga de otra cepa de *C. difficile* o por un polipéptido RBD codificado por una secuencia de polinucleótidos heteróloga de un bacteriófago que infecta *C. difficile*, y en donde el polipéptido RBD de reemplazo crea una bacteriocina modificada con un espectro bactericida diferente del creado por el polipéptido RBD nativo y en donde la bacteriocina de epm de tipo R comprende:

- (a) polipéptidos de los SEQ ID NO: 4-16, 18 y 19; o
- (b) polipéptidos de los SEQ ID NO: 66-80; o
- (c) polipéptidos que tienen una identidad de al menos 80% con (a) o (b).

Se describen moléculas de ácido nucleico aisladas que codifican bacteriocinas de elevado peso molecular (epm) de tipo R. Se describen moléculas de ácido nucleico aisladas que codifican bacteriocinas de elevado peso molecular (epm) de tipo R, en donde la molécula de ácido nucleico es de un genoma de una cepa de *Clostridium difficile*, y en donde la bacteriocina de epm de tipo R comprende un polipéptido que es idéntico en al menos 80% a un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 4-16, 18, 19 y 66-80, y la bacteriocina de epm de tipo R tiene un dominio de unión al receptor (RBD) que se une a un receptor de al menos otra cepa de *C. difficile* y por lo tanto tiene actividad bactericida contra la otra cepa o las otras cepas de *C. difficile*.

En una realización de la molécula de ácido nucleico aislada de la invención, además del polipéptido RBD heterólogo de reemplazo, la bacteriocina de epm de tipo R comprende secuencias de aminoácidos idénticas en al menos 90% a los SEQ ID NO: 4-15. En realizaciones particulares, la molécula de ácido nucleico es de un genoma de una cepa de *Clostridium difficile* seleccionada del grupo que consiste en Cd4, Cd16, Cd19108, Cd19123, Cd19126, Cd19145, y el Núm. de Acceso ATCC 43593. En algunas realizaciones, la cepa es Cd16 y la molécula de ácido nucleico incluye el SEQ ID NO: 1 o la cepa es Cd4 y la molécula de ácido nucleico incluye el SEQ ID NO: 61

En otra realización de la invención, se proporcionan bacteriocinas de tipo R aisladas codificadas por una molécula de ácido nucleico de la invención. En un aspecto, las bacteriocinas de tipo R se expresan en una bacteria productora aerobia.

En otra realización de la invención, se proporciona una bacteriocina de elevado peso molecular (epm) de tipo R aislada que tiene actividad bactericida, en donde la bacteriocina comprende una bacteriocina de tipo R de *Clostridium difficile*, en la que el polipéptido del dominio de unión al receptor nativo (RBD) ha sido reemplazado por un polipéptido RBD heterólogo de otra cepa de *C. difficile* o por un polipéptido RBD heterólogo de un bacteriófago que infecta un *C. difficile*, en donde el polipéptido RBD de reemplazo crea una bacteriocina modificada con un espectro bactericida diferente del creado por el RBD nativo, y en donde el polipéptido de la región de unión a la placa basal (BPAR) es idéntico en al menos 80% a un polipéptido compuesto por 50 o más aminoácidos contiguos de los SEQ ID NO: 16 ó 78. Se describen bacteriocinas de peso molecular elevado (epm) de tipo R aisladas que tienen actividad bactericida, en donde la bacteriocina de epm de tipo R incluye una región de unión a la placa basal (BPAR) de una primera cepa de una primera especie de bacterias del género *Clostridium*, y un dominio de unión al receptor (RBD) de una segunda cepa de la primera especie, o de una segunda especie del género *Clostridium* o de un bacteriófago que infecta una especie de *Clostridium*, o una forma modificada de un RBD, en donde la bacteriocina tiene actividad bactericida contra al menos una cepa de *Clostridium difficile*. Se describe que BPAR puede ser de una primera cepa de *Clostridium difficile* y el RBD es de una segunda cepa de *Clostridium difficile* o de un bacteriófago que infecta *Clostridium difficile*. Se describe que BPAR puede ser idéntica en al menos 80% al segmento correspondiente de uno o más de los SEQ ID NO: 16, 54-56 y 78. En otras realizaciones, el RBD es idéntico en al menos 80% al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-56. Se describe que BPAR puede ser idéntica en al menos 80% a un polipéptido que contiene 50 o más aminoácidos contiguos de los SEQ ID NO: 16 ó 78 o que contiene 50 o más aminoácidos contiguos de los SEQ ID NO: 54-56.

En otra realización de la invención, se proporcionan casetes de expresión que contienen una molécula de ácido nucleico de la invención. Los casetes de expresión pueden estar contenidos dentro de un vector de expresión, tal como un plásmido, o pueden estar contenidos dentro del cromosoma de una célula productora. En algunas realizaciones, el casete de expresión contiene un promotor heterólogo unido operativamente a la molécula de ácido nucleico que codifica la bacteriocina de tipo R. El promotor puede ser inducible, reprimible o constitutivamente activo. En un aspecto, el promotor es inducible; en otro aspecto, el promotor es reprimible. En algunas realizaciones, el promotor es inducido añadiendo o eliminando un inductor, represor o des-represor de molécula pequeña. En un aspecto, el promotor es inducido por un inductor o des-represor de molécula pequeña. En un aspecto particular, la expresión del casete se regula por medio de un gen *recA* unido operablemente que codifica una proteína RecA constitutivamente activa y bajo el control de un promotor heterólogo que responde a un inductor o des-represor de molécula pequeña.

En otra realización más de la invención, se proporcionan células productoras que contienen los casetes de expresión de la invención. El casete de expresión puede estar contenido en un vector de expresión episomal dentro de la

célula productora. Alternativamente, las células productoras pueden contener dentro de su cromosoma una molécula de ácido nucleico o casete de expresión de la invención. En ciertas realizaciones, la célula productora es una bacteria anaerobia no patógena y no obligada. En algunas realizaciones, la bacteria anaerobia no patógena y no obligada es una especie de un género de bacterias seleccionadas del grupo que consiste en *Bacillus*, *Lactobacillus*, y *Listeria*. En ciertas realizaciones, la bacteria anaerobia no patógena y no obligada es del género *Bacillus*. En algunos aspectos, la bacteria es *Bacillus subtilis*. En un aspecto particular, el *B. subtilis* carece de la agrupación de genes PBSX. En otra realización, la célula productora es una bacteria anaerobia obligada pero no patógena.

En otra realización más de la invención, se proporcionan métodos para producir una bacteriocina de epm de tipo R de la invención. El método incluye la exposición de una célula productora que contiene una secuencia de ácido nucleico de la invención unida operablemente a un promotor inducible o des-reprimible sensible a un agente inductor o represor, al agente a una concentración eficaz para inducir la expresión de la bacteriocina de tipo R y purificar la bacteriocina de tipo R expresada. En algunas realizaciones, la molécula de ácido nucleico que codifica la bacteriocina de tipo R es heteróloga para el genoma de la célula productora. En aspectos particulares, la molécula de ácido nucleico está contenida dentro del cromosoma de la célula productora o está contenida en un vector de expresión extracromosómico dentro de la célula productora. En ciertas realizaciones, la célula productora es una bacteria anaerobia no patógena y no obligada. En algunas realizaciones, la bacteria anaerobia no patógena y no obligada es una especie de un género de bacterias seleccionadas del grupo que consiste en *Bacillus*, *Lactobacillus*, y *Listeria*. En ciertas realizaciones, la bacteria anaerobia no patógena y no obligada es del género *Bacillus*. En algunos aspectos, la bacteria es *Bacillus subtilis*. En un aspecto particular, el *B. subtilis* no es lisado cuando se induce la producción de la bacteriocina de tipo R. En un aspecto adicional, el *B. subtilis* carece de la agrupación de genes PBSX.

En una realización adicional de la invención, se proporcionan métodos *in vitro* para eliminar *Clostridium difficile*, una bacteria patógena. El método incluye poner en contacto la bacteria patógena con una bacteriocina de tipo R de la invención, por medio de lo cual la bacteriocina de tipo R se une y elimina la bacteria patógena.

En otra realización de la invención se proporciona una bacteriocina de tipo R de la invención para su uso en métodos de tratamiento o prevención de una infección por *Clostridium difficile* causante de enfermedad en un animal, en donde el método incluye la administración de una cantidad bactericida de una bacteriocina de tipo R de la invención directamente a un animal que lo necesita, la administración del agente indirectamente mediante la administración de células productoras o la administración de esporas de la bacteria *C. difficile* que producen difocinas naturales pero que han sido modificadas genéticamente para no producir toxinas. En realizaciones particulares, se proporciona una célula productora de la invención para su uso en el tratamiento de una infección por *Clostridium difficile* en un animal, en donde el tratamiento comprende la administración a un animal que lo necesita de una cantidad de una célula productora de la invención para producir una cantidad bactericida de la bacteriocina, tratando de ese modo la infección. En un aspecto, el ácido nucleico que codifica la bacteriocina está bajo el control de un promotor *lac* y se administra lactosa al animal. En algunas realizaciones, el animal es un mamífero. En un aspecto, el mamífero es un ser humano.

Breve descripción de los dibujos

La Figura 1 proporciona las secuencias de ácidos nucleicos o aminoácidos de los SEQ ID NO: 1-80.

Figura 2. Actividad bactericida de las difocinas en productos aislados de *C. difficile*. Los números indicadores de las cepas se muestran a lo largo de la parte superior de las matrices; las cepas marcadas con asterisco son cepas NAP1/027/BI. Las identidades de las fuentes de *C. difficile* de las difocinas sometidas a ensayo se muestran a lo largo de los bordes izquierdos. Las cepas se adquirieron de LC Fortier (4 y 16), ATCC (43593), o RM Alden Research Lab, Culver City, CA.

La Figura 3 muestra una micrografía electrónica de barrido de dif16. Obsérvense los apéndices de fibra de la cola de tipo flor.

La Figura 4 muestra una fotografía de los resultados de un ensayo de dilución única de dif4 sobre la cepa diana Cd19135. La difocina purificada se diluyó en serie 10 veces y se aplicaron alícuotas de 5 µl de las diluciones sobre un césped de la bacteria *C. difficile* diana. Después de una incubación anaerobia a 37°C durante la noche, la eliminación de la difocina se indicó por el aclaramiento del crecimiento sobre el césped.

La Figura 5 muestra una fotografía de una SDS-PAGE teñida con plata de preparaciones de dif4 ("4") y dif16 ("16") tanto filtradas ("F") como no filtradas ("UF"). Las flechas indican bandas que fueron escindidas e identificadas por espectrometría de masas.

La Figura 6 muestra un esquema de la agrupación de genes de difocina de Cd630 y Cd16. El locus consistió en los ORF que codificaban proteínas estructurales y proteínas de ensamblaje estructurales típicas de un aparato de la cola de un fago *Myoviridae*, ORF1362 a ORF1375, indicados debajo del mapa. Flanqueando estos genes se encontraban genes que codificaban supuestas proteínas reguladoras de tipo fago. Las flechas inferiores indican la dirección en la que se transcriben los ORF y las supuestas funciones de los ORF se indican por encima del mapa.

La Figura 7 es una imagen de los ensayos de dilución única de dif16 producida por *Bacillus subtilis* BDR123-

488 (sector 1) y BDR123-491 (sector 2) sometidos a ensayo sobre un césped de la cepa 19099 de *C. difficile*.

La Figura 8 proporciona los resultados de un análisis ClustalW de secuencias de aminoácidos parciales (SEQ ID NO: 81-86, respectivamente, en orden de aparición) codificadas por el gen ORF 1374 de cada una de las 5 cepas de *C. difficile* que producían difocinas activas. Debajo de cada fila de secuencias alineadas el "*" representa las identidades de aminoácidos en esa posición codificada por los 5 genes, el ":" representa aminoácidos muy similares en esa posición codificada por los 5 genes, y el "." representa aminoácidos algo similares en esa posición codificada por los 5 genes; y el blanco en una posición dada en el alineamiento de secuencias no representa ninguna similitud de aminoácidos codificada por los 5 genes.

La Figura 9 es una imagen de los ensayos de una sola dilución de "Dif4" producida por BDR123-580 y "Dif 4-16". Esta última fue producido por *Bacillus subtilis* BDR123-587 en el que el orf1374 de dif4 (SEQ ID NO: 49) se cambió a orf1374 de dif16 (SEQ ID NO: 17). Estas dos difocinas producidas por *B. subtilis* fueron sometidas a ensayo sobre céspedes de las cepas 19137 y 19145, como se indica. La especificidad de la eliminación por "Dif 4-16" se ha cambiado de la de dif4 a la de dif16.

15 Descripción detallada de la invención

Definiciones

Según se utiliza en la presente memoria, una "bacteriocina de elevado peso molecular (epm) de tipo R" también se conoce simplemente como una "bacteriocina de tipo R" e incluye piocinas, difocinas, monocinas, enterocolitinas, meningocinas u otras bacteriocinas de elevado peso molecular (epm) de tipo R relacionadas estructural o genéticamente con la Familia de bacteriófagos *myoviridae*. Una bacteriocina de tipo R incluye versiones modificadas de piocinas de piocinas, difocinas, enterocolitinas, monocinas y meningocinas de tipo R, (Williams et al., Strauch et al., 2001; Kingsbury, 1966; Zink et al., 1995).

El término "difocina", según se utiliza en la presente memoria, se refiere a una bacteriocina de elevado peso molecular de tipo R aislada de o derivada de *Clostridium difficile* e incluye partículas nativas obtenidas de *C. difficile* así como partículas obtenidas a través de la expresión de la agrupación de genes de difocina en una célula productora no natural. Una difocina también puede ser una partícula modificada genéticamente comprendida por polipéptidos codificados por genes derivados de una o más cepas de *C. difficile*, e idénticos en 80% o más a uno o más polipéptidos de los SEQ ID NO: 2-23, 49 y 62-80.

Una bacteriocina de tipo R de la invención puede ser termolábil, resistente a ácidos suaves, resistente a tripsina, sedimentable por centrifugación a aproximadamente 65.000 xg, y resoluble por medio de microscopía electrónica (Kageyama et al., 1962; Bradley, 1967; Daw et al., 1996; Jabrane et al., 2002; Fortier et al., 2007). En muchos casos, una bacteriocina de tipo R modificada genéticamente descrita en la presente memoria tiene una o más, en cualquier combinación, de estas propiedades. Una propiedad adicional común a las bacteriocinas de tipo R descritas en la presente memoria es que no contienen ácido nucleico y por lo tanto presentan una replicación deficiente de tal manera que no pueden reproducirse después o durante la eliminación de una bacteria diana, como pueden hacer muchos bacteriófagos. Son simplemente proteínas, no organismos.

Las bacteriocinas de tipo R descritas en la presente memoria son moléculas complejas que comprenden subunidades de múltiples proteínas o polipéptidos y se asemejan a las estructuras de cola de bacteriófagos de la familia *Myoviridae*. En las bacteriocinas de tipo R naturales, las estructuras de las subunidades están codificadas por el genoma bacteriano tal como el de *C. difficile* o *P. aeruginosa* y forman bacteriocinas de tipo R que sirven como defensas naturales contra otras bacterias (Kageyama, 1975). Una bacteria diana sensible puede ser eliminada típicamente por una sola molécula de bacteriocina de tipo R (Kageyama et al., 1964; Kageyama et al., 1964a; Morse et al., 1980; Strauch et al., 2001).

Una "bacteria diana" o "bacterias diana" se refieren a una bacteria o bacterias que están unidas por una bacteriocina de tipo R de la descripción y/o cuyo crecimiento, supervivencia o replicación son inhibidos de este modo. En algunas realizaciones, la bacteria diana es del género *Clostridium*. En realizaciones particulares, la bacteria es *Clostridium difficile*. En un aspecto, se elige como diana más de una cepa de *C. difficile*. Las cepas ejemplares de *C. difficile* incluyen pero no se limitan a NAP1/BI/ribotipo 027, así como las enumerados en la Figura 2.

El término "inhibición del crecimiento" o variaciones del mismo se refieren a la disminución o detención de la velocidad de la división de una célula bacteriana o al cese de la división celular bacteriana o a la muerte de la bacteria o bacterias.

Según se utilizan en la presente memoria, un "ácido nucleico" o una "molécula de ácido nucleico" se refieren típicamente a polímeros de desoxirribonucleótidos o ribonucleótidos (puros o mixtos) en forma de hebra sencilla o doble. El término puede abarcar ácidos nucleicos que contienen análogos de nucleótidos o residuos o enlaces de cadena principal modificada, que son sintéticos, naturales y no naturales, que tienen propiedades de unión, estructurales o funcionales similares a las del ácido nucleico de referencia y que se metabolizan de una manera

similar a los nucleótidos de referencia. Los ejemplos de tales análogos incluyen, sin limitación, fosforotioatos, fosforamidatos, metilfosfonatos, metilfosfonatos quirales, 2-O-metilribonucleótidos y ácidos péptido-nucleicos (PNA). El término ácido nucleico puede, en algunos contextos, utilizarse indistintamente con gen, ADNc, ARNm, oligonucleótido y polinucleótido.

5 Una secuencia de ácido nucleico concreta también abarca variantes modificadas conservativamente de la misma (tales como sustituciones de codones degenerados) y secuencias complementarias, así como la secuencia explícitamente indicada. Específicamente, las sustituciones de codones degenerados se pueden lograr generando secuencias en las que la tercera posición ("tambaleo") de uno o más codones seleccionados (o todos) está sustituida con residuos de bases mixtas y/o desoxiinosina. De este modo, una secuencia de ácido nucleico que codifica una secuencia de proteína descrita en la presente memoria también abarca variantes modificadas de la misma como se describe en la presente memoria.

15 El término "segmento", según se utiliza en la presente memoria en referencia a una secuencia de aminoácidos, se refiere a una secuencia contigua de aminoácidos que puede tener 10, 12, 15, 20, 25, 50 ó 100 residuos de aminoácidos de longitud.

20 Según se utiliza en la presente memoria, el término "heterólogo", cuando se utiliza en referencia a porciones de una secuencia de proteína o de ácido nucleico, indica que la secuencia comprende dos o más subsecuencias que no se encuentran normalmente en la misma relación entre sí que en la naturaleza. En un ejemplo, las secuencias heterólogas provienen de diferentes especies de bacterias. En otro ejemplo, las secuencias heterólogas son de diferentes cepas de la misma especie de bacterias. En un aspecto, las secuencias heterólogas son de diferentes cepas de *C. difficile*. En otro aspecto, las secuencias heterólogas son de una bacteria y un bacteriófago o de una bacteria y una secuencia sintética no natural de ADN.

25 Los términos "polipéptido", "péptido" y "proteína" se utilizan típicamente de manera indistinta en la presente memoria para referirse a un polímero de residuos de aminoácido. Los aminoácidos pueden ser referidos en la presente memoria por sus símbolos de tres letras comúnmente conocidos o por los símbolos de una letra recomendados por la Comisión de Nomenclatura Bioquímica IUPAC-IUB.

30 Los factores de virulencia son aquellas moléculas que contribuyen a la patogenicidad de un organismo pero no necesariamente a su viabilidad general. Tras la pérdida de un factor de virulencia, el organismo es menos patógeno pero no necesariamente menos viable. Los factores de virulencia pueden tener cualquiera de las numerosas funciones, por ejemplo, regulación de la expresión génica, suministro de adherencia o movilidad, suministro de una toxina, inyección de una toxina, bombeo de agentes antibióticos, o formación de recubrimientos protectores incluyendo biopelículas.

35 Los factores de adaptabilidad son aquellas moléculas que contribuyen a la viabilidad general, la tasa de crecimiento o la competitividad del organismo en su entorno. Al perder un factor de adaptabilidad, el organismo es menos viable o competitivo y debido a este compromiso, indirectamente menos patógeno. Los factores de adaptabilidad también pueden poseer una cualquiera de las numerosas funciones, por ejemplo, adquisición de nutrientes, iones o agua, formación de componentes o protectores de membranas celulares o paredes celulares, replicación, reparación o mutagenización de ácidos nucleicos, suministro de defensa frente a u ofensa hacia insultos ambientales o competitivos.

45 El término "célula productora" según se utiliza en la presente memoria se refiere a una célula que es capaz de producir o expresar una molécula de ácido nucleico codificante de difocina y que no contiene naturalmente dicha molécula de ácido nucleico. La célula productora puede ser capaz de sobrevivir y crecer en presencia de oxígeno y está transformada con un vector que contiene una molécula de ácido nucleico que codifica la difocina, que puede estar integrado en el cromosoma de la célula productora o puede ser episomal. La célula productora puede ser una bacteria gram positiva. En ciertas realizaciones, la célula productora puede ser una bacteria del género *Bacillus*, *Lactobacillus*, *Clostridium*, o *Listeria*. En algunas realizaciones, la bacteria es una especie del género *Bacillus* seleccionada del grupo que consiste en *subtilis*, *amyloliquefaciens*, y *megaterium*. En un aspecto, la bacteria es *Bacillus subtilis*. En un aspecto particular, la célula productora es una cepa de *B. subtilis* que carece de la agrupación de genes PBSX. En otras realizaciones, la bacteria es una especie del género *Lactobacillus* seleccionada del grupo que consiste en *acidophilus*, *casei*, y *bulgaricus*. En otras realizaciones más, la bacteria es *Listeria innocua*. En otra realización, la célula productora no patógena puede ser *Escherichia coli* o del género *Clostridium*.

60 Descripción detallada de los modos de poner en práctica la descripción

Las bacteriocinas de tipo R aisladas de *C. difficile* tienen actividad bactericida.

Para someter a ensayo la actividad bactericida, se prepararon los productos lisados de dos cepas de *C. difficile* Cd4

y Cd16 (de LC Fortier) haciendo crecer las células hasta la fase semi-log bajo condiciones anaerobias estrictas desde el principio hasta el fin y exponiendo a continuación el cultivo a 3 µg/ml de mitomicina C. Después de lisar las células bacterianas, los productos lisados se concentraron y se purificaron mediante centrifugación a alta velocidad (véase el Ejemplo 1). Se demostró mediante microscopía electrónica que estas composiciones contenían partículas de fago sin cabeza concentradas (Figura 3). Después de la concentración y la purificación, se evaluaron las preparaciones para determinar la actividad bactericida mediante un método de placa de reacción mediante el cual se aplicaron muestras a un césped superpuesto de la bacteria *C. difficile* diana, bajo condiciones anaerobias estrictas. Dado que la mayoría de bacteriocinas de tipo cola de fago suelen dirigirse a cepas diferentes de las bacterias productoras, se sometieron a ensayo las cepas diana que consistían inicialmente en un panel de 29 productos aislados clínicos de *C. difficile*. Después de una incubación durante la noche, se examinaron las placas para detectar la presencia de actividad destructora, como se indica mediante un aclarado en el césped de las bacterias *C. difficile* donde se aplicaron los materiales purificados. La Figura 4 muestra un análisis típico de difocina de Cd4 (dif 4) aplicada sobre la cepa Cd19135 de *C. difficile*.

Dif4 (difocina de Cd4) y dif16 (difocina de Cd16) demostraron diferentes espectros bactericidas basados en los análisis de placas de reacción (Figura 2). Dif4 mostró actividad bactericida sobre 10 de 29 cepas aisladas de *C. difficile*, mientras que dif16 tuvo actividad contra 8 de 29 cepas. Hubo cierta superposición de cepas susceptibles a ambas difocinas, pero las dos difocinas tenían espectros bactericidas distintos. Las difocinas de Cd108 y Cd 43593 tuvieron espectros de destrucción muy similares y destruyeron al menos 9 de las 10 cepas hipervirulentas NAP1/027/BI en el panel de 29 productos aislados de *C. difficile*. Cuando se sometió a ensayo adicionalmente, dif43593 destruyó los 18 productos aislados NAP1/027/BI independientes adicionales. Por lo tanto, dif43593 destruye las 28 cepas sometidas a ensayo de *C. difficile* NAP1/027/BI, Figura 2.

Identificación del locus de difocina

Las partículas de difocina, aisladas y purificadas a partir de la cepa Cd4, se desnaturalizaron, y los componentes de la proteína se separaron mediante electroforesis en gel de poliacrilamida - dodecil sulfato de sodio (SDS PAGE) y se detectaron mediante tinción con plata (Figura 5). Se cortaron dos bandas individuales, ~200 kD y ~40 kD, y se analizaron por espectrometría de masas. Se encontró que los péptidos de la banda de 40 kD coincidían con los productos previstos de dos marcos de lectura abierta (ORF) de *C. difficile* codificados en varias cepas de *C. difficile* para las que se han secuenciado los genomas completos, incluyendo la cepa de referencia Cd630 (Número de Acceso GenBank NC_009089.1). El primero de ellos fue el ORF 1363 (SEQ ID NO: 5), una proteína de tipo fago de 39.192 Dalton. El segundo polipéptido predominante en esta banda correspondía al ORF 1371 (SEQ ID NO: 14), una proteína de la placa basal de tipo fago de 39.565 Dalton. Dado que estas proteínas eran coincidentemente muy cercanas en peso molecular, migraban en la misma banda de SDS PAGE. La banda de 200 kD produjo un polipéptido dominante que correspondía a una porción del extremo C del ORF 1374 (SEQ ID NO: 17), inmediatamente aguas abajo de los otros dos ORF de tipo fago.

Dado que estos ORF se cartografiaban en muy estrecha proximidad dentro del genoma Cd630, se analizó la región circundante. Se encontró un elemento de tipo profago entre las bases 1574593 y 1596384, que incluye los ORF 1360A - 1379 (Figura 6 y SEQ ID NO: 1 - 23). De particular importancia fue que los genes estructurales codificados en esta región correspondían sólo a los componentes de la estructura de la cola de un fago *Myoviridae* típico; no se encontraron genes para la cápside, proteína de ensamblaje de la cápside, o proteínas del portal. También estaba ausente cualquier ORF que codificara la supuesta maquinaria de replicación del ADN o de empaquetamiento del ADN. Por lo tanto, este locus fue coherente con el de una bacteriocina de tipo R (de tipo cola de fago *Myoviridae*). Flanqueando los genes estructurales se encontraban diversos ORF que codificaban supuestas proteínas reguladoras de tipo fago. Dado que varios ORF de este locus codificaban polipéptidos encontrados en las partículas de difocina y dado el hecho de que la agrupación de genes se asemejaba a la de una bacteriocina de tipo-R, los autores de la presente invención asignaron el locus de difocina a esta región.

Los genes estructurales fueron codificados sobre una hebra y transcritos en la misma dirección. La organización de los genes se asemejó a la de los genes de las piocinas de tipo R y muchos fagos *Myoviridae*. Varias de las proteínas estructurales mostraron similitud de secuencia con respecto a los bacteriófagos *C. difficile* conocidos incluyendo los fagos Φ119 y ΦC2 (Goh et al., 2007; Govind et al., 2006). La cepa de *C. difficile* Cd630 también era conocida por codificar dos profagos intactos, ambos los cuales eran conocidos por ser inducibles (Goh et al., 2007). Varios de los ORF de difocina tenían similitud de secuencia con el profago 1 y 2 de la cepa Cd630 de *C. difficile*, lo que sugiere que estos fagos *Myoviridae* *C. Difficile* y las difocinas comparten ancestros comunes.

De particular importancia fue el ORF 1374 (SEQ ID NO: 17) de la difocina. Este gen codificó un polipéptido grande que se encontraba inmediatamente aguas abajo del ORF 1373, SEQ ID NO: 16, una ubicación en la agrupación que indicaba que el ORF 1374 formaba parte de la fibra de cola de la bacteriocina de tipo R, es decir, un dominio de unión al receptor (RBD). Dado que las micrografías electrónicas de lo que se han designado difocinas revelaron una estructura de tipo flor grande en la región de fibra de la cola, era evidente que esta estructura comprendía una proteína grande. El ORF 1373 codificaba la región de anclaje a la placa basal, BPAR, de la fibra de cola de la

difocina y parecía ser una forma truncada del ORF 1373 análogo en fagos de *C. difficile* que codifica el RBD así como la BPAR, dado que tales fagos carecen de un ORF 1374. De este modo, las fibras de cola de las difocinas naturales se componen de dos proteínas, ORF1373 y ORF1374, que forman una fibra de cola unida, mientras que los bacteriófagos de *C. difficile* tienen una fibra de cola compuesta por una única proteína, un ORF1373 algo más largo que proporciona las funciones de BPAR y RBD. Con este conocimiento, los ácidos nucleicos que codifican el ORF1373 (SEQ ID NO: 16) o el ORF1374 (SEQ ID NO: 17) se emplearon en la presente memoria como sustratos desde los cuales diseñar nuevas funciones de especificidad de RBD.

Las difocinas están muy extendidas en productos aislados de *C. difficile*

Se sometió a ensayo una serie de productos aislados clínicos independientes de *C. difficile* para determinar la capacidad de producir partículas de difocina que podrían tener diferentes espectros bactericidas. Los productos aislados clínicos Cd19123, Cd19145, Cd19126, Cd19108 (de RM Alden Research Laboratory, Culver City, CA) y Cd43593 de ATCC fueron inducidos con mitomicina C en condiciones anaerobias estrictas, seguido de purificación y concentración de las partículas que se produjeron. Los materiales purificados fueron luego aplicados por separado sobre céspedes de otras cepas de *C. difficile* en la colección clínica para detectar la actividad bactericida. Los resultados mostraron que cada uno de estos productos aislados produjo partículas (difocinas denominadas dif123, dif145, dif126, dif108 y dif43593 a partir de los productos aislados clínicos Cd19123, Cd19145, Cd19126, Cd19108 y Cd43593 de ATCC, respectivamente) con diferentes espectros bactericidas (Figura 2). Dif108 y dif43593 tuvieron espectros muy similares y amplios, mientras que la cepa Cd19145 produjo partículas con un estrecho espectro bactericida. Algunos productos aislados que se sometieron a ensayo para la inducción, incluyendo Cd19113 y Cd19150, se lisaron después de la exposición a mitomicina, pero no produjeron partículas de difocina detectables.

Se reconoció en la presente memoria que la cepa de laboratorio Cd630 de genoma secuenciado codificaba un locus de difocina; y, de hecho, este genoma se utilizó en la presente memoria para identificar los genes de difocina. Sin embargo, mientras que la cepa Cd630 se lisó después de la inducción con mitomicina C, no se produjeron partículas de difocina detectables. En su lugar, se produjeron pequeñas cantidades de profago 1 y profago 2.

Identificación y clonación de una agrupación de difocina.

El Cd630 no produjo partículas de difocina detectables tras la inducción de la lisis por mitomicina C. Por consiguiente, se clonó la agrupación de genes de difocina de Cd16 en su lugar. Primero se obtuvo un borrador de secuencia genómica de Cd16. Se identificó y se anotó un locus de difocina, análogo al de Cd630. La agrupación completa de genes de difocina de la cepa productora Cd16 se clonó (SEQ ID NO: 1) en un BAC y posteriormente se demostró que incluía todos los genes necesarios para producir una difocina activa en un microorganismo no patógeno y uno que no es un anaerobio obligado. *C. difficile* es una bacteria tanto patógena como anaerobia obligatoria, lo que la convierte en una célula de producción poco práctica para las difocinas naturales o las difocinas modificadas por ingeniería de ADN recombinante. La agrupación de difocinas aisladas también puede ser modificado por tecnología de ADN recombinante para que su expresión sea regulada por un promotor heterólogo que responde a un inductor o des-represor de molécula pequeña no tóxico añadido al medio de cultivo o de inducción del microorganismo productor. Tales moléculas pequeñas incluyen pero no se limitan a tetraciclina, anhidrotetraciclina, lactosa, arabinosa, xilosa y sus análogos no metabolizados, tales como IPTG para reemplazar la lactosa.

Difocinas

Las difocinas son bacteriocinas de epm de tipo R aisladas de *Clostridium difficile*, que son bactericidas contra otras cepas de *C. difficile*. Las difocinas se pueden aislar de cepas de *C. difficile* que se han hecho crecer en condiciones anaerobias en presencia de mitomicina C. En algunas realizaciones, la difocina es de un producto aislado clínico de *C. difficile* Cd4, Cd16, Cd19123, Cd19145, Cd19126, Cd19108 o Cd43593 de ATCC (denominadas dif4, dif16, dif123, dif145, dif126, dif108 y dif43593, respectivamente). En un aspecto, la difocina es de Cd4; en otro aspecto la difocina es de Cd16.

En otra realización de la invención, se proporcionan moléculas de ácido nucleico aisladas que codifican difocinas derivadas del genoma del género de bacterias *Clostridium*. En la presente memoria se describen moléculas de ácido nucleico que contienen la agrupación de genes del SEQ ID NO: 1 o del SEQ ID NO: 61. En otros casos, la molécula de ácido nucleico codifica una difocina que incluye uno o más polipéptidos seleccionados del grupo que consiste en los SEQ ID NO: 2-23; 49, 62-80. En otras realizaciones más de la invención, la molécula de ácido nucleico codifica una difocina que comprende polipéptidos de los SEQ ID NO: 4-16, 18 y 19; o polipéptidos de los SEQ ID NO: 66-80; o polipéptidos que tienen una identidad de al menos un 80% con los polipéptidos de los SEQ ID NO: 4-16, 18 y 19 o polipéptidos de los SEQ ID NO: 66-80. En un caso, la molécula de ácido nucleico codifica los polipéptidos de los SEQ ID NO: 2-23. En otro caso, la molécula de ácido nucleico codifica los polipéptidos de los SEQ ID NO: 49 y 62-80.

También se proporcionan difocinas variantes. Las difocinas variantes incluyen aquellas difocinas que tienen una secuencia de aminoácidos que es idéntica en al menos 80% a un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 4-16, 18, 19 y 66-80. En otras realizaciones, la difocina variante tiene una secuencia de aminoácidos que es idéntica en al menos 85%, 88%, 89%, 90%, 95%, 96%, 97%, 98% o incluso 99% a un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 4-16, 18, 19 y 66-80.

En algunos casos, la difocina variante puede incluir una región de anclaje a la placa basal heteróloga (BPAR), en donde la BPAR es idéntica en al menos 80% al segmento correspondiente de uno o más de los SEQ ID NO: 16, 54-56 y 78; la BPAR es idéntica en al menos 85% al segmento correspondiente de uno o más de los SEQ ID NO: 16, 54-56 y 78, la BPAR es idéntica en al menos 89% al segmento correspondiente de uno o más de los SEQ ID NO: 16, 54 - 56, y 78; la BPAR es idéntica en al menos 90% al segmento correspondiente de uno o más de los SEQ ID NO: 16, 54-56 y 78; la BPAR es idéntica en al menos 95% al segmento correspondiente de uno o más de los SEQ ID NO: 16, 54-56 y 78; la BPAR es idéntica en al menos 98% al segmento correspondiente de uno o más de los SEQ ID NO: 16, 54-56 y 78; o la BPAR es idéntica en al menos 99% al segmento correspondiente de uno o más de los SEQ ID NO: 16, 54-56 y 78.

En otras realizaciones, la difocina variante puede incluir un dominio de unión al receptor (RBD) heterólogo, en donde el RBD es idéntico en al menos 80% al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-53, o idéntico en al menos 80% a un polipéptido que contiene la región del dominio de unión al receptor (RBD) de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56. En otra realización, el RBD es idéntico en al menos 85% al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-53, o idéntico en al menos 85% a un polipéptido que contiene la región del dominio de unión al receptor (RBD) de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56. En otra realización, el RBD es idéntico en al menos 89% al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-53, o idéntico en al menos un 85% idéntico a un polipéptido que contiene la región del dominio de unión al receptor (RBD) de un polipéptido seleccionado del grupo que consiste en SEQ ID NO: 54-56. En otra realización, el RBD es al menos un 89% idéntico al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-53, o idéntico en al menos 89% a un polipéptido que contiene la región del dominio de unión al receptor (RBD) de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56. En otra realización, el RBD es idéntico en al menos 90% al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-53, o idéntico en al menos 90% a un polipéptido que contiene la región del dominio de unión al receptor (RBD) de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56. En otra realización, el RBD es idéntico en al menos 95% al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-53, o idéntico en al menos 95% a un polipéptido que contiene la región del dominio de unión al receptor (RBD) de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56. En otra realización, el RBD es idéntico en al menos 98% al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-53, o idéntico en al menos 98% a un polipéptido que contiene la región del dominio de unión al receptor (RBD) de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56. En otra realización, el RBD es idéntico en al menos 99% al segmento correspondiente de uno o más de los SEQ ID NO: 17 y 49-53, o idéntico en al menos 99% a un polipéptido que contiene la región del dominio de unión al receptor (RBD) de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56. En algunas realizaciones, la región del dominio de unión al receptor (RBD) comprende desde el residuo de aminoácido 51 al residuo carboxi-terminal de los SEQ ID NO: 54, 55, o 56.

En otra realización, las difocinas pueden ser modificadas genéticamente para que tengan espectros bactericidas alterados fusionando la cola del fago RBD al producto del ORF 1373 de la difocina. Aunque el ORF 1374 codifica el determinante del espectro primario o RBD de las difocinas naturales, esta proteína muy grande forma complejo con la proteína ORF 1373, y la proteína ORF 1373 proporciona la BPAR, es decir, une el RBD de la proteína ORF 1374 a la estructura de la placa basal de la difocina. El ORF 1373 es análogo a, y comparte identidad de secuencia de aminoácidos con, los genes de fibra de la cola de bacteriófagos *Myoviridae* tales como Φ CD2 (SEQ ID NO: 54), Φ CD119 (SEQ ID NO: 55) y Φ CD27 (SEQ ID NO: 56), así como con las fibras de cola de las piocinas de tipo R. El ORF 1373 (p. ej., SEQ ID NO: 16 u 78) de las difocinas comparte una identidad de secuencia significativa, particularmente en los primeros 160 aminoácidos de la porción N-terminal o BPAR, con las fibras de cola del fago *myoviridae* *C. difficile*, Φ CD2 (SEQ ID NO: 54). Las fibras de cola del fago son, sin embargo, más largas que la proteína ORF 1373 de la difocina y contienen un RBD C-terminal para reconocer sus dianas bacterianas. Las proteínas ORF 1373 de las difocinas no contienen este último dominio, cuya función RBD ha sido reemplazada por un polipéptido separado, codificado por ORF1374. Por lo tanto, el ORF 1374 se puede suprimir por completo de la agrupación de difocina y un RBD de una fibra de cola de fago, tal como el de Φ CD2, se puede fusionar con la BPAR de difocina, codificada por ORF 1373, generando de este modo una difocina que tiene una fibra de cola de fago y por consiguiente, un espectro bactericida relacionado con la gama de anfitriones del fago donante. De forma importante, debido a que las regiones de homología de la secuencia de aminoácidos entre las fibras de cola del fago *C. difficile* y la proteína ORF 1373 permiten fusiones funcionales exitosas entre las dos, se pueden seleccionar variantes de la gama de anfitrión a partir de fagos de *C. difficile* mutagenizados o no mutagenizados que pueden ser fuentes de nuevos RBD para la creación de difocinas modificadas con nuevos espectros bactericidas.

En una realización de una difocina modificada, se proporciona una difocina en la que el RBD ha sido reemplazado por un RBD de otra cepa de *C. difficile* o por un RBD de un bacteriófago que infecta *C. difficile*. En un ejemplo, la

molécula de ácido nucleico comprende una secuencia que codifica los SEQ ID NO: 16 ó 78 pero no contiene el correspondiente RBD nativo (es decir, la secuencia que codifica los SEQ ID NO: 17 ó 49, respectivamente); en su lugar, el RBD nativo se reemplaza por una secuencia heteróloga que codifica un RBD. En realizaciones particulares, la molécula de ácido nucleico contiene una secuencia heteróloga que codifica un dominio de unión al receptor (RBD) de una bacteriocina de tipo R de una cepa diferente de *C. difficile*. En un aspecto, la molécula de ácido nucleico contiene una secuencia que codifica los SEQ ID NO: 16 ó 78 y una secuencia que codifica un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 49-53 o la región de unión al receptor de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56. En otro aspecto, la molécula de ácido nucleico está compuesta por una secuencia que codifica los SEQ ID NO: 2-16 y 18-23 o los SEQ ID NO: 62-80 y una secuencia heteróloga que codifica un RBD de un polipéptido seleccionado del grupo que consiste en los SEC ID NO: 17 y 49-56.

En otros casos de una difocina modificada, se puede reemplazar el RBD de la difocina por una forma modificada de un RBD nativa. Un "RBD nativo" se refiere a un RBD que tiene una secuencia de aminoácidos que es idéntica a un RBD aislado o clonado a partir de una cepa de *C. difficile* o de un bacteriófago que infecta *C. difficile*. Los RBD nativos ilustrativos de una serie de cepas de *C. difficile* incluyen los SEQ ID NO: 17 y 49-53. Los RBD nativos ilustrativos de bacteriófagos que infectan *C. difficile* incluyen los SEQ ID NO: 54-56 (p. ej., desde el residuo de aminoácido 51 al residuo carboxi terminal). En algunos casos, un RBD modificado incluye un cambio en la secuencia de aminoácidos del RBD con respecto a un RBD nativo. Los ejemplos no limitantes de un cambio en la secuencia de aminoácidos incluyen sustitución, inserción (o adición), o delección de uno o más aminoácidos. En otros casos, una difocina incluye una sustitución con, o inserción de, un RBD derivado de un organismo que diversifica la estructura desplegando un Retroelemento Generador de Diversidad (DGR), como se describe en la Solicitud de Patente publicada US 2006-0121450, publicada el 8 de Jun. de 2006.

En algunos casos, la forma modificada tiene un espectro bactericida que es diferente del de un RBD no modificado o nativo correspondiente. En casos particulares, la forma modificada es idéntica en al menos 80% al RBD nativo. En otros casos, el RBD tiene una secuencia de aminoácidos que es idéntica en al menos del 85%, 88%, 89%, 90%, 95%, 96%, 97%, 98%, o incluso 99% a un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 17 y 49-53 o la región de unión al receptor de un polipéptido seleccionado del grupo que consiste en los SEQ ID NO: 54-56 y el RBD modificado tiene un espectro bactericida que es diferente del de un RBD no modificado o nativo correspondiente.

Bacterias diana

Las cepas de *Clostridium difficile* aisladas de los pacientes varían ampliamente por electroforesis en gel de campo pulsado y en su patogenicidad. Las cepas BI/NAP1 o de ribotipo 027 que hiperproducen toxinas son particularmente virulentas como resultado de haber perdido la función del gen *tcdC* que regula negativamente el nivel de expresión de la toxina A y la toxina B (McDonald et al., 2005).

De hecho, se ha demostrado que las cepas de *C. difficile* que albergan un alelo mutante específico del gen *tcdC* se propagan epidémicamente dentro y entre los centros de salud. Estas cepas epidémicas y altamente virulentas son bacterias diana especialmente importantes que podrían eliminarse profilácticamente del tracto GI de los pacientes portadores por aplicación oral de difocinas antes o poco después del comienzo de la terapia antibiótica tradicional. Las bacterias *C. difficile* que producen niveles de tipo salvaje de toxinas A y B son también importantes patógenos diana, ya que también son potencialmente letales, particularmente para pacientes mayores de 50 años o con comorbilidades (Bartlett JG, 2002).

Factores de virulencia o accesibilidad accesibles a la superficie de redireccionamiento tales como las proteínas de la capa S, prevalentes en cepas de *C. difficile*, ya sean hiperproductoras o no, ofrecen un medio atractivo para obligar a tales patógenos a comprometer su virulencia o adaptabilidad si emergen como resistentes a la bacteriocina de tipo R dirigida. Debido a la alta especificidad del RBD de las difocinas, los organismos distintos de *C. difficile* no representan una diana, una ventaja distinta y poderosa de las difocinas ya que no causarán daño colateral a las bacterias comensales de las bacterias del tracto gastrointestinal necesarias para la función normal y la buena salud del GI.

Una "infección" hace referencia al crecimiento de bacterias, tal como en un sujeto o tejido o célula no bacteriana, en donde las bacterias podrían causar realmente o potencialmente una enfermedad o un síntoma en el sujeto, tejido o célula no bacteriana. El tratamiento de una infección puede incluir el tratamiento profiláctico con sustancias, materiales, células productoras, o esporas de bacterias *C. difficile* destoxificadas capaces de producir difocinas tales como dif43593. Los ejemplos no limitantes de los objetos tratados incluyen órganos, tejidos y células donados; equipo médico, como un respirador o una máquina de diálisis; o heridas, tales como las producidas durante o después de la cirugía. Otros usos incluyen la eliminación de bacterias diana que pueden causar problemas después de un crecimiento adicional. En realizaciones adicionales, se utiliza una bacteriocina de epm para tratar alimentos, plantas o partes cosechadas de plantas con infecciones o contaminaciones bacterianas, o para tratar incidencias

ambientales de las bacterias diana, tales como en un entorno hospitalario o comercial.

Según se describe en la presente memoria, se puede utilizar una bacteriocina antibacteriana de tipo R para inhibir el crecimiento, la supervivencia o la replicación de una bacteria particular. La bacteria puede ser una cepa patógena o perjudicial para el medio ambiente, o se puede tratar de una manera profiláctica. Un microorganismo patógeno generalmente causa enfermedad, a veces sólo en circunstancias particulares.

Preparación y uso de difocinas

Las difocinas son partículas de aproximadamente 10 millones de dalton y por lo tanto pueden aislarse y purificarse mediante centrifugación diferencial, filtración diferencial, separaciones bifásicas acuosas, precipitación con polietilenglicol (PEG) y/o cromatografía de intercambio iónico para crear agentes antibacterianos orales de calidad biofarmacéutica. Se ha encontrado que las bacteriocinas de tipo R son estables a la congelación-descongelación y se pueden secar por pulverización para crear formulaciones estables.

En algunas realizaciones de la invención, se proporcionan métodos para producir una bacteriocina de epm de tipo R. El método incluye exponer una célula productora a una secuencia de ácido nucleico que codifica una bacteriocina de epm de tipo R unida operablemente a un promotor inducible sensible a un agente inductor a una concentración que provoca la expresión de la bacteriocina de tipo R y purificar la bacteriocina de tipo R. En un aspecto, la bacteriocina de elevado peso molecular (epm) de tipo R contiene uno o más polipéptidos seleccionados del grupo que consiste en los SEQ ID NO: 2-23 o SEQ ID NO: 49 y 62-80. La molécula de ácido nucleico es heteróloga para el ácido nucleico natural de la célula productora y puede estar contenida en el cromosoma de la célula productora o puede estar contenida en un vector de expresión episomal.

Como potentes agentes antibacterianos dirigidos, se utilizarán difocinas para eliminar, o descolonizar, *C. difficile* del tracto GI inferior de los seres humanos y otros animales con el fin de prevenir CDAD. Los animales y los seres humanos tratados con antibióticos de amplio espectro tienen un alto riesgo de desarrollar CDAD potencialmente letal si han sido colonizados por *C. difficile*. La descolonización es una utilidad particularmente atractiva de las difocinas debido a su ahorro de la microbiota GI saludable. Además, las difocinas se pueden administrar directa o indirectamente a través de células productoras administradas o esporas de bacterias *C. difficile* destoxificadas capaces de producir difocinas para reducir la carga de patógenos en CDAD aguda y/o reducir la alta incidencia o recurrencia o recaída de CDAD después del tratamiento satisfactorio por otras modalidades.

Modos de administración

Las bacteriocinas de tipo R se inactivan a un pH 4,0 o inferior, la acidez de un estómago alimentado que funciona normalmente y del duodeno superior. Sin embargo, las difocinas deben atravesar el tracto GI superior para alcanzar los patógenos bacterianos elegidos como diana que colonizan predominantemente el tracto gastrointestinal inferior. Por lo tanto, las difocinas se pueden formular mediante uno o varios métodos conocidos que protegen un agente vulnerable del ácido y las proteasas del tracto GI superior y suministrar dicho agente en estado activo al tracto GI superior distal o tracto GI inferior. Además, los animales pueden ser tratados con antihistamínicos tales como cimetidina o inhibidores de la bomba de protones para prevenir la acidulación del estómago antes de la administración oral de bacteriocinas de tipo R. De este modo, la administración oral de difocinas adecuadamente formuladas, células productoras capaces de producir difocinas, o esporas de bacterias *C. difficile* destoxificadas productoras de difocina a seres humanos o animales con estómagos normales o a aquéllos en los que la acidulación ha sido evitada farmacéuticamente, permitirá el suministro a la porción colonizada del intestino y, de este modo, posibilitará la eficacia. Basándose en el tiempo de tránsito intestinal, la frecuencia de la administración oral directa o indirecta de difocinas para descolonizar personas o animales asintomáticos puede ser cada 6, cada 12, cada 18, cada 24 horas, semanalmente o mensualmente. Las difocinas también se pueden administrar a pacientes con CDAD, o recientemente "curados" de CDAD, a frecuencias iguales o mayores. Particularmente para la gestión de CDAD activo, las difocinas se pueden formular y administrar directa o indirectamente por el recto mediante supositorio, enema o perfusión colónica.

Se puede administrar una difocina modificada de la invención a cualquier sujeto afectado, diagnosticado como aquejado o sospechoso de padecer una infección o contaminación por bacterias susceptibles a la difocina. Los ejemplos no limitantes de tal sujeto incluyen especies animales (mamíferos, reptiles, anfibios, aves y peces), así como insectos, plantas y hongos. Los ejemplos representativos y no limitantes de especies de mamíferos incluyen seres humanos; primates no humanos; especies agrícolas relevantes tales como el ganado vacuno, cerdos, cabras y ovejas; roedores, tales como ratones y ratas; mamíferos de compañía, exhibición o exposición, tales como perros, gatos, cobayas, conejos y caballos; y mamíferos de trabajo, tales como perros y caballos. Los ejemplos representativos y no limitantes de especies aves incluyen pollos, patos, gansos y pájaros de compañía o para espectáculo, tales como loros y periquitos. Un sujeto animal tratado con una difocina modificada genéticamente de la descripción puede ser también un cuadrúpedo, un bípedo, un animal acuático, un vertebrado o un invertebrado, incluyendo insectos.

En algunas realizaciones, el sujeto a tratar es un niño humano u otro animal joven que todavía tiene que alcanzar la madurez. Por tanto, la descripción incluye el tratamiento de afecciones pediátricas que comprenden infección con bacterias u otro microorganismo susceptible a una difocina de la descripción.

La descripción también proporciona el tratamiento o la prevención de una infección oportunista, tal como la resultante de un crecimiento no deseable de bacterias que están presentes en la flora microbiana de un sujeto humano o de un animal no humano. Una infección oportunista puede ser el resultado de una condición inmunosuprimida en un sujeto o el resultado de un tratamiento antibiótico que altera la flora comensal del tracto genitourinario (GU) o gastrointestinal (GI). Así, la descripción también proporciona el tratamiento o la profilaxis de sujetos inmunosuprimidos y sujetos expuestos a otros agentes farmacéuticos. Se puede utilizar una difocina con su actividad antibacteriana combinada con otro agente antibacteriano o antimicrobiano, tal como un antibiótico o un agente antifúngico como ejemplos no limitantes. Un "agente antimicrobiano" es un agente o compuesto que se puede utilizar para inhibir el crecimiento de, o para destruir, organismos unicelulares. Los agentes antimicrobianos incluyen antibióticos, agentes quimioterapéuticos, anticuerpos (con o sin complemento), inhibidores químicos de la síntesis o las funciones del ADN, ARN, proteínas, lípidos o la pared celular.

En algunas realizaciones, las difocinas, las células productoras o las esporas de bacterias *C. difficile* destoxificadas capaces de producir difocinas se formulan con un excipiente, recubrimiento entérico o portador "farmacéuticamente aceptables". Tal componente es aquel que es adecuado para su uso con seres humanos, animales, y/o plantas sin efectos secundarios adversos indebidos. Los ejemplos no limitantes de efectos secundarios adversos incluyen toxicidad, irritación y/o respuesta alérgica. El excipiente o portador es típicamente uno que sea proporcional a una relación beneficio/riesgo razonable. Los portadores farmacéuticamente adecuados no limitantes incluyen soluciones, suspensiones y emulsiones acuosas o no acuosas estériles. Los ejemplos incluyen, pero no se limitan a, excipientes farmacéuticos convencionales tales como una solución salina tamponada con fosfato, agua, emulsiones tales como emulsión de aceite/agua, y diversos tipos de agentes humectantes. Los ejemplos de disolventes no acuosos son propilenglicol, polietilenglicol, aceites vegetales tales como aceite de oliva, y ésteres orgánicos inyectables tales como oleato de etilo. Los portadores acuosos incluyen agua, soluciones alcohólicas/acuosas, emulsiones o suspensiones, incluyendo solución salina y medio tamponado. Los vehículos parenterales incluyen solución de cloruro de sodio, dextrosa de Ringer, dextrosa y cloruro de sodio, solución de Ringer con lactato añadido o aceites fijos. Los vehículos intravenosos incluyen reponedores de líquidos y nutrientes, reponedores de electrolitos (tales como los basados en dextrosa de Ringer), y similares.

La formulaciones y composiciones farmacéuticas adicionales descritas en la presente memoria comprenden una difocina aislada específica para un patógeno bacteriano; una mezcla de dos, tres, cinco, diez, o veinte o más difocinas diferentes, células productoras o esporas de bacterias *C. difficile* destoxificadas capaces de producir difocinas que se dirigen al mismo patógeno bacteriano; y una mezcla de dos, tres, cinco, diez o veinte o más que se dirigen a diferentes patógenos bacterianos o diferentes cepas del mismo patógeno bacteriano.

Opcionalmente, una composición que comprende una difocina o células productoras de la descripción puede ser también secada por pulverización o liofilizada utilizando medios bien conocidos en la técnica. La reconstitución y el uso subsiguientes se pueden poner en práctica como es sabido en el campo.

Una difocina se utiliza típicamente en una cantidad o concentración que es "segura y eficaz", que se refiere a una cantidad que es suficiente para producir una respuesta terapéutica deseada sin efectos secundarios adversos indebidos como los descritos anteriormente. También se puede utilizar una difocina en una cantidad o concentración que es "terapéuticamente eficaz", que se refiere a una cantidad eficaz para producir una respuesta terapéutica deseada, tal como, pero sin limitarse a, una cantidad eficaz para reducir la velocidad de división celular bacteriana, o causar el cese de la división celular bacteriana, o causar la muerte o disminuir la tasa de crecimiento de la población de bacterias. La cantidad segura y eficaz o la cantidad terapéuticamente o profilácticamente eficaz variarán con diversos factores pero pueden ser determinadas fácilmente por el experto en la técnica sin experimentación indebida. Los ejemplos no limitantes de los factores incluyen la afección particular que se está tratando, el estado físico del sujeto, el tipo de sujeto tratado, la duración del tratamiento, la naturaleza de la terapia concurrente (si existe) y las formulaciones específicas empleadas.

Habiendo descrito ahora en general el objeto de la invención, el mismo se entenderá más fácilmente haciendo referencia a los siguientes ejemplos que se proporcionan a modo de ilustración, y no se pretende que sean limitativos de la descripción, a menos que se especifique lo contrario.

Ejemplos

1. Determinación de la actividad bactericida de las difocinas

Se hicieron crecer cultivos de *C. difficile* en condiciones anaerobias estrictas en una cámara ambiental de Forma Scientific con una atmósfera de CO₂ al 10%, H₂ al 10%, N₂ al 80%. Todos los medios, tampones y placas se

redujeron en esta atmósfera durante al menos 24 horas antes de su uso. Los cultivos fueron aplicados en estrías sobre placas de agar selectivas para *C. difficile* (BD Diagnostics, BBL Cat. 222228), y se incubaron a 37°C durante dos días. Estas placas se almacenaron en condiciones anaerobias a continuación a la temperatura ambiente como soluciones de partida.

Para inducir difocinas, se hicieron crecer bacterias *C. difficile* en cultivos líquidos utilizando medio de Brucella (Difco) a 37°C sin agitación. A una DO₆₀₀ de aproximadamente 0,2, se añadió mitomicina C a una concentración final de 3 µg/ml. Los cultivos se incubaron después durante 3 - 16 horas. La lisis bacteriana se detectó mediante un aclaramiento visual del cultivo.

Los cultivos se eliminaron de la cámara anaeróbica, y los desechos celulares se eliminaron por centrifugación a 5.000 x g. Los sobrenadantes se hicieron pasar a continuación a través de un filtro de jeringa de acetato de celulosa de 0,2 µm. El producto filtrado se centrifugó a 90.000 x g durante 2 horas para sedimentar las partículas de difocina. Los sedimentos se resuspendieron en Tris 10 mM pH 7,5, NaCl 50 mM, manitol al 3% en un volumen de cultivo original de 1/50.

Las cepas diana se hicieron crecer en caldo *Brucella* durante la noche a 37°C. Se añadieron volúmenes de cultivo de 100 µl a 5 ml de agar de recubrimiento *Brucella*, reducido, templado (agar al 0,5%), se vertieron sobre una placa de agar *Brucella* (agar al 1,5%) y se dejaron reposar. Se aplicaron muestras de 5 µl de las preparaciones de difocina sobre las placas y se dejaron secar al aire (aproximadamente 30 minutos). Las placas se incubaron después anaeróticamente a 37°C durante la noche. La actividad bactericida se determinó por un aclaramiento, o falta de crecimiento bacteriano, en la posición o el punto donde se aplicó una muestra al césped.

2. Clonación del locus de difocina Cd16

Se obtuvo una secuencia del borrador del genoma de la cepa de *C. difficile* Cd16 mediante el análisis de 454 secuencias instrumentales de ADN genómico. Se identificó el locus o la agrupación de dif16 completo (SEQ ID NO: 1) por comparación con la cepa Cd630 (véase más arriba).

Preparación de un vector BAC de cadena principal. El vector de partida fue pETcoco1 (Novagen). Este se modificó para eliminar los dos sitios XhoI con los cebadores AV1419 (SEQ ID NO: 24) y AV1420 (SEQ ID NO: 25), que tienen extremos BbsI. Para ello, se amplificó una región específica a partir del ADN de pETcoco1 con estos cebadores, y posteriormente se cortó el producto de PCR con BbsI y se ligó de nuevo al fragmento del vector pETcoco1 más grande que se había cortado previamente con XhoI. Esta ligación destruyó los dos sitios XhoI de pETcoco1. Este último plásmido se modificó posteriormente mediante una estrategia similar para destruir los sitios EcoRI utilizando los cebadores AV1416 (SEQ ID NO: 26) y AV1245 (SEQ ID NO: 27). El vector resultante se denominó **SW251**.

Preparación de un vector pUC19 para que acepte fragmentos de la agrupación de difocina. El poliligador de pUC 19 (New England BioLabs) se modificó mediante digestión con EcoRI y HindIII y ligación en los oligos AV1372 (SEQ ID NO: 28), AV1373 (SEQ ID NO: 29), AV1374 (SEQ ID NO: 30) y AV1375 (SEQ ID NO: 31). Esto cambió el poliligador a NotI-NheI-KpnI-XhoI-EcoRV-BstBI-BbsI-EcoRI-NsiI-SphI-BamHI-Ascl. Esto fue denominado **SW232**.

Clonación de la agrupación de difocina en SW232. Tres fragmentos de la agrupación de difocina, SEQ ID NO 1, se amplificaron individualmente por PCR a partir de ADN de Cd16. El fragmento 5' (SEQ ID NO: 32) se amplificó con los cebadores 1368 (SEQ ID NO: 35) y 1289 (SEQ ID NO: 36), que tenían extremos NotI y XhoI, respectivamente. El fragmento medio (SEQ ID NO: 33) se amplificó con los cebadores AV1288 (SEQ ID NO: 37) y AV1366 (SEQ ID NO: 38), que tenían extremos XhoI y EcoRI, respectivamente. El fragmento 3', SEQ ID NO: 34, se amplificó con los cebadores AV1367 (SEQ ID NO: 39) y AV1300 (SEQ ID NO: 40), que tenían extremos EcoRI y BamHI, respectivamente. Estos tres fragmentos de PCR se clonaron por separado en SW232, y se denominaron SW241, SW242 y SW243 para las porciones 5', media y 3', respectivamente.

Clonación de la agrupación de difocina en el BAC SW251. Los tres fragmentos de la agrupación de difocina (en SW241, SW242 y SW243), cada uno de los cuales ha sido ampliado por clonación en *E coli* y purificado, fueron escindidos de los vectores SW241, SW242 y SW243. SW241 se digirió con NotI y XhoI, SW242 se digirió con XhoI y EcoRI, y SW243 se digirió con EcoRI y Ascl. (Obsérvese que estos sitios Ascl fueron parte del poliligador modificado SW232 descrito anteriormente).

Estos tres fragmentos se ensamblaron en SW251 que se digirió primero con NotI y Ascl. El plásmido resultante se denominó DG461 y contenía la agrupación completa de dif16. Este se amplificó en *E coli*.

Elaboración de un vector de integración de difocina para la expresión en *B. subtilis*. Se utilizó el vector de integración de *B. subtilis*, pDR111, que incluía partes del gen *AmyE* que flanqueaba una región de clonación/promotora y un gen de resistencia a la espectinomicina.

El poliligador pDR111 se modificó digiriendo el vector con HindIII y SphI y ligando en los oligos DG1 (SEQ ID NO: 41) y DG2 (SEQ ID NO: 42). Esto añadió sitios NotI y Ascl a pDR111. La región que contenía la totalidad de la región delantera y trasera de *amyE* con el poliligador modificado se amplificó después usando los cebadores DG9 (SEQ ID NO: 43) y DG10 (SEQ ID NO: 44), que tienen ambos extremos Bsal. Este fragmento se ligó en los sitios NotI y Ascl de SW251 (dando como resultado la destrucción de los dos sitios) y se creó **DG487**. Obsérvese que había nuevos sitios NotI y Ascl introducidos en DG487 por el poliligador modificado del inserto derivado de pDR111.

Tras la ampliación de DG487 en *E. coli*, el fragmento NotI/Ascl que contenía la agrupación de difocina de DG461 se escindió después y se clonó en el sitio NotI/Ascl de DG487. Este nuevo constructo fue denominado **DG488** y fue el vector utilizado para introducir toda la agrupación del gen de difocina en *B. subtilis* (abajo).

3. Expresión de una agrupación de genes de difocina en *Bacillus subtilis*.

El productor natural de difocina es *C. difficile*, un anaerobio obligado. Si se expone incluso a trazas de oxígeno, las bacterias *C. difficile* esporulan y mueren enseguida. La posibilidad de generar incluso cantidades traza de difocinas a partir de *C. difficile* cultivado es laboriosa y ardua, y ciertamente la producción de cantidades de difocina útiles para aplicaciones profilácticas o terapéuticas no es práctica bajo las estrictas condiciones anaerobias requeridas. Por consiguiente, toda la agrupación de genes de difocina de *C. difficile* se identificó primero y luego se aisló mediante clonación molecular y se introdujo la agrupación en una bacteria aerobia gram positiva, *Bacillus subtilis*, para la modificación adicional y la producción por ingeniería genética.

El vector de integración de *Bacillus subtilis*, DG488, se elaboró como se describe en el Ejemplo 2 y contenía todo el locus de difocina de 22.827 bases (SEQ ID NO: 1). Este vector se utilizó para recombinar el locus de difocina en el genoma de *Bacillus subtilis*.

La cepa receptora de *Bacillus subtilis* fue BDR123, que tenía un marcador de resistencia al cloranfenicol insertado dentro del gen *Amy*. Cuando esta cepa se transformó con DG488, se produjo una recombinación entre las secuencias *AmyE* anteriores y posteriores en el vector y las secuencias *AmyE* genómicas. Esto dio lugar a la inserción, en el genoma BDR123, de todas las secuencias entre las regiones de *AmyE* anteriores y posteriores de DG488 incluyendo el locus de difocina y el gen de resistencia a espectinomicina. Los recombinantes satisfactorios eran resistentes a la espectinomicina pero se habían vuelto sensibles al cloranfenicol debido a la pérdida de ese marcador genómico como resultado del evento de recombinación. Esta cepa de *B. subtilis* se denominó BDR123-488.

Puesto que el locus de difocina completo se insertó en el vector DG488 (Ejemplo 2), también se insertó por lo tanto en su totalidad en BDR123-488. Este locus de difocina insertado incluía todos los genes reguladores que se requerían para la expresión normal en *Clostridium*, y los genes de partículas de difocina estructurales estaban bajo control de estos genes y/o elementos reguladores. Debido a que *Bacillus* y *Clostridium* son bacterias relacionadas, se pronosticó que estos elementos reguladores de difocina funcionarían en el fondo de *Bacillus* y, como en su estado natural, serían inducidos por daño del ADN a través de un mecanismo mediado por RecA. De hecho, este fue el caso, y se indujo la producción de partículas de difocina en BDR123-488 por contacto con el agente perjudicial para el ADN mitomicina C y la destrucción de la cepa Cd19099, Figura 7.

Los genes reguladores de difocina (ORF 1359, 1360, 1361) se localizaron en la región 5' del locus en relación con los genes estructurales. También había genes reguladores (ORF 1377 (SEQ ID NO: 20), 1378 (SEQ ID NO: 21) y 1379 (SEQ ID NO: 23)) situados aguas abajo, 3' con respecto a los genes estructurales. Para eliminar estos últimos genes reguladores se generó DG491 a partir de DG488 en una sola ligación de tres vías. Se elaboró un fragmento de PCR a partir de DG488 mediante amplificación por PCR con los cebadores DG13 (SEQ ID NO: 45) y DG14 (SEQ ID NO: 46) y el otro mediante amplificación por PCR con los cebadores DG15 (SEQ ID NO: 47) y DG16 (SEQ ID NO: 48). Ambos fragmentos de PCR se digirieron con Ascl y SphI. A continuación DG488 se digirió con SphI y los dos fragmentos de PCR digeridos se ligaron en el fragmento de vector grande de DG488 digerido con SphI para producir DG491. DG491 se transformó en BDR123 (BDR123-491) para generar *B. subtilis* recombinante que contenía la agrupación de genes de difocina que carecía de los orfs1377 (SEQ ID NO: 20), 1378 (SEQ ID NO: 21) y 1379 (SEQ ID NO: 23). La agrupación de difocina modificada que carecía de estos ORF expresaba difocinas activas tras la exposición a mitomicina C (Figura 7), al igual que la agrupación de difocina de tipo salvaje en BDR123-488.

4. Caracterización de secuencias que determinan el espectro bactericida a partir de múltiples difocinas.

Una comparación del locus de difocina Cd16 (SEQ ID NO: 1) con el de Cd630 así como de otras cepas de Cd que se han secuenciado (QCD-66c26; QCD-23m63; QCD-32g58; QCD-63q42) y Cd4 (SEQ ID NO: 61) mostró que, con una excepción, todos los marcos de lectura abiertos (SEQ ID NO: 1 y 61) compartían una identidad de secuencia de aminoácidos de 89-100%. La excepción fue el ORF 1374. Esta secuencia excepcional fue variable entre todas las difocinas secuenciadas y, aunque de tamaño similar, compartía una identidad de secuencia de 30%. La posición del ORF1374 dentro de la agrupación de difocina era compatible con la de un dominio de unión al receptor. Se

determinaron las secuencias de los ORF1374 de las difocinas activas que se habían aislado y se encontró que estas también tenían secuencias altamente variables (SEQ ID NO: 17, 49-53). Una comparación de estas secuencias se muestra en la Figura 8. Además, los espectros de las difocinas aisladas (Figura 2) reflejaban las semejanzas o disimilitudes de las secuencias de aminoácidos del ORF 1374, Figura 8. Por ejemplo, las secuencias del ORF 1374 de dif16 (SEQ ID NO: 17) y dif126 (SEQ ID NO: 52) diferían por sólo 1 aminoácido y sus espectros bactericidas eran casi idénticos. Mientras que las secuencias del ORF 1374 de dif16 (SEQ ID NO: 17) y dif108 (SEQ ID NO: 50) diferían en 188 aminoácidos, y sus espectros bactericidas eran muy disimilares con poco solapamiento. Por esta razón y porque el ORF 1374 fue la única proteína variable en la agrupación de genes, se concluyó que el ORF 1374 era el determinante de reconocimiento diana y el responsable del espectro único de cada difocina particular.

5. Clonación y expresión de Dif4 en *B. subtilis*.

El locus difocina 4 se clonó a partir de Cd4 por métodos similares a los de la difocina 16. Sin embargo, se requirieron algunas modificaciones debido a la ausencia de un sitio EcoRI dentro de la agrupación de genes dif4, SEQ ID NO: 61. El plásmido SW251 (véase el Ejemplo 2 anterior) se modificó para que tuviera un sitio XhoI en el poliligador utilizando los oligos DG211, SEQ ID NO: 57 y DG212, SEQ ID NO: 58, para introducir sitios NotI y Ascl, respectivamente. Esto creó el vector DG577.

La acumulación de difocina del ADN de Cd4 se amplificó en tres fragmentos. El primer utilizó cebadores DG210 (SEQ ID NO: 59) y AV1288 (SEQ ID NO: 37) para introducir sitios XhoI y NcoI. El segundo utilizó cebadores DG209 (SEQ ID NO: 60) y DG15 (SEQ ID NO: 47) para introducir sitios NcoI y Ascl. Estos dos fueron clonados en DG577, previamente cortado con XhoI/Ascl para crear DG578. El tercer fragmento se amplificó utilizando AV1368 (SEQ ID NO: 35) y AV1289 (SEQ ID NO: 36) para introducir sitios XhoI y NotI y se clonó en DG578, previamente cortado con XhoI y NotI para crear **DG579**. Este último constructo que contenía la agrupación dif4 (SEQ ID NO: 61) era el equivalente de DG491 para dif16, es decir, carecía de orf1377, orf1378 y orf1379, las secuencias reguladoras presumiblemente innecesarias aguas abajo de los genes estructurales para difocina. El vector de integración para introducir la agrupación de dif4 en *B. subtilis* se elaboró tomando el fragmento NotI Ascl de DG579 y clonándolo en DG487 (véase el Ejemplo 2 anterior). Este plásmido construido fue **DG580**.

Para expresar dif4 en *B. subtilis*, se recombinó DG580, que contenía el locus dif4 sin orf1377-1379 (SEQ ID NO: 61), en el genoma de *Bacillus subtilis* La cepa de *Bacillus subtilis* receptora fue BDR123, que tenía un marcador de resistencia al cloranfenicol insertado dentro del gen *AmyE*. Cuando esta cepa se transformó con DG580, se produjo la recombinación las secuencias de *AmyE* anteriores y posteriores en el vector y las secuencias de *AmyE* genómicas. Esto dio lugar a la inserción, en el genoma BDR123, de todas las secuencias entre las regiones *AmyE* anteriores y posteriores de DG580 incluyendo el locus de difocina y el gen de resistencia a espectinomicina. Los recombinantes satisfactorios eran resistentes a la espectinomicina pero se habían vuelto sensibles al cloranfenicol debido a la pérdida de dicho marcador genómico como resultado del evento de recombinación. Esta cepa de *B. subtilis* se denominó BDR123 - 580.

Este locus de dif4 integrado incluyó todos los genes reguladores requeridos para la expresión normal de difocina en *C. difficile*, y como se esperaba y como se mostró anteriormente para dif16 en *B. subtilis*, se indujo la producción de partículas de dif4 en BDR123-580 por contacto con mitomicina C, Figura 9. De este modo, los presentes ejemplos proporcionan la clonación de los loci genéticos tanto para dif4 como para dif16 y la expresión de cada uno en *B. subtilis*, un ejemplo de una bacteria de producción aerobia no patógena.

6. Orf1374 determina los espectros bactericidas de las difocinas.

El Orf 1374 codifica un polipéptido grande pronosticado (~ 200 kDa) que se demostró por medio de espectrometría de masas que formaba parte de la estructura de la difocina purificada. Cuando se comparan las agrupaciones de genes de la difocina 16 y la difocina 4, la mayoría de los productos génicos, en particular aquellos que se preveía que eran componentes estructurales, eran casi idénticos a nivel de aminoácidos. La principal diferencia de la secuencia de aminoácidos entre las dos agrupaciones es el orf1374. Por esta razón y otras razones discutidas más adelante, se especuló que este producto génico confería la especificidad de la diana de las difocinas. Para someter esto a ensayo, se sustituyó el Orf 1374 de dif4 (es decir, la secuencia que codifica el SEQ ID NO: 49) en DG580 por el Orf 1374 de Cd16 (es decir, la secuencia que codifica SEQ ID NO: 17) para crear DG587. DG587 se integró en el genoma de *B. subtilis* BDR123 para preparar un recombinante BDR123-587, como el proporcionado anteriormente para dif16 y dif4. El BDR123-587 resultante se expuso a mitomicina y el producto lísado se trató de modo que se prepararan difocinas. Las partículas de difocina resultantes tenían actividad bactericida contra la cepa de *C. difficile* 19145, que era sensible a la difocina 16, y había perdido la capacidad de destruir la cepa 19137, que era sensible a dif4 (Figura 9).

Este experimento se perfeccionó aún más. La construcción de DG587 dio como resultado el Orf1373 que era una quimera de Cd4 y Cd16. Se elaboró un constructo de modo que el Orf 1373 de DG587 se restauró para que fuera 100% idéntico al Cd4 1373 original, SEQ ID NO: 78, creando así un constructo que era un reemplazo limpio sólo del Orf 1374, SEQ ID NO: 17. Este constructo se denominó DG603. Este constructo se integró en el genoma de *B.*

subtilis BDR123 y se indujo con mitomicina C como se ha descrito anteriormente. Las partículas de difocina resultantes tenían actividad bactericida contra las cepas 19099 y 19145 y perdieron la capacidad de destruir 19137. De este modo, se determinó el espectro bactericida de las difocinas por la proteína codificada por el orf1374, y cambiando el orf1374, cambiaba el espectro bactericida de la difocina, como se demuestra en la presente memoria.

7. Célula productora sin PBSX que no se lisa cuando se activa RecA.

El profago PBSX es ubicuo en *Bacillus subtilis* de tipo salvaje. Cuando es inducido, el profago es defectuoso ya que posee una estructura de cabeza atrofiada y contiene sólo pequeños fragmentos al azar de ADN. Está bajo el control de RecA, por lo que es inducido por agentes dañinos para el ADN, p. ej. mitomicina C, y otras formas de estrés severo para la bacteria. Cuando es inducido, provoca la lisis de la bacteria y libera partículas de PBSX. Con el fin de evitar la contaminación del medio de cultivo con partículas PBSX y eliminar la lisis de la bacteria productora *Bacillus subtilis* cuando la expresión de las difocinas era regulada modificando la actividad de *recA* o *dinR/lexA*, la agrupación de genes PBSX era eliminada de las bacterias *Bacillus subtilis* BDR11.

Se construyó PBSX desactivado siguiendo el procedimiento descrito por Liu *et al.* Brevemente, utilizando los cebadores y las técnicas de PCR de extensión solapada utilizadas en el trabajo de Liu, se suprimió el gen *araR* de la cepa parental BDR11 y se reemplazó por el gen de resistencia a la neomicina/kanamicina bajo el promotor de arabinosa de *Bacillus*, *P_{araA-neo^R}*, para preparar la cepa BDG2. Esta delección del gen *araR* fue confirmada por PCR y por la atribución de resistencia a la kanamicina.

A continuación, se elaboró un constructo de ADN para eliminar el propio locus PBSX. Para elaborar este constructo, se empalmaron los siguientes cinco productos de PCR mediante PCR de extensión solapada en un producto grande: 1 kb de la secuencia 5' del gen *xyiB*, amplificado a partir de BDR11; 1 kb de la secuencia 3' del gen *xyiA* Gen, amplificado a partir de BDR11; un gen de resistencia a cloranfenicol, *cat*, amplificado a partir del plásmido pJW034; *araR*, amplificado a partir de BDR11; y finalmente, 1 kb de secuencia que contiene el gen *xyiB*, amplificado a partir de BDR11. El producto de la PCR de extensión solapada se clonó en los sitios XmaI y SpeI de pUC19. Este constructo se linealizó a continuación con SacII y se transformó en bacterias de la cepa BDG2, que se sembraron en placas de agar LB con un suplemento de 5 µg de cloranfenicol/ml. Se recogieron las colonias de esta placa y se patched sobre placas de agar LB con un suplemento de 5 µg de cloranfenicol/ml o 20 µg de kanamicina/ml. Las cepas que eran resistentes al cloranfenicol y sensibles a la kanamicina se cultivaron durante 4 horas en caldo LB sin selección de antibióticos y después se sembraron sobre placas de agar LB con un suplemento de 20 µg de kanamicina/ml. Las colonias que crecieron en estas placas se sometieron a ensayo mediante PCR de colonias para detectar la presencia de genes PBSX. La delección de la agrupación de genes PBSX se confirmó en la cepa BDG 9 mediante la secuenciación de los productos de la PCR que abarcaba el sitio de los genes PBSX en la cepa wt BD123. Un análisis posterior demostró que, a diferencia de las cepas BD123 o BDG2 de *Bacillus subtilis*, BDG9 no se lisaba ni producía partículas de PBSX en presencia de 3 µg de mitomicina C/ml.

La cepa de delección de PBSX, BDG9, se transformó con el plásmido DG580, para crear BDG27. La integración de la agrupación de difocina Cd4 fue confirmada por medio de resistencia a espectinomicina. Se desarrolló BDG27 y se indujo con mitomicina C como se ha descrito anteriormente. Después de 16 horas se recogieron las células y se lisaron con BugBuster (Novagen) para abrir las células puesto que, sin PBSX, los autores de la presente invención, esperaban que las difocinas se acumulasen intracelularmente. Después de la lisis de las células con BugBuster, los residuos fueron eliminados por centrifugación, y el sobrenadante se sometió a ensayo para determinar la actividad bactericida contra la cepa 19137. La difocina producida por BDG27 mostró actividad frente a Cd19137 pero no frente a Cd19099, demostrando así que la difocina 4 era producida en esta cepa no lítica, con PBSX suprimido.

El término "que comprende", que se usa indistintamente con "que incluye", "que contiene" o "caracterizado por", es un lenguaje inclusivo o abierto y no excluye elementos o etapas de métodos adicionales o no citados. La expresión "que consistente en" excluye cualquier elemento, etapa o ingrediente no especificados en la reivindicación. La expresión "que consiste esencialmente en" limita el alcance de una reivindicación a los materiales o etapas especificados y aquellos que no afectan materialmente a las características básicas y nuevas de la invención reivindicada.

Referencias

- Anastasio, KL, JA Soucheck, y H Sugiyama, 1971. Bacteriocinogeny and Actions of the Bacteriocin. J. of Bacteriology 107:143-149.
- Bartlett JG, Onderdonk AB, Cisneros RL, Kasper DL. 1977. Clindamycin associated colitis due to a toxin-producing species of Clostridium in hamsters. J Infect Dis. 136:701-705.
- Bartlett JG, Chang T, Taylor NS, Onderdonk AB. 1979. Colitis induced by Clostridium difficile. Rev Infect Dis 1:370-8.
- Bartlett JG. 2002. Antibiotic-associated Diarrhea. N Engl J Med 346: 334-9.
- Bartlett JG, 2007. Clostridium difficile: Old and New Observations. J Clin Gastroenterol. 41 Supl 1:S24-9.

- Benson L, Song X, Campos J, Singh N. Changing epidemiology of Clostridium difficile-associated disease in children. *Infect Control Hosp Epidemiol*. 2007; 28:1233-5.
- Blackwell, C. C. y J. A. Law. 1981. Typing of non-serogroupable Neisseria meningitidis by means of sensitivity to R-type pyocins of Pseudomonas aeruginosa;
- 5 Blackwell, C. C., F. P. Winstanley, y W. A. Telfer-Brunton. 1982. Sensitivity of thermophilic campylobacters to R-type pyocines of Pseudomonas aeruginosa. *J. Med. Microbiol*. 15:247-251.
- Bradley. Bacteriocins. *Bacteriol. Revs*. 31:230-314, 1967.
- Campagnari, A. A., R. Karalus, M. Apicella, W. Melaugh, A. J. Lesse, y B. W. Gibson. 1994. Use of pyocin to select a Haemophilus ducreyi variant defective in lipooligosaccharide biosynthesis. *Infect. Immun*. 62:2379-2386.
- 10 Coetzee, H. L., H. C. De Klerk, J. N. Coetzee, y J. A. Smit. 1968. Bacteriophage- tail-like particles associated with intra-species killing of Proteus vulgaris. *J. Gen. Virol*. 2:29-36.
- Daw, MA, y FR Falkner, 1996. Bacteriocins: nature, function and structure Review Article. *Micron* 27:467-479.
- 15 Ellison, JS y JA Kautter, 1970. Purification and Some Properties of Two Boticins. *J. of Bacteriology*, 104: 19-26.
- Filiatrault, M. J., R. S. Munson, Jr., y A. A. Campagnari. 2001. Genetic analysis of a pyocin-resistant lipooligosaccharide (LOS) mutant of Haemophilus ducreyi: restoration of full-length LOS restores pyocin sensitivity. *J. inhibition Bacteriol*. 183:5756-5761.
- 20 Fortier, LC y S Moineau, 2007. Morphological and genetic diversity of temperate phages in Clostridium difficile. *Appl Environ Microbiol*. 73:7358-7366.
- Goh, S, PF Ong, KP Song, TV Riley y BJ Chang, 2007. The complete genome sequence of Clostridium difficile phage phiC2 and comparisons to phiCD119 and inducible prophages of CD630. *Microbiology*, 153: 676-685.
- 25 Govind, R, JA Fralick, y RD Rolfe, 2006. Genomic organization and molecular characterization of Clostridium difficile bacteriophage phiCD119. *J. Bacteriol*. 188:2568-2577.
- Jabrane, A., A. Sabri, P. Compe're, P. Jacques, I. Vandenberghe, J. Van Beeumen, y P. Thenart. 2002. Characterization of serracin P, a phagetail-like bacteriocin, and its activity against Erwinia amylovora, the fire blight pathogen. *Appl. Environ. Microbiol*. 68:5704-5710.
- 30 Kageyama et al. *Life Sciences* 9:471-476, 1962.
- Kageyama, M., K. Ikeda, y F. Egami. 1964. Studies of a pyocin. III. Biological properties of the pyocin. *J. Biochem*. 55:59-64.
- Kageyama, M., K. Ikeda, y F. Egami. 1964a. Studies of a pyocin. I. Physical and chemical properties. *J. Biochem*. 55:49-53.
- 35 Kageyama, M. 1975. Bacteriocins and bacteriophages in Pseudomonas aeruginosa, p. 291-305. En T. Mitsuhashi y H. Hashimoto (ed.), *Microbial drug resistance*. University of Tokyo Press, Tokyo, Japón.
- Keel, K, JS Brazier, KW Post, S Weese y JG Songer, 2007. Prevalence of PCR Ribotypes Among Clostridium Difficile Isolates from Pigs, Calves, and Other Species *J. Clinical Microbiology*, 45: 1963-1964.
- 40 Kingsbury, D, 1966. Bacteriocin production by strains of Neisseria meningitidis." *J Bacteriol*. 91:1696-9.
- Krogh, S, M O'Reilly, N Nolan y KM Devine, 1996. The phage-like element PBSX and part of the skin element, which are resident at different locations on the Bacillus subtilis chromosome, are highly homologous. *Microbiology* 142: 2031-2040
- Liu S, Endo K, Ara K, Ozaki K, Ogasawara N. 2008. Introduction of marker-free deletions in Bacillus subtilis using the AraR repressor and the ara promoter. *Microbiology*. 154: 2562-70.
- 45 Loo VG, Poirier L, Miller MA, Oughton M, Libman MD, Michaud S, et al. A predominantly clonal multi-institutional outbreak of Clostridium difficile-associated diarrhea with high morbidity and mortality. *N Engl J Med*. 2005; 353: 2442-9.
- McDonald LC, Killgore GE, Thompson A, Owens RC Jr, Kazakova SV, SambolSP, et al. An epidemic, toxin gene-variant strain of Clostridium difficile. *N Engl J Med*. 2005; 353: 2433-41.
- 50 Morse, S. A., B. V. Jones, and P. G. Lysko. 1980. Pyocin of Neisseria gonorrhoeae: mechanism of action. *Antimicrob. Agents Chemother*. 18:416-423.
- Muto CA, Pokrywka M, Shutt K, Mendelsohn MB, Nouri K, Posey K, et al. A large outbreak of Clostridium difficile-associated disease with an unexpected proportion of deaths and colectomies at a teaching hospital following increased fluoroquinolone use. *Infect Control Hosp Epidemiol*. 2005; 26:273-80.
- 55 Nieves, BM, F Gil y FJ Castillo, 1981. Growth inhibition activity and bacteriophage and bacteriocinlike particles associated with different species of Clostridium. *Can. J. Microbiol*. 27: 216-225.
- Pépin J, Valiquette L, Alary ME, Villemure P, Pelletier A, Forget K, et al. Clostridium difficile-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. *CMAJ*. 2004;171:466-72.
- 60 Scholl, D, y DW Martin, Jr., 2008. Antibacterial efficacy of R-type pyocins towards Pseudomonas aeruginosa in a mouse peritonitis model. *Antimicrob. Agents Chemother*. 52:1647-1652.
- Scholl, D, M Cooley, SR Williams, D Gebhart, D Martin, A Bates, y R Mandrell, 2009. An Engineered R-Type Pyocin Is a Highly Specific and Sensitive Bactericidal Agent for the Food-Borne Pathogen Escherichia coli O157:H7. *Antimicrob. Agents Chemother*. 53: 3074-3080.

- 5 Strauch, E., H. Kaspar, C. Schaudinn, P. Dersch, K. Madela, C. Gewinner, S. Hertwig, J. Wecke, y B. Appel. 2001. Characterization of enterocolitacin, a phage tail-like bacteriocin, and its effect on pathogenic *Yersinia enterocolitica* strains. *Appl. Environ. Microbiol.* 67:5634-5642;
- Sunenshine, RH & LC McDonald, 2006. *Clostridium difficile*-associated disease: New challenges from an established pathogen, *Cleveland Clinic J. of Medicine*, 73: 187.
- Williams, S., D. Gebhart, D. W. Martin, y D. Scholl. 2008. Re-targeting R-type pyocins to generate novel bactericidal protein complexes. *Appl. Environ. Microbiol.* 74:3868-3876.
- 10 Wood, HE, MT Dawson, KM Devine, D J McConell, 1990. Characterization of PBSX, a Defective Prophage of *Bacillus subtilis*. *J Bacteriology* 172: 2667-2674.
- Zilberberg, MD, Tillotson, GS y McDonald, LC., 2010 *Clostridium difficile* Infections Among Hospitalized Children, United States, 1997-2006. *Emerging Infect Dis* 16: 604-609.
- Zink, R., M. J. Loessner, y S. Schere. 1995. Characterization of cryptic prophages (monocins) in *Listeria* and sequence analysis of a holin/endolysin gene. *Microbiology* 141:2577-2584.

15

LISTA DE SECUENCIAS

	<110> AVIDBIOTICS CORP.	
5	<120> DIFOCINAS Y MÉTODOS PARA SU UTILIZACIÓN	
	<130> CF127922	
	<140>	
10	<141>	
	<150> 61/349,145	
	<151> 2010-05-27	
15	<160> 86	
	<170> PatentIn versión 3.5	
	<210> 1	
20	<211> 22825	
	<212> ADN	
	<213> Clostridium difficile	
	<400> 1	
	ggccgcaata cccactacac cttcgtcatc tttaaattta agagttttta ctattgaata	60
	ataaaggtat attccagtaa aaataatctt taaatacaag aaaaataaac tctttgggta	120
	tattaaaaag ctaaaaagtg taaatataaa agcaagtaga gtacttatcc tgtaaaagaa	180
	atctatttgt gtaatgtctt tatattttat cataaacacc gaatataaaa tgatgaaaat	240
	aattgcgacg attgcatata tggtaaataa catattttca agagtacat ttgaaattac	300
	tatccactta taccacataa ttggccaaaa taatagtgtc aagaacttaa aataattatc	360
	aaacaacttt tctttataca ttcatcaaac aacctttctt aacaaaagca tatatttggt	420
	tttagaattt taaataatat gatatcatta ttatatatta atattgaatt tatagaaacc	480
	aaaatttggt aaaataaata tatagatttt actgttaagc cagttaaaat tactactatt	540
	tttattatga aattggatca aatatgtaga aatacggcaa attagttaat attaaatatt	600
	tattatttcc aagttgtaaa gactgttttt ttaatgataa aaattctaatt cttttttgaa	660
	agaaaagtaat atccacatta agtatgtctg ccatttcata aacgcaagtg atgccagaat	720
	taattatggt tattatatct tcttcagtaa ttaagaactc acaagcccat ttaaggctt	780
	tattttcgca cttatctata ataatttttg tataataatc gttataagag gatacatagt	840
	atccaaggct agtgaaatga tgtccaagtt cttcagctaa gatggatgtc aatttttttg	900
	agttttggtt taaattactg agtaatgata taattttaat accatgtttg tttatatata	960
	gcccttctaa atcacctgca atataagtgg tataatgaat tattatctct tcttgagaag	1020
	ctaattcaaa aagcttatcc aaattattca taaaaatccc cctaaaatag aatgatggt	1080
	tgcctttaa ttatattaa agagcagaaa aatagactgc tcatcatatg gtttattttt	1140
	ttttatattt atttagtaaa aattctatat aatcattaag ttgttcttgt gcttcttcag	1200
25	gtaactcttc atgtggattt tttctatgtg cagctactgt atcaatattt tccttaacta	1260

ES 2 622 280 T3

aggttcttcc aagaaggtaa tcaactgata cattaataac atcagccaat ttgtttaaaa 1320
 tgtgttcac aggaaactctg ttttctgttt catagtaccc taagactcct tgggaaacgc 1380
 ctactttttc tccaagttct ctttgagtca atccaaattc ctttctaagt tctcttaatc 1440
 ttttgcaaa cattataaca ccaccttatg tatagattat aacaaattgt tctaaaaaat 1500
 aaaactaata aaatataaaa gaatatTTTT tctaaaatct attgataaag aacaaatgat 1560
 tctatataat ctaagtgtgg aagaacaaaa tattcttaat ggtaatggag gtataaaaca 1620
 atgtttaaaa ataacttgaa atattataga aaatgcaaag gtatgacaca aattcaactt 1680
 gccagaaagg ctggaattac aaatgactat atatctcaa tagaaagagg tataaaaaat 1740
 cctggtcttc ttatggctaa gaagatttct agtatttttag aacaaaatat agaagaagtt 1800
 ttttttatac agttatagaa caatatgttc ttgaaagttg tgagattagt aaaaaactgt 1860
 gcactaaaga gattattgta aatttgaagc taataataag tatataaaaa aggagaagta 1920
 ctatggaaaa caaaaaagat atattattta aagaaacaga tgaaagatta cataattata 1980
 agtatttgga tataaagata aagaatatta atttggacat aaaaagatgt gagaatgaat 2040
 actctggatg tggagcaatg gtatatacag aaaagactag taacacatat aacataagtt 2100
 cttctgtgga aaatgaggtg ttaaaaagag aggaaagatt aagaaaatta aaaatggaaa 2160
 aagaagatat agaaatagaa aaagagaaga tagaaaatgc tctaacatgt ctaaagata 2220
 tagaaatgga attttttaat cttttttata atagtaagac aaaaaacaat atgacatata 2280
 tttctatgaa actacactta gatagaacat cttgctacaa tttaaagaaa aaaatgatata 2340
 ttaaattgag tgagatatta taaaaaatat gacaacttta caacacttta tatacactat 2400
 tgcaacacta ggcaataaaa tatgtgagat aatgttattg tgaaagaaat ccatattgaa 2460
 ggaggtgata aattgaaaag aataatatta cctaaaaata tagaagatat ttgacaggaa 2520
 taaatgagat gtatatttaa aaatgactta tatcatttat agtaagatta tcagattaag 2580
 caagaatatt tagtgatagt gtggtgatta tttgcttaa tacaaggaaa tattagaaac 2640
 aattattgag attctcaaaa aaaactttac tgaaagtatt tttattgatg atgaaagtgt 2700
 gcaaggctct gaagggtcct gttttttgt aagtatacta tcagttattt gtacacctat 2760
 aatgttaaat acgaataata aagatattgt tatctctata aaatacttac caaaaccaca 2820
 gtcaaagagt attagaatgt atgaaatttc agatgaatta aataagttat tcaacagaaa 2880
 tataaaggta acagacagaa aattaaatat aacaaagcta gaacaaagta ttaaaaaaga 2940
 agagtcaatt tatgtattga actttacaat tacaactaat tatctggata gtgtatatga 3000
 agaagatgta gtatatgaaa atatggaaga aatcaattta aatttaggag agtgatagta 3060
 tggctatag attaccaagt atcaacatat catttaagga gctagctaca actgttaaag 3120
 aacgttcagc tagaggaata attgcaatgg tgcttaaaga tgctaaggca ctaggcttta 3180

ES 2 622 280 T3

atgaaataca tgaaaaagag gatataccag ttgatttatac tgctgaaaat aaagagtata 3240
 taaathtagc tttgatggga aatgttaaca ctccaaataa attattagtt tatgtaatag 3300
 aaggagaagc agatattcaa actgcattag attttttaga gactaaggaa ttttaattatc 3360
 tatgtatgcc aaaagcagta gaagctgata agactgctat aaaaaattgg ataattaaac 3420
 ttagagatat agataaggtt aaggttaaag ctgtattagg aaaagttgta ggaaatcatg 3480
 aagggataat taatcttact acagaagatg tgttagttagg agaaaagaaa tacagtgttg 3540
 atgagtttac aagtaggggtg gctggactta tagcaggaac acctttaagt caatcagtaa 3600
 cttatactaa gcttagtgat gtagttgata tacctaagat gacgaaagtt gatgcagaat 3660
 caagggttaa taaaggagag cttatactta ttaaggaagc aggggctata agaattgcaa 3720
 gagtagtaaa ttctttaact gagttaacag aagaaaaagg agaaatgttc cagaaaataa 3780
 aaatagttga cacttttagat attatacata gtgacataag aaaggtgata atagatgact 3840
 atataggaaa ggttactaac agttatgaca acaaatgttt attgatagta gctataaaaa 3900
 gttatctaga agaattagaa aagtcagcac ttatagaatc tgattctact gttgaaatag 3960
 attttgaagc acaaaaatcg tatttaaaat caaaaggagt agatttatct tataatgacat 4020
 tacaagaaat aaaagaagct aacacaggtt ctaaagtatt tttaaaagca aaaataaaag 4080
 tacttgatgc tatggaagat atagatttat caatagaaat ataggaggat tattaatatg 4140
 gcaaatatgg aagctagaaa tgtaatgagt ggtacttggg gagaactttg gcttgatgga 4200
 aacaaagtag cagaagtaaa gaagtttcaa gcaaagatgg aatttcaaaa agaggatatt 4260
 ataatagcag gtcaaatggg tactgatatac aagtatatgg gatataaagg aaaaggttca 4320
 ataactctat accatgttag ttcaagaatg cacaagttaa ttggagaaaa gataaagaga 4380
 ggttctgaac ctagatttgt tgctatatct aaattaaatg acccagattc ttatggagca 4440
 gaaagaatag cagtaaaaaa tatagcattt gatgatttaa ctttagctga ttgggaggtt 4500
 ggagtaaaag gagagataga agcacctttc acatttactg agtatgattt tcttgatata 4560
 atttagtttt atatttggtt ttatactgat atttagtaga tatatactta ataaatntag 4620
 gtagttaata agtaaaaaag ttagttgatt gaatttgatt gataaaggag caaataataa 4680
 tgaatgaaaa tggattatca aaaaatataa acatagtaga tttactttta aatgcagata 4740
 cagaaaactt agaagacca agtactatag ttgaacttaa gagattatca actatatttg 4800
 ggcaggaatt taaagtaatg tgtagagctt taacaataag taaagatgaa gagatacaaa 4860
 atacttgtct taaaattgat gaaaatatga aaacggatat agacttaccg gagatgcaga 4920
 tgcttacaat tatagaaggt gtttgtgatt tggatggaaa gcttttattt aaaaataag 4980
 aactaatgga taaatttaag gctccaacac caaaagaatt ggcaagaaaa ctattattac 5040
 caggtgaaat taccaacctt tatagaatac ttcaagatgt tatgggttat ggtaaaaatg 5100
 cagtgataga agaggtaaaa aactaatagg gacggatacc aagactacaa taatgtacta 5160

ES 2 622 280 T3

ttattggaag aaaaaaggta taagaccgtc ccttttttat gcaatggata aaggcgaatt 5220
 aaagcttatt gaagcttttt tgcgcttaga aattgaggaa gaagttgaaa aaatgaaaca 5280
 tggatatgga gtgtgtcctt tgacaggagg tggatatgtaa tgggaaatgt gagagaagaa 5340
 ggtataaata tgtatcttac agataattac acaccaaaaa tgaaccaaat tatatcagta 5400
 actgataatt ttaggagagc aactgtggct gtttcacttt ccaactaatgt aatggctagt 5460
 agcataaaaa attctattgg aagtgcgaagt agtagagtaa acagtttaaa ttcctcgta 5520
 agaaaagttc aaactactgc tagtagtgta agttcaacta tggcaaaatt aagttctagc 5580
 ataaaatgctg tttcaggagt tattggaagt ttaaatggaa gtattatgag actagcaata 5640
 actatagcta tgattattga ttattttaat aagttgattc aaaagaaaaa tgagtttaat 5700
 tcaaatatta tgattatatt aatatttaaa gctaaaagtg atgaagtaga aaaaactaaa 5760
 aataaattac ttggaattt aaaaaagatt ggtggcaaga tttggaatat cgtaataaaa 5820
 gcaaaagata tgactaagag agtgataagt agtatcttgg gaaaattaaa acgagtagag 5880
 aaacgtcctt atcaaggaag tattaatcct aaagatatgg taagtagtgc tatggctaga 5940
 attttgccta agttaatggt gtttaaaaaat actttttgga gtggtgtaat agctataaaa 6000
 gatatggcaa gtagcattat aagtaaagta tttcccaaat tgagattggt tgcaggtaag 6060
 gtatggagtg gtgcaatagc tgtaaaggat atggcaagtg gaataacttg ttcgataaaa 6120
 gggaaagatat ctgatttgac aaatggtgct actataggtg tcgctgtgaa aaaggggtgtt 6180
 gacttacttg gtcaggaaca aaatcagaaa gttgttctag aaagtgtaat gaaaagaaat 6240
 actgaaaaaa ctagcaaaaa agatgttgat aagtattatg acagtttagt aaatatggca 6300
 aatgatacgc cttttgacc tgaagatggt gttgcaatgg gaactaaagc taaaatgatt 6360
 agtaatatta ctggtggcaa aaaagaaaaa gatataactc aagctatggt agatgttaga 6420
 gctttaaata tgaatacaag tagtgaacaa gatgtatcag cagctttctt aagtgcagca 6480
 aaaggaaata tggaaatctct taatactctg gtaggagaaa attataaaac ttttgatgaa 6540
 gcattggaag gcataagtgt aaagcagatg gggttagcta aagaaatgag taatacaata 6600
 ccaggtataa tatcaggagc tcaaacaagc attaacaatg gtttgaagag tattgttaaa 6660
 ccttttgatg atatttttag tcaaggacta aagaaaaata aaacttttat agaaagtgga 6720
 ttagggaatt tagctggctt atctgaaaaa atggctggta aaataggcaa tgtaatgaat 6780
 ggtaagataa ttattggcaa caaatatgac cagatgcaat ctagaagtgt aaaaaatgga 6840
 aaagatgttt ctgattctac tcaatatcga atttctaag aggctgaaaa gcgtaaaatg 6900
 atggttgaaa ataagcaaga acgttttgaa aatcatgcag caacaatgat agggaatgca 6960
 ccaaaagcaa ttgttaacgc aggaagtaca ctattacaaa atattgattt tacagcatta 7020
 atagattcac tacttccagt agtaaaacta gtaataaatt tactagatag tataaacaat 7080

ES 2 622 280 T3

aaatcaccaa ttgcacaagg attaataagt atatttggt caatagtaac tacagcattc 7140
 caactaatcg gacctgtagt tgaagctggt agtcctatta tcacaagaat ttttactttt 7200
 ttagtgtaat atgcacctca aataaacaat tttatagaga cactgggtgt tatttgghaa 7260
 actgtatggg agaccttagg acctctgttg gaaactggat ggaaaattat agagccaata 7320
 ttgggagctt tttttaacat attagataaa gtatgtaaaa tagttaaaga tatatgcaaa 7380
 tggtgghaaa ctatgattaa taagataaaa aatggaagca tcacaggaac agttttaaat 7440
 ctagtghaaa agagtaaaaa aaattacaaa gataatccat atgctghaac aaaggctggt 7500
 gattctggtg aagcttattc aagtaagaaa ggtaataatg catttggtatt gaactatggt 7560
 ccttataatg actatcaaac cagactccat gaaggghaaa tggttttaac taaacaagaa 7620
 gcaaatcaat atagaagcag aaaaaatggt ggaaataaa acatagctaa gtttagctgat 7680
 acaatagtga ttagagaaga agctgatata gaaaagataa catcaaaatt agttgcaagt 7740
 atccaattgg cacagttagg ggtgtctta taatggaaat gtggcttaga caagcagaag 7800
 atagatttag atttccagta tttccatctt ccttttagtat taatggaaaa gctgctgtaa 7860
 actcttctag tatactcaaa ataggtgaag tagcaacttt tgggtggtgta gctcttaaaa 7920
 gcatttcaat atcaagtttt tttccaaata aagactacac tttctgtgac tatacaggtt 7980
 ttccatcacc atatgattgt gtaaataaga tagaaaaatg gatgaaggaa ggttttatat 8040
 taagatttac aattacggaa acaaatataa atatggaagt cataattgaa gggtttagtt 8100
 atgaagaaag agatgggact cgagatgtat attttacatt agatttaaaa gagtataaaa 8160
 gaataaagat accaaaagta actccaaaac aataactatt atagataata agttgtaagt 8220
 aactgctgat agaattaaa gaaaaggcag gtgatttttt attattaaga tttgggtaca 8280
 cataaaaaac ggaagtatat atgacataac tgacatagta gacaaggat catggtcag 8340
 tgattataaa tctccatcaa ggacactaga gttttcaata atacaatcat catttgatgt 8400
 aaatttcaa caaatcgata taccaatagc tagtacagtc tgtttctatg tagatgagaa 8460
 agaactcttt agaggaatga taattaatag gtctaaagat tcaagcagta atgaaattag 8520
 ttttgtatct aaagatatgg gatttttact tacacaaagt gaagtgtcat acaattttaa 8580
 agataagtta gttgaagaca tagcaaagca agtatttgct gaaaataggc tttcagttgg 8640
 aacaatagca aagaccaatg tcaagtatac aaagatgttt ataggagtaa atggttatga 8700
 cacaataatg agtgcataa cagaggcaag taaaaagaca aagaaaaagt atatgataga 8760
 ggtaattta gataagttta atgttattga aaaaggaact gttacattaa gtgttatggt 8820
 tgaagaggya tttaatatta taataaccac cttttcggag agcatggaaa atgtaaaaaa 8880
 taaagtaata gtggtagacc agtatggaag caagattagc gaaaaaatag ataagaaat 8940
 ttttaaggaa gtaaatgtaa taatgcaaaa agtaattcag caacaagaaa atcaagatgt 9000
 agatattgat agcgagttta atgggataga aaaaagctgt tctcttaaag gttatggaga 9060

ES 2 622 280 T3

tgtaagttgt	ataactggta	gaggagtaaa	agttaaagat	tcttatacaa	agctttagg	9120
actatthttat	atagatacag	acaaacatac	ttggcaaaat	ggagaatatac	aaattgagct	9180
tgaacttaat	tttcaaaatc	ttatggatga	aaagtcagca	ggacaggatg	aacctaagga	9240
agaaaagtaat	ttagggggag	aagattatgc	aggaggaaaa	gagtttacag	cagaatttac	9300
agcttactgt	cctagaaaag	aagaagggtg	agatacagat	tgtagaaaga	aaaaacttga	9360
cccatctaaa	aaacttgccg	tgctcctatg	gttggtaaat	atgagcaaac	ttattataca	9420
aaagagtttt	taaataaaca	tcctttattg	aactatggag	atgaaatata	ggtaattaca	9480
ggagtttctg	gtcgtgatgg	agtctataaa	gtaaatgacg	taggacctgc	aataactata	9540
gaaaaaatg	gaacatacca	tatagatatt	ttatthggaa	atgthgaaaga	agctagtaaa	9600
tttggaaaga	gaaaaggaaa	aattattatt	ggtggttatt	ctggtaatgt	atctgataaa	9660
gctaaaatag	taatatcaga	ggcaaaaaaa	catctaggta	aaccttataa	atggggtgga	9720
aatggaccaa	gtagthttga	ctgthctggg	ttaatggtct	actgthttaa	aaaagthaat	9780
gthtagthtg	caagaacgtc	aatcaacaa	tctaaaaaag	gcaagaaagt	agaacaaaaa	9840
aatcttcaag	caggagattt	agthththth	cataatccag	tcagccatgt	tggattatat	9900
atagthaatg	gagaaththth	acatgctcca	caaaaagggtg	atgthagthta	aataagthaa	9960
tthaaagthga	gaagagattt	taatacagct	aggagagtht	tataaaaagga	tggthgatata	10020
atggctaatc	caataaatga	atthtatagga	ataataagag	aagaaggaaa	gthtcataat	10080
caacctthct	ththtatggg	aaaththaaag	tthaatthacca	gaththaaaa	thagagacaaa	10140
thaacatcata	tthagaaaaag	aagaththth	gathagathagth	tggthgathg	athagacagct	10200
agaaacattt	gacacagaaa	caaatcaaga	acaccagcat	gaagthaaaa	atcctththt	10260
agathaactht	gathctgggg	athatgthaat	aatgththaga	athaggcgaaa	aatthgctgt	10320
tgthaaagthaa	thggthgagct	tathaatgagth	acathattht	ctththtatagg	tgtccagag	10380
gathtatatct	taccththaac	agaagaathg	ccaathcttht	gtgaagthggc	atgggaththt	10440
gaaaaagathg	aacctaththt	agaaaaaggth	gactththaaa	thaatgaaaa	aaaaagaaagc	10500
ctthaaaagtht	tggthataca	agthgthataaa	gacaaataga	tathgacathg	agathatactc	10560
ththagaathat	gggacagagc	ththcagaact	aataggacaa	aaathatacaa	aagthcttht	10620
agaaagthgaa	gctagthagat	tcathaaaaga	ggccctthcta	athaaatccat	athatathaga	10680
agthaaacgth	aaaagthgct	actththaacag	agacgthattg	agthgcaaatg	thaaaagthatc	10740
cactatctat	ggggaggtg	aaathaaatgt	athagthgacca	gacathatgaa	gthaaataaaaa	10800
athagaaactct	tgaaaaathat	aatctthgata	ththataaagg	agaagththct	ththcthaaca	10860
acathgtht	tggaaathaat	ctagaaactth	cgaagathata	tctagaaactth	tcaagathatc	10920
athaaaatggc	thththatacaa	gacacathata	accagththct	tgathaaaaga	gthcaathgaa	10980

ES 2 622 280 T3

ttggtgtata tagaaagtta ggtacagagt caaatggaga agttgaattt attggagaga 11040
 aaggaactgt aataaataat ggcacaataa tatcatatag agatttacta tttgtagtaa 11100
 taaaagatgt aactattggt agtgaagaag gtgacaatag cccagttcaa gctctggaag 11160
 ttggttaagaa atataattta cctacaaatt gtgaatttaa actagttgat aatatactctg 11220
 gagtaacaaa gattactaac acaagaagtt ttgaaggtgg tacagatata gagacagatg 11280
 aagaactaaa agaagattt tataaaatcc aaagaaatca agctacaagt ggaaataaag 11340
 ctactatga agaatggctt ttggaagtag atggagtcta taatgttaag gtttatccaa 11400
 gatgggatgg tccaggaaca gttaaggtct tgatatttgg ggaaaataat caagctggtg 11460
 atacagaaac gattgaaaggt gtgcagcaac atatagatga agagaagcct attggaccaa 11520
 ctataacagt tgtgacacca ttaccaatag aaataagtat aagtgcagta atgaaactag 11580
 aagatggata tacattagac aatgtaaaag aatctttcct agaaagtata aatacatact 11640
 ttagagatat tagaggagag ataactata caaaagtcag gggaactactt ataaatacta 11700
 ctggtgtaca cgatttaagt aatctactta taaatggaag tacagataat ataactatta 11760
 atgaagataa aatacctagt gtaacaactg ttaattttag tgaggtggaa aatcaatgaa 11820
 gctaattgat aaactacat catttgatag aaattacatt gtagaggaga tacaaggtgc 11880
 atacgataca gaattaaata ttcttaaaga agatattgat gataccttta accaattatt 11940
 tgttgataca gcgacatggg gattagatat gtgggaagac atactctgca ttgaaaaaa 12000
 agaacttgat ttgacacaa gacgtagcaa tataaaagct aaaatgagaa gcagaggtac 12060
 tagtactatt gaagttataa aaagtatatg tgaggcatat acaaaatcag aacagatat 12120
 aaaagtttat agtgatgaat ttacattcgt attgagtttt atagcaaata actgtgacta 12180
 taaaactcct ttagattgta gcgagatgat tgaaagagta aaacctgctc acttattaca 12240
 ctatttagaa ccaataatac tagataaaag tatggtctat tgtggtggag gtatggtatg 12300
 tagtgaagag gtaaaagtcc atccataactt tgaaccaatt ataaaatgta gtgctggtgt 12360
 aaactgtgga gctggaatgt taagtagaga agaaataaag gtttatcctt taagcattaa 12420
 atgcattgaa aataattgta agattaatat agctattgca aatgatacag gcgtagaaaa 12480
 thtagtagtt tatoctaaat cggaggtggg ataattggaa gaaaaatttt atataatatt 12540
 aaccaaattt ggtagagaaa aaatagcaaa tgcaactgca ctaggagagc ttgttgatt 12600
 aaccaagttt caagttggag atagtaatgg agaattattt gagccaacag aggaacaaac 12660
 tgctttaaag aatgtagttt ggaagggaaa tataaattct ctaagaattg atgaaaaaa 12720
 tcctaattgg atagttatag agactatfff accaggaaca gttggtggat ttatgataag 12780
 agaagctgct gttctggata atgagaataa tataatagct atwggttaagt atccagagac 12840
 gtataagcca cgtgctgaag atggcagtat taaagatttg gttgtaaaaa tgattttaca 12900
 attgtccaat acttcaaattg ttacattaga agtagaccgg acggttggtt ttgtaactca 12960

ES 2 622 280 T3

aaaggatatt caagatthag atgataagtt tgataaaaat ataaaagaaa taaaagtaaa 13020
aattggagat acagatatat taactacaga ttctaaagat ttatcaggag ctataaatga 13080
ggtagttaa aaaatagaaa atatatcttt tgatgatggt ataagtggc aaatacaaac 13140
tgatatatca gtattaaaa atagctataa caaattatct gaaaaagtgc tagatatatt 13200
aatataccta gaattagagt cagaagtaac tgtagatgag gctggttatt ggtatgatac 13260
attagcaaat ggaaataaca tagtagctat agaaggcctt aagttagatt taaatagaaa 13320
atgtataaca ggtgaaattg gtaatgtgat ttttagagat gtagtattac catttagtgc 13380
aaatagagtt agatatatac atgatatgga taataacttt gttgagacaa aatctagtaa 13440
cacttattta aaagaacaaa aagatataac tctaagtaaa tattcatatg aaataagata 13500
aataaaggag gtagtactaa taatgaagca aaataaactt ttacagcgtg gtgcttattt 13560
taatgataag aacatattga ttgatgattt tgataaaaaga tataatgatt atgattttgt 13620
agaatTTTTT actggtataa gtaatagtac ctttggttta aaatcagatg gtaatttata 13680
tgcttgtggc gataatacag gttttcaact aggacttggg aaagattcgt cagagagaag 13740
gatgtttagt aaagtaaaaa ttgataatgt aaaatatgta tcttgtgggt caaacacag 13800
tgtagcagta actaaagatg gatttgcata tggagcagga acaagtaatg taggtcaatt 13860
agggtgaatt gagtctacag tatattatga atttactaag ctaccaatag atgatgtaaa 13920
aactgttgca tgtggttatg actttacatt tgtgcttaa aatgatggaa cattatattc 13980
agcaggttta aactcaagtg gtcaacttgg actaggtgat actaacaata gagctacttt 14040
tactaaagta aatatagata gtgtgaaaga tgtagtgact tataatcaat ctgtatttat 14100
cataaaaatg gatgggacag cacatgcatg tggattaaat tcaaattggc agttgggaat 14160
taatagtact ttaaaaaaaa gtgtatttaa taaaatagaa ggtatggata atgtaaaaca 14220
gatagcgtgt ggtagtagtc atacaattct tattaagaat gatggaacta tgtatactac 14280
aggctataat ggagttggtc agcttggtag aggaaataat aataattcaa ttgtatttac 14340
tctttctagt ataaataatg ttaagtatgc ttcttggga aataatcata ctatgatatt 14400
aaaatacgat aatacactgt ttagtacagg acaaaacaat tatggccaac tagccaatgc 14460
caataaagat gtagcatcaa gaaatacttt tgctaagggt aatgtagaaa atataaaaaga 14520
tattaaatgt ggttctcaat ttaatTTTTT aataaatgggt tcaaaagaga tatttgtatc 14580
tggctgtaat ttagcaggtc aacttggttc atTTTTTcat acaacttttc tgtatgagtt 14640
ttcaaattgt caatcttcaa atttagataa ttattcagggt ttattggtta atgatgatta 14700
tttatatggt acaaaggaca atagtgaatt tttaaatgta aagttaagtg ataattttca 14760
agattataag aagatagagt taacagatag caatatgttt attgttatga atgatggtac 14820
attgtatgct tgtggtttaa ataattatgg acagttagga ttgggagata ctgttaacag 14880

ES 2 622 280 T3

gtcagttatg actaaggtgg ataatagataa tgttttggat ataaaaggaa acggaaactc 14940
 aacttttgtg cttagaata atggaacatt atattcatgt ggtttaaata gtaatggaca 15000
 attgggttta agagatgaag ttaatagaaa tatatttaca aaaatagaaa tagagaatgt 15060
 aaaggatfff tgtgtaggaa gcaattatgt catagcttta aatcactcaa aagaagtata 15120
 tggatgggga aataatcctt ataataatat agaaaaaact tctaattatc catataagca 15180
 gggaaataagt aatattgaaa agatagcagc atatgattat tctgtatata tgataaacag 15240
 tgaagggaaa ctatatgttt ctggatataa ttataattat caattaggta aaggaaataa 15300
 tagtaaccaa agcaaagcat tagtatctca atgtagaaca aattcaacat cttctacatc 15360
 aaatggactt agaacgttac ctaaaataac taatgttttt cctttttatg atggttgtgc 15420
 aataattgac gaaggaggtt atgtttatftt aacaggatat catggatatt taagaacatt 15480
 aaatagcagt ccaagtatat ctgattattc aagatatgga acttttattg aggctacaaa 15540
 ttcaaatcat aatacttatt ttatacaaga gactgatttt agtgggaattg aaaaagtaat 15600
 agggatgtca aataatata ttttttttaa gaaaggaagt tcatatatta ctggatatcc 15660
 aaaaacattt ggctcaacca ttactggaca tagaagttat actagtatta attctgagag 15720
 ttctaattta ggaagtaatt ttataatata tcatagtaat tccaagttat atggaaaag 15780
 gattgcta atgtgggcaat ttgggaattc aacaaatata gatggcacia gtaactatga 15840
 tacaggatta aaagacataa aagatataat tgtaaaagga aatactgtag tagtagtaga 15900
 taaaaataac aatataatag taacaggaat gaatcagaat aacaaacttg ggatagggga 15960
 atataacaac gaaccagtaa aaaaattcac aatataact gaacaatcaa actcatttat 16020
 atttatggat gatataaaaag aaattacaac atcaagaaat acaatgttta tagtaaaaaa 16080
 tgaatggaaca gcctatgcca caggaaataa tagttctgga caattaggat taggtgacac 16140
 aataaataga aataagttca ctcagataaa ccttgataat ataaagaaa tatcaacaag 16200
 tatagatggt aacacaacat ttgcaattag aatgatgga acactatact ccacaggatt 16260
 aaataccaaa ggacaactgg gattaggtga tatagtaaat agaaatacat ttaccaaaat 16320
 aaacatcaa aatgtaagag atgttgtttt agggactact cactcgcattg caatcaaaga 16380
 tgataacaca ttatattcat gtggagaaaa cactcatggg caactgggct taggaagcga 16440
 aagcaaccat ccagacgtat tgacatttac tgtaacaat ataactaatg taagagatgt 16500
 gtactgctca gatacaaca catttattgt aaaggacaca aacattgcat attgttgtgg 16560
 atacaataat aattcacaac taggtatggg aaatactact gaccagtata gttttataaa 16620
 gtgtatggaa aatgtaaaaag aagttatacc aatgaaata aatactata taataacaat 16680
 ctataaact gcatatagta caggtttaaa tactgattat tgcttaggtc taaatagtaa 16740
 tagcaatcaa agttcatttt ctgaaattcc aatttcaaat gtagtaaaag tagctccaaa 16800
 cagaaataat gcagtaacttt tacttacaag tgaaggggat gtatatactg caggcaaatg 16860

ES 2 622 280 T3

tagtaatggt	tcaggtacag	gaagtgagac	tccagagaag	attaaaaaaaa	tagcatcaaa	16920
ggcaaaggat	attggaatga	attatagatg	tggacattat	gtaagtgata	atggagacct	16980
atatggtaca	ggttttaata	ataatggaca	attaggtggt	ggtgatgtaa	caaaaagaga	17040
tacatttata	aaaaccaata	caagagtaaa	gaaaataactt	cctttagaat	atgcaaatat	17100
agcaataaaa	gataactaatg	atatatatat	ttgtggatta	aataactatg	gacaattagg	17160
tgttggaaat	agatacgata	gtagaaataa	tgataataga	atatttaatt	ataagcatat	17220
gaattttgta	atgggtgatt	tgacatctat	taaaaacaga	cataacttta	tacttctaaa	17280
caataagata	gtgataccta	ccacaaaaga	catagattat	ggtttagtat	taggaaatth	17340
atacaaaaga	gacctttata	ctgagcttcc	atatgaagat	ataaaagaag	tatctatttc	17400
taagactcat	attattatat	tacttaatga	tggaacaatg	tatggatgtg	gtacaaacta	17460
ccatggagaa	ttattgcaag	acttgtctat	aatcaagtg	gatgaatttg	tgcaagattaa	17520
tgtatcagat	gtaaagcatg	tttcatgtgg	agataacttt	acttatttta	taaaatctga	17580
tgatagtctt	tggtctattg	gtaaaaattc	cgaatatcaa	ttaggtatag	gtcacaataa	17640
tccagttact	gaattacaaa	gaattacaac	tatatctagc	tgtaaagaag	tacattgtgg	17700
taaaaactat	acattagtag	taactacagg	taatgaatta	tttgtacaag	gatataatga	17760
taagggagct	ttaggattag	gaagcgatag	tgaaaatact	ataattaagt	tctttacaaa	17820
agcactaaca	gacataagag	aaataaaatc	ttatggaagt	gaccatatat	tagtacttaa	17880
aatgataat	tcagtatggg	ttactggaaa	aatagggat	gtatataaaa	ttgaacaacc	17940
agtagaattt	ttaaaagaat	ttactatagt	acctatttct	gaagatgtaa	atacagtaaa	18000
ggatgtactt	gcaacagaca	atacattata	tattatatca	gaagtaggaa	cgacaaatgc	18060
tgctatagaa	attactgaaa	aatcaatttc	atcaattaag	ataaaaatac	aagaccctaa	18120
taaaagatata	agtagaatag	aatgctttat	aatggtgaa	agtgtaaaat	ctgtaagtga	18180
tttaactact	gaaaaaatat	cctttgaagt	accaccagat	aaaattaaaa	taggagagaa	18240
taagatacta	tttagagctt	attgtaaagg	tgatgattta	tatgcatctt	tatttatttt	18300
taaaagagagt	actggaat	ctataattaa	agattcttat	gttatgatag	gtaatagaat	18360
gtacaaggta	gttaatacaa	catctaata	acaagatatt	acaattacac	tagatagagg	18420
acttgaagaa	gatttfaatc	ttggagacct	tatatatcaa	ttaataaata	aaactaaagt	18480
tcaagtaaaa	ataaataaat	ctgacttatt	caaagacatg	aaactagttg	aaatcaaaaa	18540
atcagactca	agttatcaag	aatctatga	attagaagaa	gccaacataa	aaagtgtca	18600
gcctaaaatc	atagtagaaa	aaggagataa	atggacagct	ataaacgtc	catctatgat	18660
tttagatata	gatgctgaaa	acaacgagcc	acaagcttaa	aatggaggtg	taaaaattgt	18720
ttaaattcga	taaaaataaa	atagaacaaa	tcaacaaggg	tagaaaagta	gaaatgcagt	18780

ES 2 622 280 T3

ataaagacat ttcagacata agtataggtc aagcaaagca agatgatgat ataacaaata 18840
 attttatagc aaatgcagaa atatatgaga tgttgттааg tcaaagttct gtcaatgaag 18900
 caagtaatat aagcactttt agtgtaagaa aatctggagg tgagagtggga atggtagaag 18960
 tatatgtagc tttaatтта agaggcagaa aaacaataga agaagtacca gcagtaatta 19020
 gagagcaagt tagaattaga tgtaaagaat tagaaatacc agttgaatag taaatttaga 19080
 ataactatgt attagttatt ttttttatgt aaagtacaag gtcttaactt taataagtaa 19140
 gccttgтact tattttttgt tatattagaa attgtatata tatttattat ttattcaatc 19200
 tataaattaa acctacaatt taaagtacag aagattaaat tgataatcct gaaaaataaa 19260
 tattgcatga tgtaagaata taacaaaaat taaagctata agtataaaaa atttagacaa 19320
 taggaggcta taatggataa attaataacc gaattgagta gtctgggggc aataggтata 19380
 ctatgtgctc tattatttaa aaatactatg caggagaaaa aagaagatag agacatgtat 19440
 aaaaaactg tagaaaattt tatagaatta tctacacaac aacaagaaat aaacaaaaat 19500
 atacttgтtc aaatgggaat aatgaaaaa gatgtagagg aaattaagga agatgttact 19560
 gatataaaag gtatgttaca aaacggтgta taacatgaaa gtagcagtag caccagatta 19620
 tatattatta gaaaagata aagtagtatt gtagatagtг ccctatttta ttgagaagga 19680
 ttttatattt taaaatatta attaaaaaa gtaataaaaa taatatataa aaataacata 19740
 taaaaattca aaaaggagtт aagcttaaat ttgattagaa aaaatcaatt ttaagacaac 19800
 tccttttttt tattaaatta ttgtctatta accaaaaatag ctatttttagc atctggatta 19860
 taacttatct gaaccatttg atttttctta acatgttcaa ggtcttcacc accataagct 19920
 atttgtaact taactgгтаа ctтaccttgт tttataatag caacgtactc ttttttacct 19980
 ttttctctaa actaatcaaa ttгccaacat aaggтттаа gttctgatac tttttactag 20040
 aatttcttat gtagaagaaa gcaccaacag caataactaa atttatgcca agtгтаaccс 20100
 aagaattgat тttaagcата gctccagгga ttattatcac gaacattaaa acgataggта 20160
 atatagcttt cttaagaagc aatttaccса ttatttcatt agcctttttc tcaggгccac 20220
 tcatagtttt tgatctagca aatgattггг cgaatttgтc tcttaagccc attttatcct 20280
 cctaatttta ataaatattt agттатаата acgagatatt acttgaaact aaaaatttac 20340
 tacatttata ttatgттгга cttttgтata aataattaca ttcaagтааа gcaaaatata 20400
 cтаattattt tatcataaaa ttataaaaa gaaaataaat gaaataaaaa tattagaaca 20460
 aagaaatgat gтаaaatcгt атcaaaaгca acataaaaaat tatttatcta ttttctcatc 20520
 tttatттггг ttatactcaa tttttcctaa atccttctct tttcatatt catgaagттt 20580
 таattcaatc ataccttcta ttttgгcttt атcataatca tttaactttc тaaagттггт 20640
 тaaaagcttt atttcattag agттtatgct аттаagтгга tagттггagg aggaatггca 20700
 aattaaatct gatttatgtg atagattatc тccattтаат агccagтcta ctgaaacatt 20760

ES 2 622 280 T3

aaatatctca gctatagatt ttaatatctc ataatttggg tttctaagt ttctctcaaa 20820
 tttacttaag ttgtcacagc ctaacatttc ttctagttca tattgtttaa ggttttttgc 20880
 tttctctaaa taaacaattc tttctcctaa agtatccata aacctctcc attcaattaa 20940
 tgtcaaaaag actttttaag atgtaaatag tttcaaatta aaggtcaaaa tgacataaaa 21000
 accattgact taaggtcaaa atgactttat aattaactta atgatacgaa tttacatcct 21060
 aattttagca caaagtaatc aaaaaatctt atttagtatt aaataaattt atatacttaa 21120
 tatgtgtaca tattaanaaa atatactaaa tagagggggg gcgtaagcta aagtaatata 21180
 aaagtaata taaatcactt agaaaggaag ttgataaatg gatgctcgaa aaaaatggat 21240
 accttttttg ggagtgcaag tcaagcaaaag acttattgaa ttaaataatga ctcaaagga 21300
 attagcgaag aaaataggtg ttaatgaaaa ctatttgcga gctattttta atggaagaag 21360
 aacaggtaaa aaatataaat catcaattta tcaattactt aatatagaat attcagaaga 21420
 tgattaataa atagtatata aagtaggtga atattcttgt gtgcaaatg gattcagatg 21480
 ggggtataga gtgtttaga gcaattgatg attttattac agcacttagt aatataaaaa 21540
 gcttaaatat ggaagatta aatactttaa ctaaatttc tagtacatgt tcaatccttc 21600
 ttaaagaggg gaattatgaa ggatgtacaa ttgtgtatag aaagatgttg gaagaattaa 21660
 aaacatgagt aatgcatttc ttaggaatat aaattataca tagaaatgta ttatattttt 21720
 caaagtactt aaactaaaat atggataaga taatctaaat attataaatg tgcttgaaat 21780
 tagactatac ttgtttttta ataactcaat atccatattt tagtaatata ctacaaaaaa 21840
 agaaggttaa tagatgatgt aaaatcgtat caaattatgt atgtttaaac ctttttatct 21900
 tcattattat tagaggaatg cttttttaag tctttatatt cagatatcct aagttcaagt 21960
 attccttcta tttttatttt atcacgttcg tttagttgtc tgtatagatt taatatcatc 22020
 attcatcat tagtaacatg taagtaactt tctttatctt cttttacact actattgaca 22080
 tttaccttct ctttaccata gagaagccag tcagtcgtaa cattaanaata atcagctatt 22140
 gacattagta tatcacaatt aggttttcta tctcctggtt catacttgcc taagttttca 22200
 aattttaaaa tatccataag tttgcgctga gtaagttttt tggagtttct caaataagca 22260
 attctttttc ctaaagtatc cacaaaatac actcctttct ttttatgagt aatgtctaaa 22320
 tgacatttga aattaanaaa atataaattt ataataaaa actactaaat taaagtctaa 22380
 atgacatttt gcttaanaaa atatgctcat aatatgattt taacatatta tagttgaaaa 22440
 tatatggttt attttgattt gtatatataa caatagattt aattgttata aaaatgtaaa 22500
 ggggtgtatg aatagattgt ataaatttat ttogataaac taagattgct ttttgattgt 22560
 ctgtaaaaaga gaaaagatt aagataaaaa tagtattata ttgtaattta tattaatcaa 22620
 ttacaaagat tttatgaatt tattcttttag ggtaanaaat ttaagaataa gataaattta 22680
 caatataata ctataacact cttttatcta gttttatttt ctttatagaa caataatatt 22740
 ataaatgcta gtagatttac acagaatact gttatataca tctgtttgaa tcctgagttt 22800
 agagtagatt gtagtgggga tccgg 22825

ES 2 622 280 T3

5 <210> 2
 <211> 65
 <212> PRT
 <213> Clostridium difficile

<400> 2
 Met Phe Lys Asn Asn Leu Lys Tyr Tyr Arg Lys Cys Lys Gly Met Thr
 1 5 10 15

Gln Ile Gln Leu Ala Arg Lys Ala Gly Ile Thr Asn Asp Tyr Ile Ser
 20 25 30

Gln Ile Glu Arg Gly Ile Lys Asn Pro Gly Leu Leu Met Ala Lys Lys
 35 40 45

Ile Ser Ser Ile Leu Glu Gln Asn Ile Glu Glu Val Phe Phe Ile Gln
 50 55 60

10 Leu
 65

<210> 3
 <211> 146
 <212> PRT
 <213> Clostridium difficile

15 <400> 3
 Met Glu Asn Lys Lys Asp Ile Leu Phe Lys Glu Thr Asp Glu Arg Leu
 1 5 10 15

His Asn Tyr Lys Tyr Leu Asp Ile Lys Ile Lys Asn Ile Asn Leu Asp
 20 25 30

Ile Lys Arg Cys Glu Asn Glu Tyr Ser Gly Cys Gly Ala Met Val Tyr
 35 40 45

Thr Glu Lys Thr Ser Asn Thr Tyr Asn Ile Ser Ser Ser Val Glu Asn
 50 55 60

Glu Val Leu Lys Arg Glu Glu Arg Leu Arg Lys Leu Lys Met Glu Lys
 65 70 75 80

Glu Asp Ile Glu Ile Glu Lys Glu Lys Ile Glu Asn Ala Leu Thr Cys
 85 90 95

Leu Asn Asp Ile Glu Met Glu Phe Phe Asn Leu Phe Tyr Asn Ser Lys
 100 105 110

ES 2 622 280 T3

Thr Lys Asn Asn Met Thr Tyr Ile Ser Met Lys Leu His Leu Asp Arg
 115 120 125

Thr Ser Cys Tyr Asn Leu Lys Lys Lys Met Ile Phe Lys Leu Ser Glu
 130 135 140

Ile Leu
 145
 <210> 4
 <211> 130
 <212> PRT
 5 <213> Clostridium difficile

<400> 4
 Asn Phe Thr Glu Ser Ile Phe Ile Asp Asp Glu Ser Val Gln Gly Ser
 1 5 10 15

Glu Gly Ser Cys Phe Phe Val Ser Ile Leu Ser Val Ile Cys Thr Pro
 20 25 30

Ile Met Leu Asn Thr Asn Asn Lys Asp Ile Val Ile Ser Ile Lys Tyr
 35 40 45

Leu Pro Lys Pro Gln Ser Lys Ser Ile Arg Met Tyr Glu Ile Ser Asp
 50 55 60

Glu Leu Asn Lys Leu Phe Asn Arg Asn Ile Lys Val Thr Asp Arg Lys
 65 70 75 80

Leu Asn Ile Thr Lys Leu Glu Gln Ser Ile Lys Lys Glu Glu Ser Ile
 85 90 95

Tyr Val Leu Asn Phe Thr Ile Thr Leu Asn Tyr Leu Asp Ser Val Tyr
 100 105 110

Glu Glu Asp Val Val Tyr Glu Asn Met Glu Glu Ile Asn Leu Asn Leu
 115 120 125

Gly Glu
 130
 10 <210> 5
 <211> 354
 <212> PRT
 <213> Clostridium difficile

15 <400> 5
 Met Ala Ile Gly Leu Pro Ser Ile Asn Ile Ser Phe Lys Glu Leu Ala
 1 5 10 15

ES 2 622 280 T3

Thr Thr Val Lys Glu Arg Ser Ala Arg Gly Ile Ile Ala Met Val Leu
 20 25 30

Lys Asp Ala Lys Ala Leu Gly Leu Asn Glu Ile His Glu Lys Glu Asp
 35 40 45

Ile Pro Val Asp Leu Ser Ala Glu Asn Lys Glu Tyr Ile Asn Leu Ala
 50 55 60

Leu Met Gly Asn Val Asn Thr Pro Asn Lys Leu Leu Val Tyr Val Ile
 65 70 75 80

Glu Gly Glu Ala Asp Ile Gln Thr Ala Leu Asp Phe Leu Glu Thr Lys
 85 90 95

Glu Phe Asn Tyr Leu Cys Met Pro Lys Ala Val Glu Ala Asp Lys Thr
 100 105 110

Ala Ile Lys Asn Trp Ile Ile Lys Leu Arg Asp Ile Asp Lys Val Lys
 115 120 125

Val Lys Ala Val Leu Gly Lys Val Val Gly Asn His Glu Gly Ile Ile
 130 135 140

Asn Phe Thr Thr Glu Asp Val Leu Val Gly Glu Lys Lys Tyr Ser Val
 145 150 155 160

Asp Glu Phe Thr Ser Arg Val Ala Gly Leu Ile Ala Gly Thr Pro Leu
 165 170 175

Ser Gln Ser Val Thr Tyr Thr Lys Leu Ser Asp Val Val Asp Ile Pro
 180 185 190

Lys Met Thr Lys Val Asp Ala Glu Ser Arg Val Asn Lys Gly Glu Leu
 195 200 205

Ile Leu Ile Lys Glu Ala Gly Ala Ile Arg Ile Ala Arg Gly Val Asn
 210 215 220

Ser Leu Thr Glu Leu Thr Glu Glu Lys Gly Glu Met Phe Gln Lys Ile
 225 230 235 240

Lys Ile Val Asp Thr Leu Asp Ile Ile His Ser Asp Ile Arg Lys Val
 245 250 255

Ile Ile Asp Asp Tyr Ile Gly Lys Val Thr Asn Ser Tyr Asp Asn Lys
 260 265 270

ES 2 622 280 T3

Cys Leu Leu Ile Val Ala Ile Lys Ser Tyr Leu Glu Glu Leu Glu Lys
 275 280 285

Ser Ala Leu Ile Glu Ser Asp Ser Thr Val Glu Ile Asp Phe Glu Ala
 290 295 300

Gln Lys Ser Tyr Leu Lys Ser Lys Gly Val Asp Leu Ser Tyr Met Thr
 305 310 315 320

Leu Gln Glu Ile Lys Glu Ala Asn Thr Gly Ser Lys Val Phe Leu Lys
 325 330 335

Ala Lys Ile Lys Val Leu Asp Ala Met Glu Asp Ile Asp Leu Ser Ile
 340 345 350

Glu Ile

<210> 6

<211> 142

5 <212> PRT

<213> Clostridium difficile

<400> 6

Met Ala Asn Met Glu Ala Arg Asn Val Met Ser Gly Thr Trp Gly Glu
 1 5 10 15

Leu Trp Leu Asp Gly Asn Lys Val Ala Glu Val Lys Lys Phe Gln Ala
 20 25 30

Lys Met Glu Phe Thr Lys Glu Asp Ile Ile Ile Ala Gly Gln Met Gly
 35 40 45

Thr Asp Thr Lys Tyr Met Gly Tyr Lys Gly Lys Gly Ser Ile Thr Leu
 50 55 60

Tyr His Val Ser Ser Arg Met His Lys Leu Ile Gly Glu Lys Ile Lys
 65 70 75 80

Arg Gly Ser Glu Pro Arg Phe Val Ala Ile Ser Lys Leu Asn Asp Pro
 85 90 95

Asp Ser Tyr Gly Ala Glu Arg Ile Ala Val Lys Asn Ile Ala Phe Asp
 100 105 110

Asp Leu Thr Leu Ala Asp Trp Glu Val Gly Val Lys Gly Glu Ile Glu
 115 120 125

Ala Pro Phe Thr Phe Thr Glu Tyr Asp Phe Leu Asp Ile Ile
 130 135 140

10

<210> 7

<211> 148

<212> PRT

<213> Clostridium difficile

15

<400> 7

ES 2 622 280 T3

Met Asn Glu Asn Gly Leu Ser Lys Asn Ile Asn Ile Val Asp Leu Leu
1 5 10 15

Leu Asn Ala Asp Thr Glu Asn Leu Glu Arg Pro Ser Thr Ile Val Glu
20 25 30

Leu Lys Arg Leu Ser Thr Ile Phe Gly Gln Glu Phe Lys Val Met Cys
35 40 45

Arg Ala Leu Thr Ile Ser Lys Asp Glu Glu Ile Gln Asn Thr Cys Leu
50 55 60

Lys Ile Asp Glu Asn Met Lys Thr Asp Ile Asp Leu Pro Glu Met Gln
65 70 75 80

Met Leu Thr Ile Ile Glu Gly Val Cys Asp Leu Asp Gly Lys Leu Leu
85 90 95

Phe Lys Asn Lys Glu Leu Met Asp Lys Phe Lys Ala Pro Thr Pro Lys
100 105 110

Glu Leu Ala Arg Lys Leu Leu Leu Pro Gly Glu Ile Thr Asn Leu Tyr
115 120 125

Arg Ile Leu Gln Asp Val Met Gly Tyr Gly Lys Asn Ala Val Ile Glu
130 135 140

Glu Val Lys Asn
145

<210> 8

<211> 55

5 <212> PRT

<213> Clostridium difficile

<400> 8

Met Tyr Tyr Tyr Trp Lys Lys Lys Gly Ile Arg Pro Ser Leu Phe Tyr
1 5 10 15

Ala Met Asp Lys Gly Glu Leu Lys Leu Ile Glu Ala Phe Phe Ala Leu
20 25 30

Glu Ile Glu Glu Glu Val Glu Lys Met Lys His Gly Tyr Gly Val Cys
35 40 45

10

Pro Leu Thr Gly Gly Gly Met
50 55

<210> 9

<211> 817

15 <212> PRT

<213> Clostridium difficile

<400> 9

ES 2 622 280 T3

Met Gly Asn Val Arg Glu Glu Gly Ile Asn Met Tyr Leu Thr Asp Asn
 1 5 10 15

Tyr Thr Pro Lys Met Asn Gln Ile Ile Ser Val Thr Asp Asn Phe Arg
 20 25 30

Arg Ala Thr Val Ala Val Ser Leu Ser Thr Asn Val Met Ala Ser Ser
 35 40 45

Ile Lys Asn Ser Ile Gly Ser Ala Ser Ser Arg Val Asn Ser Leu Asn
 50 55 60

Ser Ser Leu Arg Lys Val Gln Thr Thr Ala Ser Ser Val Ser Ser Thr
 65 70 75 80

Met Ala Lys Leu Ser Ser Ser Ile Asn Ala Val Ser Gly Val Ile Gly
 85 90 95

Ser Leu Asn Gly Ser Ile Met Arg Leu Ala Ile Thr Ile Ala Met Ile
 100 105 110

Ile Asp Tyr Phe Asn Lys Leu Ile Gln Lys Lys Asn Glu Phe Asn Ser
 115 120 125

Asn Ile Met Ile Ile Leu Ile Phe Lys Ala Lys Ser Asp Glu Val Glu
 130 135 140

Lys Thr Lys Asn Lys Leu Leu Gly Asn Leu Lys Lys Ile Gly Gly Lys
 145 150 155 160

Ile Trp Asn Ile Val Ile Lys Ala Lys Asp Met Thr Lys Arg Val Ile
 165 170 175

Ser Ser Ile Leu Gly Lys Leu Lys Arg Val Glu Lys Arg Pro Tyr Gln
 180 185 190

Gly Ser Ile Asn Leu Lys Asp Met Val Ser Ser Ala Met Ala Arg Ile
 195 200 205

Leu Pro Lys Leu Met Leu Phe Lys Asn Thr Phe Trp Ser Gly Val Ile
 210 215 220

ES 2 622 280 T3

Ala Ile Lys Asp Met Ala Ser Ser Ile Ile Ser Lys Val Phe Pro Lys
 225 230 235 240

Leu Arg Leu Phe Ala Gly Lys Val Trp Ser Gly Ala Ile Ala Val Lys
 245 250 255

Asp Met Ala Ser Gly Ile Leu Gly Ser Ile Lys Gly Lys Ile Ser Asp
 260 265 270

Leu Thr Asn Gly Ala Thr Ile Gly Val Ala Val Lys Lys Gly Val Asp
 275 280 285

Leu Leu Gly Gln Glu Gln Asn Gln Lys Val Val Leu Glu Ser Val Met
 290 295 300

Lys Arg Asn Thr Gly Lys Thr Ser Gln Lys Asp Val Asp Lys Tyr Tyr
 305 310 315 320

Asp Ser Leu Val Asn Met Ala Asn Asp Thr Pro Phe Asp Pro Glu Asp
 325 330 335

Val Val Ala Met Gly Thr Lys Ala Lys Met Ile Ser Asn Ile Thr Gly
 340 345 350

Gly Lys Lys Glu Lys Asp Ile Thr Gln Ala Met Val Asp Val Arg Ala
 355 360 365

Leu Asn Met Asn Thr Ser Ser Glu Gln Asp Val Ser Ala Ala Phe Leu
 370 375 380

Ser Ala Ala Lys Gly Asn Met Glu Ser Leu Asn Thr Leu Val Gly Glu
 385 390 395 400

Asn Tyr Lys Thr Phe Asp Glu Ala Leu Glu Gly Ile Ser Val Lys Gln
 405 410 415

Met Gly Leu Ala Lys Glu Met Ser Asn Thr Ile Pro Gly Ile Ile Ser
 420 425 430

Gly Ala Gln Thr Ser Ile Asn Asn Gly Leu Lys Ser Ile Val Lys Pro
 435 440 445

Phe Asp Asp Ile Leu Gly Gln Gly Leu Lys Lys Ile Lys Thr Phe Ile
 450 455 460

Glu Ser Gly Leu Gly Asn Leu Ala Gly Leu Ser Glu Lys Met Ala Gly
 465 470 475 480

ES 2 622 280 T3

Lys Ile Gly Asn Val Met Asn Gly Lys Ile Ile Ile Gly Asn Lys Tyr
 485 490 495
 Asp Gln Met Gln Ser Arg Ser Val Lys Asn Gly Lys Glu Phe Ser Asp
 500 505 510
 Ser Thr Gln Tyr Arg Ile Ser Asn Glu Ala Glu Lys Arg Lys Met Met
 515 520 525
 Val Glu Asn Lys Gln Glu Arg Phe Glu Asn His Ala Ala Thr Met Ile
 530 535 540
 Gly Asn Ala Pro Lys Ala Ile Val Asn Ala Gly Ser Thr Leu Leu Gln
 545 550 555 560
 Asn Ile Asp Phe Thr Ala Leu Ile Asp Ser Leu Leu Pro Val Val Asn
 565 570 575
 Leu Val Asn Asn Leu Leu Asp Ser Ile Asn Asn Lys Ser Pro Ile Ala
 580 585 590
 Gln Gly Leu Ile Ser Ile Phe Gly Thr Ile Val Thr Thr Ala Phe Gln
 595 600 605
 Leu Ile Gly Pro Val Val Glu Ala Val Ser Pro Ile Ile Thr Arg Ile
 610 615 620
 Phe Thr Phe Leu Gly Glu Tyr Ala Pro Gln Ile Asn Asn Phe Ile Glu
 625 630 635 640
 Thr Leu Gly Val Ile Trp Lys Thr Val Trp Glu Thr Leu Gly Pro Leu
 645 650 655
 Leu Glu Thr Gly Trp Lys Ile Ile Glu Pro Ile Leu Gly Ala Phe Phe
 660 665 670
 Asn Ile Leu Asp Lys Val Cys Lys Ile Val Lys Asp Ile Cys Lys Trp
 675 680 685
 Trp Gln Thr Met Ile Asn Lys Ile Lys Asn Gly Ser Ile Thr Gly Thr
 690 695 700
 Val Leu Asn Leu Val Glu Lys Ser Lys Lys Asn Tyr Lys Asp Asn Pro
 705 710 715 720
 Tyr Ala Gly Thr Lys Ala Gly Asp Ser Gly Lys Ala Tyr Ser Ser Lys
 725 730 735
 Lys Gly Asn Asn Ala Phe Gly Leu Asn Tyr Val Pro Tyr Asn Asp Tyr

ES 2 622 280 T3

740

745

750

Gln Thr Arg Leu His Glu Gly Glu Met Val Leu Thr Lys Gln Glu Ala
755 760 765

Asn Gln Tyr Arg Ser Arg Lys Asn Gly Gly Asn Ile Asn Ile Ala Lys
770 775 780

Leu Ala Asp Thr Ile Val Ile Arg Glu Glu Ala Asp Ile Glu Lys Ile
785 790 795 800

Thr Ser Lys Leu Val Ala Ser Ile Gln Leu Ala Gln Leu Gly Gly Val
805 810 815

Leu

<210> 10

<211> 140

5 <212> PRT

<213> Clostridium difficile

<400> 10

Met Glu Met Trp Leu Arg Gln Ala Glu Asp Arg Phe Arg Phe Pro Val
1 5 10 15

Phe Pro Ser Ser Phe Ser Ile Asn Gly Lys Ala Ala Val Asn Ser Ser
20 25 30

Ser Ile Leu Lys Ile Gly Glu Val Ala Thr Phe Gly Gly Val Ala Leu
35 40 45

Lys Ser Ile Ser Ile Ser Ser Phe Phe Pro Asn Lys Asp Tyr Thr Phe
50 55 60

Cys Asp Tyr Thr Gly Phe Pro Ser Pro Tyr Asp Cys Val Asn Lys Ile
65 70 75 80

Glu Lys Trp Met Lys Glu Gly Phe Ile Leu Arg Phe Thr Ile Thr Glu
85 90 95

Thr Asn Ile Asn Met Glu Val Ile Ile Glu Gly Phe Ser Tyr Glu Glu
100 105 110

Arg Asp Gly Thr Arg Asp Val Tyr Phe Thr Leu Asp Leu Lys Glu Tyr
115 120 125

Lys Arg Ile Lys Ile Pro Lys Val Thr Pro Lys Gln
130 135 140

10

<210> 11

<211> 509

<212> PRT

<213> Clostridium difficile

15

<400> 11

ES 2 622 280 T3

Met Ile Ile Asn Arg Ser Lys Asp Ser Ser Ser Asn Glu Ile Ser Phe
1 5 10 15

Val Ser Lys Asp Met Gly Phe Leu Leu Thr Gln Ser Glu Val Ser Tyr
20 25 30

Asn Phe Lys Asp Lys Leu Val Glu Asp Ile Ala Lys Gln Val Phe Ala
35 40 45

Glu Asn Arg Leu Ser Val Gly Thr Ile Ala Lys Thr Asn Val Lys Tyr
50 55 60

Thr Lys Met Phe Ile Gly Val Asn Gly Tyr Asp Thr Ile Met Ser Ala
65 70 75 80

Tyr Thr Glu Ala Ser Lys Lys Thr Lys Lys Lys Tyr Met Ile Glu Ala
85 90 95

Asn Leu Asp Lys Phe Asn Val Ile Glu Lys Gly Thr Val Thr Leu Ser
100 105 110

Val Met Phe Glu Glu Gly Phe Asn Ile Ile Asn Thr Thr Phe Ser Glu
115 120 125

Ser Met Glu Asn Val Lys Asn Lys Val Ile Val Val Asp Gln Tyr Gly
130 135 140

Ser Lys Ile Ser Glu Lys Ile Asp Asn Glu Ile Phe Lys Glu Val Asn
145 150 155 160

Val Ile Met Gln Lys Val Ile Gln Gln Gln Glu Asn Gln Asp Val Asp
165 170 175

Ile Asp Ser Glu Phe Asn Gly Ile Glu Lys Ser Cys Ser Leu Lys Gly
180 185 190

Tyr Gly Asp Val Ser Cys Ile Thr Gly Arg Gly Val Lys Val Lys Asp
195 200 205

Ser Tyr Thr Lys Leu Val Gly Leu Phe Tyr Ile Asp Thr Asp Lys His
210 215 220

Thr Trp Gln Asn Gly Glu Tyr Gln Ile Glu Leu Glu Leu Asn Phe Gln
225 230 235 240

ES 2 622 280 T3

Asn Leu Met Asp Glu Lys Ser Ala Gly Gln Asp Glu Pro Lys Glu Glu
 245 250 255

Ser Asn Leu Gly Gly Glu Asp Tyr Ala Gly Gly Lys Glu Phe Thr Ala
 260 265 270

Glu Phe Thr Ala Tyr Cys Pro Arg Lys Glu Glu Gly Gly Asp Thr Asp
 275 280 285

Cys Arg Lys Lys Lys Leu Asp Pro Ser Lys Lys Thr Cys Ala Ala Pro
 290 295 300

Met Val Gly Lys Tyr Glu Gln Thr Tyr Tyr Thr Lys Glu Phe Leu Asn
 305 310 315 320

Lys His Pro Leu Leu Asn Tyr Gly Asp Glu Ile Gln Val Ile Thr Gly
 325 330 335

Val Ser Gly Arg Asp Gly Val Tyr Lys Val Asn Asp Val Gly Pro Ala
 340 345 350

Ile Thr Ile Glu Lys Asn Gly Thr Tyr His Ile Asp Ile Leu Phe Gly
 355 360 365

Asn Val Glu Glu Ala Ser Lys Phe Gly Arg Arg Lys Gly Lys Ile Ile
 370 375 380

Ile Gly Gly Tyr Ser Gly Asn Val Ser Asp Lys Ala Lys Ile Val Ile
 385 390 395 400

Ser Glu Ala Lys Lys His Leu Gly Lys Pro Tyr Lys Trp Gly Gly Asn
 405 410 415

Gly Pro Ser Ser Phe Asp Cys Ser Gly Leu Met Val Tyr Cys Phe Lys
 420 425 430

Lys Val Asn Val Ser Leu Pro Arg Thr Ser Asn Gln Gln Ser Lys Lys
 435 440 445

Gly Lys Lys Val Glu Gln Lys Asn Leu Gln Ala Gly Asp Leu Val Phe
 450 455 460

Phe His Asn Pro Val Ser His Val Gly Leu Tyr Ile Gly Asn Gly Glu
 465 470 475 480

Phe Leu His Ala Pro Gln Lys Gly Asp Val Val Lys Ile Ser Lys Leu
 485 490 495

Ser Ser Arg Arg Asp Phe Asn Thr Ala Arg Arg Val Leu
 500 505

- 5 <210> 12
- <211> 108
- <212> PRT
- <213> Clostridium difficile

ES 2 622 280 T3

<400> 12

Met Ala Asn Pro Ile Asn Glu Phe Ile Gly Ile Ile Arg Glu Glu Gly
1 5 10 15

Lys Tyr His Asn Gln Pro Ser Phe Phe Ile Gly Lys Ile Lys Ser Lys
20 25 30

Leu Pro Asp Leu Lys Ile Glu Thr Asn Asn Ile Ile Leu Glu Lys Glu
35 40 45

Asp Ile Leu Ile Asp Ser Trp Met Ile Asp Arg Gln Leu Glu Thr Phe
50 55 60

Asp Thr Glu Thr Asn Gln Glu His Gln His Glu Val Lys Asn Pro Phe
65 70 75 80

Ile Asp Asn Phe Glu Ser Gly Asp Met Val Ile Met Phe Arg Ile Gly
85 90 95

Glu Lys Phe Ala Val Val Ser Lys Leu Val Ser Leu
100 105

<210> 13

<211> 142

<212> PRT

<213> Clostridium difficile

<400> 13

Met Ser Thr Ile Phe Pro Phe Ile Gly Val Pro Glu Asp Tyr Ile Leu
1 5 10 15

Pro Lys Thr Glu Glu Leu Pro Ile Phe Arg Glu Val Ala Trp Asp Phe
20 25 30

Glu Lys Asp Glu Pro Ile Leu Glu Lys Gly Asp Phe Lys Ile Ile Glu
35 40 45

Lys Lys Glu Ala Leu Lys Val Trp Ile Tyr Lys Cys Ile Lys Thr Asn
50 55 60

Arg Tyr Glu His Glu Ile Tyr Ser Leu Glu Tyr Gly Thr Glu Leu Ser
65 70 75 80

10 Glu Leu Ile Gly Gln Lys Tyr Thr Lys Gly Leu Thr Glu Ser Glu Ala
85 90 95

Ser Arg Phe Ile Lys Glu Ala Leu Leu Ile Asn Pro Tyr Ile Leu Glu
100 105 110

Val Asn Val Lys Ser Ala Asn Phe Asn Arg Asp Val Leu Ser Ala Asn
115 120 125

Val Lys Val Ser Thr Ile Tyr Gly Glu Val Glu Ile Asn Val
130 135 140

ES 2 622 280 T3

<210> 14
 <211> 350
 <212> PRT
 <213> Clostridium difficile

5

<400> 14
 Met Tyr Ser Asp Gln Thr Tyr Glu Val Ile Lys Asn Arg Thr Leu Glu
 1 5 10 15

 Asn Ile Asn Leu Asp Ile Tyr Lys Gly Glu Gly Ser Phe Leu Asn Asn
 20 25 30

 Met Val Ser Gly Asn Asn Leu Glu Leu Ser Lys Ile Tyr Leu Glu Leu
 35 40 45

 Ser Lys Ile His Lys Met Ala Phe Ile Gln Asp Thr Tyr Asn Gln Phe
 50 55 60

 Leu Asp Lys Arg Val Asn Glu Phe Gly Val Tyr Arg Lys Leu Gly Thr
 65 70 75 80

 Glu Ser Asn Gly Glu Val Glu Phe Ile Gly Glu Lys Gly Thr Val Ile
 85 90 95

 Asn Asn Gly Thr Ile Ile Ser Tyr Arg Asp Leu Leu Phe Val Val Ile
 100 105 110

 Lys Asp Val Thr Ile Gly Ser Glu Glu Gly Asp Asn Ser Pro Val Gln
 115 120 125

 Ala Leu Glu Val Gly Lys Lys Tyr Asn Leu Pro Thr Asn Cys Glu Phe
 130 135 140

 Lys Leu Val Asp Asn Ile Ser Gly Val Thr Lys Ile Thr Asn Thr Arg
 145 150 155 160

 Ser Phe Glu Gly Gly Thr Asp Ile Glu Thr Asp Glu Glu Leu Lys Glu
 165 170 175

ES 2 622 280 T3

Arg Phe Tyr Lys Ile Gln Arg Asn Gln Ala Thr Ser Gly Asn Lys Ala
180 185 190

His Tyr Glu Glu Trp Ala Leu Glu Val Asp Gly Val Tyr Asn Val Lys
195 200 205

Val Tyr Pro Arg Trp Asp Gly Pro Gly Thr Val Lys Val Leu Ile Phe
210 215 220

Gly Glu Asn Asn Gln Ala Val Asp Thr Glu Thr Ile Glu Arg Cys Gln
225 230 235 240

Gln His Ile Asp Glu Glu Lys Pro Ile Gly Pro Thr Ile Thr Val Val
245 250 255

Thr Pro Leu Pro Ile Glu Ile Ser Ile Ser Ala Val Met Lys Leu Glu
260 265 270

Asp Gly Tyr Thr Leu Asp Asn Val Lys Glu Ser Phe Leu Glu Ser Ile
275 280 285

Asn Thr Tyr Phe Arg Asp Ile Arg Gly Glu Ile Ile Tyr Thr Lys Val
290 295 300

Met Gly Ile Leu Ile Asn Thr Thr Gly Val His Asp Leu Ser Asn Leu
305 310 315 320

Leu Ile Asn Gly Ser Thr Asp Asn Ile Thr Ile Asn Glu Asp Lys Ile
325 330 335

Pro Ser Val Thr Thr Val Asn Phe Ser Glu Val Glu Asn Gln
340 345 350

<210> 15

<211> 232

5 <212> PRT

<213> Clostridium difficile

<400> 15

Met Lys Leu Ile Asp Lys Leu Pro Ser Phe Asp Arg Asn Tyr Ile Val
1 5 10 15

Glu Glu Ile Gln Gly Ala Tyr Asp Thr Glu Leu Asn Ile Leu Lys Glu
20 25 30

Asp Ile Asp Asp Thr Phe Asn Gln Leu Phe Val Asp Thr Ala Thr Trp
35 40 45

Gly Leu Asp Met Trp Glu Asp Ile Leu Cys Ile Glu Lys Lys Glu Leu
50 55 60

10

ES 2 622 280 T3

Asp Phe Asp Thr Arg Arg Ser Asn Ile Lys Ala Lys Met Arg Ser Arg
65 70 75 80

Gly Thr Ser Thr Ile Glu Val Ile Lys Ser Ile Cys Glu Ala Tyr Thr
85 90 95

Lys Ser Glu Thr Asp Ile Lys Val Tyr Ser Asp Glu Phe Thr Phe Val
100 105 110

Leu Ser Phe Ile Ala Asn Asn Cys Asp Tyr Lys Thr Leu Leu Asp Cys
115 120 125

Ser Glu Met Ile Glu Arg Val Lys Pro Ala His Leu Leu His Tyr Leu
130 135 140

Glu Pro Ile Ile Leu Asp Lys Ser Met Val Tyr Cys Gly Gly Gly Met
145 150 155 160

Val Cys Ser Glu Glu Val Lys Val His Pro Tyr Phe Glu Pro Ile Ile
165 170 175

Lys Cys Ser Ala Val Val Asn Cys Gly Ala Gly Met Leu Ser Arg Glu
180 185 190

Glu Ile Lys Val Tyr Pro Leu Ser Ile Lys Cys Ile Glu Asn Asn Cys
195 200 205

Lys Ile Asn Ile Ala Ile Ala Asn Asp Thr Gly Val Glu Asn Val Val
210 215 220

Val Tyr Pro Lys Ser Glu Val Val
225 230

<210> 16
<211> 328
5 <212> PRT
<213> Clostridium difficile

<400> 16
Met Glu Glu Lys Phe Tyr Ile Ile Leu Thr Lys Ile Gly Arg Glu Lys
1 5 10 15

Ile Ala Asn Ala Thr Ala Leu Gly Glu Leu Val Gly Leu Thr Lys Phe
20 25 30

Gln Val Gly Asp Ser Asn Gly Glu Tyr Tyr Glu Pro Thr Glu Glu Gln
35 40 45

Thr Ala Leu Lys Asn Val Val Trp Glu Gly Asn Ile Asn Ser Leu Arg
50 55 60

10

ES 2 622 280 T3

Ile Asp Glu Lys Asn Pro Asn Trp Ile Val Ile Glu Thr Ile Leu Pro
65 70 75 80

Gly Thr Val Gly Gly Phe Met Ile Arg Glu Ala Ala Val Leu Asp Asn
85 90 95

Glu Asn Asn Ile Ile Ala Ile Gly Lys Tyr Pro Glu Thr Tyr Lys Pro
100 105 110

Arg Ala Glu Asp Gly Ser Ile Lys Asp Leu Val Val Lys Met Ile Leu
115 120 125

Gln Leu Ser Asn Thr Ser Asn Val Thr Leu Glu Val Asp Pro Thr Leu
130 135 140

Val Phe Val Thr Gln Lys Asp Ile Gln Asp Leu Asp Asp Lys Phe Asp
145 150 155 160

Lys Asn Ile Lys Glu Ile Lys Val Lys Ile Gly Asp Thr Asp Ile Leu
165 170 175

Thr Thr Asp Ser Lys Asp Leu Ser Gly Ala Ile Asn Glu Val Val Lys
180 185 190

Lys Ile Glu Asn Ile Ser Phe Asp Asp Val Ile Ser Gly Gln Ile Gln
195 200 205

Thr Asp Ile Ser Val Leu Lys Asn Ser Tyr Asn Lys Leu Ser Glu Lys
210 215 220

Val Leu Asp Ile Leu Ile Tyr Leu Glu Leu Glu Ser Glu Val Thr Val
225 230 235 240

Asp Glu Ala Gly Tyr Trp Tyr Asp Thr Leu Ala Asn Gly Asn Asn Ile
245 250 255

Val Ala Ile Glu Gly Leu Lys Leu Asp Leu Asn Arg Lys Cys Ile Thr
260 265 270

Gly Glu Ile Gly Asn Val Ile Phe Arg Asp Val Val Leu Pro Phe Ser
275 280 285

Ala Asn Arg Val Arg Tyr Ile His Asp Met Asp Asn Asn Phe Val Glu
290 295 300

Thr Lys Ser Ser Asn Thr Tyr Leu Lys Glu Gln Lys Asp Ile Thr Leu
305 310 315 320

Ser Lys Tyr Ser Tyr Glu Ile Arg
325

- 5 <210> 17
<211> 1725
<212> PRT
<213> Clostridium difficile

ES 2 622 280 T3

<400> 17

Met Lys Gln Asn Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys
 1 5 10 15

Asn Ile Leu Ile Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe
 20 25 30

Val Glu Phe Phe Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser
 35 40 45

Asp Gly Asn Leu Tyr Ala Cys Gly Asp Asn Thr Gly Phe Gln Leu Gly
 50 55 60

Leu Gly Lys Asp Ser Ser Glu Arg Arg Met Phe Ser Lys Val Lys Ile
 65 70 75 80

Asp Asn Val Lys Tyr Val Ser Cys Gly Ser Lys His Ser Val Ala Val
 85 90 95

Thr Lys Asp Gly Phe Ala Tyr Gly Ala Gly Thr Ser Asn Val Gly Gln
 100 105 110

Leu Gly Val Ile Glu Ser Thr Val Tyr Tyr Glu Phe Thr Lys Leu Pro
 115 120 125

Ile Asp Asp Val Lys Thr Val Ala Cys Gly Tyr Asp Phe Thr Phe Val
 130 135 140

Leu Lys Asn Asp Gly Thr Leu Tyr Ser Ala Gly Leu Asn Ser Ser Gly
 145 150 155 160

Gln Leu Gly Leu Gly Asp Thr Asn Asn Arg Ala Thr Phe Thr Lys Val
 165 170 175

Asn Ile Asp Ser Val Lys Asp Val Val Thr Tyr Asn Gln Ser Val Phe
 180 185 190

Ile Ile Lys Met Asp Gly Thr Ala His Ala Cys Gly Leu Asn Ser Asn
 195 200 205

Gly Gln Leu Gly Ile Asn Ser Thr Leu Asn Lys Ser Val Phe Asn Lys
 210 215 220

ES 2 622 280 T3

Ile Glu Gly Met Asp Asn Val Lys Gln Ile Ala Cys Gly Ser Ser His
 225 230 235 240

Thr Ile Leu Ile Lys Asn Asp Gly Thr Met Tyr Thr Thr Gly Tyr Asn
 245 250 255

Gly Val Gly Gln Leu Gly Thr Gly Asn Asn Asn Asn Ser Ile Val Phe
 260 265 270

Thr Leu Ser Ser Ile Asn Asn Val Lys Tyr Ala Ser Cys Gly Asn Asn
 275 280 285

His Thr Met Ile Leu Lys Tyr Asp Asn Thr Leu Phe Ser Thr Gly Gln
 290 295 300

Asn Asn Tyr Gly Gln Leu Ala Asn Ala Asn Lys Asp Val Ala Ser Arg
 305 310 315 320

Asn Thr Phe Ala Lys Val Asn Val Glu Asn Ile Lys Asp Ile Lys Cys
 325 330 335

Gly Ser Gln Phe Asn Phe Leu Ile Asn Gly Ser Lys Glu Ile Phe Val
 340 345 350

Ser Gly Cys Asn Leu Ala Gly Gln Leu Gly Ser Phe Phe His Thr Thr
 355 360 365

Phe Leu Tyr Glu Phe Ser Asn Val Gln Ser Ser Asn Leu Asp Asn Tyr
 370 375 380

Ser Gly Leu Leu Val Asn Asp Asp Tyr Leu Tyr Val Thr Lys Asp Asn
 385 390 395 400

Ser Glu Phe Leu Asn Val Lys Leu Ser Asp Asn Phe Gln Asp Tyr Lys
 405 410 415

Lys Ile Glu Leu Thr Asp Ser Asn Met Phe Ile Val Met Asn Asp Gly
 420 425 430

Thr Leu Tyr Ala Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Leu Gly
 435 440 445

Asp Thr Val Asn Arg Ser Val Met Thr Lys Val Asp Ile Asp Asn Val
 450 455 460

Leu Asp Ile Lys Gly Asn Gly Asn Ser Thr Phe Val Leu Lys Asn Asn
 465 470 475 480

ES 2 622 280 T3

Gly Thr Leu Tyr Ser Cys Gly Leu Asn Ser Asn Gly Gln Leu Gly Leu
485 490 495

Arg Asp Glu Val Asn Arg Asn Ile Phe Thr Lys Ile Glu Ile Glu Asn
500 505 510

Val Lys Asp Phe Cys Val Gly Ser Asn Tyr Val Ile Ala Leu Asn His
515 520 525

Ser Lys Glu Val Tyr Gly Trp Gly Asn Asn Pro Tyr Asn Asn Ile Glu
530 535 540

Lys Thr Ser Asn Tyr Pro Tyr Lys Gln Gly Ile Ser Asn Ile Glu Lys
545 550 555 560

Ile Ala Ala Tyr Asp Tyr Ser Val Tyr Met Ile Asn Ser Glu Gly Lys
565 570 575

Leu Tyr Val Ser Gly Tyr Asn Tyr Asn Tyr Gln Leu Gly Lys Gly Asn
580 585 590

Asn Ser Asn Gln Ser Lys Ala Leu Val Ser Gln Cys Arg Thr Asn Ser
595 600 605

Thr Ser Ser Thr Ser Asn Gly Leu Arg Thr Leu Pro Lys Ile Thr Asn
610 615 620

Val Phe Pro Phe Tyr Asp Gly Cys Ala Ile Ile Asp Glu Gly Gly Tyr
625 630 635 640

Val Tyr Leu Thr Gly Tyr His Gly Tyr Leu Arg Thr Leu Asn Ser Ser
645 650 655

Pro Ser Ile Ser Asp Tyr Ser Arg Tyr Gly Thr Phe Ile Glu Ala Thr
660 665 670

Asn Ser Asn His Asn Thr Tyr Phe Ile Gln Glu Thr Asp Phe Ser Gly
675 680 685

Ile Glu Lys Val Ile Gly Met Ser Asn Asn Ile Leu Phe Phe Lys Lys
690 695 700

Gly Ser Ser Tyr Ile Thr Gly Tyr Pro Lys Thr Phe Gly Ser Thr Ile
705 710 715 720

Thr Gly His Arg Ser Tyr Thr Ser Ile Asn Ser Glu Ser Ser Asn Leu
725 730 735

ES 2 622 280 T3

Gly Ser Asn Phe Ile Ile Tyr His Ser Asn Ser Lys Leu Tyr Gly Lys
740 745 750

Gly Ile Ala Asn Ser Gly Gln Phe Gly Asn Ser Thr Asn Ile Asp Gly
755 760 765

Thr Ser Asn Tyr Asp Thr Gly Leu Lys Asp Ile Lys Asp Ile Ile Val
770 775 780

Lys Gly Asn Thr Val Val Val Val Asp Lys Asn Asn Asn Ile Tyr Val
785 790 795 800

Thr Gly Met Asn Gln Asn Asn Lys Leu Gly Ile Gly Glu Tyr Asn Asn
805 810 815

Glu Pro Val Lys Lys Phe Thr Asn Ile Thr Glu Gln Ser Asn Ser Phe
820 825 830

Ile Phe Met Asp Asp Ile Lys Glu Ile Thr Thr Ser Arg Asn Thr Met
835 840 845

Phe Ile Val Lys Asn Asp Gly Thr Ala Tyr Ala Thr Gly Asn Asn Ser
850 855 860

Ser Gly Gln Leu Gly Leu Gly Asp Thr Ile Asn Arg Asn Lys Phe Thr
865 870 875 880

Gln Ile Asn Leu Asp Asn Ile Lys Lys Ile Ser Thr Ser Ile Asp Gly
885 890 895

Asn Thr Thr Phe Ala Ile Arg Asn Asp Gly Thr Leu Tyr Ser Thr Gly
900 905 910

Leu Asn Thr Lys Gly Gln Leu Gly Leu Gly Asp Ile Val Asn Arg Asn
915 920 925

Thr Phe Thr Lys Val Asn Ile Gln Asn Val Arg Asp Val Val Leu Gly
930 935 940

Thr Thr His Ser His Ala Ile Lys Asp Asp Asn Thr Leu Tyr Ser Cys
945 950 955 960

Gly Glu Asn Thr His Gly Gln Leu Gly Leu Gly Ser Glu Ser Asn His
965 970 975

Pro Asp Val Leu Thr Phe Thr Val Asn Asn Ile Thr Asn Val Arg Asp
980 985 990

Val Tyr Cys Ser Asp Thr Thr Thr Phe Ile Val Lys Asp Thr Asn Ile

ES 2 622 280 T3

995			1000					1005						
Ala	Tyr	Cys	Cys	Gly	Tyr	Asn	Asn	Asn	Ser	Gln	Leu	Gly	Met	Gly
1010						1015					1020			
Asn	Thr	Thr	Asp	Gln	Tyr	Ser	Phe	Ile	Lys	Cys	Met	Glu	Asn	Val
1025						1030					1035			
Lys	Glu	Val	Ile	Pro	Asn	Glu	Ile	Asn	Thr	Tyr	Ile	Ile	Thr	Ile
1040						1045					1050			
Tyr	Asn	Thr	Ala	Tyr	Ser	Thr	Gly	Leu	Asn	Thr	Asp	Tyr	Cys	Leu
1055						1060					1065			
Gly	Leu	Asn	Ser	Asn	Ser	Asn	Gln	Ser	Ser	Phe	Ser	Glu	Ile	Pro
1070						1075					1080			
Ile	Ser	Asn	Val	Val	Lys	Val	Ala	Pro	Asn	Arg	Asn	Asn	Ala	Val
1085						1090					1095			
Leu	Leu	Leu	Thr	Ser	Glu	Gly	Asp	Val	Tyr	Thr	Ala	Gly	Lys	Cys
1100						1105					1110			
Ser	Asn	Gly	Ser	Gly	Thr	Gly	Ser	Glu	Thr	Pro	Glu	Lys	Ile	Lys
1115						1120					1125			
Lys	Ile	Ala	Ser	Lys	Ala	Lys	Asp	Ile	Gly	Met	Asn	Tyr	Arg	Cys
1130						1135					1140			
Gly	His	Tyr	Val	Ser	Asp	Asn	Gly	Asp	Leu	Tyr	Gly	Thr	Gly	Phe
1145						1150					1155			
Asn	Asn	Asn	Gly	Gln	Leu	Gly	Val	Gly	Asp	Val	Thr	Lys	Arg	Asp
1160						1165					1170			
Thr	Phe	Ile	Lys	Thr	Asn	Thr	Arg	Val	Lys	Lys	Ile	Leu	Pro	Leu
1175						1180					1185			
Glu	Tyr	Ala	Asn	Ile	Ala	Ile	Lys	Asp	Thr	Asn	Asp	Ile	Tyr	Ile
1190						1195					1200			
Cys	Gly	Leu	Asn	Asn	Tyr	Gly	Gln	Leu	Gly	Val	Gly	Asn	Arg	Tyr
1205						1210					1215			
Asp	Ser	Arg	Asn	Asn	Asp	Asn	Arg	Ile	Phe	Asn	Tyr	Lys	His	Met
1220						1225					1230			
Asn	Phe	Val	Met	Gly	Asp	Leu	Thr	Ser	Ile	Lys	Asn	Arg	His	Asn
1235						1240					1245			

ES 2 622 280 T3

Phe Ile Leu Leu Asn Asn Lys Ile Val Ile Pro Thr Thr Lys Asp
 1250 1255 1260
 Ile Asp Tyr Gly Leu Val Leu Gly Asn Leu Tyr Lys Gly Asp Leu
 1265 1270 1275
 Tyr Thr Glu Leu Pro Tyr Glu Asp Ile Lys Glu Val Ser Ile Ser
 1280 1285 1290
 Lys Thr His Ile Ile Ile Leu Leu Asn Asp Gly Thr Met Tyr Gly
 1295 1300 1305
 Cys Gly Thr Asn Tyr His Gly Glu Leu Leu Gln Asp Leu Ser Ile
 1310 1315 1320
 Asn Gln Val Asp Glu Phe Val Gln Ile Asn Val Ser Asp Val Lys
 1325 1330 1335
 His Val Ser Cys Gly Asp Asn Phe Thr Tyr Phe Ile Lys Ser Asp
 1340 1345 1350
 Asp Ser Leu Trp Ser Ile Gly Lys Asn Ser Glu Tyr Gln Leu Gly
 1355 1360 1365
 Ile Gly His Asn Asn Pro Val Thr Glu Leu Gln Arg Ile Thr Thr
 1370 1375 1380
 Ile Ser Ser Cys Lys Glu Val His Cys Gly Lys Asn Tyr Thr Leu
 1385 1390 1395
 Val Val Thr Thr Gly Asn Glu Leu Phe Val Gln Gly Tyr Asn Asp
 1400 1405 1410
 Lys Gly Ala Leu Gly Leu Gly Ser Asp Ser Glu Asn Thr Ile Ile
 1415 1420 1425
 Lys Phe Phe Thr Lys Ala Leu Thr Asp Ile Arg Glu Ile Lys Ser
 1430 1435 1440
 Tyr Gly Ser Asp His Ile Leu Val Leu Lys Asn Asp Asn Ser Val
 1445 1450 1455
 Trp Val Thr Gly Lys Asn Arg Asp Val Tyr Lys Ile Glu Gln Pro
 1460 1465 1470
 Val Glu Phe Leu Lys Glu Phe Thr Ile Val Pro Ile Ser Glu Asp
 1475 1480 1485

ES 2 622 280 T3

Val Asn Thr Val Lys Asp Val Leu Ala Thr Asp Asn Thr Leu Tyr
 1490 1495 1500

Ile Ile Ser Glu Val Gly Thr Thr Asn Ala Ala Ile Glu Ile Thr
 1505 1510 1515

Glu Lys Ser Ile Ser Ser Ile Lys Ile Lys Ile Gln Asp Pro Asn
 1520 1525 1530

Lys Asp Ile Ser Arg Ile Glu Met Leu Ile Asn Gly Glu Ser Val
 1535 1540 1545

Lys Ser Val Ser Asp Leu Thr Thr Glu Lys Ile Ser Phe Glu Val
 1550 1555 1560

Pro Pro Asp Lys Ile Lys Ile Gly Glu Asn Lys Ile Leu Phe Arg
 1565 1570 1575

Ala Tyr Cys Lys Gly Asp Asp Leu Tyr Ala Ser Leu Phe Ile Phe
 1580 1585 1590

Lys Glu Ser Thr Gly Asn Ser Ile Ile Lys Asp Ser Tyr Val Met
 1595 1600 1605

Ile Gly Asn Arg Met Tyr Lys Val Val Asn Thr Thr Ser Asn Glu
 1610 1615 1620

Gln Asp Ile Thr Ile Thr Leu Asp Arg Gly Leu Glu Glu Asp Leu
 1625 1630 1635

Asn Leu Gly Asp Pro Ile Tyr Gln Leu Ile Asn Lys Thr Lys Val
 1640 1645 1650

Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp Met Lys Leu
 1655 1660 1665

Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu Ile Tyr Glu
 1670 1675 1680

Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys Ile Ile Val
 1685 1690 1695

Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro Ser Met Ile
 1700 1705 1710

Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln Ala
 1715 1720 1725

<210> 18
 <211> 98
 <212> PRT
 <213> Clostridium difficile

5

<400> 18

ES 2 622 280 T3

Met Gln Tyr Lys Asp Ile Ser Asp Ile Ser Ile Gly Gln Ala Lys Gln
 1 5 10 15

Asp Asp Asp Ile Thr Asn Asn Phe Ile Ala Asn Ala Glu Ile Tyr Glu
 20 25 30

Met Leu Leu Ser Gln Ser Ser Val Asn Glu Ala Ser Asn Ile Ser Thr
 35 40 45

Phe Ser Val Arg Lys Ser Gly Gly Glu Ser Gly Met Val Glu Val Tyr
 50 55 60

Val Ala Leu Ile Leu Arg Gly Arg Lys Thr Ile Glu Glu Val Pro Ala
 65 70 75 80

Val Ile Arg Glu Gln Val Arg Ile Arg Cys Lys Glu Leu Glu Ile Pro
 85 90 95

Val Glu

<210> 19

<211> 86

5 <212> PRT

<213> Clostridium difficile

<400> 19

Met Asp Lys Leu Ile Thr Glu Leu Ser Ser Leu Gly Ala Ile Gly Ile
 1 5 10 15

Leu Cys Ala Leu Leu Phe Lys Asn Thr Met Gln Glu Lys Lys Glu Asp
 20 25 30

Arg Asp Met Tyr Lys Lys Thr Val Glu Asn Phe Ile Glu Leu Ser Thr
 35 40 45

Gln Gln Gln Glu Ile Asn Lys Asn Ile Leu Val Gln Met Gly Ile Met
 50 55 60

Lys Thr Asp Val Glu Glu Ile Lys Glu Asp Val Thr Asp Ile Lys Gly
 65 70 75 80

Met Leu Gln Asn Gly Val
 85

10

<210> 20

<211> 151

<212> PRT

<213> Clostridium difficile

15

<400> 20

ES 2 622 280 T3

Met Gly Leu Arg Asp Lys Phe Ala Gln Ser Phe Ala Arg Ser Lys Thr
1 5 10 15

Met Ser Gly Pro Glu Lys Lys Ala Asn Glu Ile Met Gly Lys Leu Leu
20 25 30

Leu Lys Lys Ala Ile Leu Pro Ile Val Leu Met Phe Val Ile Ile Ile
35 40 45

Ala Gly Ala Met Leu Lys Ile Asn Ser Trp Val Thr Leu Gly Ile Asn
50 55 60

Leu Val Ile Ala Val Gly Ala Phe Phe Tyr Ile Arg Asn Ser Ser Lys
65 70 75 80

Lys Tyr Gln Asn Phe Lys Pro Tyr Val Gly Asn Leu Ile Ser Leu Glu
85 90 95

Lys Lys Gly Lys Lys Glu Tyr Val Ala Ile Ile Lys Gln Gly Lys Leu
100 105 110

Pro Val Lys Leu Gln Ile Ala Tyr Gly Gly Glu Asp Leu Glu His Val
115 120 125

Lys Lys Asn Gln Met Val Gln Ile Ser Tyr Asn Pro Asp Ala Lys Ile
130 135 140

Ala Ile Leu Val Asn Arg Gln
145 150

<210> 21
<211> 139
5 <212> PRT
<213> Clostridium difficile

<400> 21
Met Asp Thr Leu Gly Glu Arg Ile Val Tyr Leu Arg Lys Ala Lys Asn
1 5 10 15

Leu Lys Gln Tyr Glu Leu Glu Glu Met Leu Gly Cys Asp Asn Leu Ser
20 25 30

Lys Phe Glu Arg Asn Ile Arg Lys Pro Asn Tyr Glu Ile Leu Lys Ser
35 40 45

Ile Ala Glu Ile Phe Asn Val Ser Val Asp Trp Leu Leu Asn Gly Asp
50 55 60

10

ES 2 622 280 T3

Asn Leu Ser His Lys Ser Asp Leu Ile Cys Asp Ser Ser Ser Asn Tyr
65 70 75 80

Pro Leu Asn Ser Ile Asn Ser Asn Glu Ile Lys Leu Leu Asn Asn Phe
85 90 95

Arg Lys Leu Asn Asp Tyr Asp Lys Ala Lys Ile Glu Gly Met Ile Glu
100 105 110

Leu Lys Leu His Glu Tyr Glu Lys Glu Lys Asp Leu Gly Lys Ile Glu
115 120 125

Tyr Asn Lys Asn Lys Asp Glu Lys Ile Asp Lys
130 135

<210> 22

<211> 69

5 <212> PRT

<213> Clostridium difficile

<400> 22

Met Asp Ala Arg Lys Lys Trp Ile Pro Phe Leu Gly Val Gln Val Lys
1 5 10 15

Gln Arg Leu Ile Glu Leu Asn Met Thr Gln Arg Glu Leu Ala Lys Lys
20 25 30

Ile Gly Val Asn Glu Asn Tyr Leu Ser Ala Ile Leu Asn Gly Arg Arg
35 40 45

Thr Gly Lys Lys Tyr Lys Ser Ser Ile Tyr Gln Leu Leu Asn Ile Glu
50 55 60

Tyr Ser Glu Asp Asp
65

10

<210> 23

<211> 132

<212> PRT

<213> Clostridium difficile

15

<400> 23

Val Asp Thr Leu Gly Lys Arg Ile Ala Tyr Leu Arg Asn Ser Lys Lys
1 5 10 15

Leu Thr Gln Arg Lys Leu Met Asp Ile Leu Lys Phe Glu Asn Leu Gly
20 25 30

Lys Tyr Glu Thr Gly Asp Arg Lys Pro Asn Cys Asp Ile Leu Met Ser
35 40 45

ES 2 622 280 T3

Ile Ala Asp Tyr Phe Asn Val Thr Thr Asp Trp Leu Leu Tyr Gly Lys
 50 55 60

Glu Lys Val Asn Val Asn Ser Ser Val Lys Glu Asp Lys Glu Asp Tyr
 65 70 75 80

Leu His Val Thr Asn Asp Glu Met Met Ile Leu Asn Leu Tyr Arg Gln
 85 90 95

Leu Asn Glu Arg Asp Lys Ile Lys Ile Glu Gly Ile Leu Glu Leu Lys
 100 105 110

Ile Ser Glu Tyr Lys Asp Leu Lys Lys His Ser Ser Asn Asn Asn Glu
 115 120 125

Asp Lys Met Val
 130

- 5 <210> 24
 <211> 40
 <212> ADN
 <213> Secuencia Artificial
- 10 <220>
 <221> fuente
 <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"
- 15 <400> 24
 ttcttgaag accatcgaag caccaccacc accaccactg 40
- 20 <210> 25
 <211> 41
 <212> ADN
 <213> Secuencia Artificial
- 25 <220>
 <221> fuente
 <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"
- 30 <400> 25
 tttttgaag acaatcgaag ggcttcgccc tgtcgctcga c 41
- 35 <210> 26
 <211> 38
 <212> ADN
 <213> Secuencia Artificial
- 40 <220>
 <221> fuente
 <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"
- 45 <400> 26
 ttccttgaag acctaattg gggcaatccc gcaaggag 38
- 50 <210> 27
 <211> 39
 <212> ADN
 <213> Secuencia Artificial
- 55 <220>
 <221> fuente

ES 2 622 280 T3

	<223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
5	<400> 27 ccccttgaag acccaattc gtaggcaat gaaagacgg	39
	<210> 28 <211> 59 <212> ADN <213> Secuencia Artificial	
10	<220> <221> fuente <223>/nota = "Descripción de la secuencia artificial: oligonucleótido sintético"	
15	<400> 28 aattggggcc gcagctcgct agcggtagct cgaggatc ttcgaagaag acacatccg	59
20	<210> 29 <211> 36 <212> ADN <213> Secuencia Artificial	
25	<220> <221> fuente <223>/nota = "Descripción de la secuencia artificial: oligonucleótido sintético"	
30	<400> 29 aattccggga tgcagcctc taggatccgg cgcgcc	36
35	<210> 30 <211> 55 <212> ADN <213> Secuencia Artificial	
40	<220> <221> fuente <223>/nota = "Descripción de la secuencia artificial: oligonucleótido sintético"	
45	<400> 30 agctggcgcg ccggatccta gaggcatgca tccggaatt cggatgtgc ttctt	55
50	<210> 31 <211> 40 <212> ADN <213> Secuencia Artificial	
55	<220> <221> fuente <223>/nota = "Descripción de la secuencia artificial: oligonucleótido sintético"	
	<400> 31 cgaagatc ctcgaggtag cgctagcgag ctgcggccgc	40
	<210> 32 <211> 8119 <212> ADN <213> Clostridium difficile	
	<400> 32	

ES 2 622 280 T3

ggccgcaata cccactacac cttcgtcatc tttaaattta agagttttta ctattgaata	60
ataaaggat attccagtaa aaataatctt taaatacaag aaaaataaac tctttgggta	120
tattaaaaag ctaaaaagtg taaatataaa agcaagtaga gtacttatcc tgtaaaagaa	180
atctatttgt gtaatgtctt tatattttat cataaacacc gaatataaaa tgatgaaaat	240
aattgcgacg attgcatata tggtaaataa catattttca agagtacat ttgaaattac	300
tatccactta taccacataa ttggccaaaa taatagtgtc aagaacttaa aataattatc	360
aaacaacttt tctttataca ttcatcaaac aacctttctt acaaaaagca tatatttggt	420
tttagaattt taaataatat gatatcatta ttatatatta atattgaatt tatagaaacc	480
aaaatttggt aaaataaata tatagatttt actgttaagc cagttaaaat tactactatt	540
tttattatga aattggatca aatatgtaga aatacggcaa attagttaat attaaatatt	600
tattatttcc aagttgtaaa gactgttttt ttaatgataa aaattctaatt cttttttgaa	660
agaaagtaat atccacatta agtatgtctg ccatttcata aacgcaagtg atgccagaat	720
taattatggt tattatatct tcttcagtaa ttaagaactc acaagccat ttaaggctt	780
tattttcgca cttatctata ataatttttg tataataatc gttataagag gatacatagt	840
atccaaggct agtgaaatga tgtccaagtt cttcagctaa gatggatgtc aatttttttg	900
agttttggtt taaattactg agtaatgata taattttaat accatgtttg tttatatata	960
gcccttctaa atcacctgca atataagtgg tataatgaat tattatctct tcttgagaag	1020
ctaattcaaa aagcttatcc aaattattca taaaaatccc cctaaaatag aatgtatggt	1080
tgcctttaa ttatattaaa agagcagaaa aatagactgc tcatcatatg gtttattttt	1140
ttttatattt atttagtaaa aattctatat aatcattaag ttgttcttgt gcttcttcag	1200
gtaactcttc atgtggattt tttctatgtg cagctactgt atcaatattt tccttaacta	1260
aggttcttcc aagaaggtaa tcaactgata cattaatac atcagccaat ttgtttaaaa	1320
tgtgttcac aggaaatctg ttttctgttt catagtaccc taagactctt tgggaaacgc	1380
ctactttttc tccaagttct ctttgagtca atccaaattc ctttctaagt tctcttaatc	1440
ttttggcaaa cattataaca ccacctatg tatagattat aacaaattgt tctaaaaaat	1500
aaaactaata aaatataaaa gaatattttt tctaaaatct attgataaag aacaaatgat	1560

ES 2 622 280 T3

tctatataat ctaagtgtgg aagaacaaaa tattcttaaat ggtaatggag gtataaaaca 1620
atgtttaaaa ataacttgaa atattataga aaatgcaaag gtatgacaca aattcaactt 1680
gccagaaagg ctggaattac aaatgactat atatctcaa tagaaagagg tataaaaaat 1740
cctggtcttc ttatggctaa gaagatttct agtatttttag aacaaaatat agaagaagtt 1800
ttttttatac agttatagaa caatatgttc ttgaaagttg tgagattagt aaaaaactgt 1860
gcactaaaga gattattgta aatttgaagc taataataag tatataaaaa aggagaagta 1920
ctatggaaaa caaaaaagat atattattta aagaaacaga tgaaagatta cataattata 1980
agtatttgga tataaagata aagaatatta atttggacat aaaaagatgt gagaatgaat 2040
actctggatg tggagcaatg gtatatacag aaaagactag taacacatat aacataagtt 2100
cttctgtgga aaatgaggtg ttaaaaagag aggaaagatt aagaaaatta aaaatggaaa 2160
aagaagatat agaaatagaa aaagagaaga tagaaaatgc tctaacatgt ctaaagata 2220
tagaaatgga attttttaaat cttttttata atagtaagac aaaaaacaat atgacatata 2280
tttctatgaa actacactta gatagaacat cttgctacaa tttaaagaaa aaaatgatat 2340
ttaaattgag tgagatatta taaaaaatat gacaacttta caacacttta tatacactat 2400
tgcaacacta ggcaataaaa tatgtgagat aatgttattg tgaaagaaat ccatattgaa 2460
ggaggtgata aattgaaaag aataatatta cctaaaaata tagaagatat ttgacaggaa 2520
taaattgagat gtatatttaa aaatgactta tatcatttat agtaagatta tcagattaag 2580
caagaatatt tagtgatagt gtggtgatta tttgcttaa tacaaggaaa tattagaaac 2640
aattattgag attctcaaaa aaaactttac tgaaagtatt tttattgatg atgaaagtgt 2700
gcaaggctct gaagggtctt gttttttgt aagtatacta tcagttattt gtacacctat 2760
aatgttaaat acgaataata aagatattgt tatctctata aaatacttac caaaaccaca 2820
gtcaaagagt attagaatgt atgaaatttc agatgaatta aataagttat tcaacagaaa 2880
tataaaggta acagacagaa aattaaatat aacaaagcta gaacaaagta ttaaaaaaga 2940
agagtcaatt tatgtattga actttacaat tacaactaat tatctggata gtgtatatga 3000
agaagatgta gtatatgaaa atatggaaga aatcaattta aatttaggag agtgatagta 3060
tggctatagg attaccaagt atcaacatat catttaagga gctagctaca actgttaaag 3120
aacgttcagc tagaggaata attgcaatgg tgcttaaaga tgctaaggca ctaggcttta 3180
atgaaataca tgaaaaagag gatataccag ttgatttatc tgctgaaaat aaagagtata 3240
taaatttagc tttgatggga aatgttaaca ctccaaataa attattagtt tatgtaatag 3300
aaggagaagc agatattcaa actgcattag attttttaga gactaaggaa ttaattatc 3360
tatgtatgcc aaaagcagta gaagctgata agactgctat aaaaaattgg ataattaaac 3420
ttagagatat agataaggtt aaggttaaag ctgtattagg aaaagttgta ggaaatcatg 3480
aagggataat taattttact acagaagatg tgttagttgg agaaaagaaa tacagtgttg 3540

ES 2 622 280 T3

atgagtttac aagtaggggtg gctggactta tagcaggaac acctttaagt caatcagtaa 3600
 cttatactaa gcttagtgat gtagttgata tacctaagat gacgaaagtt gatgcagaat 3660
 caagggttaa taaaggagag cttatactta ttaaggaagc aggggctata agaattgcaa 3720
 gaggagtaaa ttctttaact gagttaacag aagaaaaagg agaatgttc cagaaaataa 3780
 aaatagttga cactttagat attatacata gtgacataag aaaggtgata atagatgact 3840
 atataggaaa ggttactaac agttatgaca acaaatgttt attgatagta gctataaaaa 3900
 gttatttaga agaattagaa aagtcagcac ttatagaatc tgattctact gttgaaatag 3960
 attttgaagc acaaaaatcg tatttaaaat caaaaggagt agatttatct tatatgacat 4020
 tacaagaaat aaaagaagct aacacaggtt ctaaagtatt tttaaaagca aaaataaaag 4080
 tacttgatgc tatggaagat atagatttat caatagaaat ataggaggat tattaatatg 4140
 gcaaatatgg aagctagaaa tgtaatgagt ggtacttggg gagaactttg gcttgatgga 4200
 aacaaagtag cagaagtaaa gaagtttcaa gcaaagatgg aatttacaaa agaggatatt 4260
 ataatagcag gtcaaatggg tactgataca aagtatatgg gatataaagg aaaaggttca 4320
 ataactctat accatgtagg ttcaagaatg cacaagttaa ttggagaaaa gataaagaga 4380
 ggttctgaac cttagattgt tgctatatct aaattaaatg acccagattc ttatggagca 4440
 gaaagaatag cagtaaaaaa tatagcattt gatgatttaa ctttagctga ttgggaggtt 4500
 ggagtaaaaag gagagataga agcaccttc acatttactg agtatgattt tcttgatata 4560
 attagtttt atatttgggtt ttatactgat attagtaga tatatactta ataaatttag 4620
 gtagttaata agtaaaaaag ttagttgatt gaatttgatt gataaaggag caaataataa 4680
 tgaatgaaaa tggattatca aaaaatataa acatagtaga ttactttta aatgcagata 4740
 cagaaaactt agaagacca agtactatag ttgaacttaa gagattatca actatatttg 4800
 ggcaggaatt taaagtaatg tgtagagctt taacaataag taaagatgaa gagatacaaa 4860
 atacttgtct taaaattgat gaaaatatga aaacggatat agacttaccg gagatgcaga 4920
 tgcttacaat tatagaaggt gtttgtgatt tggatggaaa gcttttattt aaaaataag 4980
 aactaatgga taaatttaag gctccaacac caaaagaatt ggcaagaaaa ctattattac 5040
 caggtgaaat taccaaccta tatagaatac ttcaagatgt tatgggttat ggtaaaaatg 5100
 cagtgataga agaggtaaaa aactaatag gacggatacc aagactacaa taatgtacta 5160
 ttattggaag aaaaaggta taagaccgtc cttttttat gcaatggata aaggcgaatt 5220
 aaagcttatt gaagcttttt tgccttaga aattgaggaa gaagttgaaa aaatgaaaca 5280
 tggatatgga gtgtgcctt tgacaggag tggatgtaa tgggaaatgt gagagaagaa 5340
 ggtataaata tgtatcttac agataattac acaccaaaaa tgaaccaa at tatatcagta 5400
 actgataatt ttaggagagc aactgtggct gtttcacttt ccactaatgt aatggctagt 5460

ES 2 622 280 T3

agcataaaaa attctattgg aagtgcaagt agtagagtaa acagtttaaa ttcctcgtta 5520
 agaaaagttc aaactactgc tagtagtgta agttcaacta tggcaaaatt aagttctagc 5580
 ataaatgctg tttcaggagt tattggaagt ttaaattgaa gtattatgag actagcaata 5640
 actatagcta tgattattga ttattttaat aagttgattc aaaagaaaaa tgagtttaat 5700
 tcaaataatta tgattatatt aatatttaaa gctaaaagtg atgaagtaga aaaaactaaa 5760
 aataaattac ttggaattt aaaaaagatt ggtggcaaga tttggaatat cgtaataaaa 5820
 gcaaaaagata tgactaagag agtgataagt agtatcttgg gaaaattaaa acgagtagag 5880
 aaacgtcctt atcaaggaag tattaatctt aaagatatgg taagtagtgc tatggctaga 5940
 attttgcccta agttaatggt gtttaaaaat actttttgga gtggtgtaat agctataaaa 6000
 gatatggcaa gtagcattat aagtaaagta tttcccaaat tgagattggt tgcaggtaag 6060
 gtatggagtg gtgcaatagc tgtaaaggat atggcaagtg gaataacttg ttcgataaaa 6120
 gggaaagatat ctgatttgac aaatgggtgct actataggtg tgcgtgtgaa aaaggggtgt 6180
 gacttacttg gtcaggaaca aaatcagaaa gttgttctag aaagtgtaat gaaaagaaat 6240
 actggaaaaa ctagccaaaa agatggtgat aagtattatg acagtttagt aaatatggca 6300
 aatgatacgc cttttgacct tgaagatggt gttgcaatgg gaactaaagc taaaatgatt 6360
 agtaataatta ctggtggcaa aaaagaaaaa gatataactc aagctatggt agatgttaga 6420
 gctttaaata tgaatacaag tagtgaacaa gatgtatcag cagctttctt aagtgcagca 6480
 aaaggaataa tggaaatctct taatactctg gtaggagaaa attataaaac ttttgatgaa 6540
 gcattggaag gcataagtgt aaagcagatg gggtagcta aagaaatgag taatacaata 6600
 ccaggtataa tatcaggagc tcaacaagc attacaatg gtttgaagag tattgttaaa 6660
 cctttgatg atatttttag tcaaggacta aagaaaataa aaacttttat agaaagtgga 6720
 ttagggaatt tagctggcct atctgaaaaa atggctggta aaataggcaa tgtaatgaat 6780
 ggtaagataa ttattggcaa caaatatgac cagatgcaat ctagaagtgt aaaaaatgga 6840
 aaagagtttt ctgattctac tcaatatcga atttctaag aggctgaaaa gcgtaaaatg 6900
 atggttgaaa ataagcaaga acgttttgaa aatcatgcag caacaatgat agggaatgca 6960
 ccaaaaagcaa ttgttaacgc aggaagtaca ctattacaaa atattgattt tacagcatta 7020
 atagattcac tacttccagt agtaaaacta gtaaataatt tactagatag tataaacaat 7080
 aaatcaccaa ttgcacaagg attaataagt atatttggtta caatagtaac tacagcattc 7140
 caactaatcg gacctgtagt tgaagctggt agtcctatta tcacaagaat ttttactttt 7200
 ttaggtgaat atgcacctca aataaacaat tttatagaga cactgggtgt tatttgghaa 7260
 actgtatggg agaccttagg acctctgttg gaaactggat ggaaaattat agagccaata 7320
 ttgggagcct tttttaacat attagataaa gtatgtaaaa tagttaaaga tatatgcaaa 7380
 tgggtggcaaa ctatgattaa taagataaaa aatggaagca tcacaggaac agttttaaat 7440

ES 2 622 280 T3

ctagtggaaa agagtaaaaa aaattacaaa gataatccat atgctggaac aaaggctggt 7500
gattctggta aagcttattc aagtaagaaa ggtaataatg catttggatt gaactatggt 7560
ccttataatg actatcaaac cagactccat gaaggtgaaa tggttttaac taaacaagaa 7620
gcaaatcaat atagaagcag aaaaaatggt ggaaatataa acatagctaa gttagctgat 7680
acaatagtga ttagagaaga agctgatata gaaaagataa catcaaaatt agttgcaagt 7740
atccaattgg cacagttagg ggggtgtctta taatggaaat gtggcttaga caagcagaag 7800
atagatttag atttccagta tttccatctt cctttagtat taatggaaaa gctgctgtaa 7860
actcttctag tataactcaa ataggtgaag tagcaacttt tgggtggtgta gctcttaaaa 7920
gcatttcaat atcaagtttt tttccaaata aagactacac tttctgtgac tatacagggt 7980
ttccatcacc atatgattgt gtaaataaga tagaaaaatg gatgaaggaa ggttttatat 8040
taagatttac aattacggaa acaaatataa atatggaagt cataattgaa gggtttagtt 8100
atgaagaaag agatgggac 8119

<210> 33
<211> 7686
5 <212> ADN
<213> Clostridium difficile

<400> 33
tcgagatgta tattttacat tagatttaaa agagtataaa agaataaaga taccaaaaagt 60
aactccaaaa caataactat tatagataat aagttgtaag taactgctga tagaattaaa 120
tgaaaaggca ggtgattttt tattattaag atttgggtac acataaaaaa cggaagtata 180
tatgacataa ctgacatagt agacaaggta tcatggtcag gtgattataa atctccatca 240
aggacactag agttttcaat aatacaatca tcatgtgatg taaatttcca acaaatcgat 300
ataccaatag ctagtacagt ctgtttctat gtagatgaga aagaactctt tagaggaatg 360
ataattaata ggtctaaaga ttcaagcagt aatgaaatta gttttgtatc taaagatatg 420
ggatttttac ttacacaaag tgaagtgtca tacaatttta aagataagtt agttgaagac 480
atagcaaagc aagtatttgc tgaaaatagg ctttcagttg gaacaatagc aaagaccaat 540
gtcaagtata caaagatggt tataggagta aatggttatg acacaataat gagtgcatat 600
acagaggcaa gtaaaaagac aaagaaaaag tatatgatag aggctaattt agataagttt 660
aatgttattg aaaaaggaac tgttacatta agtgttatgt ttgaagaggg atttaaatatt 720
ataaatacca ctttttcgga gagcatggaa aatgtaaaaa ataaagtaat agtggtagac 780
cagtatggaa gcaagattag cgaaaaata gataatgaaa tttttaagga agtaaatgta 840
ataatgcaaa aagtaattca gcaacaagaa aatcaagatg tagatattga tagcgagttt 900
aatgggatag aaaaagctg ttctcttaaa ggttatggag atgtaagttg tataactggt 960
agaggagtaa aagttaaaga ttcttataca aagctttagt gactatttta tatagataca 1020

10

ES 2 622 280 T3

gacaaacata cttggcaaaa tggagaatat caaatgagc ttgaacttaa ttttcaaaat 1080
cttatggatg aaaagtcagc aggacaggat gaacctaaagg aagaaagtaa tttaggggga 1140
gaagattatg caggaggaaa agagtttaca gcagaattta cagcttactg tcctagaaaa 1200
gaagaagggtg gagatacaga ttgtagaag aaaaaacttg acccatctaa aaaacttgcg 1260
ctgctcctat ggttggtaaa tatgagcaaa cttattatac aaaagagttt ttaataaac 1320
atcctttatt gaactatgga gatgaaatac aggtaattac aggagtttct ggtcgtgatg 1380
gagtctataa agtaaatgac gtaggacctg caataactat agaaaaaat ggaacatacc 1440
atatagatat tttatttggg aatgttgaag aagctagtaa atttggaaga agaaaaggaa 1500
aaattattat tgggtggtat tctggtaatg tatctgataa agctaaaata gtaatatcag 1560
aggcaaaaaa acatctagggt aaaccttata aatgggggtg aaatggacca agtagttttg 1620
actgttctgg tttaatggtc tactgtttta aaaaagttaa tgttagtttg ccaagaacgt 1680
caaatcaaca atctaaaaa ggcaagaag tagaacaaaa aaatcttcaa gcaggagatt 1740
tagtattttt tcataatcca gtcagccatg ttggattata tataggtaat ggagaatttt 1800
tacatgctcc acaaaaagggt gatgtagtta aaataagtaa gttaagtagt agaagagatt 1860
ttaatacagc taggagagta ttataaaagg atggtgatat aatggctaata ccaataaatg 1920
aatttatagg aataataaga gaagaaggaa agtatcataa tcaaccttct tttttattgg 1980
aaaattaaaa gtaaattacc agatttaaaa atagagacaa ataacatcat attagaaaaa 2040
gaagatattt tgatagatag ttggatgatt gatagacagc tagaaacatt tgacacagaa 2100
acaaatcaag aacaccagca tgaagtaaaa aatcctttta tagataactt tgaatctggg 2160
gatatggtaa taatgtttag aataggcgaa aaatttgctg ttgtaagtaa gttggtgagc 2220
ttataatgag tacaatattt ccttttatag gtgtcccaga ggattatatac ttacctaaaa 2280
cagaagaatt gccaatcttt cgtgaagtgg catgggattt tgaaaaagat gaacctattt 2340
tagaaaaagg tgactttaaa ataattgaaa aaaaaagaag ccttaaaagt ttggatatac 2400
aagtgtataa agacaaatag atatgaacat gagatatact ctttagaata tgggacagag 2460
ctttcagaac taataggaca aaaatataca aaaggtotta cagaaagtga agctagtaga 2520
ttcataaaaag aggcccttct aataaatcca tatatattag aagtaaactg aaaaagtgct 2580
aactttaaca gagacgtatt gagtgcaaat gtaaaagtat ccactatcta tggggagggtg 2640
gaaataaatg tatagtgacc agacatatga agtaataaaa aatagaactc ttgaaatat 2700
taatcttgat atttataaag gagaaggttc ttttctaaac aacatggtat ctggaaataa 2760
tctagaactt tcgaagatat atctagaact ttcaaagata cataaaatgg cttttataca 2820
agacacatat aaccagtttc ttgataaaag agtcaatgaa tttggtgat atagaaagtt 2880
aggtagagag tcaaatggag aagttgaatt tattggagag aaaggaactg taataataa 2940

ES 2 622 280 T3

tggcacaata atatcatata gagatttact atttgtagta ataaaagatg taactattgg 3000
 tagtgaagaa ggtgacaata gcccgattca agctctggaa gttggaaga aatataattt 3060
 acctacaaat tgtgaattta aactagttga taatatatct ggagtaacaa agattactaa 3120
 cacaagaagt tttgaagggtg gtacagatat agagacagat gaagaactaa aagaagatt 3180
 ttataaaatc caaagaaatc aagctacaag tggaataaa gctcactatg aagaatgggc 3240
 tttggaagta gatggagtct ataatgttaa ggtttatcca agatgggatg gtccaggaac 3300
 agttaagggtc ttgatatttg gggaaaataa tcaagctggt gatacagaaa cgattgaaag 3360
 gtgtcagcaa catatagatg aagagaagcc tattggacca actataacag ttgtgacacc 3420
 attaccaata gaaataagta taagtgcagt aatgaaacta gaagatggat atacattaga 3480
 caatgtaaaa gaatctttcc tagaaagtat aaatacatac tttagagata ttagaggaga 3540
 gataatctat acaaaagtca tgggaatact tataaatact actgggtgtac acgatttaag 3600
 taatctactt ataaatggaa gtacagataa tataactatt aatgaagata aaatacctag 3660
 tgtaacaact gttaatttta gtgaggtgga aatcaatga agctaattga taaactacca 3720
 tcatttgata gaaattacat tgttagaggag atacaagggtg catacgatac agaattaaat 3780
 attcttaaag aagatattga tgataccttt aaccaattat ttgttgatac agcgacatgg 3840
 ggattagata tgtgggaaga catactctgc attgaaaaaa aagaacttga ttttgacaca 3900
 agacgtagca atataaaagc taaaatgaga agcagaggta ctagtactat tgaagttata 3960
 aaaagtatat gtgaggcata taaaaatca gaaacagata taaaagtta tagtgatgaa 4020
 tttacattcg tattgagttt tatagcaaat aactgtgact ataaaactct tttagattgt 4080
 agcgagatga ttgaaagagt aaaacctgct cacttattac actattttaga accaataata 4140
 ctagataaaa gtatggctta ttgtggtgga ggtatggtat gtagtgaaga ggtaaaagtt 4200
 catccatact ttgaaccaat tataaaatgt agtgctggtg taaactgtgg agctggaatg 4260
 ttaagtagag aagaaataaa ggtttatcct ttaagcatta aatgcattga aaataattgt 4320
 aagattaata tagctattgc aaatgataca ggcgtagaaa atgtagtagt ttatcctaaa 4380
 tcggaggtgg tataattgga agaaaaattt tatataatat taaccaaatt tggtagagaa 4440
 aaaatagcaa atgcaactgc actaggagag cttgttggat taaccaagtt tcaagttgga 4500
 gatagtaatg gagaatatta tgagccaaca gaggaacaaa ctgctttaa gaatgtagtt 4560
 tgggaaggaa atataaattc tctaagaatt gatgaaaaaa atcctaattg gatagttata 4620
 gagactattt taccaggaac agttggtgga tttatgataa gagaagctgc tgttctggat 4680
 aatgagaata atataatagc tatwgtaag tatccagaga cgtataagcc acgtgctgaa 4740
 gatggcagta ttaaagattt ggttgtaaaa atgattttac aattgtccaa tacttcaaat 4800
 gttacattag aagtagaccc gacgttggtt tttgtaactc aaaaggatat tcaagattta 4860
 gatgataagt ttgataaaaa tataaaagaa ataaaagtaa aaattggaga tacagatata 4920

ES 2 622 280 T3

ttaactacag attctaaaga tttatcagga gctataaatg aggtagttaa aaaaatagaa 4980
 aatataatcctt ttgatgatgt tataagtggg caaatacaaa ctgatataatc agtattaaaa 5040
 aatagctata acaaattatc tgaaaaagtg ctgatataat taatatacct agaattagag 5100
 tcagaagtaa ctgtagatga ggctggttat tggtagata cattagcaaa tggaaataac 5160
 atagtagcta tagaagggt taagttagat ttaaataagaa aatgtataac aggtgaaatt 5220
 ggtaatgtga tttttagaga tgtagtatta ccatttagtg caaatagagt tagatatata 5280
 catgatatgg ataataaact tgttgagaca aaatctagta acacttattt aaaagaacaa 5340
 aaagatataa ctctaagtaa atattcatat gaaataagat aaataaagga ggtagtacta 5400
 ataatgaagc aaaataaact tttacagcgt ggtgcttatt ttaatgataa gaacatattg 5460
 attgatgatt ttgataaaaag atataatgat tatgattttg tagaattttt tactggata 5520
 agtaatagta cctttggttt aaaatcagat ggtaatttat atgcttggg cgataatata 5580
 ggttttcaac taggacttgg aaaagattcg tcagagagaa ggatgtttag taaagtaaaa 5640
 attgataatg taaaatatgt atcttgggt tcaaaacaca gtgtagcagt aactaaagat 5700
 ggatttgc atggagcagg aacaagtaat gtaggtaat taggtgtaat tgagtctaca 5760
 gtatattatg aatttactaa gctaccaata gatgatgtaa aaactggtgc atgtggttat 5820
 gactttacat ttgtgcttaa aatgatgga acattatatt cagcaggttt aaactcaagt 5880
 ggtcaacttg gactagtgga tactaacaat agagctactt ttactaaagt aaatatagat 5940
 agtgtgaaag atgtagtgac ttataatcaa tctgtattta tcataaaaat ggatgggaca 6000
 gcacatgcat gtggattaaa ttcaaatggg cagttgggaa ttaatagtac tttaaataaa 6060
 agtgtattta ataaaataga aggtatggat aatgtaaaac agatagcgtg tggtagtagt 6120
 catacaattc ttattaagaa tgatggaact atgtatacta caggctataa tggagttggt 6180
 cagcttggta caggaataa taataattca attgtattta ctctttctag tataaataat 6240
 gttaagtatg cttcttggg aaataatcat actatgatat taaaatacga taatacactg 6300
 tttagtagag gacaaaacaa ttatgggtcaa ctagccaatg ccaataaaga tgtagcatca 6360
 agaaataactt ttgctaagggt taatgtagaa aatataaaag atattaaatg tggttctcaa 6420
 ttttaattttt taataaatgg ttcaaaagag atatttgtat ctggctgtaa tttagcaggt 6480
 caacttgggt cattttttca tacaactttt ctgtatgagt tttcaaatgt gcaatcttca 6540
 aatttagata attattcagg tttattgggt aatgatgatt atttatatgt tacaaggac 6600
 aatagtgaat ttttaaagt aaagttaagt gataattttc aagattataa gaagatagag 6660
 ttaacagata gcaatatggt tattgttatg aatgatggta cattgtatgc ttgtggttta 6720
 aataattatg gacagttagg attgggagat actgttaaca ggtcagttat gactaagggtg 6780
 gatatagata atgttttggg tataaaagga aacggaaact caacttttgt gcttaagaat 6840

ES 2 622 280 T3

aatggaacat tatattcatg tggtttaaat agtaatggac aattggggtt aagagatgaa 6900
 gttaatagaa atatatttac aaaaatagaa atagagaatg taaaggattt ttgtgtagga 6960
 agcaattatg tcatagcttt aaatcactca aaagaagtat atggatgggg aaataatcct 7020
 tataataata tagaaaaaac ttctaattat ccatataagc agggaataag taatattgaa 7080
 aagatagcag catatgatta ttctgtatat atgataaaca gtgaaggga actatatggt 7140
 tctggataca attataatta tcaattaggt aaaggaaata atagtaacca aagcaaagca 7200
 ttagtatctc aatgtagaac aaattcaaca tcttctacat caaatggact tagaacgtta 7260
 cctaaaataa ctaatgtttt tcctttttat gatggttgtg caataattga cgaaggagggt 7320
 tatgtttatt taacaggata tcatggatat ttaagaacat taaatagcag tccaagtata 7380
 tctgattatt caagatatgg aacttttatt gaggctacaa attcaaatca taatacttat 7440
 tttatacaag agactgattt tagtggaatt gaaaaagtaa tagggatgtc aaataatata 7500
 ttatTTTTTA agaaaggaag ttcatatatt actggatata caaaaacatt tggctcaacc 7560
 attactggac atagaagtta tactagtatt aattctgaga gttctaattt aggaagtaat 7620
 tttataatat atcatagtaa ttccaagtta tatggaaaag ggattgctaa tagtgggcaa 7680
 tttggg 7686

<210> 34
 <211> 7020
 <212> ADN
 <213> Clostridium difficile

5

<400> 34
 aattcaacaa atatagatgg cacaagtaac tatgatacag gattaaaaga cataaaagat 60
 ataattgtaa aaggaaatac tgtagtagta gtagataaaa ataacaatat atatgtaaca 120
 ggaatgaatc agaataacaa acttgggata ggggaatata acaacgaacc agtaaaaaaa 180
 ttcacaaata taactgaaca atcaaactca tttatattta tggatgatata aaaagaaatt 240
 acaacatcaa gaaatacaat gtttatagta aaaaatgatg gaacagccta tgccacagga 300
 aataatagtt ctggacaatt aggattaggt gacacaataa atagaaataa gttcactcag 360
 ataaaccttg ataataataa gaaaatatca acaagtatag atggtaacac aacatttgca 420
 attagaaatg atggaacact atactccaca ggattaaata ccaaaggaca actgggatta 480
 ggtgatatag taaatagaaa tacatttacc aaagtaaaca tccaaaatgt aagagatggt 540
 gttttagga ctactcactc gcatgcaatc aaagatgata acacattata ttcatgtgga 600
 gaaaacactc atgggcaact gggcttagga agcgaagca accatccaga cgtattgaca 660
 tttactgtaa acaatataac taatgtaaga gatgtgtact gctcagatac aacaacattt 720
 attgtaaagg acacaaacat tgcatttgt tgtggataca ataataattc acaactaggt 780
 atgggaaata ctactgacca gtatagtttt ataaagtgta tggaaaatgt aaaagaagtt 840

10

ES 2 622 280 T3

ataccaaatg aaataaatac ctatataata acaatctata atactgcata tagtacaggt 900
 ttaaatactg attattgctt aggtctaaat agtaatatgca atcaaagttc attttctgaa 960
 attccaattt caaatgtagt aaaagtagct ccaaacagaa ataatgcagt acttttactt 1020
 acaagtgaag gggatgtata tactgtaggc aaatgtagta atgggttcagg tacaggaagt 1080
 gagactccag agaagattaa aaaaatagca tcaaaggcaa aggatattgg aatgaattat 1140
 agatgtggac attatgtaag tgataatgga gacctatatg gtacagggtt taataataat 1200
 ggacaattag gtgttggtga tgtaacaaaa agagatacat ttataaaaa caatacaaga 1260
 gtaaagaaaa tacttccttt agaatatgca aatatagcaa taaaagatac taatgatata 1320
 tatatttgtag gattaaataa ctatggacaa ttaggtgttg gaaatagata cgatagtaga 1380
 aataatgata atagaatatt taattataag catatgaatt ttgtaatggg tgatttgaca 1440
 tctatataaa acagacataa ctttatactt ctaacaata agatagtgat acctaccaca 1500
 aaagacatag attatgggtt agtattagga aatttataca aaggagacct ttatactgag 1560
 cttccatatg aagatataaa agaagtatct atttctaaga ctcatattat tatattactt 1620
 aatgatggaa caatgtatgg atgtggtaca aactaccatg gagaattatt gcaagacttg 1680
 tctataaatc aagtggatga atttgtgcag attaatgtat cagatgtaaa gcatgtttca 1740
 tgtggagata actttactta ttttataaaa tctgatgata gtctttggtc tattggtaaa 1800
 aatccgaat atcaattagg tataggtcac aataatccag ttactgaatt acaaagaatt 1860
 acaactatat ctagctgtaa agaagtacat tgtggtaaaa actatacatt agtagtaact 1920
 acaggtaatg aattatttgt acaaggatat aatgataagg gagctttagg attaggaagc 1980
 gatagtgaaa atactataat taagttcttt acaaaagcac taacagacat aagagaaata 2040
 aaatcttatg gaagtgacca tatattagta cttaaaaatg ataattcagt atgggttact 2100
 ggaaaaaata gggatgtata taaaattgaa caaccagtag aatttttaa agaatttact 2160
 atagtaccta tttctgaaga tgtaaataca gtaaaggatg tacttgcaac agacaatata 2220
 ttatatatta tatcagaagt aggaacgaca aatgctgcta tagaaattac tgaaaaatca 2280
 atttcatcaa ttaagataaa aatacaagac cctaataaag atataagtag aatagaaatg 2340
 cttataaatg gtgaaagtgt aaaatctgta agtgatttaa ctactgaaaa aatatocttt 2400
 gaagtaccac cagataaaa taaaatagga gagaataaga tactatttag agcttattgt 2460
 aaaggtgatg atttatatgc atctttatct atttttaaag agagtactgg aaattctata 2520
 attaaagatt cttatgttat gatagtaat agaatgtaca aggtagttaa tacaacatct 2580
 aatgaacaag atattacaat tacactagat agaggacttg aagaagattt aaatcttgg 2640
 gacctatat atcaattaat aaataaaact aaagttcaag taaaaataaa taaatctgac 2700
 ttattcaaag acatgaaact agttgaaatc aaaaaatcag actcaagtta tcaagaaatc 2760
 tatgaattag aagaagccaa cataaaaagt gctcagccta aaatcatagt agaaaaagga 2820

ES 2 622 280 T3

gataaatgga cagctataaa acgtccatct atgattttta gatatgatgc tgaaaacaac 2880
gagccacaag cttaaaatgg aggtgtaaaa attgttttaa ttcgataaaa ataaaataga 2940
acaaatcaaa caaggtagaa aagtagaaat gcagtataaa gacatttcag acataagtat 3000
aggtcaagca aagcaagatg atgatataac aaataatfff atagcaaatg cagaaatata 3060
tgagatgttg ttaagtcaaa gttctgtcaa tgaagcaagt aatataagca cttttagtgt 3120
aagaaaatct ggaggtgaga gtggaatggg agaagtatat gtagctttaa ttttaagagg 3180
cagaaaaaca atagaagaag taccagcagt aattagagag caagttagaa ttagatgtaa 3240
agaattagaa ataccagttg aatagtaaat ttagaataac tatgtattag ttatfttttt 3300
tatgtaaagt acaaggtcct aactttaata agtaagcctt gtacttattt tttgttatat 3360
tagaaattgt atatatattt attatttatt caatctataa attaaaccta caatttaag 3420
tacagaagat taaattgata atcctgaaaa tataatattg catgatgtaa gaatataaca 3480
aaaattaaag ctataagtat aaaaaattta gacaatagga ggctataatg gataaattaa 3540
taaccgaatt gagtagtctg ggggcaatag gtatactatg tgctctatta tttaaaaata 3600
ctatgcagga gaaaaaagaa gatagagaca tgtataaaaa aactgtagaa aattttatag 3660
aattatctac acaacaacaa gaaataaaca aaaatatact tgttcaaatg ggaataatga 3720
aaacagatgt agaggaaatt aaggaagatg ttactgatat aaaaggtatg ttacaaaacg 3780
gtgtataaca tgaaagtagc agtagacca gattatatat tattaggaaa agataaagta 3840
gtattgtaga tagtgcccta ttttattgag aaggatttta tattttaaaa tattaattaa 3900
aaaaagtaat aaaaataata tataaaaaa acatataaaa attcaaaaag gagttaagct 3960
taaatttgat tagaaaaaat caattttaag acaactcctt ttttttatta aattattgtc 4020
tattaaccaa aatagctatt ttagcatctg gattataact tatctgaacc atttgatttt 4080
tcttaacatg ttcaaggtct tcaccacat aagctatttg taacttaact ggtaacttac 4140
cttgttttat aatagcaacg tactcttttt tacctttttc tctaaactaa tcaaattgcc 4200
aacataaggt ttaaagttct gatacttttt actagaattt ctatgtaga agaaagcacc 4260
aacagcaata actaaattta tgccaagtgt aaccaagaa ttgattttaa gcatagctcc 4320
agcgattatt atcacgaaca ttaaaacgat aggtaatata gctttcttaa gaagcaattt 4380
accattatt tcattagcct ttttctcagg gccactcata gtttttgatc tagcaaatga 4440
ttgcgcgaat ttgtctctta agcccatttt atcctcctaa ttttaataaa tatttagtta 4500
taataacgag atattacttg aaactaaaaa ttactacat ttatattatg tttgactttt 4560
gtataaataa ttacattcaa gtaaagcaaa atatactaat tattttatca taaaattata 4620
aaaaagaaa taaatgaaat aaaaatatta gaacaaagaa atgatgtaaa atcgtatcaa 4680
aagcaacata aaaattattt atctatftttc tcatctttat ttttgttata ctcaattttt 4740

ES 2 622 280 T3

cctaaatcct tctctttttc atattcatga agttttaatt caatcatacc ttctattttg 4800
 gctttatcat aatcatttaa ctttctaag ttgtttaaaa gctttatttc attagagttt 4860
 atgctattaa gtggatagtt tgaggaggaa tcgcaaatta aatctgattt atgtgataga 4920
 ttatctccat ttaatagcca gtctactgaa acattaaata tctcagctat agattttaat 4980
 atttcataat ttggttttct aatgtttctc tcaaatttac ttaagttgtc acagcctaac 5040
 atttcttcta gttcatattg ttttaaggtt ttgtcttttc tcaaataaac aattctttct 5100
 cctaaagtat ccataaacac tctccattca attaattgtca aaaagacttt ttaagatgta 5160
 aatagtttca aattaaaggt caaaatgaca taaaaacat tgacttaag tcaaatgac 5220
 tttataatta acttaatgat acgaatttac atcctaattt tagcacaaaag taatcaaaaa 5280
 atcttattta gtattaaata aatttatata cttaatatgt gtacatatta aaaatatata 5340
 ctaaatagag ggggtgcgta agctaaagta atataaaagt aatataaat cacttagaaa 5400
 ggaagttgat aaatggatgc tcgaaaaaaa tggatacctt ttttgggagt gcaagtcaag 5460
 caaagactta ttgaattaa tatgactcaa aggaattag cgaagaaaat aggtgttaat 5520
 gaaaactatt tgcagctat tttaaatgga agaagaacag gtaaaaaata taaatcatca 5580
 atttatcaat tacttaatat agaatttca gaagatgatt aataaatagt atataaagta 5640
 ggtgaatatt cttgtgtgca aattggattc agatggggtt atagagtgtt gtagagcaat 5700
 tgatgatttt attacagcac ttagtaatat aaaaagctta aatatggaaa gattaaatac 5760
 tttactaaa tattctagta catgttcaat ccttcttaa gaggggaatt atgaaggatg 5820
 tacaattgtg tatagaaaga tgttgaaga attaaaaaca tgagtaatgc atttcttagg 5880
 aatataaatt atacatagaa atgtattata ttttcaaag tacttaaaact aaaatatgga 5940
 taagataatc taaatattat aaatgtgctt gaaattagac tatacttgtt tttaaataat 6000
 ccaatatcca tatttttagta atatactaca aaaaaagaag gttaatagat gatgtaaaat 6060
 cgtatcaaat tatgtatggt taaaccattt tatcttcatt attattagag gaatgctttt 6120
 ttaagtcttt atattcagat atcttaagtt caagtattcc ttctattttt attttatcac 6180
 gttcgttttag ttgtctgtat agatttaata tcatcatttc atcattagta acatgtaagt 6240
 aatcttcttt atcttctttt aactactat tgacatttac cttctcttta ccatagagaa 6300
 gccagtcagt cgtaacatta aaataatcag ctattgacat tagtatatca caattaggtt 6360
 ttctatctcc tgtttcatac ttgcctaagt tttcaaattt taaaatatcc ataagtttgc 6420
 gctgagtaag ttttttggag tttctcaaat aagcaattct ttttctaaa gtatccacaa 6480
 aatacactcc tttcttttta tgagtaatgt ctaaagaca ttgaaatta aaaatatata 6540
 aatttataat ataaaactac taaattaaag tctaaatgac attttgctta aattaatatg 6600
 ctcataatat gattttaaca tattatagtt gaaaatatat ggtttatfff gatttgata 6660
 tataacaata gatttaattg ttataaaaat gtaaaggggt gtatgaatag attgtataaa 6720
 tttatttoga taaactaaga ttgctttttg attgtctgta aaagagaaaa agattaagat 6780
 aaaaatagta ttatattgta atttatatta atcaattaca aagattttat gaatttattc 6840
 tttagggtaa aatatttaag aataagataa atttacaata taatactata aactctttt 6900
 atctagtttt attttcttta tagaacaata atattataaa tgctagtaga tttacacaga 6960
 atactgttat atacatctgt ttgaatcctg agtttagagt agattgtagt gtggatccgg 7020

ES 2 622 280 T3

5	<210> 35 <211> 35 <212> ADN <213> Secuencia Artificial	
	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
10	<400> 35 tttttgCGG cGcaatacc cactacacct tcgtc	35
15	<210> 36 <211> 26 <212> ADN <213> Secuencia Artificial	
20	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
	<400> 36 tatacatctc gagtcccatc tctttc	26
25	<210> 37 <211> 25 <212> ADN <213> Secuencia Artificial	
30	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
35	<400> 37 gaagaaagag atgggactcg agatg	25
40	<210> 38 <211> 22 <212> ADN <213> Secuencia Artificial	
45	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
	<400> 38 ctgtgccaat ctatattgt tg	22
50	<210> 39 <211> 22 <212> ADN <213> Secuencia Artificial	
55	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
60	<400> 39 ggaaaaggga ttgctaataag tg	22
	<210> 40 <211> 38 <212> ADN <213> Secuencia Artificial	

	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
5	<400> 40 tccccggat ccacactaca atctactcta aactcagg	38
	<210> 41 <211> 34 <212> ADN <213> Secuencia Artificial	
10	<220> <221> fuente <223>/nota = "Descripción de la secuencia artificial: oligonucleótido sintético"	
15	<400> 41 ggcgcgccac tagtaccggt gccatggcgg ccgc	34
20	<210> 42 <211> 42 <212> ADN <213> Secuencia Artificial	
25	<220> <221> fuente <223>/nota = "Descripción de la secuencia artificial: oligonucleótido sintético"	
30	<400> 42 agctgcggcc gccatggcac cgttactagt ggcgcgccca tg	42
	<210> 43 <211> 37 <212> ADN <213> Secuencia Artificial	
35	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
40	<400> 43 ttccttggtc tcacgcgaac aaaattctcc agtctc	37
45	<210> 44 <211> 37 <212> ADN <213> Secuencia Artificial	
50	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
55	<400> 44 ttccttggtc tcaggccgtc gcgactaaga aaatgcc	37
60	<210> 45 <211> 20 <212> ADN <213> Secuencia Artificial	
	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	

ES 2 622 280 T3

	<400> 45 gtgagcggat aacaattccc	20
5	<210> 46 <211> 20 <212> ADN <213> Secuencia Artificial	
10	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
15	<400> 46 agattgtagt gtggatccgg	20
20	<210> 47 <211> 34 <212> ADN <213> Secuencia Artificial	
25	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
30	<400> 47 tccttcggcg cgctcaaat ttaagcttaa ctcc	34
35	<210> 48 <211> 20 <212> ADN <213> Secuencia Artificial	
40	<220> <221> fuente <223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"	
45	<400> 48 ttagggact actcactcgc	20
50	<210> 49 <211> 1773 <212> PRT <213> Clostridium difficile	
55	<400> 49	

ES 2 622 280 T3

Met Lys Arg Thr Lys Leu Leu Gln Arg Gly Asn Phe Phe Gly Asp Lys
 1 5 10 15

Asn Met Val Val Asp Glu Phe Asp Glu Gly Tyr Asp Asn Tyr Asp Phe
 20 25 30

Ile Asn Phe Phe Thr Gly Cys Cys Asn Tyr Thr Phe Gly Leu Lys Asn
 35 40 45

Asn Asn Ile Leu Tyr Gly Cys Gly Asp Asn Ser Asn Phe Gln Leu Gly
 50 55 60

Leu Gly Glu Asp Asn Thr Thr Arg Lys Leu Phe Thr Lys Ile Pro Asn
 65 70 75 80

Ile Ser Thr Asn Ile Lys Lys Val Ala Cys Gly Glu Ser His Ala Val
 85 90 95

Ile Leu Thr Ser Asp Gly Glu Leu Leu Val Ala Gly Ile Asn Thr Asp
 100 105 110

Gly Gln Met Gly Leu Gly Leu Glu Lys Val Gly Lys Thr Val Ser Thr
 115 120 125

Phe Glu Lys Val Pro Glu Ile Lys Gly Val Lys Asp Ile Ala Cys Gly
 130 135 140

Leu Gln Ser Thr Tyr Leu Leu Tyr Asn Asp Gly Thr Leu Tyr Val Ala
 145 150 155 160

Gly Asn Asn Leu Tyr Gly Gln Leu Gly Leu Gly Thr Asn Gly Ala Ser
 165 170 175

Ala Asn Val Asn Thr Phe Thr Lys Val Asp Val Asp Asn Val Lys Ala
 180 185 190

ES 2 622 280 T3

Val Phe Ser Tyr Asn Lys Ser Ala Phe Ile Ile Lys Asn Asp Asn Lys
195 200 205

Cys Tyr Ser Thr Gly Phe Asn Asn Gln Gly Gln Leu Gly Leu Gly Asp
210 215 220

Lys Asn Asn Arg Asp Leu Phe Ser Leu Val Ser Ile Asn Asp Val Lys
225 230 235 240

Thr Ile Ala Cys Gly Ser Glu His Thr Val Leu Met Thr Tyr Asn Asn
245 250 255

Asp Ile Tyr Gly Cys Gly Lys Glu Lys Cys Phe Gly Asn Ala Leu Gln
260 265 270

Ser Ser Leu Phe Thr Lys Ile Glu Glu Val Asn Ile Lys Thr Ile Ala
275 280 285

Cys Gly His Gly Asn Thr Met Leu Ile Asp Asn Lys Gly Thr Leu Lys
290 295 300

Val Ala Gly Asn Asn Asp Ile Tyr Gln Leu Gly Ile Ala Asn Tyr Ser
305 310 315 320

Glu Asn Ile Asp Asn Ser Phe Ile Asp Leu Lys Asn Ile Val Ala Lys
325 330 335

Asn Ile Phe Ile Gly Leu Ser His Ser Ile Leu Ile Asp Ser Asn Asn
340 345 350

Asp Ser Tyr Cys Thr Gly Asp Asn Thr Tyr Gly Gln Leu Gly Ser Phe
355 360 365

Phe Asp Asp Met His Ile Val Glu Phe Lys Lys Met Asp Ser Glu Lys
370 375 380

Tyr Ser Tyr Ser Asn Tyr Ile Asn Leu Ile Lys Ser Glu Asp Lys Leu
385 390 395 400

Thr Leu Leu Lys Glu Glu Met Glu Ile Lys Asp Ile Glu Leu Pro Leu
405 410 415

Asp Ile His Ser Val Arg Asp Val Val Phe Ser Pro Tyr Cys Thr Leu
420 425 430

Val Ile Leu Gly Asn Gly Asp Val Tyr Gly Leu Gly Asn Asn Arg Tyr
435 440 445

ES 2 622 280 T3

Lys Gly Met Gly Ser Asp Leu Pro Ser Gln Leu Asn Glu Leu Thr Lys
 450 455 460

Leu Ser Ile Ser Asn Val Lys Ser Ile Val Ala Ser Lys Asn Ile Ser
 465 470 475 480

Gly Gly Ile Phe Tyr Ile Lys Asn Asp Asp Thr Cys Tyr Tyr Ser Gly
 485 490 495

Pro Asn Ser Asn Ser Ile Ala Gly Val Leu Pro Ser Asn Ser Asp Val
 500 505 510

Phe Lys Lys Ile Ser Ile Asp Asn Val Lys Lys Val Val Ile Asn Thr
 515 520 525

Asp Leu Ser Asn Trp Phe Ser Leu Ile Val Thr Asn Asn Lys Gln Ile
 530 535 540

Tyr Thr Ser Gly Lys Ser Ser Ser Tyr Val Asn Gly Leu Ser Asn Ala
 545 550 555 560

Leu Ile Ser Gln Tyr Thr Glu Ile Ser Leu Ser Asn Val Thr Asp Ala
 565 570 575

Tyr Ser Ser Tyr Asn Ala Thr Phe Ile Val Val Asp Glu Lys Lys Val
 580 585 590

Tyr Ala Thr Gly Ile Asn Thr Asn Tyr Leu Leu Gly Phe Ser Thr Ser
 595 600 605

Asp Gly Ser Asn Val Asn Leu Gly Leu Leu Ser Asp Trp Tyr Tyr Ile
 610 615 620

Asn Ile Ser Gly Ser Ser Tyr Ser Arg Val Ser Cys Thr Asn Asn Ile
 625 630 635 640

Thr Lys Ile Asn Asn Ile Ile Ile Tyr Glu Tyr Val Thr Val Phe Cys
 645 650 655

Thr Asn Ile Gly Ser Phe Leu Thr Gly Tyr His Gly Thr Ser Trp Thr
 660 665 670

Lys Pro Thr Asp Ser Ser Tyr Arg Val Gln Tyr Gln Gly Ile Ser Tyr
 675 680 685

Ala Gly Tyr Leu Asp Ser Tyr Ile Tyr Asn Tyr Tyr Pro Thr Arg Cys
 690 695 700

ES 2 622 280 T3

Thr Gln Ser Ser Ser Ser Thr Thr Phe Ala Tyr Leu Tyr Asn Gly Glu
 705 710 715 720
 Ser Ser Ser Asn Leu Lys Asn Val Asn Pro Asp Asn Leu Leu Ile Ser
 725 730 735
 Gly Gly Ser Ser Tyr Ile His Gln Tyr Gly Arg Asn Tyr Leu Asn Asn
 740 745 750
 Gln Ser Ser Asn Asn Ile Ala Ala Ser Asn Ile Asn Ser Gly Pro Ile
 755 760 765
 Thr Ser Asp Lys Ala Ile Phe Leu Tyr Lys Ala Leu Leu Tyr Leu Ser
 770 775 780
 Ser Asn Thr Leu Tyr Gly Phe Gly Asn Ile Ser Glu Ser Ala Lys Glu
 785 790 795 800
 Leu Asp Val Ser Asp Thr Gln Asp Gly Tyr Asn Ala Thr Asn Tyr Lys
 805 810 815
 Lys Val Met Lys Asn Ile Lys Asn Ile Phe Ile Pro Pro Tyr Asp Leu
 820 825 830
 Ser Arg Asp Lys Thr Arg Phe Ala Ile Leu Thr Asp Lys Ser Leu Phe
 835 840 845
 Ile Cys Gly Tyr Asn Ser Lys Gly Thr His Gly Ile Ser Val Asn Ser
 850 855 860
 Ser Leu Asn Leu Asn Asn Lys Ile Asn Tyr Asn Lys Lys Asn Ser Ser
 865 870 875 880
 Ser Glu Ile Ser Ser Asn Ile Gln Glu Ile Tyr Ser His Ser Lys Ser
 885 890 895
 Thr Tyr Leu Leu Thr Asn Asn Asn Met Leu Tyr Ser Val Gly Leu Asn
 900 905 910
 Asp Val Gly Gln Leu Gly Val Gly Asp Glu Ile Asn Arg Lys Val Phe
 915 920 925
 Thr Lys Ile Asn Ile Asp Asn Ile Lys Ser Ile Asn Val Asn Arg Phe
 930 935 940
 Thr Asp Asn Ser Lys His Ala Phe Ala Ile Lys Asn Asp Asn Thr Cys
 945 950 955 960
 Tyr Ala Val Gly Leu Asn Asn Ser Gly Gln Leu Gly Ile Gly Asp Asn

ES 2 622 280 T3

		965						970						975	
Val	Asn	Arg	Asn	Ile	Phe	Thr	Lys	Ile	Asn	Val	Glu	Asn	Val	Lys	Tyr
			980					985						990	
Val	Ala	Val	Tyr	Gly	Asn	Thr	Ser	Leu	Leu	Leu	Thr	Asn	Asp	Gly	Leu
		995					1000					1005			
Leu	Tyr	Gly	Ala	Gly	Asn	Asn	Gly	Lys	Gly	Gln	Leu	Gly	Leu	Gly	
	1010					1015					1020				
Asp	Thr	Thr	Ser	Arg	Asn	Ile	Phe	Thr	Arg	Ile	Pro	Ile	Asn	Gly	
	1025					1030					1035				
Val	Arg	Asp	Val	Tyr	Leu	Cys	Asn	Asp	Val	Ser	Ile	Ile	Val	Lys	
	1040					1045					1050				
Asn	Asp	Asn	Thr	Cys	Tyr	Val	Cys	Gly	Leu	Val	Asn	Gly	Tyr	Phe	
	1055					1060					1065				
Gly	Phe	Thr	Glu	Gly	Ser	Ile	Ser	Thr	Phe	Thr	Lys	Ile	Asn	Ile	
	1070					1075					1080				
Glu	Asn	Val	Lys	Ser	Val	Val	Thr	Ala	Gly	Ser	Glu	Ala	Thr	Phe	
	1085					1090					1095				
Phe	Ile	Thr	Asn	Asp	Asn	Met	Ile	Tyr	Thr	Thr	Gly	Lys	Lys	Glu	
	1100					1105					1110				
Arg	Val	Phe	Phe	Ser	Thr	Glu	Thr	Asn	Asp	Ile	Lys	Gly	Ile	Arg	
	1115					1120					1125				
Val	Ile	Asn	Asn	Ile	Ile	Asn	Ala	Lys	Lys	Ile	Val	Val	Asn	Gly	
	1130					1135					1140				
Tyr	Thr	Ser	Ala	Ile	Leu	Thr	Asn	Asp	Asn	Lys	Leu	Phe	Val	Gly	
	1145					1150					1155				
Gly	Leu	Ser	Gly	Tyr	Gly	Ser	Ile	Ala	Asn	Asn	Asn	Asn	Thr	Asn	
	1160					1165					1170				
Ser	Val	Glu	Asp	Val	Lys	Asp	Val	Phe	Val	Thr	Ala	Asn	Asn	Thr	
	1175					1180					1185				
Leu	Tyr	Ile	Asp	Asn	Asn	Asn	Asn	Leu	Ile	Ser	Ser	Gly	Arg	Asp	
	1190					1195					1200				
Thr	Tyr	Gly	Ile	Ser	Asp	Glu	Ser	Tyr	Arg	Asp	Met	Ser	Val	Pro	
	1205					1210					1215				

ES 2 622 280 T3

Tyr Tyr Lys Val Ser Ile Lys Lys Asp Val Asp Thr Val Phe Ser
 1220 1225 1230

 Ser Tyr Asn Thr Ile Phe Ile Lys Asp Ile Tyr Gly Lys Phe Tyr
 1235 1240 1245

 Ser Ser Thr Arg Asp Asn Arg Tyr Asn His Leu Gly Ile His His
 1250 1255 1260

 Arg Tyr Asp Asn Asp Lys Asn Glu Ala Leu Glu Gly Ser Leu His
 1265 1270 1275

 Ser Tyr Phe Lys Thr Asp Asn Thr Ser Asp Lys Ile Val Phe Asn
 1280 1285 1290

 Lys Lys Asn Glu Lys Leu Val Met Phe Asn Asp Lys Tyr Ile Lys
 1295 1300 1305

 Thr Asn Asn Lys Tyr Ile Asn Tyr Lys Asn Ile Phe Lys Asp Asn
 1310 1315 1320

 Phe Lys Tyr Thr Ser Ile Ile Leu Pro Phe Glu Val Ser Asp Ile
 1325 1330 1335

 Asp Ile Ser Lys Thr His Ser Leu Ala Val Ala Lys Asp Gly Lys
 1340 1345 1350

 Leu Tyr Gly Ile Gly Ser Asn Ser Tyr Lys Glu Ile Asn Gln Thr
 1355 1360 1365

 Leu Glu Asp Ile Glu Leu Leu Thr Leu Thr Glu Val Asn Ile Ser
 1370 1375 1380

 Asp Val Lys Lys Val Ala Cys Gly Asp Asn Tyr Ser Tyr Ile Ile
 1385 1390 1395

 Lys Thr Asp Asn Thr Leu Trp Ser Tyr Gly Lys Asn Thr Glu Tyr
 1400 1405 1410

 Gln Leu Gly Val Gly His Asn Asn Asp Val Arg Glu Leu Gln Lys
 1415 1420 1425

 Val Thr Gly Leu Pro Ser Val Lys Asp Ile Ser Ile Tyr Asn Ser
 1430 1435 1440

 Met Thr Leu Val Leu Thr Asn Glu Gly Glu Leu Tyr Ala Gln Gly
 1445 1450 1455

ES 2 622 280 T3

Tyr Asn Thr Asn Gly Leu Phe Gly Leu Gly Glu Ser Glu Lys Asp
 1460 1465 1470
 Lys Ile Ile Arg Thr Phe Thr Lys Val Leu Thr Asn Val Lys Glu
 1475 1480 1485
 Ile Lys Ser His Asn Asp Asp His Ile Leu Val Ile Lys Asn Asp
 1490 1495 1500
 Asn Ser Leu Trp Ile Thr Gly Lys Asn Lys Ser Met Tyr Lys Ile
 1505 1510 1515
 Ser Ile Ser Ile Thr Asp Leu Tyr Glu Phe Thr Lys Ile Pro Ile
 1520 1525 1530
 Pro Glu His Leu Asn Asp Ile Leu Asp Ile Glu Leu Ser Asp Asp
 1535 1540 1545
 Thr Ile Tyr Met Ile Thr Lys Val Asp Thr Ser Lys Ala Ser Ile
 1550 1555 1560
 Glu Ile Val Glu Lys Ser Ile Ser Gln Val Arg Val Val Val Gln
 1565 1570 1575
 Asp Pro Asn Asn Val Ile Glu Lys Leu Glu Met Phe Ile Asn Asp
 1580 1585 1590
 Glu Leu Ile Ser Thr Lys Thr Asn Leu Glu Ile Asn Ser Ile Ile
 1595 1600 1605
 Phe Glu Ile Pro Gln Asn Lys Ile Val Leu Gly Glu Asn Lys Ile
 1610 1615 1620
 Leu Ile Lys Ala Ser Ser Pro Thr Gly Asp Leu Tyr Ser Ser Met
 1625 1630 1635
 Phe Ile Phe Lys Ser Glu Thr Gly Leu Lys Val Lys Lys Asp Ser
 1640 1645 1650
 Ile Leu Met Ile Asn Asn Lys Val Tyr Ser Ile Ile Asn Ile Thr
 1655 1660 1665
 Glu Asn Asn Thr Asp Leu Ile Val Thr Leu Asn Glu Gly Leu Lys
 1670 1675 1680
 Asp Asp Met Met Glu Asn Asn Pro Ile Tyr Gln Leu Ile Asn Lys
 1685 1690 1695

ES 2 622 280 T3

Thr Lys Val Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp
 1700 1705 1710

Met Lys Leu Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu
 1715 1720 1725

Ile Tyr Glu Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys
 1730 1735 1740

Ile Ile Val Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro
 1745 1750 1755

Ser Met Ile Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln Ala
 1760 1765 1770

<210> 50
 <211> 1743
 <212> PRT
 <213> Clostridium difficile

5

<400> 50
 Met Lys Gln Asn Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys
 1 5 10 15

Asn Ile Leu Ile Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe
 20 25 30

Val Glu Phe Phe Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser
 35 40 45

Asp Gly Asn Leu Tyr Ala Cys Gly Asp Asn Thr Gly Phe Gln Leu Gly
 50 55 60

Leu Gly Lys Asp Ser Ser Glu Arg Arg Met Phe Ser Lys Val Lys Ile
 65 70 75 80

Asp Asn Val Lys Tyr Val Ser Cys Gly Ser Lys His Ser Val Ala Val
 85 90 95

Thr Lys Asp Gly Phe Ala Tyr Gly Ala Gly Thr Ser Asn Val Gly Gln
 100 105 110

Leu Gly Val Ile Glu Ser Thr Val Tyr Tyr Glu Phe Thr Lys Leu Pro
 115 120 125

Ile Asp Asp Val Lys Thr Val Ala Cys Gly Tyr Asp Phe Thr Phe Val
 130 135 140

Leu Lys Asn Asp Gly Thr Leu Tyr Ser Ala Gly Leu Asn Ser Ser Gly
 145 150 155 160

10

ES 2 622 280 T3

Gln Leu Gly Leu Gly Asp Thr Asn Asn Arg Val Thr Phe Thr Lys Val
165 170 175

Asn Ile Asp Ser Val Lys Asp Val Val Thr Tyr Asn Gln Ser Val Phe
180 185 190

Ile Ile Lys Met Asp Gly Thr Ala His Ala Cys Gly Leu Asn Ser Asn
195 200 205

Gly Gln Leu Gly Ile Asn Ser Thr Leu Asn Lys Ser Val Phe Asn Lys
210 215 220

Ile Glu Gly Met Asp Asn Val Lys Gln Ile Ala Cys Gly Ser Ser His
225 230 235 240

Thr Ile Leu Ile Lys Asn Asp Gly Thr Met Tyr Thr Thr Gly Ser Asn
245 250 255

Gly Tyr Gly Gln Leu Gly Thr Gly Asn Asn Asn Asn Ser Ile Val Phe
260 265 270

Thr Leu Ser Ser Ile Asn Asn Val Lys Tyr Ala Ser Cys Gly Asn Asn
275 280 285

His Thr Met Ile Leu Lys Tyr Asp Asn Thr Leu Phe Ser Thr Gly Gln
290 295 300

Asn Asn Tyr Gly Gln Leu Ala Asn Ala Asn Lys Asp Val Ala Ser Arg
305 310 315 320

Asn Thr Phe Val Lys Val Asn Val Glu Asn Ile Lys Asp Ile Lys Cys
325 330 335

Gly Ser Gln Phe Asn Phe Leu Ile Asn Gly Ser Lys Glu Ile Phe Val
340 345 350

Ser Gly Cys Asn Leu Ala Gly Gln Leu Gly Ser Phe Phe His Thr Thr
355 360 365

Phe Leu Tyr Glu Phe Ser Lys Val Gln Ser Ser Asn Leu Asp Asn Tyr
370 375 380

Ser Gly Leu Leu Val Asn Asp Asp Tyr Leu Tyr Val Thr Lys Asp Asn
385 390 395 400

Ser Glu Phe Leu Asn Val Lys Leu Ser Asp Asn Phe Gln Asp Tyr Lys
405 410 415

ES 2 622 280 T3

Lys Ile Glu Leu Thr Asp Asn Asn Met Phe Ile Val Met Asn Asp Gly
 420 425 430

Thr Leu Tyr Ala Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Leu Gly
 435 440 445

Asp Thr Val Asn Arg Ser Val Met Thr Lys Val Asp Ile Asp Asn Val
 450 455 460

Leu Asp Ile Lys Gly Asn Gly Asn Ser Thr Phe Val Leu Lys Asn Asn
 465 470 475 480

Gly Thr Leu Tyr Ser Cys Gly Tyr Asn Ser Ser Gly Ile Leu Gly Leu
 485 490 495

Lys Asp Asn Thr Asn Arg Asn Ile Phe Thr Lys Ile Glu Ile Glu Asn
 500 505 510

Ile Lys Glu Phe Cys Val Glu Ser Asn Tyr Ile Val Ala Leu Asn His
 515 520 525

Ser Lys Glu Leu Tyr Gly Trp Gly Asn Gln Ser Tyr Ile Val Tyr Gly
 530 535 540

Asp Asn Arg Asn Tyr Pro Tyr Lys Asp Thr Arg Val Ser Asn Val Glu
 545 550 555 560

Lys Ile Ala Thr Trp Ser Asp Thr Leu Tyr Ile Leu Asp Ser Thr Gly
 565 570 575

Ala Thr Lys Thr Ile Gly Tyr Ser Tyr Asn Gly Ser Gly Gly Tyr Pro
 580 585 590

Ala Pro Ser Ser Ser Ser Thr Tyr Arg Glu Gly Gly Tyr Ile Asn Lys
 595 600 605

Asn Thr Ser Tyr Arg Thr Leu Glu Phe Tyr Asn Thr Ser Lys Thr Lys
 610 615 620

Leu Val Asn Leu Phe Ala Phe Tyr Asn Gly Cys Val Phe Val Asp Glu
 625 630 635 640

Asn Gly Leu Ala Tyr Cys Ile Gly Glu Asn Asn Ile Asn Phe Arg Gly
 645 650 655

Gly Ser Thr Thr Asn Glu Asn Asn Ser Leu Arg Phe Ile Asn Asn Ser
 660 665 670

Gly Val Tyr Tyr Thr Asn Thr Asp Gly Thr Asp Tyr Thr Cys Tyr Gln

ES 2 622 280 T3

		675					680								685
Trp	Thr	Tyr	Lys	Leu	Ile	Arg	Cys	Ser	Ile	Phe	Asp	Ser	Pro	Gln	Asn
	690					695					700				
Ile	Ile	Gly	Asn	Ser	Lys	Asn	Ile	Leu	Tyr	Leu	Ser	Lys	Asn	Asn	Ser
705					710					715					720
Thr	Phe	Lys	Cys	Thr	Gly	Asn	Cys	Ile	Thr	Tyr	Gly	Ile	Asn	Ser	Gln
				725					730					735	
Asn	Trp	Tyr	Ser	Tyr	Phe	Ser	Asp	Ser	Ser	Asn	Gly	Ala	Ile	Ala	Leu
			740					745					750		
Gly	Asn	Glu	Phe	Ile	Leu	Lys	Asn	Tyr	Ser	Gly	Glu	Cys	Leu	Leu	Lys
		755					760					765			
Gly	Tyr	Gly	Lys	Ala	Thr	Asn	Gly	Glu	Phe	Gly	Asn	Ser	Thr	Asn	Ile
	770					775					780				
Ser	Ser	Ile	Ser	Asn	Tyr	Asp	Thr	Gly	Leu	Lys	Asp	Ile	Lys	Asp	Ile
785				790						795					800
Ile	Val	Lys	Asn	Asn	Thr	Val	Val	Val	Val	Asp	Lys	Asn	Asn	Asn	Ile
			805						810					815	
Tyr	Val	Thr	Gly	Ala	Asn	Gln	Phe	Asn	Lys	Leu	Gly	Ile	Gly	Glu	Tyr
			820					825					830		
Asn	Asn	Gln	Pro	Ile	Arg	Lys	Phe	Thr	Asn	Ile	Thr	Glu	Gln	Ser	Asn
		835					840					845			
Ser	Phe	Ile	Phe	Met	Asp	Asp	Ile	Lys	Glu	Ile	Thr	Thr	Ser	Arg	Asn
	850					855					860				
Thr	Met	Phe	Ile	Val	Lys	Asn	Asp	Gly	Thr	Ala	Tyr	Ala	Thr	Gly	Asn
865					870					875					880
Asn	Ser	Ser	Gly	Gln	Leu	Gly	Leu	Gly	Asp	Thr	Ile	Asn	Arg	Asn	Lys
				885					890					895	
Phe	Thr	Gln	Ile	Asn	Leu	Asp	Asn	Ile	Lys	Lys	Ile	Ser	Thr	Ser	Ile
		900						905					910		
Asp	Gly	Asn	Thr	Thr	Phe	Ala	Ile	Arg	Asn	Asp	Gly	Thr	Leu	Tyr	Ser
		915					920					925			
Thr	Gly	Leu	Asn	Thr	Lys	Gly	Gln	Leu	Gly	Leu	Gly	Asp	Ile	Val	Asn
	930					935						940			

ES 2 622 280 T3

Arg Asn Thr Phe Thr Lys Val Asn Ile Gln Asn Val Arg Asp Val Val
 945 950 955 960

Leu Gly Thr Thr His Ser His Ala Ile Lys Asp Asp Asn Thr Leu Tyr
 965 970 975

Ser Cys Gly Glu Asn Thr His Gly Gln Leu Gly Leu Gly Ser Glu Ser
 980 985 990

Asn His Pro Asp Val Leu Thr Phe Thr Val Asn Asn Ile Thr Asn Val
 995 1000 1005

Arg Asp Val Tyr Cys Ser Asp Thr Thr Thr Phe Ile Val Lys Asp
 1010 1015 1020

Thr Asn Ile Ala Tyr Cys Cys Gly Tyr Asn Asn Asn Ser Gln Leu
 1025 1030 1035

Gly Met Gly Asn Thr Thr Asp Gln Tyr Ser Phe Ile Lys Cys Met
 1040 1045 1050

Glu Asn Val Lys Glu Val Ile Pro Asn Glu Ile Asn Thr Tyr Ile
 1055 1060 1065

Ile Thr Ile Tyr Asn Thr Ala Tyr Ser Thr Gly Leu Asn Thr Asp
 1070 1075 1080

Tyr Cys Leu Gly Leu Asn Ser Asn Ser Asn Gln Ser Ser Phe Ser
 1085 1090 1095

Glu Ile Pro Ile Ser Asn Val Val Lys Val Ala Pro Asn Arg Asn
 1100 1105 1110

Asn Ala Val Leu Leu Leu Thr Ser Glu Gly Asp Val Tyr Thr Ala
 1115 1120 1125

Gly Lys Cys Ser Asn Gly Ser Gly Thr Gly Ser Glu Thr Pro Glu
 1130 1135 1140

Lys Ile Lys Lys Ile Ala Ser Lys Ala Lys Asp Ile Gly Met Asn
 1145 1150 1155

Tyr Arg Cys Gly His Tyr Val Ser Asp Asn Gly Asp Leu Tyr Gly
 1160 1165 1170

Thr Gly Phe Asn Asp Cys Gly Gln Leu Gly Val Gly Asn Val Thr
 1175 1180 1185

ES 2 622 280 T3

Lys Arg Asp Thr Phe Ile Lys Thr Asn Thr Arg Val Lys Lys Ile
 1190 1195 1200

Leu Pro Leu Glu Tyr Ala Asn Ile Ala Ile Lys Asp Thr Asn Asp
 1205 1210 1215

Ile Tyr Ile Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Val Gly
 1220 1225 1230

Asn Arg Tyr Asp Ser Arg Asn Asn Asp Asn Arg Ile Phe Asn Tyr
 1235 1240 1245

Lys His Met Asn Phe Val Met Gly Asp Leu Thr Ser Ile Lys Asn
 1250 1255 1260

Arg His Asn Phe Ile Leu Leu Asn Asn Lys Ile Val Ile Pro Thr
 1265 1270 1275

Thr Lys Asp Ile Asp Tyr Gly Leu Val Leu Gly Asn Leu Tyr Lys
 1280 1285 1290

Gly Asp Leu Tyr Thr Glu Leu Pro Tyr Glu Asp Ile Lys Glu Val
 1295 1300 1305

Ser Ile Ser Lys Thr His Ile Ile Ile Leu Leu Asn Asp Gly Thr
 1310 1315 1320

Met Tyr Gly Cys Gly Thr Asn Tyr His Gly Glu Leu Leu Gln Asp
 1325 1330 1335

Leu Ser Ile Asn Gln Val Asp Glu Phe Val Gln Ile Asn Val Ser
 1340 1345 1350

Asp Val Lys His Val Ser Cys Gly Asp Asn Phe Thr Tyr Phe Ile
 1355 1360 1365

Lys Ser Asp Asp Ser Leu Trp Ser Ile Gly Lys Asn Ser Glu Tyr
 1370 1375 1380

Gln Leu Gly Ile Gly His Asn Asn Pro Val Thr Glu Leu Gln Arg
 1385 1390 1395

Ile Thr Thr Ile Ser Ser Cys Lys Glu Val His Cys Gly Lys Asn
 1400 1405 1410

Tyr Thr Leu Val Val Thr Thr Ser Asn Glu Leu Phe Val Gln Gly
 1415 1420 1425

ES 2 622 280 T3

Tyr Asn Asp Lys Gly Ala Leu Gly Leu Gly Ser Asp Ser Glu Asn
 1430 1435 1440

 Thr Ile Ile Lys Phe Phe Thr Lys Ala Leu Thr Asp Ile Arg Glu
 1445 1450 1455

 Ile Lys Ser Tyr Gly Ser Asp His Ile Leu Val Leu Lys Asn Asp
 1460 1465 1470

 Asn Ser Val Trp Val Thr Gly Lys Asn Arg Asp Val Tyr Lys Ile
 1475 1480 1485

 Glu Gln Pro Val Glu Phe Leu Lys Glu Phe Thr Ile Val Pro Ile
 1490 1495 1500

 Ser Glu Asp Val Asn Thr Val Lys Asp Val Leu Ala Thr Asp Asn
 1505 1510 1515

 Thr Leu Tyr Ile Ile Ser Glu Val Gly Thr Thr Asn Ala Ala Ile
 1520 1525 1530

 Glu Ile Thr Glu Lys Ser Ile Ser Ser Ile Lys Ile Lys Ile Gln
 1535 1540 1545

 Asp Pro Asn Lys Asp Ile Ser Arg Ile Glu Met Leu Ile Asn Gly
 1550 1555 1560

 Glu Ser Val Lys Ser Val Ser Asp Leu Ile Thr Glu Lys Ile Ser
 1565 1570 1575

 Phe Glu Val Pro Pro Asp Lys Ile Lys Ile Gly Glu Asn Lys Ile
 1580 1585 1590

 Leu Phe Arg Ala Tyr Cys Lys Gly Asp Asp Leu Tyr Ala Ser Leu
 1595 1600 1605

 Phe Ile Phe Lys Glu Ser Thr Gly Asn Ser Ile Ile Lys Asp Ser
 1610 1615 1620

 Tyr Val Met Ile Gly Asn Arg Met Tyr Lys Val Val Asn Thr Thr
 1625 1630 1635

 Ser Asn Glu Gln Asp Ile Thr Ile Thr Leu Asp Arg Gly Leu Glu
 1640 1645 1650

 Glu Asp Leu Asn Leu Gly Asp Pro Ile Tyr Gln Leu Ile Asn Lys
 1655 1660 1665

 Thr Lys Val Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp

ES 2 622 280 T3

1670 1675 1680

Met Lys Leu Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu
1685 1690 1695

Ile Tyr Glu Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys
1700 1705 1710

Ile Ile Val Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro
1715 1720 1725

Ser Met Ile Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln Ala
1730 1735 1740

<210> 51

<211> 1773

<212> PRT

<213> Clostridium difficile

5

<400> 51

Met Lys Arg Thr Lys Leu Leu Gln Arg Gly Asn Phe Phe Gly Asp Lys
1 5 10 15

Asn Met Val Val Asp Glu Phe Asp Glu Gly Tyr Asp Asn Tyr Asp Phe
20 25 30

Ile Asn Phe Phe Thr Gly Cys Cys Asn Tyr Thr Phe Gly Leu Lys Asn
35 40 45

Asn Asn Ile Leu Tyr Gly Cys Gly Asp Asn Ser Asn Phe Gln Leu Gly
50 55 60

Leu Gly Glu Asp Asn Thr Thr Arg Lys Leu Phe Thr Lys Ile Pro Asn
65 70 75 80

Ile Ser Thr Asn Ile Lys Lys Val Ala Cys Gly Glu Ser His Ala Val
85 90 95

Ile Leu Thr Ser Asp Gly Glu Leu Leu Val Ala Gly Ile Asn Thr Asp
100 105 110

Gly Gln Met Gly Leu Gly Leu Glu Lys Val Gly Lys Thr Val Ser Thr
115 120 125

Phe Glu Lys Val Pro Glu Ile Lys Gly Val Lys Asp Ile Ala Cys Gly
130 135 140

Leu Gln Ser Thr Tyr Leu Leu Tyr Asn Asp Gly Thr Leu Tyr Val Ala
145 150 155 160

10

ES 2 622 280 T3

Gly Asn Asn Leu Tyr Gly Gln Leu Gly Leu Gly Thr Asn Gly Ala Ser
 165 170 175

Ala Asn Val Asn Thr Phe Thr Lys Val Asp Val Asp Asn Val Lys Ala
 180 185 190

Val Phe Ser Tyr Asn Lys Ser Ala Phe Ile Ile Lys Asn Asp Asn Lys
 195 200 205

Cys Tyr Ser Thr Gly Phe Asn Asn Gln Gly Gln Leu Gly Leu Gly Asp
 210 215 220

Lys Asn Asn Arg Asp Leu Phe Ser Leu Val Ser Ile Asn Asp Val Lys
 225 230 235 240

Thr Ile Ala Cys Gly Ser Glu His Thr Val Leu Met Thr Tyr Asn Asn
 245 250 255

Asp Ile Tyr Gly Cys Gly Lys Glu Lys Cys Phe Gly Asn Ala Leu Gln
 260 265 270

Ser Ser Leu Phe Thr Lys Ile Glu Glu Val Asn Ile Lys Thr Ile Ala
 275 280 285

Cys Gly His Gly Asn Thr Met Leu Ile Asp Asn Lys Gly Thr Leu Lys
 290 295 300

Val Ala Gly Asn Asn Asp Ile Tyr Gln Leu Gly Ile Ala Asn Tyr Ser
 305 310 315 320

Glu Asn Ile Asp Asn Ser Phe Ile Asp Leu Lys Asn Ile Val Ala Lys
 325 330 335

Asn Ile Phe Ile Gly Leu Ser His Ser Ile Leu Ile Asp Ser Asn Asn
 340 345 350

Asp Ser Tyr Cys Thr Gly Asp Asn Thr Tyr Gly Gln Leu Gly Ser Phe
 355 360 365

Phe Asp Asp Met His Ile Val Glu Phe Lys Lys Met Asp Ser Glu Lys
 370 375 380

Tyr Ser Tyr Ser Asn Tyr Ile Asn Leu Ile Lys Ser Glu Asp Lys Leu
 385 390 395 400

Thr Leu Leu Lys Glu Glu Met Glu Ile Lys Asp Ile Glu Leu Pro Leu
 405 410 415

Asp Ile His Ser Val Arg Asp Val Val Phe Ser Pro Tyr Cys Thr Leu

ES 2 622 280 T3

420 425 430
 Val Ile Leu Gly Asn Gly Asp Val Tyr Gly Leu Gly Asn Asn Arg Tyr
 435 440 445
 Lys Gly Met Gly Ser Asp Leu Pro Ser Gln Leu Asn Glu Leu Thr Lys
 450 455 460
 Leu Ser Ile Ser Asn Val Lys Ser Ile Val Ala Ser Lys Asn Ile Ser
 465 470 475 480
 Gly Gly Ile Phe Tyr Ile Lys Asn Asp Asp Thr Cys Tyr Tyr Ser Gly
 485 490 495
 Pro Asn Ser Asn Ser Ile Ala Gly Val Leu Pro Ser Asn Ser Asp Val
 500 505 510
 Phe Lys Lys Ile Ser Ile Asp Asn Val Lys Lys Val Val Ile Asn Thr
 515 520 525
 Asp Leu Ser Asn Trp Phe Ser Leu Ile Val Thr Asn Asn Lys Gln Ile
 530 535 540
 Tyr Thr Ser Gly Lys Ser Ser Ser Tyr Val Asn Gly Leu Ser Asn Ala
 545 550 555 560
 Leu Ile Ser Gln Tyr Thr Glu Ile Ser Leu Ser Asn Val Thr Asp Ala
 565 570 575
 Tyr Ser Ser Tyr Asn Ala Thr Phe Ile Val Val Asp Glu Lys Lys Val
 580 585 590
 Tyr Ala Thr Gly Ile Asn Thr Asn Tyr Leu Leu Gly Phe Ser Thr Ser
 595 600 605
 Asp Gly Ser Asn Val Asn Leu Gly Leu Leu Ser Asp Trp Tyr Tyr Ile
 610 615 620
 Asn Ile Ser Gly Ser Ser Tyr Ser Arg Val Ser Cys Thr Asn Asn Ile
 625 630 635 640
 Thr Lys Ile Asn Asn Ile Ile Ile Tyr Glu Tyr Val Thr Val Phe Cys
 645 650 655
 Thr Asn Ile Gly Ser Phe Leu Thr Gly Tyr His Gly Thr Ser Trp Thr
 660 665 670
 Lys Pro Thr Asp Ser Ser Tyr Arg Val Gln Tyr Gln Gly Ile Ser Tyr
 675 680 685

ES 2 622 280 T3

Ala Gly Tyr Leu Asp Ser Tyr Ile Tyr Asn Tyr Tyr Pro Thr Arg Cys
690 695 700

Thr Gln Ser Ser Ser Ser Thr Thr Phe Ala Tyr Leu Tyr Asn Gly Glu
705 710 715 720

Ser Ser Ser Asn Leu Lys Asn Val Asn Pro Asp Asn Leu Leu Ile Ser
725 730 735

Gly Gly Ser Ser Tyr Ile His Gln Tyr Gly Arg Asn Tyr Leu Asn Asn
740 745 750

Gln Ser Ser Asn Asn Ile Ala Ala Ser Asn Ile Asn Ser Gly Pro Ile
755 760 765

Thr Ser Asp Lys Ala Ile Phe Leu Tyr Lys Ala Leu Leu Tyr Leu Ser
770 775 780

Ser Asn Thr Leu Tyr Gly Phe Gly Asn Ile Ser Glu Ser Ala Lys Glu
785 790 795 800

Leu Asp Val Ser Asp Thr Gln Asp Gly Tyr Asn Ala Thr Asn Tyr Lys
805 810 815

Lys Val Met Lys Asn Ile Lys Asn Ile Phe Ile Pro Pro Tyr Asp Leu
820 825 830

Ser Arg Asp Lys Thr Arg Phe Ala Ile Leu Thr Asp Lys Ser Leu Phe
835 840 845

Ile Cys Gly Tyr Asn Ser Lys Gly Thr His Gly Ile Ser Val Asn Ser
850 855 860

Ser Leu Asn Leu Asn Asn Lys Ile Asn Tyr His Lys Lys Asn Ser Ser
865 870 875 880

Ser Glu Ile Ser Ser Asn Ile Gln Glu Ile Tyr Ser His Ser Lys Ser
885 890 895

Thr Tyr Leu Leu Thr Asn Asn Asn Met Leu Tyr Ser Val Gly Leu Asn
900 905 910

Asp Val Gly Gln Leu Gly Val Gly Asp Glu Ile Asn Arg Lys Val Phe
915 920 925

Thr Lys Ile Asn Ile Asp Asn Ile Lys Ser Ile Asn Val Asn Arg Phe
930 935 940

ES 2 622 280 T3

Thr Asp Asn Ser Lys His Ala Phe Ala Ile Lys Asn Asp Asn Thr Cys
 945 950 955 960

Tyr Ala Val Gly Leu Asn Asn Ser Gly Gln Leu Gly Ile Gly Asp Asn
 965 970 975

Val Asn Arg Asn Ile Phe Thr Lys Ile Asn Val Glu Asn Val Lys Tyr
 980 985 990

Val Ala Val Tyr Gly Asn Thr Ser Leu Leu Leu Thr Asn Asp Gly Leu
 995 1000 1005

Leu Tyr Gly Ala Gly Asn Asn Gly Lys Gly Gln Leu Gly Leu Gly
 1010 1015 1020

Asp Thr Thr Ser Arg Asn Ile Phe Thr Arg Ile Pro Ile Asn Gly
 1025 1030 1035

Val Arg Asp Val Tyr Leu Cys Asn Asp Val Ser Ile Ile Val Lys
 1040 1045 1050

Asn Asp Asn Thr Cys Tyr Val Cys Gly Leu Val Asn Gly Tyr Phe
 1055 1060 1065

Gly Phe Thr Glu Gly Ser Ile Ser Thr Phe Thr Lys Ile Asn Ile
 1070 1075 1080

Glu Asn Val Lys Ser Val Val Thr Ala Gly Ser Glu Ala Thr Phe
 1085 1090 1095

Phe Ile Thr Asn Asp Asn Met Ile Tyr Thr Thr Gly Lys Lys Glu
 1100 1105 1110

Arg Val Phe Phe Ser Thr Glu Thr Asn Asp Ile Lys Gly Ile Arg
 1115 1120 1125

Val Ile Asn Asn Ile Ile Asn Ala Lys Lys Ile Val Val Asn Gly
 1130 1135 1140

Tyr Thr Ser Ala Ile Leu Thr Asn Asp Asn Lys Leu Phe Val Gly
 1145 1150 1155

Gly Leu Ser Gly Tyr Gly Ser Ile Ala Asn Asn Asn Asn Thr Asn
 1160 1165 1170

Ser Val Glu Asp Val Lys Asp Val Phe Val Thr Ala Asn Asn Thr
 1175 1180 1185

ES 2 622 280 T3

Leu Tyr Ile Asp Asn Asn Asn Asn Leu Ile Ser Ser Gly Arg Asp
 1190 1195 1200
 Thr Tyr Gly Ile Ser Asp Glu Ser Tyr Arg Asp Met Ser Val Pro
 1205 1210 1215
 Tyr Tyr Lys Val Ser Ile Lys Lys Asp Val Asp Thr Val Phe Ser
 1220 1225 1230
 Ser Tyr Asn Thr Ile Phe Ile Lys Asp Ile Tyr Gly Lys Phe Tyr
 1235 1240 1245
 Ser Ser Thr Arg Asp Asn Arg Tyr Asn His Leu Gly Ile His His
 1250 1255 1260
 Arg Tyr Asp Asn Asp Lys Asn Glu Ala Leu Glu Gly Ser Leu His
 1265 1270 1275
 Ser Tyr Phe Lys Thr Asp Asn Thr Ser Asp Lys Ile Val Phe Asn
 1280 1285 1290
 Lys Lys Asn Glu Lys Leu Val Met Phe Asn Asp Lys Tyr Ile Lys
 1295 1300 1305
 Thr Asn Asn Lys Tyr Ile Asn Tyr Lys Asn Ile Phe Lys Asp Asn
 1310 1315 1320
 Phe Lys Tyr Thr Ser Ile Ile Leu Pro Phe Glu Val Ser Asp Ile
 1325 1330 1335
 Asp Ile Ser Lys Thr His Ser Leu Ala Val Ala Lys Asp Gly Lys
 1340 1345 1350
 Leu Tyr Gly Ile Gly Ser Asn Ser Tyr Lys Glu Ile Asn Gln Thr
 1355 1360 1365
 Leu Glu Asp Ile Glu Leu Leu Thr Leu Thr Glu Val Asn Ile Ser
 1370 1375 1380
 Asp Val Lys Lys Val Ala Cys Gly Asp Asn Tyr Ser Tyr Ile Ile
 1385 1390 1395
 Lys Thr Asp Asn Thr Leu Trp Ser Tyr Gly Lys Asn Thr Glu Tyr
 1400 1405 1410
 Gln Leu Gly Val Gly His Asn Asn Asp Val Arg Glu Leu Gln Lys
 1415 1420 1425
 Val Thr Gly Leu Pro Ser Val Lys Asp Ile Ser Ile Tyr Asn Ser

ES 2 622 280 T3

1430						1435									1440
Met	Thr	Leu	Val	Leu	Thr	Asn	Glu	Gly	Glu	Leu	Tyr	Ala	Gln	Gly	
1445						1450					1455				
Tyr	Asn	Thr	Asn	Gly	Leu	Phe	Gly	Leu	Gly	Glu	Ser	Glu	Lys	Asp	
1460						1465					1470				
Lys	Ile	Ile	Arg	Thr	Phe	Thr	Lys	Val	Leu	Thr	Asn	Val	Lys	Glu	
1475						1480					1485				
Ile	Lys	Ser	His	Asn	Asp	Asp	His	Ile	Leu	Val	Ile	Lys	Asn	Asp	
1490						1495					1500				
Asn	Ser	Leu	Trp	Ile	Thr	Gly	Lys	Asn	Lys	Ser	Met	Tyr	Lys	Ile	
1505						1510					1515				
Ser	Ile	Ser	Ile	Thr	Asp	Leu	Tyr	Glu	Phe	Thr	Lys	Ile	Pro	Ile	
1520						1525					1530				
Pro	Glu	His	Leu	Asn	Asp	Ile	Leu	Asp	Ile	Glu	Leu	Ser	Asp	Asp	
1535						1540					1545				
Thr	Ile	Tyr	Met	Ile	Thr	Lys	Val	Asp	Thr	Ser	Lys	Ala	Ser	Ile	
1550						1555					1560				
Glu	Ile	Val	Glu	Lys	Ser	Ile	Ser	Gln	Val	Arg	Val	Val	Val	Gln	
1565						1570					1575				
Asp	Pro	Asn	Asn	Val	Ile	Glu	Lys	Leu	Glu	Met	Phe	Ile	Asn	Asp	
1580						1585					1590				
Glu	Leu	Ile	Ser	Thr	Lys	Thr	Asn	Leu	Glu	Ile	Asn	Ser	Ile	Ile	
1595						1600					1605				
Phe	Glu	Ile	Pro	Gln	Asn	Lys	Ile	Val	Leu	Gly	Glu	Asn	Lys	Ile	
1610						1615					1620				
Leu	Ile	Lys	Ala	Ser	Ser	Pro	Thr	Gly	Asp	Leu	Tyr	Ser	Ser	Met	
1625						1630					1635				
Phe	Ile	Phe	Lys	Ser	Glu	Thr	Gly	Leu	Lys	Val	Lys	Lys	Asp	Ser	
1640						1645					1650				
Ile	Leu	Met	Ile	Asn	Asn	Lys	Val	Tyr	Ser	Ile	Ile	Asn	Ile	Thr	
1655						1660					1665				
Glu	Asn	Asn	Thr	Asp	Leu	Ile	Val	Thr	Leu	Asn	Glu	Gly	Leu	Lys	
1670						1675					1680				

ES 2 622 280 T3

Asp Asp Met Met Glu Asn Asn Pro Ile Tyr Gln Leu Ile Asn Lys
1685 1690 1695

Thr Lys Val Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp
1700 1705 1710

Met Lys Leu Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu
1715 1720 1725

Ile Tyr Glu Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys
1730 1735 1740

Ile Ile Val Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro
1745 1750 1755

Ser Met Ile Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln Ala
1760 1765 1770

<210> 52
<211> 1725
<212> PRT
<213> Clostridium difficile

5

<400> 52
Met Lys Gln Asn Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys
1 5 10 15

Asn Ile Leu Ile Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe
20 25 30

Val Glu Phe Phe Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser
35 40 45

Asp Gly Asn Leu Tyr Ala Cys Gly Asp Asn Thr Gly Phe Pro Leu Gly
50 55 60

Leu Gly Lys Asp Ser Ser Glu Arg Arg Met Phe Ser Lys Val Lys Ile
65 70 75 80

Asp Asn Val Lys Tyr Val Ser Cys Gly Ser Lys His Ser Val Ala Val
85 90 95

Thr Lys Asp Gly Phe Ala Tyr Gly Ala Gly Thr Ser Asn Val Gly Gln
100 105 110

Leu Gly Val Ile Glu Ser Thr Val Tyr Tyr Glu Phe Thr Lys Leu Pro
115 120 125

Ile Asp Asp Val Lys Thr Val Ala Cys Gly Tyr Asp Phe Thr Phe Val

10

ES 2 622 280 T3

130						135						140					
Leu	Lys	Asn	Asp	Gly	Thr	Leu	Tyr	Ser	Ala	Gly	Leu	Asn	Ser	Ser	Gly		
145					150					155					160		
Gln	Leu	Gly	Leu	Gly	Asp	Thr	Asn	Asn	Arg	Ala	Thr	Phe	Thr	Lys	Val		
				165					170					175			
Asn	Ile	Asp	Ser	Val	Lys	Asp	Val	Val	Thr	Tyr	Asn	Gln	Ser	Val	Phe		
			180					185					190				
Ile	Ile	Lys	Met	Asp	Gly	Thr	Ala	His	Ala	Cys	Gly	Leu	Asn	Ser	Asn		
		195					200					205					
Gly	Gln	Leu	Gly	Ile	Asn	Ser	Thr	Leu	Asn	Lys	Ser	Val	Phe	Asn	Lys		
	210					215					220						
Ile	Glu	Gly	Met	Asp	Asn	Val	Lys	Gln	Ile	Ala	Cys	Gly	Ser	Ser	His		
225					230					235					240		
Thr	Ile	Leu	Ile	Lys	Asn	Asp	Gly	Thr	Met	Tyr	Thr	Thr	Gly	Tyr	Asn		
				245					250					255			
Gly	Val	Gly	Gln	Leu	Gly	Thr	Gly	Asn	Asn	Asn	Asn	Ser	Ile	Val	Phe		
			260					265					270				
Thr	Leu	Ser	Ser	Ile	Asn	Asn	Val	Lys	Tyr	Ala	Ser	Cys	Gly	Asn	Asn		
		275					280					285					
His	Thr	Met	Ile	Leu	Lys	Tyr	Asp	Asn	Thr	Leu	Phe	Ser	Thr	Gly	Gln		
	290					295					300						
Asn	Asn	Tyr	Gly	Gln	Leu	Ala	Asn	Ala	Asn	Lys	Asp	Val	Ala	Ser	Arg		
305					310					315					320		
Asn	Thr	Phe	Ala	Lys	Val	Asn	Val	Glu	Asn	Ile	Lys	Asp	Ile	Lys	Cys		
				325					330					335			
Gly	Ser	Gln	Phe	Asn	Phe	Leu	Ile	Asn	Gly	Ser	Lys	Glu	Ile	Phe	Val		
			340					345					350				
Ser	Gly	Cys	Asn	Leu	Ala	Gly	Gln	Leu	Gly	Ser	Phe	Phe	His	Thr	Thr		
		355					360					365					
Phe	Leu	Tyr	Glu	Phe	Ser	Asn	Val	Gln	Ser	Ser	Asn	Leu	Asp	Asn	Tyr		
	370					375					380						
Ser	Gly	Leu	Leu	Val	Asn	Asp	Asp	Tyr	Leu	Tyr	Val	Thr	Lys	Asp	Asn		
385					390					395					400		

ES 2 622 280 T3

Ser Glu Phe Leu Asn Val Lys Leu Ser Asp Asn Phe Gln Asp Tyr Lys
405 410 415

Lys Ile Glu Leu Thr Asp Ser Asn Met Phe Ile Val Met Asn Asp Gly
420 425 430

Thr Leu Tyr Ala Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Leu Gly
435 440 445

Asp Thr Val Asn Arg Ser Val Met Thr Lys Val Asp Ile Asp Asn Val
450 455 460

Leu Asp Ile Lys Gly Asn Gly Asn Ser Thr Phe Val Leu Lys Asn Asn
465 470 475 480

Gly Thr Leu Tyr Ser Cys Gly Leu Asn Ser Asn Gly Gln Leu Gly Leu
485 490 495

Arg Asp Glu Val Asn Arg Asn Ile Phe Thr Lys Ile Glu Ile Glu Asn
500 505 510

Val Lys Asp Phe Cys Val Gly Ser Asn Tyr Val Ile Ala Leu Asn His
515 520 525

Ser Lys Glu Val Tyr Gly Trp Gly Asn Asn Pro Tyr Asn Asn Ile Glu
530 535 540

Lys Thr Ser Asn Tyr Pro Tyr Lys Gln Gly Ile Ser Asn Ile Glu Lys
545 550 555 560

Ile Ala Ala Tyr Asp Tyr Ser Val Tyr Met Ile Asn Ser Glu Gly Lys
565 570 575

Leu Tyr Val Ser Gly Tyr Asn Tyr Asn Tyr Gln Leu Gly Lys Gly Asn
580 585 590

Asn Ser Asn Gln Ser Lys Ala Leu Val Ser Gln Cys Arg Thr Asn Ser
595 600 605

Thr Ser Ser Thr Ser Asn Gly Leu Arg Thr Leu Pro Lys Ile Thr Asn
610 615 620

Val Phe Pro Phe Tyr Asp Gly Cys Ala Ile Ile Asp Glu Gly Gly Tyr
625 630 635 640

Val Tyr Leu Thr Gly Tyr His Gly Tyr Leu Arg Thr Leu Asn Ser Ser
645 650 655

ES 2 622 280 T3

Pro Ser Ile Ser Asp Tyr Ser Arg Tyr Gly Thr Phe Ile Glu Ala Thr
660 665 670

Asn Ser Asn His Asn Thr Tyr Phe Ile Gln Glu Thr Asp Phe Ser Gly
675 680 685

Ile Glu Lys Val Ile Gly Met Ser Asn Asn Ile Leu Phe Phe Lys Lys
690 695 700

Gly Ser Ser Tyr Ile Thr Gly Tyr Pro Lys Thr Phe Gly Ser Thr Ile
705 710 715 720

Thr Gly His Arg Ser Tyr Thr Ser Ile Asn Ser Glu Ser Ser Asn Leu
725 730 735

Gly Ser Asn Phe Ile Ile Tyr His Ser Asn Ser Lys Leu Tyr Gly Lys
740 745 750

Gly Ile Ala Asn Ser Gly Gln Phe Gly Asn Ser Thr Asn Ile Asp Gly
755 760 765

Thr Ser Asn Tyr Asp Thr Gly Leu Lys Asp Ile Lys Asp Ile Ile Val
770 775 780

Lys Gly Asn Thr Val Val Val Val Asp Lys Asn Asn Asn Ile Tyr Val
785 790 795 800

Thr Gly Met Asn Gln Asn Asn Lys Leu Gly Ile Gly Glu Tyr Asn Asn
805 810 815

Glu Pro Val Lys Lys Phe Thr Asn Ile Thr Glu Gln Ser Asn Ser Phe
820 825 830

Ile Phe Met Asp Asp Ile Lys Glu Ile Thr Thr Ser Arg Asn Thr Met
835 840 845

Phe Ile Val Lys Asn Asp Gly Thr Ala Tyr Ala Thr Gly Asn Asn Ser
850 855 860

Ser Gly Gln Leu Gly Leu Gly Asp Thr Ile Asn Arg Asn Lys Phe Thr
865 870 875 880

Gln Ile Asn Leu Asp Asn Ile Lys Lys Ile Ser Thr Ser Ile Asp Gly
885 890 895

Asn Thr Thr Phe Ala Ile Arg Asn Asp Gly Thr Leu Tyr Ser Thr Gly
900 905 910

ES 2 622 280 T3

Leu Asn Thr Lys Gly Gln Leu Gly Leu Gly Asp Ile Val Asn Arg Asn
 915 920 925
 Thr Phe Thr Lys Val Asn Ile Gln Asn Val Arg Asp Val Val Leu Gly
 930 935 940
 Thr Thr His Ser His Ala Ile Lys Asp Asp Asn Thr Leu Tyr Ser Cys
 945 950 955 960
 Gly Glu Asn Thr His Gly Gln Leu Gly Leu Gly Ser Glu Ser Asn His
 965 970 975
 Pro Asp Val Leu Thr Phe Thr Val Asn Asn Ile Thr Asn Val Arg Asp
 980 985 990
 Val Tyr Cys Ser Asp Thr Thr Thr Phe Ile Val Lys Asp Thr Asn Ile
 995 1000 1005
 Ala Tyr Cys Cys Gly Tyr Asn Asn Asn Ser Gln Leu Gly Met Gly
 1010 1015 1020
 Asn Thr Thr Asp Gln Tyr Ser Phe Ile Lys Cys Met Glu Asn Val
 1025 1030 1035
 Lys Glu Val Ile Pro Asn Glu Ile Asn Thr Tyr Ile Ile Thr Ile
 1040 1045 1050
 Tyr Asn Thr Ala Tyr Ser Thr Gly Leu Asn Thr Asp Tyr Cys Leu
 1055 1060 1065
 Gly Leu Asn Ser Asn Ser Asn Gln Ser Ser Phe Ser Glu Ile Pro
 1070 1075 1080
 Ile Ser Asn Val Val Lys Val Ala Pro Asn Arg Asn Asn Ala Val
 1085 1090 1095
 Leu Leu Leu Thr Ser Glu Gly Asp Val Tyr Thr Ala Gly Lys Cys
 1100 1105 1110
 Ser Asn Gly Ser Gly Thr Gly Ser Glu Thr Pro Glu Lys Ile Lys
 1115 1120 1125
 Lys Ile Ala Ser Lys Ala Lys Asp Ile Gly Met Asn Tyr Arg Cys
 1130 1135 1140
 Gly His Tyr Val Ser Asp Asn Gly Asp Leu Tyr Gly Thr Gly Phe
 1145 1150 1155
 Asn Asn Asn Gly Gln Leu Gly Val Gly Asp Val Thr Lys Arg Asp

ES 2 622 280 T3

1160		1165		1170
Thr Phe Ile Lys Thr Asn Thr Arg Val Lys Lys Ile Leu Pro Leu 1175		1180		1185
Glu Tyr Ala Asn Ile Ala Ile Lys Asp Thr Asn Asp Ile Tyr Ile 1190		1195		1200
Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Val Gly Asn Arg Tyr 1205		1210		1215
Asp Ser Arg Asn Asn Asp Asn Arg Ile Phe Asn Tyr Lys His Met 1220		1225		1230
Asn Phe Val Met Gly Asp Leu Thr Ser Ile Lys Asn Arg His Asn 1235		1240		1245
Phe Ile Leu Leu Asn Asn Lys Ile Val Ile Pro Thr Thr Lys Asp 1250		1255		1260
Ile Asp Tyr Gly Leu Val Leu Gly Asn Leu Tyr Lys Gly Asp Leu 1265		1270		1275
Tyr Thr Glu Leu Pro Tyr Glu Asp Ile Lys Glu Val Ser Ile Ser 1280		1285		1290
Lys Thr His Ile Ile Ile Leu Leu Asn Asp Gly Thr Met Tyr Gly 1295		1300		1305
Cys Gly Thr Asn Tyr His Gly Glu Leu Leu Gln Asp Leu Ser Ile 1310		1315		1320
Asn Gln Val Asp Glu Phe Val Gln Ile Asn Val Ser Asp Val Lys 1325		1330		1335
His Val Ser Cys Gly Asp Asn Phe Thr Tyr Phe Ile Lys Ser Asp 1340		1345		1350
Asp Ser Leu Trp Ser Ile Gly Lys Asn Ser Glu Tyr Gln Leu Gly 1355		1360		1365
Ile Gly His Asn Asn Pro Val Thr Glu Leu Gln Arg Ile Thr Thr 1370		1375		1380
Ile Ser Ser Cys Lys Glu Val His Cys Gly Lys Asn Tyr Thr Leu 1385		1390		1395
Val Val Thr Thr Gly Asn Glu Leu Phe Val Gln Gly Tyr Asn Asp 1400		1405		1410

ES 2 622 280 T3

Lys Gly Ala Leu Gly Leu Gly Ser Asp Ser Glu Asn Thr Ile Ile
 1415 1420 1425
 Lys Phe Phe Thr Lys Ala Leu Thr Asp Ile Arg Glu Ile Lys Ser
 1430 1435 1440
 Tyr Gly Ser Asp His Ile Leu Val Leu Lys Asn Asp Asn Ser Val
 1445 1450 1455
 Trp Val Thr Gly Lys Asn Arg Asp Val Tyr Lys Ile Glu Gln Pro
 1460 1465 1470
 Val Glu Phe Leu Lys Glu Phe Thr Ile Val Pro Ile Ser Glu Asp
 1475 1480 1485
 Val Asn Thr Val Lys Asp Val Leu Ala Thr Asp Asn Thr Leu Tyr
 1490 1495 1500
 Ile Ile Ser Glu Val Gly Thr Thr Asn Ala Ala Ile Glu Ile Thr
 1505 1510 1515
 Glu Lys Ser Ile Ser Ser Ile Lys Ile Lys Ile Gln Asp Pro Asn
 1520 1525 1530
 Lys Asp Ile Ser Arg Ile Glu Met Leu Ile Asn Gly Glu Ser Val
 1535 1540 1545
 Lys Ser Val Ser Asp Leu Thr Thr Glu Lys Ile Ser Phe Glu Val
 1550 1555 1560
 Pro Pro Asp Lys Ile Lys Ile Gly Glu Asn Lys Ile Leu Phe Arg
 1565 1570 1575
 Ala Tyr Cys Lys Gly Asp Asp Leu Tyr Ala Ser Leu Phe Ile Phe
 1580 1585 1590
 Lys Glu Ser Thr Gly Asn Ser Ile Ile Lys Asp Ser Tyr Val Met
 1595 1600 1605
 Ile Gly Asn Arg Met Tyr Lys Val Val Asn Thr Thr Ser Asn Glu
 1610 1615 1620
 Gln Asp Ile Thr Ile Thr Leu Asp Arg Gly Leu Glu Glu Asp Leu
 1625 1630 1635
 Asn Leu Gly Asp Pro Ile Tyr Gln Leu Ile Asn Lys Thr Lys Val
 1640 1645 1650

ES 2 622 280 T3

Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp Met Lys Leu
1655 1660 1665

Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu Ile Tyr Glu
1670 1675 1680

Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys Ile Ile Val
1685 1690 1695

Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro Ser Met Ile
1700 1705 1710

Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln Ala
1715 1720 1725

<210> 53

<211> 1743

5 <212> PRT

<213> Clostridium difficile

<400> 53

Met Lys Gln Asn Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys
1 5 10 15

Asn Ile Leu Ile Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe
20 25 30

Val Glu Phe Phe Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser
35 40 45

Asp Gly Asn Leu Tyr Ala Cys Gly Asn Asn Thr Gly Phe Pro Leu Gly
50 55 60

Leu Gly Lys Asp Ser Ser Glu Arg Arg Met Phe Ser Lys Val Lys Ile
65 70 75 80

Asp Asn Val Lys Tyr Val Ser Cys Gly Ser Lys His Ser Val Ala Val
85 90 95

Thr Lys Asp Gly Phe Ala Tyr Gly Ala Gly Thr Ser Asn Val Gly Gln
100 105 110

Leu Gly Val Ile Glu Ser Thr Val Tyr Tyr Glu Phe Thr Lys Leu Pro
115 120 125

Ile Asp Asp Val Lys Thr Val Ala Cys Gly Tyr Asp Phe Thr Phe Val
130 135 140

Leu Lys Asn Asp Gly Thr Leu Tyr Ser Ala Gly Leu Asn Ser Ser Gly
145 150 155 160

10

ES 2 622 280 T3

Gln Leu Gly Leu Gly Asp Thr Asn Asn Arg Ala Thr Phe Thr Lys Val
 165 170 175

Asn Ile Asp Ser Val Lys Asp Val Val Thr Tyr Asn Gln Ser Val Phe
 180 185 190

Ile Ile Lys Met Asp Gly Thr Ala His Ala Cys Gly Leu Asn Ser Asn
 195 200 205

Gly Gln Leu Gly Ile Asn Ser Thr Leu Asn Lys Ser Val Phe Asn Lys
 210 215 220

Ile Glu Gly Met Asp Asn Val Lys Gln Ile Ala Cys Gly Ser Ser His
 225 230 235 240

Thr Ile Leu Ile Lys Asn Asp Gly Thr Met Tyr Thr Thr Gly Tyr Asn
 245 250 255

Gly Val Gly Gln Leu Gly Thr Gly Asn Asn Asn Asn Ser Ile Val Phe
 260 265 270

Thr Leu Ser Ser Ile Asn Asn Val Lys Tyr Ala Ser Cys Gly Asn Asn
 275 280 285

His Thr Met Ile Leu Lys Tyr Asp Asn Thr Leu Phe Ser Thr Gly Gln
 290 295 300

Asn Thr Tyr Gly Gln Leu Ala Asn Ala Asn Lys Asp Val Ala Ser Arg
 305 310 315 320

Asn Thr Phe Ala Lys Val Asn Val Glu Asn Ile Lys Asp Ile Lys Cys
 325 330 335

Gly Ser Gln Phe Asn Phe Leu Ile Asn Gly Ser Lys Glu Ile Phe Val
 340 345 350

Ser Gly Cys Asn Leu Ala Gly Gln Leu Gly Ser Phe Phe His Thr Thr
 355 360 365

Phe Leu Tyr Glu Phe Ser Lys Val Gln Ser Ser Asn Leu Asp Asn Tyr
 370 375 380

Ser Gly Leu Leu Val Asn Asp Asp Tyr Leu Tyr Val Thr Lys Asp Asn
 385 390 395 400

Ser Glu Phe Leu Asn Val Lys Leu Ser Asp Asn Phe Gln Asp Tyr Lys
 405 410 415

ES 2 622 280 T3

Lys Ile Glu Leu Thr Asp Asn Asn Met Phe Ile Val Met Asn Asp Gly
 420 425 430

Thr Leu Tyr Ala Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Leu Gly
 435 440 445

Asp Thr Val Asn Arg Ser Val Met Thr Lys Val Asp Ile Asp Asn Val
 450 455 460

Leu Asp Ile Lys Gly Asn Gly Asn Ser Thr Phe Val Leu Lys Asn Asn
 465 470 475 480

Gly Thr Leu Tyr Ser Cys Gly Tyr Asn Ser Ser Gly Ile Leu Gly Leu
 485 490 495

Lys Asp Asn Thr Asn Arg Asn Ile Phe Thr Lys Ile Glu Ile Glu Asn
 500 505 510

Val Lys Ala Phe Cys Val Glu Ser Asn Tyr Ile Val Val Leu Asn His
 515 520 525

Ser Lys Glu Leu Tyr Gly Trp Gly Asn Glu Ser Tyr Ile Val Tyr Gly
 530 535 540

Asn Ser Arg Asn Tyr Pro Tyr Lys Asp Thr Arg Val Ser Asn Val Glu
 545 550 555 560

Lys Ile Ala Thr Trp Ser Asp Thr Leu Tyr Ile Leu Asp Ser Thr Gly
 565 570 575

Ala Thr Lys Thr Ile Gly Tyr Ser Tyr Asn Gly Ser Gly Gly Tyr Pro
 580 585 590

Ala Pro Ser Ser Ser Ser Thr Tyr Arg Asp Gly Gly Tyr Ile Asn Lys
 595 600 605

Asn Thr Ser Tyr Arg Thr Leu Glu Phe Tyr Asn Thr Ser Lys Thr Lys
 610 615 620

Leu Val Asn Leu Phe Ala Phe Tyr Asn Gly Cys Val Phe Val Asp Glu
 625 630 635 640

Asn Gly Leu Ala Tyr Cys Ile Gly Glu Asn Asn Ile Asn Phe Arg Gly
 645 650 655

Asn Ser Thr Thr Asn Glu Asn Asn Ser Leu Arg Phe Ile Asn Asn Ser
 660 665 670

ES 2 622 280 T3

Gly Val Tyr Tyr Thr Asn Thr Asp Gly Thr Asp Tyr Thr Cys Tyr Gln
675 680 685

Trp Thr Tyr Lys Leu Ile Arg Cys Ser Ile Phe Asp Ser Pro Gln Asn
690 695 700

Ile Ile Gly Asn Ser Lys Asn Ile Leu Tyr Leu Ser Lys Asn Asn Ser
705 710 715 720

Thr Phe Lys Cys Thr Gly Asn Cys Ile Thr Tyr Gly Ile Asn Ser Gln
725 730 735

Asn Trp Tyr Ser Tyr Phe Ser Asp Ser Ser Asn Gly Ala Ile Ala Leu
740 745 750

Gly Asn Glu Phe Ile Leu Lys Asn Tyr Ser Gly Glu Cys Leu Leu Lys
755 760 765

Gly Tyr Gly Lys Ala Thr Asn Gly Glu Phe Gly Asn Ser Thr Asn Ile
770 775 780

Ser Ser Ile Ser Asn Tyr Asp Thr Gly Leu Lys Asp Ile Lys Asp Ile
785 790 795 800

Ile Val Lys Asn Asn Thr Val Val Val Val Asp Lys Asn Asn Asn Ile
805 810 815

Tyr Val Thr Gly Ala Asn Gln Phe Asn Lys Leu Gly Ile Gly Glu Tyr
820 825 830

Asn Asn Gln Pro Ile Lys Lys Phe Thr Asn Ile Thr Glu Gln Ser Asn
835 840 845

Ser Phe Ile Phe Met Asp Asp Ile Lys Glu Ile Thr Thr Ser Arg Asn
850 855 860

Thr Met Phe Ile Val Lys Asn Asp Gly Thr Ala Tyr Ala Thr Gly Asn
865 870 875 880

Asn Ser Ser Gly Gln Leu Gly Leu Gly Asp Thr Ile Asn Arg Asn Lys
885 890 895

Phe Thr Gln Ile Asn Leu Asp Asn Ile Lys Lys Ile Ser Thr Ser Ile
900 905 910

Asp Gly Asn Thr Thr Phe Ala Ile Arg Asn Asp Gly Thr Leu Tyr Ser
915 920 925

Thr Gly Leu Asn Thr Lys Gly Gln Leu Gly Leu Gly Asp Ile Val Asn

ES 2 622 280 T3

Lys Arg Asp Thr Phe Ile Lys Thr Asn Thr Arg Val Lys Lys Ile
 1190 1195 1200

Leu Pro Leu Glu Tyr Ala Asn Ile Ala Ile Lys Asp Thr Asn Asp
 1205 1210 1215

Ile Tyr Ile Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Val Gly
 1220 1225 1230

Asn Arg Tyr Asp Ser Arg Asn Asn Asp Asn Arg Ile Phe Asn Tyr
 1235 1240 1245

Lys His Met Asn Phe Val Met Gly Asp Leu Thr Ser Ile Lys Asn
 1250 1255 1260

Arg His Asn Phe Ile Leu Leu Asn Asn Lys Ile Val Ile Pro Thr
 1265 1270 1275

Thr Lys Asp Ile Asp Tyr Gly Leu Val Leu Gly Asn Leu Tyr Lys
 1280 1285 1290

Gly Asp Leu Tyr Thr Glu Leu Pro Tyr Glu Asp Ile Lys Glu Val
 1295 1300 1305

Ser Ile Ser Lys Thr His Ile Ile Ile Leu Leu Asn Asp Gly Thr
 1310 1315 1320

Met Tyr Gly Cys Gly Thr Asn Tyr His Gly Glu Leu Leu Gln Asp
 1325 1330 1335

Leu Ser Ile Asn Gln Val Asp Glu Phe Val Gln Ile Asn Val Ser
 1340 1345 1350

Asp Val Lys His Val Ser Cys Gly Asp Asn Phe Thr Tyr Phe Ile
 1355 1360 1365

Lys Ser Asp Asp Ser Leu Trp Ser Ile Gly Lys Asn Ser Glu Tyr
 1370 1375 1380

Gln Leu Gly Ile Gly His Asn Asn Pro Val Thr Glu Leu Gln Arg
 1385 1390 1395

Ile Thr Thr Ile Ser Ser Cys Lys Glu Val His Cys Gly Lys Asn
 1400 1405 1410

Tyr Thr Leu Val Val Thr Thr Gly Asn Glu Leu Phe Val Gln Gly
 1415 1420 1425

ES 2 622 280 T3

Tyr Asn Asp Lys Gly Ala Leu Gly Leu Gly Ser Asp Ser Glu Asn
 1430 1435 1440
 Thr Ile Ile Lys Phe Phe Thr Lys Ala Leu Thr Asp Ile Arg Glu
 1445 1450 1455
 Ile Lys Ser Tyr Gly Ser Asp His Ile Leu Val Leu Lys Asn Asp
 1460 1465 1470
 Asn Ser Val Trp Val Thr Gly Lys Asn Arg Asp Val Tyr Lys Ile
 1475 1480 1485
 Glu Gln Pro Val Glu Phe Leu Lys Glu Phe Thr Ile Ile Pro Ile
 1490 1495 1500
 Ser Glu Asp Val Asn Thr Val Lys Asp Val Leu Ala Thr Asp Asn
 1505 1510 1515
 Thr Leu Tyr Ile Ile Ser Glu Val Gly Thr Thr Asn Ala Ala Ile
 1520 1525 1530
 Glu Ile Thr Glu Lys Ser Ile Ser Ser Ile Lys Ile Lys Ile Gln
 1535 1540 1545
 Asp Pro Asn Lys Asp Ile Ser Arg Ile Glu Met Leu Ile Asn Gly
 1550 1555 1560
 Glu Ser Val Lys Ser Val Ser Asp Leu Ile Thr Glu Lys Ile Ser
 1565 1570 1575
 Phe Glu Val Pro Pro Asp Lys Ile Lys Ile Gly Glu Asn Lys Ile
 1580 1585 1590
 Leu Phe Arg Ala Tyr Cys Lys Gly Asp Asp Leu Tyr Ala Ser Leu
 1595 1600 1605
 Phe Ile Phe Lys Glu Ser Thr Gly Asn Ser Ile Ile Lys Asp Ser
 1610 1615 1620
 Tyr Val Met Ile Gly Asn Arg Met Tyr Lys Val Val Asn Thr Thr
 1625 1630 1635
 Ser Asn Glu Gln Asp Ile Thr Ile Thr Leu Asp Arg Gly Leu Glu
 1640 1645 1650
 Glu Asp Leu Asn Leu Gly Asp Pro Ile Tyr Gln Leu Ile Asn Lys
 1655 1660 1665

ES 2 622 280 T3

Thr Lys Val Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp
1670 1675 1680

Met Lys Leu Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu
1685 1690 1695

Ile Tyr Glu Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys
1700 1705 1710

Ile Ile Val Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro
1715 1720 1725

Ser Met Ile Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln Ala
1730 1735 1740

<210> 54

<211> 341

5 <212> PRT

<213> fago de Clostridium

<400> 54

Met Ala Ile Asp Lys Ser Tyr Tyr Thr Ile Ile Thr Asp Val Gly Lys
1 5 10 15

Ala Lys Ile Ala Asn Ala Ser Val Thr Gly Asn Lys Val Gly Phe Val
20 25 30

Lys Ile Gln Leu Gly Asp Gly Gly Gly Ser Glu Tyr Thr Pro Thr Glu
35 40 45

Ser Gln Thr Ala Leu Lys Asn Val Val Trp Glu Gly Asn Ile Gly Asn
50 55 60

Thr Thr Thr Asp Glu Thr Ala Pro Asn Cys Ile Ile Leu Glu Ser Leu
65 70 75 80

Ile Pro Ser Ser Val Gly Gly Phe Met Ile Arg Glu Ile Gly Tyr Leu
85 90 95

Asp Asp Glu Asn Asn Leu Ile Ala Ile Ser Lys Tyr Lys Glu Cys Tyr
100 105 110

Lys Pro Ser Ile Glu Gln Gly Ala Val Val Asp Met Lys Val Lys Thr
115 120 125

Val Leu Ile Val Ser Asn Val Asn Asn Ile Glu Leu Lys Ile Asp Pro
130 135 140

Thr Ile Ile Phe Ala Thr Leu Lys Asp Ile Gln Asp Leu Glu Thr Lys
145 150 155 160

10

ES 2 622 280 T3

Ile Gly Thr Val Asn Thr Lys Ile Asp Thr Thr Lys Thr Glu Leu Thr
 165 170 175

Ser Asn Ile Glu Thr Thr Lys Thr Glu Leu Asn Thr Arg Ile Asp Thr
 180 185 190

Glu Asn Glu Lys Gln Asn Ile Lys Ile Asp Gln Leu Ile Ala Gly Gly
 195 200 205

Ser Asn Val Ala Ser Thr Gln Ile Ile Thr Ile Asp Asp Trp Val Glu
 210 215 220

Asp Ala Glu Asn Gly Phe Lys Ala Thr Val Thr His Ser Leu Leu Thr
 225 230 235 240

Gln Arg Ile Val Val Asn Ile Ile Asp Ala Thr Thr Lys Glu Asn Val
 245 250 255

Val Thr Asn Phe Lys Ile Ile Asp Asp Asn Ser Ile Glu Ile Arg Ser
 260 265 270

Glu Val Lys Val Glu Leu Asn Val Tyr Val Ile Asn Gly Asn Ala Glu
 275 280 285

Thr His Phe Ile Asn Ala Thr Val Asp Asp Asn Arg Val Ser Glu Met
 290 295 300

Thr Thr Tyr Ser Ser Lys Lys Ile Glu Asp Arg Leu Val Asn Ile Glu
 305 310 315 320

Glu Lys Val Asn Gly Gly Leu Ser Asn Ile Ala Thr Ser Val Asn Glu
 325 330 335

Leu Ile Thr Tyr Cys
 340

<210> 55

<211> 132

5 <212> PRT

<213> fago de Clostridium

<400> 55

Met Ala Glu Gln Gln Tyr Phe Thr Leu Val Thr Asp Ile Gly Lys Ala
 1 5 10 15

Ala Ile Ala Asn Ala Ser Val Thr Gly Glu Lys Val Asp Phe Ala Lys
 20 25 30

Ile Lys Val Gly Asp Gly Gly Gly Ser Ser Tyr Thr Pro Asn Glu Ser
 35 40 45

10

ES 2 622 280 T3

Gln Thr Ala Leu Lys Asn Val Val Trp Glu Ser Thr Leu Glu His Ala
50 55 60

Gln Val Asp Lys Asp Asn Pro Asn Trp Val Val Ile Gln Lys Phe Ile
65 70 75 80

Pro Gly Asp Val Gly Gly Phe Glu Ile Arg Glu Val Gly Leu Phe Asp
85 90 95

Ser Lys Asp Gln Leu Leu Ala Val Ser Ser Tyr Pro Thr Thr Tyr Lys
100 105 110

Pro Glu Ser Arg Phe Gly Asp Cys Lys Arg Thr Ile Asn Lys Ser Asn
115 120 125

Ile Ser Cys Ile
130

<210> 56

<211> 205

5 <212> PRT

<213> fago de Clostridium

<400> 56

Met Pro Asn Glu Leu Asn Phe Asn Asn Glu Ile Glu Glu Tyr Leu Ile
1 5 10 15

Thr Thr Pro Ala His Ala Asn Glu Phe Asn Asn Arg Gln Gln Lys Leu
20 25 30

Leu Asp Asn Asp Lys Tyr Leu Asn Asn Lys Ile Asp Thr Thr Lys Thr
35 40 45

Glu Leu Asn Thr Arg Ile Asp Thr Glu Asn Glu Lys Gln Asn Ile Lys
50 55 60

Ile Asp Gln Leu Ile Ala Gly Gly Ser Asn Val Ala Tyr Thr Gln Arg
65 70 75 80

Val Ala Ile Asp Asp Trp Val Glu Asp Ala Glu Asn Gly Phe Lys Ala
85 90 95

Thr Val Thr His Ser Leu Leu Thr Gln Arg Ile Val Val Asn Ile Ile
100 105 110

Asp Ala Thr Thr Lys Glu Asn Val Val Thr Asn Phe Lys Ile Ile Asp
115 120 125

Asp Asn Ser Ile Glu Ile Arg Ser Glu Thr Arg Ser Glu Leu Asn Val

10

ES 2 622 280 T3

130

135

140

Tyr Val Ile Asn Gly Asn Ala Glu Thr His Phe Ile Asn Ala Thr Val
145 150 155 160

Asp Asp Asn Arg Val Ser Glu Met Thr Thr Tyr Ser Ser Lys Lys Ile
165 170 175

Glu Asp Arg Leu Val Asn Ile Glu Glu Lys Val Asn Gly Gly Leu Ser
180 185 190

Asn Ile Ala Thr Ser Val Asn Glu Leu Ile Thr Tyr Cys
195 200 205

<210> 57

<211> 14

5 <212> ADN

<213> Secuencia Artificial

<220>

<221> fuente

10 <223>/nota = "Descripción de la secuencia artificial: oligonucleótido sintético"

<400> 57

ggccgcctcg aggg

14

15 <210> 58

<211> 14

<212> ADN

<213> Secuencia Artificial

20 <220>

<221> fuente

<223>/nota = "Descripción de la secuencia artificial: oligonucleótido sintético"

<400> 58

25 cgcgccctcg aggc

14

<210> 59

<211> 20

<212> ADN

30 <213> Secuencia Artificial

<220>

<221> fuente

<223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"

35

<400> 59

tgaagtacca tggatccag

20

<210> 60

<211> 20

<212> ADN

<213> Secuencia Artificial

45 <220>

<221> fuente

<223>/nota = "Descripción de la Secuencia Artificial: Cebador sintético"

<400> 60

actggatacc atgtacttc

20

ES 2 622 280 T3

<210> 61
 <211> 19897
 <212> ADN
 <213> Clostridium difficile

5

<400> 61
 gcaataccba ctacaccttc gtcactctta aatttaagag tttttactat tgaataataa 60
 aggtatattc cagtaaaaaat aatcttttaa tacaagaaaa ataaactctt tgggtatatt 120
 aaaaagctaa aaagtgtaaa tataaaagca agtagagtac ttatcctgta aaagaaatct 180
 atttgtgtaa tgtctttata ttttatcata aacaccgaat ataaaatgat gaaaacaatt 240
 gcgacgattg catatatggg aaataacata ttttcaagag taccatttga aattactatc 300
 cacttatacc acataattgg ccaaaataat agtgctaaga acttaaaata attatcaaac 360
 aacttttctt tatacattca tcaaacacc tttcttaaca aaagcatata tttgttttta 420
 gaattttaaa taatatgata tcattattat atattaatat tgaatttata gaaacccaaa 480
 tttgttaaaa taaatatata gattttactg ttaagccagt taaaattact actattttta 540
 ttatgaaatt ggatcaaata tgtagaaata cggcaaatta gttaatatta aatatttatt 600
 atttccaagt tgtaagact gtttttttaa tgatagaaat tctaactctt tttgaaagaa 660
 agtaatatcc acattaagta tgtctgccat ttcataaacg caagtgatgc cagagttaat 720
 tatgtttatt atatcttctt cagtaattaa gaactcaca gcccatttta aggctttatt 780
 ttcgcactta tctataataa tttttgtata ataatcgta taagaggata catagtatcc 840
 aaggctagtg aatgatgtc caagttcttc agctaagatg gatgtcaatt tttttgagtt 900
 ttgttttaaa ttactgagta atgatataat tttaatacca tgtttgttta tatatagtcc 960
 ttctaaatca cctgcaatat aagtgtata atgaattatt atctcttctt gagaagctaa 1020
 ttcaaaaagc ttatccaaat tattcataaa aatcccccta aaatagaatg tatgtttgcc 1080
 tttaaattat attaaaagag cagaaaaata gactgctcat catatggttt atttttttt 1140
 atatttattt agtaaaaatt ctatataatc attaagttgt tcttgtgctt cttcaggtaa 1200
 ctcttcatgt ggattttttc tatgtgcagc tactgtatca atattttcct taactaaggt 1260
 tcttccaaga aggtaatcaa ctgatacatt aaatacatca gccaatat ttaaaatgtg 1320
 ttcacagga aatctgtttt ctgtttcata gtaccctaag actctttggg aaacgcctac 1380
 tttttctcca agttctctt gagtcaatcc aaattcctt ctaagttctc ttaatctttt 1440
 ggcaaacatt ataacaccac cttatgtata gattataaca aattgttcta aaaaataaaa 1500

ES 2 622 280 T3

ctaataaaat ataaaagaat attttttctt aaaatctatt gataaagaac aaataattct 1560
 atataatcta agtgaggaag aacaaaatat tcttaatagt aatggaggta taaaacaatg 1620
 tttaaaata acttgaaata ttatagaaaa tgcaaaggta tgacacaaat tcaacttgcc 1680
 agaaaggctg gaattacaaa tgattatata tctcaaatag aaagaggat aaaaaatcct 1740
 ggacttctta tggctaagaa gatttctagt attttagaac aaaatataga agaagttttt 1800
 tttatacagt tatagaacaa tatgttcttg aaagttgtga gattagtaaa aaactgtgca 1860
 ctaaagagat tattgtaaat ttgaagctaa taataagtat ataaaaaagg ggaagtacta 1920
 tggaaaacaa aaaagatata ttatttaaag aaacagataa aagattacat aattataagt 1980
 atttggatat aaagataaag aatattaact tggacataaa aagatgtgag aatgaatact 2040
 ctggatgtgg agcaatggta tatacagaaa agactagtaa cacatataac ataagctctt 2100
 ctgtggaaaa tgaggtgtta aaaagagagg aaagattaag aaaattaaaa atggaaaaag 2160
 aagatataga aatagaaaaa gagaagatag aaaatgctct aacgtgtcta aatgatatag 2220
 aatggaatt ttttaacttt ttttataata gtaagacaaa aaacaatatg acatatattt 2280
 ctatgaaact acacttagat agaacatctt gctacaatth aaagaaaaag atgatattta 2340
 aattgagtga gatattataa aaaataggac aatthtaca cactttatat acaccattgc 2400
 aacaataggc aataaaaat gtgagataat gttattgtga aagaaatcca tattgaagga 2460
 ggtgatagat tgaaaagaat aatattacct aaaaatatag aagatacttg acaggaataa 2520
 atgagatata tatttaaaaa tgacttatat catttatagt aagattatca gattaagcaa 2580
 gaatatttag tgatagtgtg gtgattatth gcttaaac aaagaaatat tagaaacaat 2640
 tattgagatt ctcaaaaaa actttactga aagtatthtt attgatgatg aaagtgtgca 2700
 aggctctgaa gggctctgtt tttttgtaag tatactatca gttatttgta cacctgtaat 2760
 gttaaatagc aataacaaag atattgttat ctctataaaa tacttaccaa aaccacagtc 2820
 aaagagtatt agaatgtatg aaatthcaga tgaattaaat aagctatthta acagaaatat 2880
 aaagtaaca gacagaaat taaatataac aaagctagaa caaagtatta aaaagaaga 2940
 gtcaattht gtattgaact ttacatthac actaaactat ctggatagtg tatatgaaga 3000
 agatgtagta tatgaaaata tgaaagaaat caatthaaat ttaggagagt gatagtatgg 3060
 ctataggatt accaagtatc aacatatcat ttaaggagct agctacaact gttaaagaac 3120
 gttcagctag aggaataatt gcaatggtac ttaaagatgc taaggcacta ggtcttaatg 3180
 aaatacatga aaaagaggat ataccagttg atthtctgc tgaaaataaa gaatatataa 3240
 atthtagctth gatgggaaat gttaacactc caaataaatt atthagthtatt gtaatagaag 3300
 gagaagcaga tattcaaacg gcattagatt tthtagagac taaggaaatth aattatctat 3360
 gtatgcaaaa agcagtagaa gctgataaga ctgctataaa aaatthgata atthaaactta 3420

ES 2 622 280 T3

gagatataga taaggttaag gttaaagctg tattaggaaa agttgtagga aatcatgaag 3480
ggataattaa ttttactaca gaagatgtgt tagttggaga aaagaatac agtgttgatg 3540
agtttacaag caggggtggct ggacttatag caggtacacc tttaaagtcaa tcagtaactt 3600
atactaaact tagtgatgta gtcgatatac ctaagatgac gaaagttgat gcagaatcaa 3660
gggtaataa aggagagctt atacttatta aagaagcag agctataaga attgctagag 3720
gagtaaattc tttaaactgaa ttaacagcag aaaaaggaga aatgttccag aaaataaaaa 3780
tagttgacac tttagatatt atacatagtg acataagaaa ggtgataata gatgactata 3840
taggaaaggt tactaacagt tatgacaaca aatgtttatt gatagtagct ataaaaagtt 3900
attdagaaga attagaaaaa tcagcactta tagaatctga ttctactggt gaaatagatt 3960
ttgaagcaca aaaatcgtat ttaaaatcaa aaggagtaga tttatcttat atgacattac 4020
aagaaataaa agaagctaac acaggttcta aagtatTTTT aaaagcaaaa ataaaagtac 4080
ttgatgctat ggaagatata gattttcaa tagaaatata ggaggattat taatatggca 4140
aatatggaag ctagaaatgt aatgagtggg acttggggag aactttggct tgatggaaac 4200
aaagtagcag aagtaaagaa atttcaagca aagatggaat ttacaaaaga agatattata 4260
atagcaggtc aaatgggtac tgatacaaag tatatgggat ataaaggaaa aggctcaata 4320
actctatacc acgttagttc aagaatgcac aagttaattg gagaaaagat aaagagaggt 4380
tctgaaccta gatttggttc tatatcaaaa ttaaatgacc cagattctta tggagcagaa 4440
agaatagcag taaaaaatat agcatttgat gatttaactt tagctgattg ggaggttga 4500
gtaaaaggag agatagaagc tcctttcaca tttactgagt atgattttct tgatataatt 4560
tagttttata tttagtttta tactgatatt tagtaagtat atacttaata aattcagata 4620
gttaataagt aaaaaagtta gttgattgaa tttgattgat aaaggagcaa ataataatga 4680
gtgaaaatgg attatcaaaa aatataaaca tagtagattt acttttaaat tcagatacag 4740
aaaacttaga aagaccaagt actatagttg aacttaagag attatcaact atatttgggc 4800
aggaatttaa agtaatgtgt agagctttaa caataagtaa agatgaagaa atacaaaata 4860
cttgtcttaa aattgatgaa aatagaaaa cggatataga cttaccggag atgcagatgc 4920
ttacaattat agaaggtggt tgtgatttgg atggaaagct tttatttaaa aataaggagc 4980
taatggataa atttaaggct ccaacaccaa aagaattagc aagaaaatta ttattaccag 5040
gtgaaattac caacttatat agaatacttc aagatgttat gggttatggt aaaaatgcag 5100
tgatagaaga ggtaaaaaac taatagggac ggataccag agtacaataa tgtactata 5160
ttggaagaaa aaaggtataa gaccgtccct tttttatgca atggataaag gcgaattaa 5220
gcttattgaa gcttttttcg ccttagaaat tgaggaagaa gttgaaaaaa tgaaacatgg 5280
atatggagtg tgccttttga caggaggtgg tatgtaatgg gaaatgtgag agaagaaggt 5340
ataaatatgt accttacaga taattacaca ccaaaaatga atcaaattat atcagtaact 5400

ES 2 622 280 T3

gataatthta ggagagcaac tgtggctggt tcaactttcca ctaatgtaat ggctagtagc 5460
ataaaaaatt ctattggaag tgcaagtaat agagtaaaca gtttaaattc ctcgtaaga 5520
aaagttcaaa ctactgctag tagtgtaagt tcaactatga caaaattaag ttctagcata 5580
aatgctgtht caggagttat tggagttta aatggaagta ttatgagact agcaataact 5640
atagctatga ttattgatta thhtaataag ttgattcaaa agaaaaatga gthtaattca 5700
aatattatga ttatattaat thhtaagct aaaagtgatg aagtagaaaa aactaaaaat 5760
aaattacttg gaaattthaa aaagattggt ggcaagattt ggaatatcgt aataaaagca 5820
aaagatatga ctaagagagt gataagtagt atcttgggaa aathaaaaca agtagagaaa 5880
cgtccttatac aaggaagtat taatctthaa gatatggtga gtagtgctat gggtagaatt 5940
ttgcctaagt taatgthtgt thhaaataact thhtggagtg gtgtaatagc tataaaagat 6000
atggcaagtg gcattataag thaaagtattt cccaattga gattatttgc aggtaaggta 6060
tggagtggtg caatagctgt aaagatatg gcaagtggaa tacttggthc gataaaaggg 6120
aagatatctg atthgacaaa tgggtgctact ataggtgtcg ctgtgaaaaa ggggtthgat 6180
ttacttggthc aggaacaaaa tcagaaagth gthctagaaa gtgtaatgaa aagaaatact 6240
ggaaaagtha atcaaataga thgtgatgat thtatggca gthtagtaag aatggcaaat 6300
gatacgcctt thgacctga agatgthtgt gcaatgggaa cthaaagctaa aatgattagt 6360
aatattactg thggcaaaaa agaaaaagat thaaactcaag ctatgtaga thtagagct 6420
thaaatatga atacaagtag thgacaagat gthctagcag ththctthaa thgagcaaaa 6480
ggaaacatgg aatctctthaa tactctggta ggagaaaatt thaaaactth thgatgaagca 6540
thggaaggca thagtthaaa gcagatgggg thtagctaaag aatgagthaa thacaatcca 6600
ggtataatat caggagctca aacaagcatt acaaatggct thgagagtht thththaaact 6660
ththgatgata ththtagthca aggactthaaag thhaathaaaa ththththtaga aagthgatta 6720
ggcaattthg ctggcttatac thgaaaaatg gctggthaaa thaggcaatgt aatgaaatgg 6780
aagataatta thggcaacaa atatgaccag atgcaatcta gaagtthaaa aatggaaaa 6840
gagththctg atthctactca atatcgaatt thctaatgag ctgaaaagcg thaaatgatg 6900
gthgaaaata agcaagaacg ththgaaaat catgagcaa caatgatagg gaatgacca 6960
aaagcaattg thaacgcag aggtacacta thacaaaata thgattthac agcattaata 7020
gattcattac thccagtagt aaacttagth aataattthac thagatagtht thacaataaa 7080
thaccaattg cacaaggatt aataagthata ththgthacaa thagthactac agcattcca 7140
ctaatcggac ctgtagthga agctgthtagt cctattatca caagaattth tactththta 7200
ggthgaaatg cacctcaaat thacaattth atagagacac thgggtgtht thggaaaact 7260
gthtgggaga ccttaggacc thtgtthgaa actggatgga thattataga gccaatattg 7320

ES 2 622 280 T3

ggagcttttt ttaacatatt agataaagta tgtaaaatag ttaaagatat atgtaaatgg 7380
 tggcaaaacta tgattaataa gataaaaaat ggaagcatca caggaacagt tttaaatcta 7440
 gtggaaaaga gtaaaaaaaa ttacaaagat aatccatatg ctggaacaaa ggctggtgat 7500
 tctggtaaag cttattcagc taagaaaggt aataatgcat ttggattgaa ctatgttctc 7560
 tataatgact atcaaaccag actccatgaa ggtgaaatgg ttttaactaa acaagaagca 7620
 aatcaatata gaagcagaaa aatggtgga aatataaaca tagctaagtt agctgataca 7680
 atagtgatta gagaagaagc tgatatagaa aagataacat caaaattagt tgcaagtatc 7740
 caattggcac agttaggggg tgtcttataa tggaaatgtg gcttagacaa gctgaagata 7800
 gatttagatt tccagtattt ccatcttctc ttagtattaa tggaaaagct gctgtaaact 7860
 cttctagtat actcaaaata ggtgaaatag caacttttgg tgggttagct cttaaaagca 7920
 tttcaatatac aagttttttt ccaataaag actatacttt ctgtgactat acaggttttc 7980
 catcaccata tgattgtgta aataagatag aaaaatggat gaaggaaggt tttatattaa 8040
 gatttacaat tacggaaaca aatataaata tggaaatcat aattgaaggg tttagttagt 8100
 aagaaagaga tgggactcga gatgtatatt ttacattaga tttaaaagag tataaaagaa 8160
 taaagatacc aaaagtaact ccaaaacaat aactattata gataataagt tataaataac 8220
 tgctgataga attaaatgaa aaggcaggtg attttttatt attaagattt gggtagacat 8280
 aaaaaacgga agtatatatg acataactga catagtagac aaggtatcat ggtcaggtga 8340
 ttataaatct ccatcaagga cactagagtt ttcaataata caatcatcat ttgatgtaa 8400
 tttccaacaa atcgatatac caatagctag tacagctctgt ttctatgtag atgagaaaga 8460
 actctttaga ggaatgataa ttaataggtc taaagattca agcagtaatg aaattagttt 8520
 tgatctaaa gatatgggat ttttacttac acaaagtga gtgtcataca attttaaaga 8580
 taagttagtt gaagacatag caaagcaagt atttgctgaa aataggcttt cagttggaat 8640
 aatagcaaag accaatgtca agtatacaaa gatgtttata ggagtaaag gttatgacac 8700
 aataatgagt gcatatacag aagcaagtaa aaagacaaag aaaaagtata tgatagaggc 8760
 caatttagat aagtttaatg ttattgaaaa aggaactggt acattaagtg ttatgtttga 8820
 agagggattt aatattataa ataccactt ttcggagagc atggaaaatg taaaaataa 8880
 agtaatagtg gtagaccagt atggaagcaa gattagcgaa aaaatagata atgaaatttt 8940
 taaggaagta aatgtaataa tgcaaaaagt aattcagcaa caagaaaatc aagatgtaga 9000
 tattgatagc gagtttaatg ggatagaaaa aagctgttct cttaaagggt atggagatgt 9060
 aagttgtata actggtagag gagtaaaagt taaagattct tatacaaagc ttgtaggact 9120
 attttatata gatacagaca aacatacttg gcaaaatgga gaatatcaaa ttgagcttga 9180
 acttaatttt caaaatctta tggatgaaaa gtcagcagga caggatgaac ctaaggaaga 9240
 aagtaattta gggggagaag attatgcagg aggaaaagag tttacagcag aatttacagc 9300

ES 2 622 280 T3

ttactgtcct	agaaaagaag	aaggtggaga	tacagattgt	agaaagaaaa	aacttgaccc	9360
atctaaaaaa	acttgcgctg	ctcctatggt	tggtaaatat	gagcaaactt	attatacaaa	9420
agagttttta	aataaacatc	ctttattaaa	ctatggagat	gaaatacag	taattacagg	9480
agtttctggt	cgtgatggag	tctataaagt	aatgacgta	ggacctgcaa	taactataga	9540
aaagaatgga	acataccata	tagatatttt	atttgaaaat	gttgaagaag	ctagtaaatt	9600
tggaagaaga	aaaggaaaaa	ttattattgg	tggttattct	ggtaatgtat	ctgataaagc	9660
taaaaatagta	atatcagaag	caaaaaaaca	tctaggtaaa	ccttataaat	ggggtggaaa	9720
tggaccaagt	agttttgact	gttctggttt	aatggtctat	tgttttaaaa	aagttaatgt	9780
tagtttgcca	agaacgtcaa	atcaacaatc	taaaaaaggc	aagaaagtag	aacaaaaaaa	9840
tcttcaagca	ggagatttag	tattttttca	taatccagtc	agccatggtg	gattatatat	9900
aggtaatgga	gaatttttac	atgctccaca	aaaaggatg	gtagttaaaa	taagtaagtt	9960
aagtagtaga	agagatttta	acacagctag	gagagtatta	taaaaggatg	gtgatataat	10020
ggctaatacca	ataaatgaat	ttataggaat	aataagagaa	gaaggaaagt	atcataatca	10080
accttctttt	tttattggaa	aaattaaaag	taaattacca	gatttaaaaa	tagagacaaa	10140
taacatcata	ttagaaaaag	aagatatttt	gatagatagt	tggatgattg	atagacagct	10200
agaaacattt	gacacagaaa	caaatcaaga	acaccagcat	gaagtaaaaa	atccttttat	10260
agataacttt	gaatctgggg	atatggtaat	aatgtttaga	ataggcgaaa	aatttgctgt	10320
tgtaagtaga	ttggtgagct	tataatgagt	acaatatttc	cttttatagg	tgtcccagag	10380
gattatattc	tacctaaaac	agaagaattg	ccaatctttc	gtgaagtggc	atgggatttt	10440
gaaaaagatg	aacctatttt	agaaaaagg	gactttaaaa	taattgaaaa	aaaagaagcc	10500
ttaaaagttt	ggatatacaa	gtgtataaag	acaaatagat	atgaacatga	gatatactct	10560
ttagaatatg	ggacagagct	ttcagaacta	ataggacaaa	aatatacaaaa	aggtcttaca	10620
gaaagtgaag	ctagtagatt	cataaaagag	gcccttctaa	taaatccata	tatattagaa	10680
gtaaacgtaa	aaagtgctaa	ctttaacaga	gacatattga	gtgcaaatgt	aaaagtatcc	10740
actatctatg	gggaggtgga	aataaatgta	tagtgaccag	acatatgaag	taataaaaaa	10800
tagaactcct	gaaaatatta	atcttgatat	ttataaagga	gaaggttcct	ttctaaacaa	10860
catggtatct	ggaataatc	tagaactttc	gaagatatat	ctagaacttt	caaagatgca	10920
taaaatggct	tttatacaag	acacatataa	ccagtttctt	gataaaagag	tcaatgaatt	10980
tggtgtatat	agaaagttag	gtacagagtc	aatggagaa	gttgaattta	ttggagagaa	11040
aggtactgta	ataaataatg	gcacaataat	atcatataga	gatttactat	ttgtagtaat	11100
aaaagatgta	actattggta	gtgaagaagg	tgacaatagc	ccagttcaag	ctctggaagt	11160
tggtaaagaa	tataatttac	ctacaaattg	tgaatttaaa	ctagttgata	atatatctgg	11220

ES 2 622 280 T3

agtaacaaag attactaaca caagaagttt tgaagtggt acagatatag agacagatga 11280
 agaactaaaa gaaagatttt ataaaatcca aagaaatcaa gctacaagtg gaaataaagc 11340
 tcactatgaa gaatgggctt tggaagtaga tggagtctat aatgttaag tttatccaag 11400
 atgggatggt ccggaacag ttaaggtctt gatatttggg aaaaataatc aagctgttga 11460
 tacagaaaca attgaaaggt gtcagcaaca tatagatgaa gagaagccta ttggaccaac 11520
 tataacagtt gtgacaccat taccaataga aataagtata agtgcagtaa tgaaactaga 11580
 agatggatat acattagaca atgtaaaaga atctttccta gaaagtataa atacatactt 11640
 tagagatatt agaggagaga taatctatac aaaagtaatg ggaatactta taaatactac 11700
 tgggttacac gatttaagta acctacttat aaatggaagt acagataata taactattaa 11760
 tgaagataaa atacctagtg taacaactgt taattttagt gaggtggaaa atcaatgaag 11820
 ctaattgata aactaccatc atttgataga aattacattg tagaggagat acaaggtgca 11880
 tacgatacag aattaaatat tcttaaagaa gatattgatg atacctttaa ccaattattt 11940
 gttgacactg caacatgggg attagatag tgggaagaca tactctgcat tgaaaaaaaa 12000
 gaacttgatt ttgacacaag acgtagcaat ataaaagcta aaatgagaag cagaggtact 12060
 agtactattg aagttataaa aagtatatgt gaggcataata caaaatcaga aacagatata 12120
 aaagtttata gtgatgaatt tacattcgtt ttgagtttta tagcaaataa ctgtgactat 12180
 aaaactcttt tagattgtag cgatatgatt gaaagagtaa aacctgctca cttattacac 12240
 tatttagaac caataatact agataaaaagt atggtctatt gtggtggagg tatggtatgt 12300
 agtgaagagg taaaagttca tccatacttt gaaccaatta taaaatgtag tgctgttgta 12360
 aactgtggag ctggaatgat aagtagagaa gaaataaagg tttatccttt aagcattaaa 12420
 tgcattgaaa ataattgtaa gattaatata gctattgcaa atgatacagg tgtagaaaat 12480
 gtagtagttt atcctaaatc ggaggtggta taattggaag aaaaatttta tataatatta 12540
 accaaaattg gtagagaaaa aatagcaaat gcaactgcac taggagagct tgttgatta 12600
 accaagtttc aagttggaga tagtaatgga gaatattatg agccaacaga ggaacaaact 12660
 gctttaaaga atgtagtttg ggaaggaaat ataaattctc taagaattga tgaaaaaaaa 12720
 cctaattgga tagttataga gactatttta ccaggaacag ttggtggatt tatgataaga 12780
 gaagctgctg ttctggataa tgagaataat ataatagcta taggtaagta tccagagacg 12840
 tataagccac gtgctgaaga tggcagtatt aaagatttgg ttgtaaaaat gatthtacia 12900
 ttgtccaata ctccaatgt tacattagaa gtagaccga cgttggtttt tgtaactcaa 12960
 aaggatattc aagatttaga tgataagttt gataaaaata taaaagaaat aaaagtaaaa 13020
 attggcgaag aactcttctc tacagaagct aaaaacttat caggagctat aatgaggtta 13080
 gtagaaaaaa ttaaaaatat atctattgat gatgtaatag gaggtcaaat acaactgaa 13140
 ctatctgtat taaaaaatag ttacaataaa ttatctgaaa aagtattaga tatcttaata 13200

ES 2 622 280 T3

tacttagaat tagagtcaga aatagatgta gatgaagctg gatattggta tgatacctta 13260
actaatgcta aaaacataat agctatagaa ggccttaagt tagatttaaa tagaaagtgt 13320
ataactggag aacttggtag tgttacattt aagaatgtgg tgctaccatt taatgcaaat 13380
agagttagat atatacatga aatggataat aactttgttg aaacaaaatc taatagggca 13440
tattcaattg gtcagacaga tataacttta aataaatatt cgtatgaaat aagataatta 13500
ggaggttttt ataatgaaaa gaactaaact acttcaaaga ggtaatttct ttggcgataa 13560
aaatagtgta gttgatgaat ttgatgaagg gtatgataat tatgacttta ttaatttttt 13620
tactggatgt tgtaactata catttggctc aaaaaataat aatatcttgt atggatgtgg 13680
agataatagt aactttcaac ttggattggg agaagacaat acaacaagaa aattatttac 13740
gaaaaacca aatatatcta ccaatattaa aaaagttgca tgtggagaat ctcatgcagt 13800
tatacttact tcagatggag aattacttgt cgcaggtata aatacagatg gtcaaatggg 13860
attgggatta gaaaaagtag ggaaaacagt ttctacattt gagaaggttc cagaaataaa 13920
aggcgtaaag gatattgcat gtggacttca atcaacatat cttttataca atgatggaac 13980
tttatatggt gctggaaata atttgtatgg tcaattaggt ctaggaaacta atggagcatc 14040
tgcaaatgta aatacattta caaaagtaga tgttgacaat gtaaaggctg tattttcata 14100
taataaatca gcttttataa taaagaatga caataaatgc tatttctactg gttttaataa 14160
tcaaggtcaa ctaggtttag gagataagaa taatagagat ttatttagtt tagtttctat 14220
taatgatggt aagactatag cttgtggttc tgaacacact gtgttaatga cgtataataa 14280
tgatatatat ggttggtaa aggaaaaatg ttttgaaat gcacttcaat catcactatt 14340
tactaagata gaagaagtaa atataaaaac tattgcatgt ggtcatggta acactatgct 14400
tatagataac aaaggtactt taaaggttgc tggaaataat gatatatatc agttaggat 14460
agcaaattac tctgagaata tagataatc atttatagat ttaaaaaata ttgtagctaa 14520
gaatattttc attggtttat cacatagcat actaattgat tcaaataatg attcatattg 14580
tacaggagat aatacttatg gacaattagg ttcgtttttt gatgatatgc acattgtaga 14640
atthaagaaa atggatagtg aaaaatatag ttatagtaat tatataaatt taattaaatc 14700
tgaggataaa ttaactttat taaaagaaga aatggaaata aaggatattg aacttccact 14760
agatatacat tctgtaagag atgtcgtttt tagtccttat tgtactctgg ttattttagg 14820
gaatggagat gtatatggtc taggaaataa tagatacaaa ggaatgggtt ctgacttacc 14880
aagtcaatta aatgagttga caaaattaag tatctctaata gtaaagtcta tagtagcatc 14940
aaaaaatatt tctggaggaa tattctacat taaaaatgat gatacttgtt attattctgg 15000
accaaatagt aactcaatag caggtgttct tccttctaata tcagatgtat ttaagaaaat 15060
atctatagat aatgtaaaaa aagttgttat aaatactgat ttatcaaact ggttttcatt 15120

ES 2 622 280 T3

aattgtaact aataataagc aaatatacac ttctggaag agttcaagtt atgттаатгг 15180
acttagtaat gcattaataa gtcaatatac tgagattagc cttagtaatg taactgatgc 15240
ttatagttca tataatgcaa cttttattgt agttgatgaa aaaaaggtat atgcaactgg 15300
tataaataca aattacctgt taggttttag tacttctgat ggatctaатg taaatctagg 15360
tttattaagt gattggтatt atataaatat atcagggtca agttatagta gagtttcatg 15420
cacgaataat attactaaaa ttaataaatat tatcatatat gagtatgtaa ctgtattttg 15480
tacaacatt ggatcttttc taactggata ccatggтact tcatggacaa aaccaactga 15540
ttcaagctat agagttcaat atcagggaaт ttcatatgca ggatatcttg attcttatat 15600
atataattat tатoctacaa gatgtacaca atcatcatct tctacaactt ttgcttattt 15660
atataatggg gaatcgтcaa gтаatttaaa aaatgтcaat ccagataaтт tacttatttc 15720
tgгaggттca tcttatatac atcaatatgg aaggaattat cttaacaatc aatcatctaa 15780
taatattgca gcatctaata taaattcagg tctattacc tctgataaag ccatattttt 15840
atataaagct ctattgtatt tatcttctaa cacgctatat ggttttggga atatatctga 15900
aagtgcaaaa gaactagatg tttcagatac acaagatgga tataatgcca cтаattataa 15960
aaagгтаатg aaaaataaa aaaatatatt tатacctcct tatgatttaa gtagagataa 16020
aactagattt gcaatattaa ctgataagag cttatttata tgtggatata actctaaggg 16080
tacgcatggt atatcagtta atagtagttt aaatttaaat aataagataa attacaataa 16140
aaagaatagc agtagtgaaa tatcttctaa tatacaagaa atatatagcc attcaaagtc 16200
tacatattta ttaactaata атаатггct ttacagtgtt ggтттаатg atgtaggtca 16260
attaggagtg ggagatgaga taaatagaaa ggtatttact aaaataaata ttgataaatat 16320
aaaatctata aatgтааата gatttactga caatagтаа catgcatttg cgataaaaaa 16380
tgataatacc tgttatgctg ttggтттаа таattctggt cagttaggaa taggagataa 16440
tgтааатаgaa aatataттta ctaaaataaa tgttgaaaat gтаааатаг tagctgtata 16500
tgгaaacaca tctctattat таactaatga tggтctттta tatggagcag gтааатагг 16560
aaaaggacag ttaggattgg gtgatactac aagtaggaat atatttacac gtataoctat 16620
aaatggтgtt agagatgtat atctatgтаa tgatgtatca atcattgtta aaaatgataa 16680
tacatgctat gtatgtggac ttgтааатгг ctattttggg tttactgaag gaagtataag 16740
tacatttaca aaaataaata ttgagaatgt aaaatctgтт gtgacagcag gaagtgaagc 16800
tacatttttt атааааатг аcaatatgat ttatactaca gggaaaaaag agagggtatt 16860
cttttcaaca gagactaatg atataaaggg gatacgagta атааатаата ttataaatgc 16920
aaaaaaaata gtagttaatg gatatacttc agccatttta acaaatgaca атааactatt 16980
tgttgagggt cttagtggat atggaagtat agcaataat aataatacaa atagtgtgga 17040
agatgtтааа gatgттттг таacagctaa таatacactt tatatagata атааатааа 17100

ES 2 622 280 T3

tttgatatca	tcaggtagag	atacttatgg	tatatctgat	gaatcttata	gggatatgtc	17160
agttccatat	tataaagtat	ctataaagaa	agatgttgat	actgtatfff	ctagttacaa	17220
tactatattt	attaaagata	tatatggaaa	atfctattct	tcaacaagag	ataatagata	17280
taatcattta	ggtattcacc	atagatatga	taatgataaa	aatgaagctc	ttgaaggffc	17340
cctacattca	tattttaaaa	cagataaacac	atcagataaa	atagttfctta	ataagaaaaa	17400
tgaaaagcta	gtaatgttta	atgataagta	tataaaaaaca	aataataagt	atataaatta	17460
taaaaacata	tttaaagata	atfcttaagta	tacttcaata	atattgccat	ttgaggatc	17520
tgatattgat	atatcaaaaa	cacattcatt	ggctgttgct	aaggatggca	agttatatgg	17580
aataggaagt	aattcatata	aagaaattaa	tcaaacctct	gaagatatag	aattattaac	17640
tcttactgaa	gtaaatatat	cagatgtcaa	aaaagttgct	tgtggagata	actactccta	17700
tattattaag	acagataata	ctctatggct	atatggaaag	aatactgagt	accaattggg	17760
agttggccac	aataatgatg	taagagagtt	acaaaaggtt	actggattac	cttctgttaa	17820
agatataagt	atatataact	caatgacact	tgtfcttaact	aatgagggag	agttgtacgc	17880
tcaagggtac	aatacaaatg	gattatfctg	actaggagaa	agtgaaaaag	ataagataat	17940
aagaactfff	actaaagtat	taactaatgt	taaagaaatt	aagtcacata	atgatgacca	18000
catactagta	attaaaaatg	ataatagtct	atggataact	ggtaaaaata	aatctatgta	18060
taaaaatct	atatcaatta	ctgattfctta	tgaattfct	aaaataccta	ttctgaaca	18120
tctaaatgat	atfcttagata	tagagctfct	agatgataca	atatacatga	taacaaaagt	18180
agatacaagt	aaagcatcta	tagaaatagt	tgaaaaatca	atatctcaag	tgagagttgt	18240
agtacaagac	cctaataatg	ttatagaaaa	acttgaaatg	tttataaatg	atgaattaat	18300
atctactaag	actaatfctg	aaataaatag	cattatattt	gagataccac	aaaataaaat	18360
agtattagga	gaaaataaga	tactgattaa	agccagtagt	cctacaggcg	atfctatfct	18420
aagtatgttt	atattfctaaat	cagaaaacagg	gctfctaaagta	aaaaaggatt	ctatfctfctaaat	18480
gataaacaat	aaagtatatt	caatcataaa	cattactgaa	aataaactg	acttaatagt	18540
aacattaaat	gagggattaa	aggatgatat	gatggaaaac	aatcctatat	atcaattaat	18600
aaataaaact	aaagttcaag	taaaaataaa	taaatctgac	ttattcaag	acatgaaact	18660
agttgaaatc	aaaaaatcag	actcaagtta	ccaagaaatc	tatgaattag	aagaagccaa	18720
cataaaaagt	gctcagccta	aatcatagt	agaaaaagga	gataaatgga	cagctataaa	18780
acgtccatct	atgattfctta	gatatgatgc	tgaaaacaac	gagccacaag	ctfctaaatgg	18840
aggtgtgaaa	atgattfctta	ttcgataaaa	ataaaataga	acaaatcaaa	caaggtagaa	18900
aagtagaaat	gcagtataaa	gacattfctcag	acataagtat	aggtcaagta	aagcaagatg	18960
atgatataac	aaataatfff	atagcaaatg	tagaaatata	tgagatgttg	ttaaatcaaa	19020

ES 2 622 280 T3

gttctgtcaa tgaagcaagt aatataagca cttttagtgt aagaaaatct ggaggtgaga 19080
 gtggaatggt agaagtatat gtagctttaa ttttaagagg caaaaaaaca atagaagaag 19140
 taccagcagt aattagagag caagttagaa ttagatgtaa agaattagaa ataccagttg 19200
 aatagtaaat ttagaataac tatgtattag ttatTTTTTT tatgtaaagt acaaggtcct 19260
 aactttaata agtaagcctt gtacttattt tttgttacat tagaacttgt atatatattt 19320
 attattttatt caatctataa attacaccta caatttaaag tacagaagat taaattgata 19380
 atcctgaaaa tataatattg catgatgtaa gaatacaaca aaaattaaag ctataagtat 19440
 aaaaaattta gacaatagga ggctataatg gataaattaa taaccgaatt gagtagtcta 19500
 ggggcaatag gtatactatg tgctctatta tttaaaaata ctatgcagga gaaaaaagaa 19560
 gatagagaca tgtataaaaa aactgtagaa aattttatag aattatctac acaacaacaa 19620
 gaaataaaca aaaatatact tgttcaaag ggtataatga aacagatgt agaggaaatt 19680
 aaggaagatg ttactgatat aaaaggtatg ttacaaaatg gtgtataaca tgaaagagta 19740
 gcaccagatt atatatgtt agggaaagat aaagtagtat thtagatagt tcactatttt 19800
 attgagaagg atttaattt taaaatatta attaaaaaaaa gtaataaaaa taacatataa 19860
 aaattaaaaa aggagttaag cttaaatttg aggcgcg 19897

<210> 62
 <211> 168
 <212> PRT
 <213> Clostridium difficile

5

<400> 62
 Met Asn Asn Leu Asp Lys Leu Phe Glu Leu Ala Ser Gln Glu Glu Ile
 1 5 10 15
 Ile Ile His Tyr Thr Thr Tyr Ile Ala Gly Asp Leu Glu Gly Leu Tyr
 20 25 30
 Ile Asn Lys His Gly Ile Lys Ile Ile Ser Leu Leu Ser Asn Leu Lys
 35 40 45
 Gln Asn Ser Lys Lys Leu Thr Ser Ile Leu Ala Glu Glu Leu Gly His
 50 55 60
 His Phe Thr Ser Leu Gly Tyr Tyr Val Ser Ser Tyr Asn Asp Tyr Tyr
 65 70 75 80
 Thr Lys Ile Ile Ile Asp Lys Cys Glu Asn Lys Ala Leu Lys Trp Ala
 85 90 95
 Cys Glu Phe Leu Ile Thr Glu Glu Asp Ile Ile Asn Ile Ile Asn Ser
 100 105 110

10

ES 2 622 280 T3

Gly Ile Thr Cys Val Tyr Glu Met Ala Asp Ile Leu Asn Val Asp Ile
 115 120 125

Thr Phe Phe Gln Lys Arg Leu Glu Phe Leu Ser Leu Lys Lys Gln Ser
 130 135 140

Leu Gln Leu Gly Asn Asn Lys Tyr Leu Ile Leu Thr Asn Leu Pro Tyr
 145 150 155 160

Phe Tyr Ile Phe Asp Pro Ile Ser
 165

<210> 63
 <211> 106
 <212> PRT
 <213> Clostridium difficile

5

<400> 63
 Met Phe Ala Lys Arg Leu Arg Glu Leu Arg Lys Glu Phe Gly Leu Thr
 1 5 10 15

Gln Arg Glu Leu Gly Glu Lys Val Gly Val Ser Gln Arg Val Leu Gly
 20 25 30

Tyr Tyr Glu Thr Glu Asn Arg Phe Pro Asp Glu His Ile Leu Asn Lys
 35 40 45

Leu Ala Asp Val Phe Asn Val Ser Val Asp Tyr Leu Leu Gly Arg Thr
 50 55 60

Leu Val Lys Glu Asn Ile Asp Thr Val Ala Ala His Arg Lys Asn Pro
 65 70 75 80

His Glu Glu Leu Pro Glu Glu Ala Gln Glu Gln Leu Asn Asp Tyr Ile
 85 90 95

Glu Phe Leu Leu Asn Lys Tyr Lys Lys Lys
 100 105

10

<210> 64
 <211> 65
 <212> PRT
 <213> Clostridium difficile

15

<400> 64
 Met Phe Lys Asn Asn Leu Lys Tyr Tyr Arg Lys Cys Lys Gly Met Thr
 1 5 10 15

Gln Ile Gln Leu Ala Arg Lys Ala Gly Ile Thr Asn Asp Tyr Ile Ser
 20 25 30

Gln Ile Glu Arg Gly Ile Lys Asn Pro Gly Leu Leu Met Ala Lys Lys

ES 2 622 280 T3

35

40

45

Ile Ser Ser Ile Leu Glu Gln Asn Ile Glu Glu Val Phe Phe Ile Gln
50 55 60

Leu
65

<210> 65

<211> 146

5 <212> PRT

<213> Clostridium difficile

<400> 65

Met Glu Asn Lys Lys Asp Ile Leu Phe Lys Glu Thr Asp Lys Arg Leu
1 5 10 15

His Asn Tyr Lys Tyr Leu Asp Ile Lys Ile Lys Asn Ile Asn Leu Asp
20 25 30

Ile Lys Arg Cys Glu Asn Glu Tyr Ser Gly Cys Gly Ala Met Val Tyr
35 40 45

Thr Glu Lys Thr Ser Asn Thr Tyr Asn Ile Ser Ser Ser Val Glu Asn
50 55 60

Glu Val Leu Lys Arg Glu Glu Arg Leu Arg Lys Leu Lys Met Glu Lys
65 70 75 80

Glu Asp Ile Glu Ile Glu Lys Glu Lys Ile Glu Asn Ala Leu Thr Cys
85 90 95

Leu Asn Asp Ile Glu Met Glu Phe Phe Asn Leu Phe Tyr Asn Ser Lys
100 105 110

Thr Lys Asn Asn Met Thr Tyr Ile Ser Met Lys Leu His Leu Asp Arg
115 120 125

Thr Ser Cys Tyr Asn Leu Lys Lys Lys Met Ile Phe Lys Leu Ser Glu
130 135 140

Ile Leu
145

10

<210> 66

<211> 147

<212> PRT

<213> Clostridium difficile

15

<400> 66

Leu Leu Lys Tyr Lys Glu Ile Leu Glu Thr Ile Ile Glu Ile Leu Lys
1 5 10 15

ES 2 622 280 T3

Lys Asn Phe Thr Glu Ser Ile Phe Ile Asp Asp Glu Ser Val Gln Gly
 20 25 30

Ser Glu Gly Ser Cys Phe Phe Val Ser Ile Leu Ser Val Ile Cys Thr
 35 40 45

Pro Val Met Leu Asn Thr Asn Asn Lys Asp Ile Val Ile Ser Ile Lys
 50 55 60

Tyr Leu Pro Lys Pro Gln Ser Lys Ser Ile Arg Met Tyr Glu Ile Ser
 65 70 75 80

Asp Glu Leu Asn Lys Leu Phe Asn Arg Asn Ile Lys Val Thr Asp Arg
 85 90 95

Lys Leu Asn Ile Thr Lys Leu Glu Gln Ser Ile Lys Lys Glu Glu Ser
 100 105 110

Ile Tyr Val Leu Asn Phe Thr Phe Thr Leu Asn Tyr Leu Asp Ser Val
 115 120 125

Tyr Glu Glu Asp Val Val Tyr Glu Asn Met Lys Glu Ile Asn Leu Asn
 130 135 140

Leu Gly Glu
 145

<210> 67

<211> 354

5 <212> PRT

<213> Clostridium difficile

<400> 67

Met Ala Ile Gly Leu Pro Ser Ile Asn Ile Ser Phe Lys Glu Leu Ala
 1 5 10 15

Thr Thr Val Lys Glu Arg Ser Ala Arg Gly Ile Ile Ala Met Val Leu
 20 25 30

Lys Asp Ala Lys Ala Leu Gly Leu Asn Glu Ile His Glu Lys Glu Asp
 35 40 45

Ile Pro Val Asp Leu Ser Ala Glu Asn Lys Glu Tyr Ile Asn Leu Ala
 50 55 60

Leu Met Gly Asn Val Asn Thr Pro Asn Lys Leu Leu Val Tyr Val Ile
 65 70 75 80

Glu Gly Glu Ala Asp Ile Gln Thr Ala Leu Asp Phe Leu Glu Thr Lys

10

ES 2 622 280 T3

85 90 95

Glu Phe Asn Tyr Leu Cys Met Pro Lys Ala Val Glu Ala Asp Lys Thr
 100 105 110

Ala Ile Lys Asn Trp Ile Ile Lys Leu Arg Asp Ile Asp Lys Val Lys
 115 120 125

Val Lys Ala Val Leu Gly Lys Val Val Gly Asn His Glu Gly Ile Ile
 130 135 140

Asn Phe Thr Thr Glu Asp Val Leu Val Gly Glu Lys Lys Tyr Ser Val
 145 150 155 160

Asp Glu Phe Thr Ser Arg Val Ala Gly Leu Ile Ala Gly Thr Pro Leu
 165 170 175

Ser Gln Ser Val Thr Tyr Thr Lys Leu Ser Asp Val Val Asp Ile Pro
 180 185 190

Lys Met Thr Lys Val Asp Ala Glu Ser Arg Val Asn Lys Gly Glu Leu
 195 200 205

Ile Leu Ile Lys Glu Ala Gly Ala Ile Arg Ile Ala Arg Gly Val Asn
 210 215 220

Ser Leu Thr Glu Leu Thr Ala Glu Lys Gly Glu Met Phe Gln Lys Ile
 225 230 235 240

Lys Ile Val Asp Thr Leu Asp Ile Ile His Ser Asp Ile Arg Lys Val
 245 250 255

Ile Ile Asp Asp Tyr Ile Gly Lys Val Thr Asn Ser Tyr Asp Asn Lys
 260 265 270

Cys Leu Leu Ile Val Ala Ile Lys Ser Tyr Leu Glu Glu Leu Glu Lys
 275 280 285

Ser Ala Leu Ile Glu Ser Asp Ser Thr Val Glu Ile Asp Phe Glu Ala
 290 295 300

Gln Lys Ser Tyr Leu Lys Ser Lys Gly Val Asp Leu Ser Tyr Met Thr
 305 310 315 320

Leu Gln Glu Ile Lys Glu Ala Asn Thr Gly Ser Lys Val Phe Leu Lys
 325 330 335

Ala Lys Ile Lys Val Leu Asp Ala Met Glu Asp Ile Asp Leu Ser Ile
 340 345 350

Glu Ile

- 5 <210> 68
- <211> 142
- <212> PRT
- <213> Clostridium difficile

ES 2 622 280 T3

<400> 68

Met Ala Asn Met Glu Ala Arg Asn Val Met Ser Gly Thr Trp Gly Glu
1 5 10 15

Leu Trp Leu Asp Gly Asn Lys Val Ala Glu Val Lys Lys Phe Gln Ala
20 25 30

Lys Met Glu Phe Thr Lys Glu Asp Ile Ile Ile Ala Gly Gln Met Gly
35 40 45

Thr Asp Thr Lys Tyr Met Gly Tyr Lys Gly Lys Gly Ser Ile Thr Leu
50 55 60

Tyr His Val Ser Ser Arg Met His Lys Leu Ile Gly Glu Lys Ile Lys
65 70 75 80

Arg Gly Ser Glu Pro Arg Phe Val Ala Ile Ser Lys Leu Asn Asp Pro
85 90 95

Asp Ser Tyr Gly Ala Glu Arg Ile Ala Val Lys Asn Ile Ala Phe Asp
100 105 110

Asp Leu Thr Leu Ala Asp Trp Glu Val Gly Val Lys Gly Glu Ile Glu
115 120 125

Ala Pro Phe Thr Phe Thr Glu Tyr Asp Phe Leu Asp Ile Ile
130 135 140

<210> 69

5 <211> 148

<212> PRT

<213> Clostridium difficile

<400> 69

Met Ser Glu Asn Gly Leu Ser Lys Asn Ile Asn Ile Val Asp Leu Leu
1 5 10 15

Leu Asn Ser Asp Thr Glu Asn Leu Glu Arg Pro Ser Thr Ile Val Glu
20 25 30

10 Leu Lys Arg Leu Ser Thr Ile Phe Gly Gln Glu Phe Lys Val Met Cys
35 40 45

ES 2 622 280 T3

Arg Ala Leu Thr Ile Ser Lys Asp Glu Glu Ile Gln Asn Thr Cys Leu
50 55 60

Lys Ile Asp Glu Asn Met Lys Thr Asp Ile Asp Leu Pro Glu Met Gln
65 70 75 80

Met Leu Thr Ile Ile Glu Gly Val Cys Asp Leu Asp Gly Lys Leu Leu
85 90 95

Phe Lys Asn Lys Glu Leu Met Asp Lys Phe Lys Ala Pro Thr Pro Lys
100 105 110

Glu Leu Ala Arg Lys Leu Leu Leu Pro Gly Glu Ile Thr Asn Leu Tyr
115 120 125

Arg Ile Leu Gln Asp Val Met Gly Tyr Gly Lys Asn Ala Val Ile Glu
130 135 140

Glu Val Lys Asn
145

<210> 70
<211> 55
5 <212> PRT
<213> Clostridium difficile

<400> 70
Met Tyr Tyr Tyr Trp Lys Lys Lys Gly Ile Arg Pro Ser Leu Phe Tyr
1 5 10 15

Ala Met Asp Lys Gly Glu Leu Lys Leu Ile Glu Ala Phe Phe Ala Leu
20 25 30

Glu Ile Glu Glu Glu Val Glu Lys Met Lys His Gly Tyr Gly Val Cys
35 40 45

Pro Leu Thr Gly Gly Gly Met
50 55

10 <210> 71
<211> 817
<212> PRT
<213> Clostridium difficile

15 <400> 71
Met Gly Asn Val Arg Glu Glu Gly Ile Asn Met Tyr Leu Thr Asp Asn
1 5 10 15

Tyr Thr Pro Lys Met Asn Gln Ile Ile Ser Val Thr Asp Asn Phe Arg
20 25 30

ES 2 622 280 T3

Arg Ala Thr Val Ala Val Ser Leu Ser Thr Asn Val Met Ala Ser Ser
 35 40 45

Ile Lys Asn Ser Ile Gly Ser Ala Ser Asn Arg Val Asn Ser Leu Asn
 50 55 60

Ser Ser Leu Arg Lys Val Gln Thr Thr Ala Ser Ser Val Ser Ser Thr
 65 70 75 80

Met Thr Lys Leu Ser Ser Ser Ile Asn Ala Val Ser Gly Val Ile Gly
 85 90 95

Ser Leu Asn Gly Ser Ile Met Arg Leu Ala Ile Thr Ile Ala Met Ile
 100 105 110

Ile Asp Tyr Phe Asn Lys Leu Ile Gln Lys Lys Asn Glu Phe Asn Ser
 115 120 125

Asn Ile Met Ile Ile Leu Ile Phe Lys Ala Lys Ser Asp Glu Val Glu
 130 135 140

Lys Thr Lys Asn Lys Leu Leu Gly Asn Leu Lys Lys Ile Gly Gly Lys
 145 150 155 160

Ile Trp Asn Ile Val Ile Lys Ala Lys Asp Met Thr Lys Arg Val Ile
 165 170 175

Ser Ser Ile Leu Gly Lys Leu Lys Gln Val Glu Lys Arg Pro Tyr Gln
 180 185 190

Gly Ser Ile Asn Leu Lys Asp Met Val Ser Ser Ala Met Gly Arg Ile
 195 200 205

Leu Pro Lys Leu Met Leu Phe Lys Asn Thr Phe Trp Ser Gly Val Ile
 210 215 220

Ala Ile Lys Asp Met Ala Ser Gly Ile Ile Ser Lys Val Phe Pro Lys
 225 230 235 240

Leu Arg Leu Phe Ala Gly Lys Val Trp Ser Gly Ala Ile Ala Val Lys
 245 250 255

Asp Met Ala Ser Gly Ile Leu Gly Ser Ile Lys Gly Lys Ile Ser Asp
 260 265 270

Leu Thr Asn Gly Ala Thr Ile Gly Val Ala Val Lys Lys Gly Val Asp
 275 280 285

Leu Leu Gly Gln Glu Gln Asn Gln Lys Val Val Leu Glu Ser Val Met

ES 2 622 280 T3

Asn Ile Asp Phe Thr Ala Leu Ile Asp Ser Leu Leu Pro Val Val Asn
565 570 575

Leu Val Asn Asn Leu Leu Asp Ser Ile Asn Asn Lys Ser Pro Ile Ala
580 585 590

Gln Gly Leu Ile Ser Ile Phe Gly Thr Ile Val Thr Thr Ala Phe Gln
595 600 605

Leu Ile Gly Pro Val Val Glu Ala Val Ser Pro Ile Ile Thr Arg Ile
610 615 620

Phe Thr Phe Leu Gly Glu Tyr Ala Pro Gln Ile Asn Asn Phe Ile Glu
625 630 635 640

Thr Leu Gly Val Ile Trp Lys Thr Val Trp Glu Thr Leu Gly Pro Leu
645 650 655

Leu Glu Thr Gly Trp Lys Ile Ile Glu Pro Ile Leu Gly Ala Phe Phe
660 665 670

Asn Ile Leu Asp Lys Val Cys Lys Ile Val Lys Asp Ile Cys Lys Trp
675 680 685

Trp Gln Thr Met Ile Asn Lys Ile Lys Asn Gly Ser Ile Thr Gly Thr
690 695 700

Val Leu Asn Leu Val Glu Lys Ser Lys Lys Asn Tyr Lys Asp Asn Pro
705 710 715 720

Tyr Ala Gly Thr Lys Ala Gly Asp Ser Gly Lys Ala Tyr Ser Gly Lys
725 730 735

Lys Gly Asn Asn Ala Phe Gly Leu Asn Tyr Val Pro Tyr Asn Asp Tyr
740 745 750

Gln Thr Arg Leu His Glu Gly Glu Met Val Leu Thr Lys Gln Glu Ala
755 760 765

Asn Gln Tyr Arg Ser Arg Lys Asn Gly Gly Asn Ile Asn Ile Ala Lys
770 775 780

Leu Ala Asp Thr Ile Val Ile Arg Glu Glu Ala Asp Ile Glu Lys Ile
785 790 795 800

Thr Ser Lys Leu Val Ala Ser Ile Gln Leu Ala Gln Leu Gly Gly Val
805 810 815

Leu

- 5 <210> 72
- <211> 140
- <212> PRT
- <213> Clostridium difficile

ES 2 622 280 T3

<400> 72

Met Glu Met Trp Leu Arg Gln Ala Glu Asp Arg Phe Arg Phe Pro Val
1 5 10 15

Phe Pro Ser Ser Phe Ser Ile Asn Gly Lys Ala Ala Val Asn Ser Ser
20 25 30

Ser Ile Leu Lys Ile Gly Glu Ile Ala Thr Phe Gly Gly Val Ala Leu
35 40 45

Lys Ser Ile Ser Ile Ser Ser Phe Phe Pro Asn Lys Asp Tyr Thr Phe
50 55 60

Cys Asp Tyr Thr Gly Phe Pro Ser Pro Tyr Asp Cys Val Asn Lys Ile
65 70 75 80

Glu Lys Trp Met Lys Glu Gly Phe Ile Leu Arg Phe Thr Ile Thr Glu
85 90 95

Thr Asn Ile Asn Met Glu Val Ile Ile Glu Gly Phe Ser Tyr Glu Glu
100 105 110

Arg Asp Gly Thr Arg Asp Val Tyr Phe Thr Leu Asp Leu Lys Glu Tyr
115 120 125

Lys Arg Ile Lys Ile Pro Lys Val Thr Pro Lys Gln
130 135 140

<210> 73

5 <211> 509

<212> PRT

<213> Clostridium difficile

<400> 73

Met Ile Ile Asn Arg Ser Lys Asp Ser Ser Ser Asn Glu Ile Ser Phe
1 5 10 15

Val Ser Lys Asp Met Gly Phe Leu Leu Thr Gln Ser Glu Val Ser Tyr
20 25 30

10 Asn Phe Lys Asp Lys Leu Val Glu Asp Ile Ala Lys Gln Val Phe Ala
35 40 45

ES 2 622 280 T3

Glu Asn Arg Leu Ser Val Gly Ile Ile Ala Lys Thr Asn Val Lys Tyr
 50 55 60
 Thr Lys Met Phe Ile Gly Val Asn Gly Tyr Asp Thr Ile Met Ser Ala
 65 70 75 80
 Tyr Thr Glu Ala Ser Lys Lys Thr Lys Lys Lys Tyr Met Ile Glu Ala
 85 90 95
 Asn Leu Asp Lys Phe Asn Val Ile Glu Lys Gly Thr Val Thr Leu Ser
 100 105 110
 Val Met Phe Glu Glu Gly Phe Asn Ile Ile Asn Thr Thr Phe Ser Glu
 115 120 125
 Ser Met Glu Asn Val Lys Asn Lys Val Ile Val Val Asp Gln Tyr Gly
 130 135 140
 Ser Lys Ile Ser Glu Lys Ile Asp Asn Glu Ile Phe Lys Glu Val Asn
 145 150 155 160
 Val Ile Met Gln Lys Val Ile Gln Gln Gln Glu Asn Gln Asp Val Asp
 165 170 175
 Ile Asp Ser Glu Phe Asn Gly Ile Glu Lys Ser Cys Ser Leu Lys Gly
 180 185 190
 Tyr Gly Asp Val Ser Cys Ile Thr Gly Arg Gly Val Lys Val Lys Asp
 195 200 205
 Ser Tyr Thr Lys Leu Val Gly Leu Phe Tyr Ile Asp Thr Asp Lys His
 210 215 220
 Thr Trp Gln Asn Gly Glu Tyr Gln Ile Glu Leu Glu Leu Asn Phe Gln
 225 230 235 240
 Asn Leu Met Asp Glu Lys Ser Ala Gly Gln Asp Glu Pro Lys Glu Glu
 245 250 255
 Ser Asn Leu Gly Gly Glu Asp Tyr Ala Gly Gly Lys Glu Phe Thr Ala
 260 265 270
 Glu Phe Thr Ala Tyr Cys Pro Arg Lys Glu Glu Gly Gly Asp Thr Asp
 275 280 285
 Cys Arg Lys Lys Lys Leu Asp Pro Ser Lys Lys Thr Cys Ala Ala Pro
 290 295 300
 Met Val Gly Lys Tyr Glu Gln Thr Tyr Tyr Thr Lys Glu Phe Leu Asn

ES 2 622 280 T3

Leu Pro Asp Leu Lys Ile Glu Thr Asn Asn Ile Ile Leu Glu Lys Glu
 35 40 45

Asp Ile Leu Ile Asp Ser Trp Met Ile Asp Arg Gln Leu Glu Thr Phe
 50 55 60

Asp Thr Glu Thr Asn Gln Glu His Gln His Glu Val Lys Asn Pro Phe
 65 70 75 80

Ile Asp Asn Phe Glu Ser Gly Asp Met Val Ile Met Phe Arg Ile Gly
 85 90 95

Glu Lys Phe Ala Val Val Ser Lys Leu Val Ser Leu
 100 105

<210> 75
 <211> 142
 5 <212> PRT
 <213> Clostridium difficile

<400> 75
 Met Ser Thr Ile Phe Pro Phe Ile Gly Val Pro Glu Asp Tyr Ile Leu
 1 5 10 15

Pro Lys Thr Glu Glu Leu Pro Ile Phe Arg Glu Val Ala Trp Asp Phe
 20 25 30

Glu Lys Asp Glu Pro Ile Leu Glu Lys Gly Asp Phe Lys Ile Ile Glu
 35 40 45

Lys Lys Glu Ala Leu Lys Val Trp Ile Tyr Lys Cys Ile Lys Thr Asn
 50 55 60

Arg Tyr Glu His Glu Ile Tyr Ser Leu Glu Tyr Gly Thr Glu Leu Ser
 65 70 75 80

Glu Leu Ile Gly Gln Lys Tyr Thr Lys Gly Leu Thr Glu Ser Glu Ala
 85 90 95

Ser Arg Phe Ile Lys Glu Ala Leu Leu Ile Asn Pro Tyr Ile Leu Glu
 100 105 110

Val Asn Val Lys Ser Ala Asn Phe Asn Arg Asp Ile Leu Ser Ala Asn
 115 120 125

Val Lys Val Ser Thr Ile Tyr Gly Glu Val Glu Ile Asn Val
 130 135 140

10 <210> 76
 <211> 350
 <212> PRT
 <213> Clostridium difficile

15 <400> 76

ES 2 622 280 T3

Met Tyr Ser Asp Gln Thr Tyr Glu Val Ile Lys Asn Arg Thr Leu Glu
1 5 10 15

Asn Ile Asn Leu Asp Ile Tyr Lys Gly Glu Gly Ser Phe Leu Asn Asn
20 25 30

Met Val Ser Gly Asn Asn Leu Glu Leu Ser Lys Ile Tyr Leu Glu Leu
35 40 45

Ser Lys Met His Lys Met Ala Phe Ile Gln Asp Thr Tyr Asn Gln Phe
50 55 60

Leu Asp Lys Arg Val Asn Glu Phe Gly Val Tyr Arg Lys Leu Gly Thr
65 70 75 80

Glu Ser Asn Gly Glu Val Glu Phe Ile Gly Glu Lys Gly Thr Val Ile
85 90 95

Asn Asn Gly Thr Ile Ile Ser Tyr Arg Asp Leu Leu Phe Val Val Ile
100 105 110

Lys Asp Val Thr Ile Gly Ser Glu Glu Gly Asp Asn Ser Pro Val Gln
115 120 125

Ala Leu Glu Val Gly Lys Lys Tyr Asn Leu Pro Thr Asn Cys Glu Phe
130 135 140

Lys Leu Val Asp Asn Ile Ser Gly Val Thr Lys Ile Thr Asn Thr Arg
145 150 155 160

Ser Phe Glu Gly Gly Thr Asp Ile Glu Thr Asp Glu Glu Leu Lys Glu
165 170 175

Arg Phe Tyr Lys Ile Gln Arg Asn Gln Ala Thr Ser Gly Asn Lys Ala
180 185 190

His Tyr Glu Glu Trp Ala Leu Glu Val Asp Gly Val Tyr Asn Val Lys
195 200 205

Val Tyr Pro Arg Trp Asp Gly Pro Gly Thr Val Lys Val Leu Ile Phe
210 215 220

Gly Lys Asn Asn Gln Ala Val Asp Thr Glu Thr Ile Glu Arg Cys Gln
225 230 235 240

Gln His Ile Asp Glu Glu Lys Pro Ile Gly Pro Thr Ile Thr Val Val
245 250 255

ES 2 622 280 T3

Thr Pro Leu Pro Ile Glu Ile Ser Ile Ser Ala Val Met Lys Leu Glu
260 265 270

Asp Gly Tyr Thr Leu Asp Asn Val Lys Glu Ser Phe Leu Glu Ser Ile
275 280 285

Asn Thr Tyr Phe Arg Asp Ile Arg Gly Glu Ile Ile Tyr Thr Lys Val
290 295 300

Met Gly Ile Leu Ile Asn Thr Thr Gly Val His Asp Leu Ser Asn Leu
305 310 315 320

Leu Ile Asn Gly Ser Thr Asp Asn Ile Thr Ile Asn Glu Asp Lys Ile
325 330 335

Pro Ser Val Thr Thr Val Asn Phe Ser Glu Val Glu Asn Gln
340 345 350

<210> 77

<211> 232

5 <212> PRT

<213> Clostridium difficile

<400> 77

Met Lys Leu Ile Asp Lys Leu Pro Ser Phe Asp Arg Asn Tyr Ile Val
1 5 10 15

Glu Glu Ile Gln Gly Ala Tyr Asp Thr Glu Leu Asn Ile Leu Lys Glu
20 25 30

Asp Ile Asp Asp Thr Phe Asn Gln Leu Phe Val Asp Thr Ala Thr Trp
35 40 45

Gly Leu Asp Met Trp Glu Asp Ile Leu Cys Ile Glu Lys Lys Glu Leu
50 55 60

Asp Phe Asp Thr Arg Arg Ser Asn Ile Lys Ala Lys Met Arg Ser Arg
65 70 75 80

Gly Thr Ser Thr Ile Glu Val Ile Lys Ser Ile Cys Glu Ala Tyr Thr
85 90 95

Lys Ser Glu Thr Asp Ile Lys Val Tyr Ser Asp Glu Phe Thr Phe Val
100 105 110

Leu Ser Phe Ile Ala Asn Asn Cys Asp Tyr Lys Thr Leu Leu Asp Cys
115 120 125

Ser Asp Met Ile Glu Arg Val Lys Pro Ala His Leu Leu His Tyr Leu

10

ES 2 622 280 T3

130

135

140

Glu Pro Ile Ile Leu Asp Lys Ser Met Val Tyr Cys Gly Gly Gly Met
145 150 155 160

Val Cys Ser Glu Glu Val Lys Val His Pro Tyr Phe Glu Pro Ile Ile
165 170 175

Lys Cys Ser Ala Val Val Asn Cys Gly Ala Gly Met Ile Ser Arg Glu
180 185 190

Glu Ile Lys Val Tyr Pro Leu Ser Ile Lys Cys Ile Glu Asn Asn Cys
195 200 205

Lys Ile Asn Ile Ala Ile Ala Asn Asp Thr Gly Val Glu Asn Val Val
210 215 220

Val Tyr Pro Lys Ser Glu Val Val
225 230

<210> 78

<211> 327

5 <212> PRT

<213> Clostridium difficile

<400> 78

Leu Glu Glu Lys Phe Tyr Ile Ile Leu Thr Lys Ile Gly Arg Glu Lys
1 5 10 15

Ile Ala Asn Ala Thr Ala Leu Gly Glu Leu Val Gly Leu Thr Lys Phe
20 25 30

Gln Val Gly Asp Ser Asn Gly Glu Tyr Tyr Glu Pro Thr Glu Glu Gln
35 40 45

Thr Ala Leu Lys Asn Val Val Trp Glu Gly Asn Ile Asn Ser Leu Arg
50 55 60

Ile Asp Glu Lys Asn Pro Asn Trp Ile Val Ile Glu Thr Ile Leu Pro
65 70 75 80

Gly Thr Val Gly Gly Phe Met Ile Arg Glu Ala Ala Val Leu Asp Asn
85 90 95

Glu Asn Asn Ile Ile Ala Ile Gly Lys Tyr Pro Glu Thr Tyr Lys Pro
100 105 110

Arg Ala Glu Asp Gly Ser Ile Lys Asp Leu Val Val Lys Met Ile Leu
115 120 125

10

ES 2 622 280 T3

Gln Leu Ser Asn Thr Ser Asn Val Thr Leu Glu Val Asp Pro Thr Leu
 130 135 140

Val Phe Val Thr Gln Lys Asp Ile Gln Asp Leu Asp Asp Lys Phe Asp
 145 150 155 160

Lys Asn Ile Lys Glu Ile Lys Val Lys Ile Gly Glu Glu Leu Leu Ser
 165 170 175

Thr Glu Ala Lys Asn Leu Ser Gly Ala Ile Asn Glu Val Val Glu Lys
 180 185 190

Ile Lys Asn Ile Ser Ile Asp Asp Val Ile Gly Gly Gln Ile Gln Thr
 195 200 205

Glu Leu Ser Val Leu Lys Asn Ser Tyr Asn Lys Leu Ser Glu Lys Val
 210 215 220

Leu Asp Ile Leu Ile Tyr Leu Glu Leu Glu Ser Glu Ile Asp Val Asp
 225 230 235 240

Glu Ala Gly Tyr Trp Tyr Asp Thr Leu Thr Asn Ala Lys Asn Ile Ile
 245 250 255

Ala Ile Glu Gly Leu Lys Leu Asp Leu Asn Arg Lys Cys Ile Thr Gly
 260 265 270

Glu Leu Gly Ser Val Thr Phe Lys Asn Val Val Leu Pro Phe Asn Ala
 275 280 285

Asn Arg Val Arg Tyr Ile His Glu Met Asp Asn Asn Phe Val Glu Thr
 290 295 300

Lys Ser Asn Arg Ala Tyr Ser Ile Gly Gln Thr Asp Ile Thr Leu Asn
 305 310 315 320

Lys Tyr Ser Tyr Glu Ile Arg
 325

<210> 79
 <211> 98
 5 <212> PRT
 <213> Clostridium difficile

<400> 79
 Met Gln Tyr Lys Asp Ile Ser Asp Ile Ser Ile Gly Gln Val Lys Gln
 1 5 10 15

Asp Asp Asp Ile Thr Asn Asn Phe Ile Ala Asn Val Glu Ile Tyr Glu
 20 25 30

10

ES 2 622 280 T3

Met Leu Leu Asn Gln Ser Ser Val Asn Glu Ala Ser Asn Ile Ser Thr
 35 40 45

Phe Ser Val Arg Lys Ser Gly Gly Glu Ser Gly Met Val Glu Val Tyr
 50 55 60

Val Ala Leu Ile Leu Arg Gly Lys Lys Thr Ile Glu Glu Val Pro Ala
 65 70 75 80

Val Ile Arg Glu Gln Val Arg Ile Arg Cys Lys Glu Leu Glu Ile Pro
 85 90 95

Val Glu

<210> 80

<211> 86

5 <212> PRT

<213> Clostridium difficile

<400> 80

Met Asp Lys Leu Ile Thr Glu Leu Ser Ser Leu Gly Ala Ile Gly Ile
 1 5 10 15

Leu Cys Ala Leu Leu Phe Lys Asn Thr Met Gln Glu Lys Lys Glu Asp
 20 25 30

Arg Asp Met Tyr Lys Lys Thr Val Glu Asn Phe Ile Glu Leu Ser Thr
 35 40 45

Gln Gln Gln Glu Ile Asn Lys Asn Ile Leu Val Gln Met Gly Ile Met
 50 55 60

Lys Thr Asp Val Glu Glu Ile Lys Glu Asp Val Thr Asp Ile Lys Gly
 65 70 75 80

Met Leu Gln Asn Gly Val
 85

10

<210> 81

<211> 1802

<212> PRT

<213> Clostridium difficile

15

<400> 81

Met Lys Gln Asn Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys
 1 5 10 15

Asn Ile Leu Ile Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe
 20 25 30

ES 2 622 280 T3

Val Glu Phe Phe Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser
 35 40 45

Asp Gly Asn Leu Tyr Ala Cys Gly Asp Asn Thr Gly Met Lys Gln Asn
 50 55 60

Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys Asn Ile Leu Ile
 65 70 75 80

Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe Val Glu Phe Phe
 85 90 95

Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser Asp Gly Asn Leu
 100 105 110

Tyr Ala Cys Gly Asp Asn Thr Gly Phe Gln Leu Gly Leu Gly Lys Asp
 115 120 125

Ser Ser Glu Arg Arg Met Phe Ser Lys Val Lys Ile Asp Asn Val Lys
 130 135 140

Tyr Val Ser Cys Gly Ser Lys His Ser Val Ala Val Thr Lys Asp Gly
 145 150 155 160

Phe Ala Tyr Gly Ala Gly Thr Ser Asn Val Gly Gln Leu Gly Val Ile
 165 170 175

Glu Ser Thr Val Tyr Tyr Glu Phe Thr Lys Leu Pro Ile Asp Asp Val
 180 185 190

Lys Thr Val Ala Cys Gly Tyr Asp Phe Thr Phe Val Leu Lys Asn Asp
 195 200 205

Gly Thr Leu Tyr Ser Ala Gly Leu Asn Ser Ser Gly Gln Leu Gly Leu
 210 215 220

Gly Asp Thr Asn Asn Arg Val Thr Phe Thr Lys Val Asn Ile Asp Ser
 225 230 235 240

Val Lys Asp Val Val Thr Tyr Asn Gln Ser Val Phe Ile Ile Lys Met
 245 250 255

Asp Gly Thr Ala His Ala Cys Gly Leu Asn Ser Asn Gly Gln Leu Gly
 260 265 270

Ile Asn Ser Thr Leu Asn Lys Ser Val Phe Asn Lys Ile Glu Gly Met
 275 280 285

Asp Asn Val Lys Gln Ile Ala Cys Gly Ser Ser His Thr Ile Leu Ile

ES 2 622 280 T3

290		295		300																
Lys	Asn	Asp	Gly	Thr	Met	Tyr	Thr	Thr	Gly	Ser	Asn	Gly	Tyr	Gly	Gln					
305					310					315					320					
Leu	Gly	Thr	Gly	Asn	Asn	Asn	Asn	Ser	Ile	Val	Phe	Thr	Leu	Ser	Ser					
				325					330					335						
Ile	Asn	Asn	Val	Lys	Tyr	Ala	Ser	Cys	Gly	Asn	Asn	His	Thr	Met	Ile					
			340					345					350							
Leu	Lys	Tyr	Asp	Asn	Thr	Leu	Phe	Ser	Thr	Gly	Gln	Asn	Asn	Tyr	Gly					
		355					360						365							
Gln	Leu	Ala	Asn	Ala	Asn	Lys	Asp	Val	Ala	Ser	Arg	Asn	Thr	Phe	Val					
	370					375					380									
Lys	Val	Asn	Val	Glu	Asn	Ile	Lys	Asp	Ile	Lys	Cys	Gly	Ser	Gln	Phe					
385					390					395					400					
Asn	Phe	Leu	Ile	Asn	Gly	Ser	Lys	Glu	Ile	Phe	Val	Ser	Gly	Cys	Asn					
				405					410					415						
Leu	Ala	Gly	Gln	Leu	Gly	Ser	Phe	Phe	His	Thr	Thr	Phe	Leu	Tyr	Glu					
			420					425					430							
Phe	Ser	Lys	Val	Gln	Ser	Ser	Asn	Leu	Asp	Asn	Tyr	Ser	Gly	Leu	Leu					
		435					440					445								
Val	Asn	Asp	Asp	Tyr	Leu	Tyr	Val	Thr	Lys	Asp	Asn	Ser	Glu	Phe	Leu					
	450					455					460									
Asn	Val	Lys	Leu	Ser	Asp	Asn	Phe	Gln	Asp	Tyr	Lys	Lys	Ile	Glu	Leu					
465					470					475				480						
Thr	Asp	Asn	Asn	Met	Phe	Ile	Val	Met	Asn	Asp	Gly	Thr	Leu	Tyr	Ala					
				485					490					495						
Cys	Gly	Leu	Asn	Asn	Tyr	Gly	Gln	Leu	Gly	Leu	Gly	Asp	Thr	Val	Asn					
			500					505					510							
Arg	Ser	Val	Met	Thr	Lys	Val	Asp	Ile	Asp	Asn	Val	Leu	Asp	Ile	Lys					
		515					520					525								
Gly	Asn	Gly	Asn	Ser	Thr	Phe	Val	Leu	Lys	Asn	Asn	Gly	Thr	Leu	Tyr					
	530					535						540								
Ser	Cys	Gly	Tyr	Asn	Ser	Ser	Gly	Ile	Leu	Gly	Leu	Lys	Asp	Asn	Thr					
545					550					555				560						

ES 2 622 280 T3

Asn Arg Asn Ile Phe Thr Lys Ile Glu Ile Glu Asn Ile Lys Glu Phe
 565 570 575
 Cys Val Glu Ser Asn Tyr Ile Val Ala Leu Asn His Ser Lys Glu Leu
 580 585 590
 Tyr Gly Trp Gly Asn Gln Ser Tyr Ile Val Tyr Gly Asp Asn Arg Asn
 595 600 605
 Tyr Pro Tyr Lys Asp Thr Arg Val Ser Asn Val Glu Lys Ile Ala Thr
 610 615 620
 Trp Ser Asp Thr Leu Tyr Ile Leu Asp Ser Thr Gly Ala Thr Lys Thr
 625 630 635 640
 Ile Gly Tyr Ser Tyr Asn Gly Ser Gly Tyr Pro Ala Pro Ser Ser
 645 650 655
 Ser Ser Thr Tyr Arg Glu Gly Gly Tyr Ile Asn Lys Asn Thr Ser Tyr
 660 665 670
 Arg Thr Leu Glu Phe Tyr Asn Thr Ser Lys Thr Lys Leu Val Asn Leu
 675 680 685
 Phe Ala Phe Tyr Asn Gly Cys Val Phe Val Asp Glu Asn Gly Leu Ala
 690 695 700
 Tyr Cys Ile Gly Glu Asn Asn Ile Asn Phe Arg Gly Gly Ser Thr Thr
 705 710 715 720
 Asn Glu Asn Asn Ser Leu Arg Phe Ile Asn Asn Ser Gly Val Tyr Tyr
 725 730 735
 Thr Asn Thr Asp Gly Thr Asp Tyr Thr Cys Tyr Gln Trp Thr Tyr Lys
 740 745 750
 Leu Ile Arg Cys Ser Ile Phe Asp Ser Pro Gln Asn Ile Ile Gly Asn
 755 760 765
 Ser Lys Asn Ile Leu Tyr Leu Ser Lys Asn Asn Ser Thr Phe Lys Cys
 770 775 780
 Thr Gly Asn Cys Ile Thr Tyr Gly Ile Asn Ser Gln Asn Trp Tyr Ser
 785 790 795 800
 Tyr Phe Ser Asp Ser Ser Asn Gly Ala Ile Ala Leu Gly Asn Glu Phe
 805 810 815

ES 2 622 280 T3

Ile Leu Lys Asn Tyr Ser Gly Glu Cys Leu Leu Lys Gly Tyr Gly Lys
820 825 830

Ala Thr Asn Gly Glu Phe Gly Asn Ser Thr Asn Ile Ser Ser Ile Ser
835 840 845

Asn Tyr Asp Thr Gly Leu Lys Asp Ile Lys Asp Ile Ile Val Lys Asn
850 855 860

Asn Thr Val Val Val Val Asp Lys Asn Asn Asn Ile Tyr Val Thr Gly
865 870 875 880

Ala Asn Gln Phe Asn Lys Leu Gly Ile Gly Glu Tyr Asn Asn Gln Pro
885 890 895

Ile Arg Lys Phe Thr Asn Ile Thr Glu Gln Ser Asn Ser Phe Ile Phe
900 905 910

Met Asp Asp Ile Lys Glu Ile Thr Thr Ser Arg Asn Thr Met Phe Ile
915 920 925

Val Lys Asn Asp Gly Thr Ala Tyr Ala Thr Gly Asn Asn Ser Ser Gly
930 935 940

Gln Leu Gly Leu Gly Asp Thr Ile Asn Arg Asn Lys Phe Thr Gln Ile
945 950 955 960

Asn Leu Asp Asn Ile Lys Lys Ile Ser Thr Ser Ile Asp Gly Asn Thr
965 970 975

Thr Phe Ala Ile Arg Asn Asp Gly Thr Leu Tyr Ser Thr Gly Leu Asn
980 985 990

Thr Lys Gly Gln Leu Gly Leu Gly Asp Ile Val Asn Arg Asn Thr Phe
995 1000 1005

Thr Lys Val Asn Ile Gln Asn Val Arg Asp Val Val Leu Gly Thr
1010 1015 1020

Thr His Ser His Ala Ile Lys Asp Asp Asn Thr Leu Tyr Ser Cys
1025 1030 1035

Gly Glu Asn Thr His Gly Gln Leu Gly Leu Gly Ser Glu Ser Asn
1040 1045 1050

His Pro Asp Val Leu Thr Phe Thr Val Asn Asn Ile Thr Asn Val
1055 1060 1065

ES 2 622 280 T3

Arg Asp Val Tyr Cys Ser Asp Thr Thr Thr Phe Ile Val Lys Asp
 1070 1075 1080

Thr Asn Ile Ala Tyr Cys Cys Gly Tyr Asn Asn Asn Ser Gln Leu
 1085 1090 1095

Gly Met Gly Asn Thr Thr Asp Gln Tyr Ser Phe Ile Lys Cys Met
 1100 1105 1110

Glu Asn Val Lys Glu Val Ile Pro Asn Glu Ile Asn Thr Tyr Ile
 1115 1120 1125

Ile Thr Ile Tyr Asn Thr Ala Tyr Ser Thr Gly Leu Asn Thr Asp
 1130 1135 1140

Tyr Cys Leu Gly Leu Asn Ser Asn Ser Asn Gln Ser Ser Phe Ser
 1145 1150 1155

Glu Ile Pro Ile Ser Asn Val Val Lys Val Ala Pro Asn Arg Asn
 1160 1165 1170

Asn Ala Val Leu Leu Leu Thr Ser Glu Gly Asp Val Tyr Thr Ala
 1175 1180 1185

Gly Lys Cys Ser Asn Gly Ser Gly Thr Gly Ser Glu Thr Pro Glu
 1190 1195 1200

Lys Ile Lys Lys Ile Ala Ser Lys Ala Lys Asp Ile Gly Met Asn
 1205 1210 1215

Tyr Arg Cys Gly His Tyr Val Ser Asp Asn Gly Asp Leu Tyr Gly
 1220 1225 1230

Thr Gly Phe Asn Asp Cys Gly Gln Leu Gly Val Gly Asn Val Thr
 1235 1240 1245

Lys Arg Asp Thr Phe Ile Lys Thr Asn Thr Arg Val Lys Lys Ile
 1250 1255 1260

Leu Pro Leu Glu Tyr Ala Asn Ile Ala Ile Lys Asp Thr Asn Asp
 1265 1270 1275

Ile Tyr Ile Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Val Gly
 1280 1285 1290

Asn Arg Tyr Asp Ser Arg Asn Asn Asp Asn Arg Ile Phe Asn Tyr
 1295 1300 1305

Lys His Met Asn Phe Val Met Gly Asp Leu Thr Ser Ile Lys Asn

ES 2 622 280 T3

1310	1315	1320
Arg His Asn Phe Ile Leu Leu Asn Asn Lys Ile Val Ile Pro Thr 1325 1330 1335		
Thr Lys Asp Ile Asp Tyr Gly Leu Val Leu Gly Asn Leu Tyr Lys 1340 1345 1350		
Gly Asp Leu Tyr Thr Glu Leu Pro Tyr Glu Asp Ile Lys Glu Val 1355 1360 1365		
Ser Ile Ser Lys Thr His Ile Ile Ile Leu Leu Asn Asp Gly Thr 1370 1375 1380		
Met Tyr Gly Cys Gly Thr Asn Tyr His Gly Glu Leu Leu Gln Asp 1385 1390 1395		
Leu Ser Ile Asn Gln Val Asp Glu Phe Val Gln Ile Asn Val Ser 1400 1405 1410		
Asp Val Lys His Val Ser Cys Gly Asp Asn Phe Thr Tyr Phe Ile 1415 1420 1425		
Lys Ser Asp Asp Ser Leu Trp Ser Ile Gly Lys Asn Ser Glu Tyr 1430 1435 1440		
Gln Leu Gly Ile Gly His Asn Asn Pro Val Thr Glu Leu Gln Arg 1445 1450 1455		
Ile Thr Thr Ile Ser Ser Cys Lys Glu Val His Cys Gly Lys Asn 1460 1465 1470		
Tyr Thr Leu Val Val Thr Thr Ser Asn Glu Leu Phe Val Gln Gly 1475 1480 1485		
Tyr Asn Asp Lys Gly Ala Leu Gly Leu Gly Ser Asp Ser Glu Asn 1490 1495 1500		
Thr Ile Ile Lys Phe Phe Thr Lys Ala Leu Thr Asp Ile Arg Glu 1505 1510 1515		
Ile Lys Ser Tyr Gly Ser Asp His Ile Leu Val Leu Lys Asn Asp 1520 1525 1530		
Asn Ser Val Trp Val Thr Gly Lys Asn Arg Asp Val Tyr Lys Ile 1535 1540 1545		
Glu Gln Pro Val Glu Phe Leu Lys Glu Phe Thr Ile Val Pro Ile 1550 1555 1560		

ES 2 622 280 T3

Ser Glu Asp Val Asn Thr Val Lys Asp Val Leu Ala Thr Asp Asn
 1565 1570 1575

Thr Leu Tyr Ile Ile Ser Glu Val Gly Thr Thr Asn Ala Ala Ile
 1580 1585 1590

Glu Ile Thr Glu Lys Ser Ile Ser Ser Ile Lys Ile Lys Ile Gln
 1595 1600 1605

Asp Pro Asn Lys Asp Ile Ser Arg Ile Glu Met Leu Ile Asn Gly
 1610 1615 1620

Glu Ser Val Lys Ser Val Ser Asp Leu Ile Thr Glu Lys Ile Ser
 1625 1630 1635

Phe Glu Val Pro Pro Asp Lys Ile Lys Ile Gly Glu Asn Lys Ile
 1640 1645 1650

Leu Phe Arg Ala Tyr Cys Lys Gly Asp Asp Leu Tyr Ala Ser Leu
 1655 1660 1665

Phe Ile Phe Lys Glu Ser Thr Gly Asn Ser Ile Ile Lys Asp Ser
 1670 1675 1680

Tyr Val Met Ile Gly Asn Arg Met Tyr Lys Val Val Asn Thr Thr
 1685 1690 1695

Ser Asn Glu Gln Asp Ile Thr Ile Thr Leu Asp Arg Gly Leu Glu
 1700 1705 1710

Glu Asp Leu Asn Leu Gly Asp Pro Ile Tyr Gln Leu Ile Asn Lys
 1715 1720 1725

Thr Lys Val Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp
 1730 1735 1740

Met Lys Leu Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu
 1745 1750 1755

Ile Tyr Glu Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys
 1760 1765 1770

Ile Ile Val Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro
 1775 1780 1785

Ser Met Ile Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln
 1790 1795 1800

<210> 82
 <211> 1742
 <212> PRT
 <213> Clostridium difficile

5

<400> 82

ES 2 622 280 T3

Met Lys Gln Asn Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys
 1 5 10 15

Asn Ile Leu Ile Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe
 20 25 30

Val Glu Phe Phe Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser
 35 40 45

Asp Gly Asn Leu Tyr Ala Cys Gly Asn Asn Thr Gly Phe Pro Leu Gly
 50 55 60

Leu Gly Lys Asp Ser Ser Glu Arg Arg Met Phe Ser Lys Val Lys Ile
 65 70 75 80

Asp Asn Val Lys Tyr Val Ser Cys Gly Ser Lys His Ser Val Ala Val
 85 90 95

Thr Lys Asp Gly Phe Ala Tyr Gly Ala Gly Thr Ser Asn Val Gly Gln
 100 105 110

Leu Gly Val Ile Glu Ser Thr Val Tyr Tyr Glu Phe Thr Lys Leu Pro
 115 120 125

Ile Asp Asp Val Lys Thr Val Ala Cys Gly Tyr Asp Phe Thr Phe Val
 130 135 140

Leu Lys Asn Asp Gly Thr Leu Tyr Ser Ala Gly Leu Asn Ser Ser Gly
 145 150 155 160

Gln Leu Gly Leu Gly Asp Thr Asn Asn Arg Ala Thr Phe Thr Lys Val
 165 170 175

Asn Ile Asp Ser Val Lys Asp Val Val Thr Tyr Asn Gln Ser Val Phe
 180 185 190

Ile Ile Lys Met Asp Gly Thr Ala His Ala Cys Gly Leu Asn Ser Asn
 195 200 205

Gly Gln Leu Gly Ile Asn Ser Thr Leu Asn Lys Ser Val Phe Asn Lys
 210 215 220

Ile Glu Gly Met Asp Asn Val Lys Gln Ile Ala Cys Gly Ser Ser His
 225 230 235 240

ES 2 622 280 T3

Thr Ile Leu Ile Lys Asn Asp Gly Thr Met Tyr Thr Thr Gly Tyr Asn
 245 250 255

Gly Val Gly Gln Leu Gly Thr Gly Asn Asn Asn Ser Ile Val Phe
 260 265 270

Thr Leu Ser Ser Ile Asn Asn Val Lys Tyr Ala Ser Cys Gly Asn Asn
 275 280 285

His Thr Met Ile Leu Lys Tyr Asp Asn Thr Leu Phe Ser Thr Gly Gln
 290 295 300

Asn Thr Tyr Gly Gln Leu Ala Asn Ala Asn Lys Asp Val Ala Ser Arg
 305 310 315 320

Asn Thr Phe Ala Lys Val Asn Val Glu Asn Ile Lys Asp Ile Lys Cys
 325 330 335

Gly Ser Gln Phe Asn Phe Leu Ile Asn Gly Ser Lys Glu Ile Phe Val
 340 345 350

Ser Gly Cys Asn Leu Ala Gly Gln Leu Gly Ser Phe Phe His Thr Thr
 355 360 365

Phe Leu Tyr Glu Phe Ser Lys Val Gln Ser Ser Asn Leu Asp Asn Tyr
 370 375 380

Ser Gly Leu Leu Val Asn Asp Asp Tyr Leu Tyr Val Thr Lys Asp Asn
 385 390 395 400

Ser Glu Phe Leu Asn Val Lys Leu Ser Asp Asn Phe Gln Asp Tyr Lys
 405 410 415

Lys Ile Glu Leu Thr Asp Asn Asn Met Phe Ile Val Met Asn Asp Gly
 420 425 430

Thr Leu Tyr Ala Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Leu Gly
 435 440 445

Asp Thr Val Asn Arg Ser Val Met Thr Lys Val Asp Ile Asp Asn Val
 450 455 460

Leu Asp Ile Lys Gly Asn Gly Asn Ser Thr Phe Val Leu Lys Asn Asn
 465 470 475 480

Gly Thr Leu Tyr Ser Cys Gly Tyr Asn Ser Ser Gly Ile Leu Gly Leu
 485 490 495

ES 2 622 280 T3

Lys Asp Asn Thr Asn Arg Asn Ile Phe Thr Lys Ile Glu Ile Glu Asn
 500 505 510

Val Lys Ala Phe Cys Val Glu Ser Asn Tyr Ile Val Val Leu Asn His
 515 520 525

Ser Lys Glu Leu Tyr Gly Trp Gly Asn Glu Ser Tyr Ile Val Tyr Gly
 530 535 540

Asn Ser Arg Asn Tyr Pro Tyr Lys Asp Thr Arg Val Ser Asn Val Glu
 545 550 555 560

Lys Ile Ala Thr Trp Ser Asp Thr Leu Tyr Ile Leu Asp Ser Thr Gly
 565 570 575

Ala Thr Lys Thr Ile Gly Tyr Ser Tyr Asn Gly Ser Gly Gly Tyr Pro
 580 585 590

Ala Pro Ser Ser Ser Ser Thr Tyr Arg Asp Gly Gly Tyr Ile Asn Lys
 595 600 605

Asn Thr Ser Tyr Arg Thr Leu Glu Phe Tyr Asn Thr Ser Lys Thr Lys
 610 615 620

Leu Val Asn Leu Phe Ala Phe Tyr Asn Gly Cys Val Phe Val Asp Glu
 625 630 635 640

Asn Gly Leu Ala Tyr Cys Ile Gly Glu Asn Asn Ile Asn Phe Arg Gly
 645 650 655

Asn Ser Thr Thr Asn Glu Asn Asn Ser Leu Arg Phe Ile Asn Asn Ser
 660 665 670

Gly Val Tyr Tyr Thr Asn Thr Asp Gly Thr Asp Tyr Thr Cys Tyr Gln
 675 680 685

Trp Thr Tyr Lys Leu Ile Arg Cys Ser Ile Phe Asp Ser Pro Gln Asn
 690 695 700

Ile Ile Gly Asn Ser Lys Asn Ile Leu Tyr Leu Ser Lys Asn Asn Ser
 705 710 715 720

Thr Phe Lys Cys Thr Gly Asn Cys Ile Thr Tyr Gly Ile Asn Ser Gln
 725 730 735

Asn Trp Tyr Ser Tyr Phe Ser Asp Ser Ser Asn Gly Ala Ile Ala Leu
 740 745 750

ES 2 622 280 T3

Gly Asn Glu Phe Ile Leu Lys Asn Tyr Ser Gly Glu Cys Leu Leu Lys
755 760 765

Gly Tyr Gly Lys Ala Thr Asn Gly Glu Phe Gly Asn Ser Thr Asn Ile
770 775 780

Ser Ser Ile Ser Asn Tyr Asp Thr Gly Leu Lys Asp Ile Lys Asp Ile
785 790 795 800

Ile Val Lys Asn Asn Thr Val Val Val Val Asp Lys Asn Asn Asn Ile
805 810 815

Tyr Val Thr Gly Ala Asn Gln Phe Asn Lys Leu Gly Ile Gly Glu Tyr
820 825 830

Asn Asn Gln Pro Ile Lys Lys Phe Thr Asn Ile Thr Glu Gln Ser Asn
835 840 845

Ser Phe Ile Phe Met Asp Asp Ile Lys Glu Ile Thr Thr Ser Arg Asn
850 855 860

Thr Met Phe Ile Val Lys Asn Asp Gly Thr Ala Tyr Ala Thr Gly Asn
865 870 875 880

Asn Ser Ser Gly Gln Leu Gly Leu Gly Asp Thr Ile Asn Arg Asn Lys
885 890 895

Phe Thr Gln Ile Asn Leu Asp Asn Ile Lys Lys Ile Ser Thr Ser Ile
900 905 910

Asp Gly Asn Thr Thr Phe Ala Ile Arg Asn Asp Gly Thr Leu Tyr Ser
915 920 925

Thr Gly Leu Asn Thr Lys Gly Gln Leu Gly Leu Gly Asp Ile Val Asn
930 935 940

Arg Asn Thr Phe Thr Lys Val Asn Ile Gln Asn Val Arg Asp Val Val
945 950 955 960

Leu Gly Thr Thr His Ser His Ala Ile Lys Asp Asp Asn Thr Leu Tyr
965 970 975

Ser Cys Gly Glu Asn Thr His Gly Gln Leu Gly Leu Gly Ser Glu Ser
980 985 990

Asn His Pro Asp Val Leu Thr Phe Thr Val Asn Asn Ile Thr Asn Val
995 1000 1005

Arg Asp Val Tyr Cys Ser Asp Thr Thr Thr Phe Ile Val Lys Asp

ES 2 622 280 T3

1010	1015	1020
Thr Asn Ile Ala Tyr Cys Cys Gly Tyr Asn Asn Asn Ser Gln Leu 1025 1030 1035		
Gly Met Gly Asn Thr Thr Asp Gln Tyr Ser Phe Ile Lys Cys Met 1040 1045 1050		
Glu Asn Val Lys Glu Val Ile Pro Asn Glu Ile Asn Thr Tyr Ile 1055 1060 1065		
Ile Thr Ile Tyr Asn Thr Ala Tyr Ser Thr Gly Leu Asn Thr Asp 1070 1075 1080		
Tyr Cys Leu Gly Leu Asn Ser Asn Ser Asn Gln Ser Ser Phe Ser 1085 1090 1095		
Glu Ile Pro Ile Ser Asn Val Val Lys Val Ala Pro Asn Arg Asn 1100 1105 1110		
Asn Ala Val Leu Leu Leu Thr Ser Glu Gly Asp Val Tyr Thr Ala 1115 1120 1125		
Gly Lys Cys Ser Asn Gly Ser Gly Thr Gly Ser Glu Thr Pro Glu 1130 1135 1140		
Lys Ile Lys Lys Ile Ala Ser Lys Ala Lys Asp Ile Gly Met Asn 1145 1150 1155		
Tyr Arg Cys Gly His Tyr Val Ser Asp Asn Gly Asp Leu Tyr Gly 1160 1165 1170		
Thr Gly Phe Asn Asp Cys Gly Gln Leu Gly Val Gly Asp Val Thr 1175 1180 1185		
Lys Arg Asp Thr Phe Ile Lys Thr Asn Thr Arg Val Lys Lys Ile 1190 1195 1200		
Leu Pro Leu Glu Tyr Ala Asn Ile Ala Ile Lys Asp Thr Asn Asp 1205 1210 1215		
Ile Tyr Ile Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Val Gly 1220 1225 1230		
Asn Arg Tyr Asp Ser Arg Asn Asn Asp Asn Arg Ile Phe Asn Tyr 1235 1240 1245		
Lys His Met Asn Phe Val Met Gly Asp Leu Thr Ser Ile Lys Asn 1250 1255 1260		

ES 2 622 280 T3

Arg His Asn Phe Ile Leu Leu Asn Asn Lys Ile Val Ile Pro Thr
 1265 1270 1275

Thr Lys Asp Ile Asp Tyr Gly Leu Val Leu Gly Asn Leu Tyr Lys
 1280 1285 1290

Gly Asp Leu Tyr Thr Glu Leu Pro Tyr Glu Asp Ile Lys Glu Val
 1295 1300 1305

Ser Ile Ser Lys Thr His Ile Ile Ile Leu Leu Asn Asp Gly Thr
 1310 1315 1320

Met Tyr Gly Cys Gly Thr Asn Tyr His Gly Glu Leu Leu Gln Asp
 1325 1330 1335

Leu Ser Ile Asn Gln Val Asp Glu Phe Val Gln Ile Asn Val Ser
 1340 1345 1350

Asp Val Lys His Val Ser Cys Gly Asp Asn Phe Thr Tyr Phe Ile
 1355 1360 1365

Lys Ser Asp Asp Ser Leu Trp Ser Ile Gly Lys Asn Ser Glu Tyr
 1370 1375 1380

Gln Leu Gly Ile Gly His Asn Asn Pro Val Thr Glu Leu Gln Arg
 1385 1390 1395

Ile Thr Thr Ile Ser Ser Cys Lys Glu Val His Cys Gly Lys Asn
 1400 1405 1410

Tyr Thr Leu Val Val Thr Thr Gly Asn Glu Leu Phe Val Gln Gly
 1415 1420 1425

Tyr Asn Asp Lys Gly Ala Leu Gly Leu Gly Ser Asp Ser Glu Asn
 1430 1435 1440

Thr Ile Ile Lys Phe Phe Thr Lys Ala Leu Thr Asp Ile Arg Glu
 1445 1450 1455

Ile Lys Ser Tyr Gly Ser Asp His Ile Leu Val Leu Lys Asn Asp
 1460 1465 1470

Asn Ser Val Trp Val Thr Gly Lys Asn Arg Asp Val Tyr Lys Ile
 1475 1480 1485

Glu Gln Pro Val Glu Phe Leu Lys Glu Phe Thr Ile Ile Pro Ile
 1490 1495 1500

ES 2 622 280 T3

Ser Glu Asp Val Asn Thr Val Lys Asp Val Leu Ala Thr Asp Asn
 1505 1510 1515

Thr Leu Tyr Ile Ile Ser Glu Val Gly Thr Thr Asn Ala Ala Ile
 1520 1525 1530

Glu Ile Thr Glu Lys Ser Ile Ser Ser Ile Lys Ile Lys Ile Gln
 1535 1540 1545

Asp Pro Asn Lys Asp Ile Ser Arg Ile Glu Met Leu Ile Asn Gly
 1550 1555 1560

Glu Ser Val Lys Ser Val Ser Asp Leu Ile Thr Glu Lys Ile Ser
 1565 1570 1575

Phe Glu Val Pro Pro Asp Lys Ile Lys Ile Gly Glu Asn Lys Ile
 1580 1585 1590

Leu Phe Arg Ala Tyr Cys Lys Gly Asp Asp Leu Tyr Ala Ser Leu
 1595 1600 1605

Phe Ile Phe Lys Glu Ser Thr Gly Asn Ser Ile Ile Lys Asp Ser
 1610 1615 1620

Tyr Val Met Ile Gly Asn Arg Met Tyr Lys Val Val Asn Thr Thr
 1625 1630 1635

Ser Asn Glu Gln Asp Ile Thr Ile Thr Leu Asp Arg Gly Leu Glu
 1640 1645 1650

Glu Asp Leu Asn Leu Gly Asp Pro Ile Tyr Gln Leu Ile Asn Lys
 1655 1660 1665

Thr Lys Val Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp
 1670 1675 1680

Met Lys Leu Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu
 1685 1690 1695

Ile Tyr Glu Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys
 1700 1705 1710

Ile Ile Val Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro
 1715 1720 1725

Ser Met Ile Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln
 1730 1735 1740

<210> 83
 <211> 1724
 <212> PRT
 <213> Clostridium difficile

5

<400> 83

ES 2 622 280 T3

Met Lys Gln Asn Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys
 1 5 10 15

Asn Ile Leu Ile Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe
 20 25 30

Val Glu Phe Phe Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser
 35 40 45

Asp Gly Asn Leu Tyr Ala Cys Gly Asp Asn Thr Gly Phe Gln Leu Gly
 50 55 60

Leu Gly Lys Asp Ser Ser Glu Arg Arg Met Phe Ser Lys Val Lys Ile
 65 70 75 80

Asp Asn Val Lys Tyr Val Ser Cys Gly Ser Lys His Ser Val Ala Val
 85 90 95

Thr Lys Asp Gly Phe Ala Tyr Gly Ala Gly Thr Ser Asn Val Gly Gln
 100 105 110

Leu Gly Val Ile Glu Ser Thr Val Tyr Tyr Glu Phe Thr Lys Leu Pro
 115 120 125

Ile Asp Asp Val Lys Thr Val Ala Cys Gly Tyr Asp Phe Thr Phe Val
 130 135 140

Leu Lys Asn Asp Gly Thr Leu Tyr Ser Ala Gly Leu Asn Ser Ser Gly
 145 150 155 160

Gln Leu Gly Leu Gly Asp Thr Asn Asn Arg Ala Thr Phe Thr Lys Val
 165 170 175

Asn Ile Asp Ser Val Lys Asp Val Val Thr Tyr Asn Gln Ser Val Phe
 180 185 190

Ile Ile Lys Met Asp Gly Thr Ala His Ala Cys Gly Leu Asn Ser Asn
 195 200 205

Gly Gln Leu Gly Ile Asn Ser Thr Leu Asn Lys Ser Val Phe Asn Lys
 210 215 220

Ile Glu Gly Met Asp Asn Val Lys Gln Ile Ala Cys Gly Ser Ser His
 225 230 235 240

ES 2 622 280 T3

Thr Ile Leu Ile Lys Asn Asp Gly Thr Met Tyr Thr Thr Gly Tyr Asn
 245 250 255
 Gly Val Gly Gln Leu Gly Thr Gly Asn Asn Asn Ser Ile Val Phe
 260 265 270
 Thr Leu Ser Ser Ile Asn Asn Val Lys Tyr Ala Ser Cys Gly Asn Asn
 275 280 285
 His Thr Met Ile Leu Lys Tyr Asp Asn Thr Leu Phe Ser Thr Gly Gln
 290 295 300
 Asn Asn Tyr Gly Gln Leu Ala Asn Ala Asn Lys Asp Val Ala Ser Arg
 305 310 315 320
 Asn Thr Phe Ala Lys Val Asn Val Glu Asn Ile Lys Asp Ile Lys Cys
 325 330 335
 Gly Ser Gln Phe Asn Phe Leu Ile Asn Gly Ser Lys Glu Ile Phe Val
 340 345 350
 Ser Gly Cys Asn Leu Ala Gly Gln Leu Gly Ser Phe Phe His Thr Thr
 355 360 365
 Phe Leu Tyr Glu Phe Ser Asn Val Gln Ser Ser Asn Leu Asp Asn Tyr
 370 375 380
 Ser Gly Leu Leu Val Asn Asp Asp Tyr Leu Tyr Val Thr Lys Asp Asn
 385 390 395 400
 Ser Glu Phe Leu Asn Val Lys Leu Ser Asp Asn Phe Gln Asp Tyr Lys
 405 410 415
 Lys Ile Glu Leu Thr Asp Ser Asn Met Phe Ile Val Met Asn Asp Gly
 420 425 430
 Thr Leu Tyr Ala Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Leu Gly
 435 440 445
 Asp Thr Val Asn Arg Ser Val Met Thr Lys Val Asp Ile Asp Asn Val
 450 455 460
 Leu Asp Ile Lys Gly Asn Gly Asn Ser Thr Phe Val Leu Lys Asn Asn
 465 470 475 480
 Gly Thr Leu Tyr Ser Cys Gly Leu Asn Ser Asn Gly Gln Leu Gly Leu
 485 490 495

ES 2 622 280 T3

Arg Asp Glu Val Asn Arg Asn Ile Phe Thr Lys Ile Glu Ile Glu Asn
500 505 510

Val Lys Asp Phe Cys Val Gly Ser Asn Tyr Val Ile Ala Leu Asn His
515 520 525

Ser Lys Glu Val Tyr Gly Trp Gly Asn Asn Pro Tyr Asn Asn Ile Glu
530 535 540

Lys Thr Ser Asn Tyr Pro Tyr Lys Gln Gly Ile Ser Asn Ile Glu Lys
545 550 555 560

Ile Ala Ala Tyr Asp Tyr Ser Val Tyr Met Ile Asn Ser Glu Gly Lys
565 570 575

Leu Tyr Val Ser Gly Tyr Asn Tyr Asn Tyr Gln Leu Gly Lys Gly Asn
580 585 590

Asn Ser Asn Gln Ser Lys Ala Leu Val Ser Gln Cys Arg Thr Asn Ser
595 600 605

Thr Ser Ser Thr Ser Asn Gly Leu Arg Thr Leu Pro Lys Ile Thr Asn
610 615 620

Val Phe Pro Phe Tyr Asp Gly Cys Ala Ile Ile Asp Glu Gly Gly Tyr
625 630 635 640

Val Tyr Leu Thr Gly Tyr His Gly Tyr Leu Arg Thr Leu Asn Ser Ser
645 650 655

Pro Ser Ile Ser Asp Tyr Ser Arg Tyr Gly Thr Phe Ile Glu Ala Thr
660 665 670

Asn Ser Asn His Asn Thr Tyr Phe Ile Gln Glu Thr Asp Phe Ser Gly
675 680 685

Ile Glu Lys Val Ile Gly Met Ser Asn Asn Ile Leu Phe Phe Lys Lys
690 695 700

Gly Ser Ser Tyr Ile Thr Gly Tyr Pro Lys Thr Phe Gly Ser Thr Ile
705 710 715 720

Thr Gly His Arg Ser Tyr Thr Ser Ile Asn Ser Glu Ser Ser Asn Leu
725 730 735

Gly Ser Asn Phe Ile Ile Tyr His Ser Asn Ser Lys Leu Tyr Gly Lys
740 745 750

Gly Ile Ala Asn Ser Gly Gln Phe Gly Asn Ser Thr Asn Ile Asp Gly

ES 2 622 280 T3

Asn Thr Thr Asp Gln Tyr Ser Phe Ile Lys Cys Met Glu Asn Val
 1025 1030 1035

Lys Glu Val Ile Pro Asn Glu Ile Asn Thr Tyr Ile Ile Thr Ile
 1040 1045 1050

Tyr Asn Thr Ala Tyr Ser Thr Gly Leu Asn Thr Asp Tyr Cys Leu
 1055 1060 1065

Gly Leu Asn Ser Asn Ser Asn Gln Ser Ser Phe Ser Glu Ile Pro
 1070 1075 1080

Ile Ser Asn Val Val Lys Val Ala Pro Asn Arg Asn Asn Ala Val
 1085 1090 1095

Leu Leu Leu Thr Ser Glu Gly Asp Val Tyr Thr Ala Gly Lys Cys
 1100 1105 1110

Ser Asn Gly Ser Gly Thr Gly Ser Glu Thr Pro Glu Lys Ile Lys
 1115 1120 1125

Lys Ile Ala Ser Lys Ala Lys Asp Ile Gly Met Asn Tyr Arg Cys
 1130 1135 1140

Gly His Tyr Val Ser Asp Asn Gly Asp Leu Tyr Gly Thr Gly Phe
 1145 1150 1155

Asn Asn Asn Gly Gln Leu Gly Val Gly Asp Val Thr Lys Arg Asp
 1160 1165 1170

Thr Phe Ile Lys Thr Asn Thr Arg Val Lys Lys Ile Leu Pro Leu
 1175 1180 1185

Glu Tyr Ala Asn Ile Ala Ile Lys Asp Thr Asn Asp Ile Tyr Ile
 1190 1195 1200

Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Val Gly Asn Arg Tyr
 1205 1210 1215

Asp Ser Arg Asn Asn Asp Asn Arg Ile Phe Asn Tyr Lys His Met
 1220 1225 1230

Asn Phe Val Met Gly Asp Leu Thr Ser Ile Lys Asn Arg His Asn
 1235 1240 1245

Phe Ile Leu Leu Asn Asn Lys Ile Val Ile Pro Thr Thr Lys Asp
 1250 1255 1260

ES 2 622 280 T3

Ile Asp Tyr Gly Leu Val Leu Gly Asn Leu Tyr Lys Gly Asp Leu
1265 1270 1275

Tyr Thr Glu Leu Pro Tyr Glu Asp Ile Lys Glu Val Ser Ile Ser
1280 1285 1290

Lys Thr His Ile Ile Ile Leu Leu Asn Asp Gly Thr Met Tyr Gly
1295 1300 1305

Cys Gly Thr Asn Tyr His Gly Glu Leu Leu Gln Asp Leu Ser Ile
1310 1315 1320

Asn Gln Val Asp Glu Phe Val Gln Ile Asn Val Ser Asp Val Lys
1325 1330 1335

His Val Ser Cys Gly Asp Asn Phe Thr Tyr Phe Ile Lys Ser Asp
1340 1345 1350

Asp Ser Leu Trp Ser Ile Gly Lys Asn Ser Glu Tyr Gln Leu Gly
1355 1360 1365

Ile Gly His Asn Asn Pro Val Thr Glu Leu Gln Arg Ile Thr Thr
1370 1375 1380

Ile Ser Ser Cys Lys Glu Val His Cys Gly Lys Asn Tyr Thr Leu
1385 1390 1395

Val Val Thr Thr Gly Asn Glu Leu Phe Val Gln Gly Tyr Asn Asp
1400 1405 1410

Lys Gly Ala Leu Gly Leu Gly Ser Asp Ser Glu Asn Thr Ile Ile
1415 1420 1425

Lys Phe Phe Thr Lys Ala Leu Thr Asp Ile Arg Glu Ile Lys Ser
1430 1435 1440

Tyr Gly Ser Asp His Ile Leu Val Leu Lys Asn Asp Asn Ser Val
1445 1450 1455

Trp Val Thr Gly Lys Asn Arg Asp Val Tyr Lys Ile Glu Gln Pro
1460 1465 1470

Val Glu Phe Leu Lys Glu Phe Thr Ile Val Pro Ile Ser Glu Asp
1475 1480 1485

Val Asn Thr Val Lys Asp Val Leu Ala Thr Asp Asn Thr Leu Tyr
1490 1495 1500

ES 2 622 280 T3

Ile Ile Ser Glu Val Gly Thr Thr Asn Ala Ala Ile Glu Ile Thr
 1505 1510 1515

Glu Lys Ser Ile Ser Ser Ile Lys Ile Lys Ile Gln Asp Pro Asn
 1520 1525 1530

Lys Asp Ile Ser Arg Ile Glu Met Leu Ile Asn Gly Glu Ser Val
 1535 1540 1545

Lys Ser Val Ser Asp Leu Thr Thr Glu Lys Ile Ser Phe Glu Val
 1550 1555 1560

Pro Pro Asp Lys Ile Lys Ile Gly Glu Asn Lys Ile Leu Phe Arg
 1565 1570 1575

Ala Tyr Cys Lys Gly Asp Asp Leu Tyr Ala Ser Leu Phe Ile Phe
 1580 1585 1590

Lys Glu Ser Thr Gly Asn Ser Ile Ile Lys Asp Ser Tyr Val Met
 1595 1600 1605

Ile Gly Asn Arg Met Tyr Lys Val Val Asn Thr Thr Ser Asn Glu
 1610 1615 1620

Gln Asp Ile Thr Ile Thr Leu Asp Arg Gly Leu Glu Glu Asp Leu
 1625 1630 1635

Asn Leu Gly Asp Pro Ile Tyr Gln Leu Ile Asn Lys Thr Lys Val
 1640 1645 1650

Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp Met Lys Leu
 1655 1660 1665

Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu Ile Tyr Glu
 1670 1675 1680

Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys Ile Ile Val
 1685 1690 1695

Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro Ser Met Ile
 1700 1705 1710

Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln
 1715 1720

<210> 84
 <211> 1724
 <212> PRT
 <213> Clostridium difficile
 <400> 84

5

ES 2 622 280 T3

Met Lys Gln Asn Lys Leu Leu Gln Arg Gly Ala Tyr Phe Asn Asp Lys
1 5 10 15

Asn Ile Leu Ile Asp Asp Phe Asp Lys Arg Tyr Asn Asp Tyr Asp Phe
20 25 30

Val Glu Phe Phe Thr Gly Ile Ser Asn Ser Thr Phe Gly Leu Lys Ser
35 40 45

Asp Gly Asn Leu Tyr Ala Cys Gly Asp Asn Thr Gly Phe Pro Leu Gly
50 55 60

Leu Gly Lys Asp Ser Ser Glu Arg Arg Met Phe Ser Lys Val Lys Ile
65 70 75 80

Asp Asn Val Lys Tyr Val Ser Cys Gly Ser Lys His Ser Val Ala Val
85 90 95

Thr Lys Asp Gly Phe Ala Tyr Gly Ala Gly Thr Ser Asn Val Gly Gln
100 105 110

Leu Gly Val Ile Glu Ser Thr Val Tyr Tyr Glu Phe Thr Lys Leu Pro
115 120 125

Ile Asp Asp Val Lys Thr Val Ala Cys Gly Tyr Asp Phe Thr Phe Val
130 135 140

Leu Lys Asn Asp Gly Thr Leu Tyr Ser Ala Gly Leu Asn Ser Ser Gly
145 150 155 160

Gln Leu Gly Leu Gly Asp Thr Asn Asn Arg Ala Thr Phe Thr Lys Val
165 170 175

Asn Ile Asp Ser Val Lys Asp Val Val Thr Tyr Asn Gln Ser Val Phe
180 185 190

Ile Ile Lys Met Asp Gly Thr Ala His Ala Cys Gly Leu Asn Ser Asn
195 200 205

Gly Gln Leu Gly Ile Asn Ser Thr Leu Asn Lys Ser Val Phe Asn Lys
210 215 220

Ile Glu Gly Met Asp Asn Val Lys Gln Ile Ala Cys Gly Ser Ser His
225 230 235 240

Thr Ile Leu Ile Lys Asn Asp Gly Thr Met Tyr Thr Thr Gly Tyr Asn
245 250 255

ES 2 622 280 T3

Gly Val Gly Gln Leu Gly Thr Gly Asn Asn Asn Asn Ser Ile Val Phe
 260 265 270

Thr Leu Ser Ser Ile Asn Asn Val Lys Tyr Ala Ser Cys Gly Asn Asn
 275 280 285

His Thr Met Ile Leu Lys Tyr Asp Asn Thr Leu Phe Ser Thr Gly Gln
 290 295 300

Asn Asn Tyr Gly Gln Leu Ala Asn Ala Asn Lys Asp Val Ala Ser Arg
 305 310 315 320

Asn Thr Phe Ala Lys Val Asn Val Glu Asn Ile Lys Asp Ile Lys Cys
 325 330 335

Gly Ser Gln Phe Asn Phe Leu Ile Asn Gly Ser Lys Glu Ile Phe Val
 340 345 350

Ser Gly Cys Asn Leu Ala Gly Gln Leu Gly Ser Phe Phe His Thr Thr
 355 360 365

Phe Leu Tyr Glu Phe Ser Asn Val Gln Ser Ser Asn Leu Asp Asn Tyr
 370 375 380

Ser Gly Leu Leu Val Asn Asp Asp Tyr Leu Tyr Val Thr Lys Asp Asn
 385 390 395 400

Ser Glu Phe Leu Asn Val Lys Leu Ser Asp Asn Phe Gln Asp Tyr Lys
 405 410 415

Lys Ile Glu Leu Thr Asp Ser Asn Met Phe Ile Val Met Asn Asp Gly
 420 425 430

Thr Leu Tyr Ala Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Leu Gly
 435 440 445

Asp Thr Val Asn Arg Ser Val Met Thr Lys Val Asp Ile Asp Asn Val
 450 455 460

Leu Asp Ile Lys Gly Asn Gly Asn Ser Thr Phe Val Leu Lys Asn Asn
 465 470 475 480

Gly Thr Leu Tyr Ser Cys Gly Leu Asn Ser Asn Gly Gln Leu Gly Leu
 485 490 495

Arg Asp Glu Val Asn Arg Asn Ile Phe Thr Lys Ile Glu Ile Glu Asn
 500 505 510

Val Lys Asp Phe Cys Val Gly Ser Asn Tyr Val Ile Ala Leu Asn His

ES 2 622 280 T3

Lys Gly Asn Thr Val Val Val Val Asp Lys Asn Asn Asn Ile Tyr Val
 785 790 795 800
 Thr Gly Met Asn Gln Asn Asn Lys Leu Gly Ile Gly Glu Tyr Asn Asn
 805 810 815
 Glu Pro Val Lys Lys Phe Thr Asn Ile Thr Glu Gln Ser Asn Ser Phe
 820 825 830
 Ile Phe Met Asp Asp Ile Lys Glu Ile Thr Thr Ser Arg Asn Thr Met
 835 840 845
 Phe Ile Val Lys Asn Asp Gly Thr Ala Tyr Ala Thr Gly Asn Asn Ser
 850 855 860
 Ser Gly Gln Leu Gly Leu Gly Asp Thr Ile Asn Arg Asn Lys Phe Thr
 865 870 875 880
 Gln Ile Asn Leu Asp Asn Ile Lys Lys Ile Ser Thr Ser Ile Asp Gly
 885 890 895
 Asn Thr Thr Phe Ala Ile Arg Asn Asp Gly Thr Leu Tyr Ser Thr Gly
 900 905 910
 Leu Asn Thr Lys Gly Gln Leu Gly Leu Gly Asp Ile Val Asn Arg Asn
 915 920 925
 Thr Phe Thr Lys Val Asn Ile Gln Asn Val Arg Asp Val Val Leu Gly
 930 935 940
 Thr Thr His Ser His Ala Ile Lys Asp Asp Asn Thr Leu Tyr Ser Cys
 945 950 955 960
 Gly Glu Asn Thr His Gly Gln Leu Gly Leu Gly Ser Glu Ser Asn His
 965 970 975
 Pro Asp Val Leu Thr Phe Thr Val Asn Asn Ile Thr Asn Val Arg Asp
 980 985 990
 Val Tyr Cys Ser Asp Thr Thr Thr Phe Ile Val Lys Asp Thr Asn Ile
 995 1000 1005
 Ala Tyr Cys Cys Gly Tyr Asn Asn Asn Ser Gln Leu Gly Met Gly
 1010 1015 1020
 Asn Thr Thr Asp Gln Tyr Ser Phe Ile Lys Cys Met Glu Asn Val
 1025 1030 1035

ES 2 622 280 T3

Lys Glu Val Ile Pro Asn Glu Ile Asn Thr Tyr Ile Ile Thr Ile
 1040 1045 1050
 Tyr Asn Thr Ala Tyr Ser Thr Gly Leu Asn Thr Asp Tyr Cys Leu
 1055 1060 1065
 Gly Leu Asn Ser Asn Ser Asn Gln Ser Ser Phe Ser Glu Ile Pro
 1070 1075 1080
 Ile Ser Asn Val Val Lys Val Ala Pro Asn Arg Asn Asn Ala Val
 1085 1090 1095
 Leu Leu Leu Thr Ser Glu Gly Asp Val Tyr Thr Ala Gly Lys Cys
 1100 1105 1110
 Ser Asn Gly Ser Gly Thr Gly Ser Glu Thr Pro Glu Lys Ile Lys
 1115 1120 1125
 Lys Ile Ala Ser Lys Ala Lys Asp Ile Gly Met Asn Tyr Arg Cys
 1130 1135 1140
 Gly His Tyr Val Ser Asp Asn Gly Asp Leu Tyr Gly Thr Gly Phe
 1145 1150 1155
 Asn Asn Asn Gly Gln Leu Gly Val Gly Asp Val Thr Lys Arg Asp
 1160 1165 1170
 Thr Phe Ile Lys Thr Asn Thr Arg Val Lys Lys Ile Leu Pro Leu
 1175 1180 1185
 Glu Tyr Ala Asn Ile Ala Ile Lys Asp Thr Asn Asp Ile Tyr Ile
 1190 1195 1200
 Cys Gly Leu Asn Asn Tyr Gly Gln Leu Gly Val Gly Asn Arg Tyr
 1205 1210 1215
 Asp Ser Arg Asn Asn Asp Asn Arg Ile Phe Asn Tyr Lys His Met
 1220 1225 1230
 Asn Phe Val Met Gly Asp Leu Thr Ser Ile Lys Asn Arg His Asn
 1235 1240 1245
 Phe Ile Leu Leu Asn Asn Lys Ile Val Ile Pro Thr Thr Lys Asp
 1250 1255 1260
 Ile Asp Tyr Gly Leu Val Leu Gly Asn Leu Tyr Lys Gly Asp Leu
 1265 1270 1275

ES 2 622 280 T3

Tyr Thr Glu Leu Pro Tyr Glu Asp Ile Lys Glu Val Ser Ile Ser
 1280 1285 1290
 Lys Thr His Ile Ile Ile Leu Leu Asn Asp Gly Thr Met Tyr Gly
 1295 1300 1305
 Cys Gly Thr Asn Tyr His Gly Glu Leu Leu Gln Asp Leu Ser Ile
 1310 1315 1320
 Asn Gln Val Asp Glu Phe Val Gln Ile Asn Val Ser Asp Val Lys
 1325 1330 1335
 His Val Ser Cys Gly Asp Asn Phe Thr Tyr Phe Ile Lys Ser Asp
 1340 1345 1350
 Asp Ser Leu Trp Ser Ile Gly Lys Asn Ser Glu Tyr Gln Leu Gly
 1355 1360 1365
 Ile Gly His Asn Asn Pro Val Thr Glu Leu Gln Arg Ile Thr Thr
 1370 1375 1380
 Ile Ser Ser Cys Lys Glu Val His Cys Gly Lys Asn Tyr Thr Leu
 1385 1390 1395
 Val Val Thr Thr Gly Asn Glu Leu Phe Val Gln Gly Tyr Asn Asp
 1400 1405 1410
 Lys Gly Ala Leu Gly Leu Gly Ser Asp Ser Glu Asn Thr Ile Ile
 1415 1420 1425
 Lys Phe Phe Thr Lys Ala Leu Thr Asp Ile Arg Glu Ile Lys Ser
 1430 1435 1440
 Tyr Gly Ser Asp His Ile Leu Val Leu Lys Asn Asp Asn Ser Val
 1445 1450 1455
 Trp Val Thr Gly Lys Asn Arg Asp Val Tyr Lys Ile Glu Gln Pro
 1460 1465 1470
 Val Glu Phe Leu Lys Glu Phe Thr Ile Val Pro Ile Ser Glu Asp
 1475 1480 1485
 Val Asn Thr Val Lys Asp Val Leu Ala Thr Asp Asn Thr Leu Tyr
 1490 1495 1500
 Ile Ile Ser Glu Val Gly Thr Thr Asn Ala Ala Ile Glu Ile Thr
 1505 1510 1515
 Glu Lys Ser Ile Ser Ser Ile Lys Ile Lys Ile Gln Asp Pro Asn

ES 2 622 280 T3

Asn Met Val Val Asp Glu Phe Asp Glu Gly Tyr Asp Asn Tyr Asp Phe
 20 25 30
 Ile Asn Phe Phe Thr Gly Cys Cys Asn Tyr Thr Phe Gly Leu Lys Asn
 35 40 45
 Asn Asn Ile Leu Tyr Gly Cys Gly Asp Asn Ser Asn Phe Gln Leu Gly
 50 55 60
 Leu Gly Glu Asp Asn Thr Thr Arg Lys Leu Phe Thr Lys Ile Pro Asn
 65 70 75 80
 Ile Ser Thr Asn Ile Lys Lys Val Ala Cys Gly Glu Ser His Ala Val
 85 90 95
 Ile Leu Thr Ser Asp Gly Glu Leu Leu Val Ala Gly Ile Asn Thr Asp
 100 105 110
 Gly Gln Met Gly Leu Gly Leu Glu Lys Val Gly Lys Thr Val Ser Thr
 115 120 125
 Phe Glu Lys Val Pro Glu Ile Lys Gly Val Lys Asp Ile Ala Cys Gly
 130 135 140
 Leu Gln Ser Thr Tyr Leu Leu Tyr Asn Asp Gly Thr Leu Tyr Val Ala
 145 150 155 160
 Gly Asn Asn Leu Tyr Gly Gln Leu Gly Leu Gly Thr Asn Gly Ala Ser
 165 170 175
 Ala Asn Val Asn Thr Phe Thr Lys Val Asp Val Asp Asn Val Lys Ala
 180 185 190
 Val Phe Ser Tyr Asn Lys Ser Ala Phe Ile Ile Lys Asn Asp Asn Lys
 195 200 205
 Cys Tyr Ser Thr Gly Phe Asn Asn Gln Gly Gln Leu Gly Leu Gly Asp
 210 215 220
 Lys Asn Asn Arg Asp Leu Phe Ser Leu Val Ser Ile Asn Asp Val Lys
 225 230 235 240
 Thr Ile Ala Cys Gly Ser Glu His Thr Val Leu Met Thr Tyr Asn Asn
 245 250 255
 Asp Ile Tyr Gly Cys Gly Lys Glu Lys Cys Phe Gly Asn Ala Leu Gln
 260 265 270
 Ser Ser Leu Phe Thr Lys Ile Glu Glu Val Asn Ile Lys Thr Ile Ala

ES 2 622 280 T3

275	280	285															
Cys Gly His Gly Asn Thr Met Leu Ile Asp Asn Lys Gly Thr Leu Lys																	
290		295															
305	Val Ala Gly Asn Asn Asp Ile Tyr Gln Leu Gly Ile Ala Asn Tyr Ser																
		310															
			315														
Glu Asn Ile Asp Asn Ser Phe Ile Asp Leu Lys Asn Ile Val Ala Lys																	
		325															
			330														
				335													
Asn Ile Phe Ile Gly Leu Ser His Ser Ile Leu Ile Asp Ser Asn Asn																	
		340															
			345														
				350													
Asp Ser Tyr Cys Thr Gly Asp Asn Thr Tyr Gly Gln Leu Gly Ser Phe																	
		355															
			360														
				365													
Phe Asp Asp Met His Ile Val Glu Phe Lys Lys Met Asp Ser Glu Lys																	
		370															
			375														
				380													
Tyr Ser Tyr Ser Asn Tyr Ile Asn Leu Ile Lys Ser Glu Asp Lys Leu																	
		385															
			390														
				395													
Thr Leu Leu Lys Glu Glu Met Glu Ile Lys Asp Ile Glu Leu Pro Leu																	
		405															
			410														
				415													
Asp Ile His Ser Val Arg Asp Val Val Phe Ser Pro Tyr Cys Thr Leu																	
		420															
			425														
				430													
Val Ile Leu Gly Asn Gly Asp Val Tyr Gly Leu Gly Asn Asn Arg Tyr																	
		435															
			440														
				445													
Lys Gly Met Gly Ser Asp Leu Pro Ser Gln Leu Asn Glu Leu Thr Lys																	
		450															
			455														
				460													
Leu Ser Ile Ser Asn Val Lys Ser Ile Val Ala Ser Lys Asn Ile Ser																	
		465															
			470														
				475													
Gly Gly Ile Phe Tyr Ile Lys Asn Asp Asp Thr Cys Tyr Tyr Ser Gly																	
		485															
			490														
				495													
Pro Asn Ser Asn Ser Ile Ala Gly Val Leu Pro Ser Asn Ser Asp Val																	
		500															
			505														
				510													
Phe Lys Lys Ile Ser Ile Asp Asn Val Lys Lys Val Val Ile Asn Thr																	
		515															
			520														
				525													
Asp Leu Ser Asn Trp Phe Ser Leu Ile Val Thr Asn Asn Lys Gln Ile																	
		530															
			535														
				540													

ES 2 622 280 T3

Tyr Thr Ser Gly Lys Ser Ser Ser Tyr Val Asn Gly Leu Ser Asn Ala
 545 550 555 560

 Leu Ile Ser Gln Tyr Thr Glu Ile Ser Leu Ser Asn Val Thr Asp Ala
 565 570 575

 Tyr Ser Ser Tyr Asn Ala Thr Phe Ile Val Val Asp Glu Lys Lys Val
 580 585 590

 Tyr Ala Thr Gly Ile Asn Thr Asn Tyr Leu Leu Gly Phe Ser Thr Ser
 595 600 605

 Asp Gly Ser Asn Val Asn Leu Gly Leu Leu Ser Asp Trp Tyr Tyr Ile
 610 615 620

 Asn Ile Ser Gly Ser Ser Tyr Ser Arg Val Ser Cys Thr Asn Asn Ile
 625 630 635 640

 Thr Lys Ile Asn Asn Ile Ile Ile Tyr Glu Tyr Val Thr Val Phe Cys
 645 650 655

 Thr Asn Ile Gly Ser Phe Leu Thr Gly Tyr His Gly Thr Ser Trp Thr
 660 665 670

 Lys Pro Thr Asp Ser Ser Tyr Arg Val Gln Tyr Gln Gly Ile Ser Tyr
 675 680 685

 Ala Gly Tyr Leu Asp Ser Tyr Ile Tyr Asn Tyr Tyr Pro Thr Arg Cys
 690 695 700

 Thr Gln Ser Ser Ser Ser Thr Thr Phe Ala Tyr Leu Tyr Asn Gly Glu
 705 710 715 720

 Ser Ser Ser Asn Leu Lys Asn Val Asn Pro Asp Asn Leu Leu Ile Ser
 725 730 735

 Gly Gly Ser Ser Tyr Ile His Gln Tyr Gly Arg Asn Tyr Leu Asn Asn
 740 745 750

 Gln Ser Ser Asn Asn Ile Ala Ala Ser Asn Ile Asn Ser Gly Pro Ile
 755 760 765

 Thr Ser Asp Lys Ala Ile Phe Leu Tyr Lys Ala Leu Leu Tyr Leu Ser
 770 775 780

 Ser Asn Thr Leu Tyr Gly Phe Gly Asn Ile Ser Glu Ser Ala Lys Glu
 785 790 795 800

ES 2 622 280 T3

Leu Asp Val Ser Asp Thr Gln Asp Gly Tyr Asn Ala Thr Asn Tyr Lys
 805 810 815
 Lys Val Met Lys Asn Ile Lys Asn Ile Phe Ile Pro Pro Tyr Asp Leu
 820 825 830
 Ser Arg Asp Lys Thr Arg Phe Ala Ile Leu Thr Asp Lys Ser Leu Phe
 835 840 845
 Ile Cys Gly Tyr Asn Ser Lys Gly Thr His Gly Ile Ser Val Asn Ser
 850 855 860
 Ser Leu Asn Leu Asn Asn Lys Ile Asn Tyr Asn Lys Lys Asn Ser Ser
 865 870 875 880
 Ser Glu Ile Ser Ser Asn Ile Gln Glu Ile Tyr Ser His Ser Lys Ser
 885 890 895
 Thr Tyr Leu Leu Thr Asn Asn Asn Met Leu Tyr Ser Val Gly Leu Asn
 900 905 910
 Asp Val Gly Gln Leu Gly Val Gly Asp Glu Ile Asn Arg Lys Val Phe
 915 920 925
 Thr Lys Ile Asn Ile Asp Asn Ile Lys Ser Ile Asn Val Asn Arg Phe
 930 935 940
 Thr Asp Asn Ser Lys His Ala Phe Ala Ile Lys Asn Asp Asn Thr Cys
 945 950 955 960
 Tyr Ala Val Gly Leu Asn Asn Ser Gly Gln Leu Gly Ile Gly Asp Asn
 965 970 975
 Val Asn Arg Asn Ile Phe Thr Lys Ile Asn Val Glu Asn Val Lys Tyr
 980 985 990
 Val Ala Val Tyr Gly Asn Thr Ser Leu Leu Leu Thr Asn Asp Gly Leu
 995 1000 1005
 Leu Tyr Gly Ala Gly Asn Asn Gly Lys Gly Gln Leu Gly Leu Gly
 1010 1015 1020
 Asp Thr Thr Ser Arg Asn Ile Phe Thr Arg Ile Pro Ile Asn Gly
 1025 1030 1035
 Val Arg Asp Val Tyr Leu Cys Asn Asp Val Ser Ile Ile Val Lys
 1040 1045 1050

ES 2 622 280 T3

Asn Asp Asn Thr Cys Tyr Val Cys Gly Leu Val Asn Gly Tyr Phe
 1055 1060 1065
 Gly Phe Thr Glu Gly Ser Ile Ser Thr Phe Thr Lys Ile Asn Ile
 1070 1075 1080
 Glu Asn Val Lys Ser Val Val Thr Ala Gly Ser Glu Ala Thr Phe
 1085 1090 1095
 Phe Ile Thr Asn Asp Asn Met Ile Tyr Thr Thr Gly Lys Lys Glu
 1100 1105 1110
 Arg Val Phe Phe Ser Thr Glu Thr Asn Asp Ile Lys Gly Ile Arg
 1115 1120 1125
 Val Ile Asn Asn Ile Ile Asn Ala Lys Lys Ile Val Val Asn Gly
 1130 1135 1140
 Tyr Thr Ser Ala Ile Leu Thr Asn Asp Asn Lys Leu Phe Val Gly
 1145 1150 1155
 Gly Leu Ser Gly Tyr Gly Ser Ile Ala Asn Asn Asn Asn Thr Asn
 1160 1165 1170
 Ser Val Glu Asp Val Lys Asp Val Phe Val Thr Ala Asn Asn Thr
 1175 1180 1185
 Leu Tyr Ile Asp Asn Asn Asn Asn Leu Ile Ser Ser Gly Arg Asp
 1190 1195 1200
 Thr Tyr Gly Ile Ser Asp Glu Ser Tyr Arg Asp Met Ser Val Pro
 1205 1210 1215
 Tyr Tyr Lys Val Ser Ile Lys Lys Asp Val Asp Thr Val Phe Ser
 1220 1225 1230
 Ser Tyr Asn Thr Ile Phe Ile Lys Asp Ile Tyr Gly Lys Phe Tyr
 1235 1240 1245
 Ser Ser Thr Arg Asp Asn Arg Tyr Asn His Leu Gly Ile His His
 1250 1255 1260
 Arg Tyr Asp Asn Asp Lys Asn Glu Ala Leu Glu Gly Ser Leu His
 1265 1270 1275
 Ser Tyr Phe Lys Thr Asp Asn Thr Ser Asp Lys Ile Val Phe Asn
 1280 1285 1290
 Lys Lys Asn Glu Lys Leu Val Met Phe Asn Asp Lys Tyr Ile Lys

ES 2 622 280 T3

1295		1300		1305
Thr Asn	Asn Lys Tyr Ile	Asn Tyr Lys Asn Ile	Phe Lys Asp Asn	
1310		1315	1320	
Phe Lys	Tyr Thr Ser Ile	Ile Leu Pro Phe Glu Val	Ser Asp Ile	
1325		1330	1335	
Asp Ile	Ser Lys Thr His	Ser Leu Ala Val Ala Lys	Asp Gly Lys	
1340		1345	1350	
Leu Tyr	Gly Ile Gly Ser	Asn Ser Tyr Lys Glu Ile	Asn Gln Thr	
1355		1360	1365	
Leu Glu	Asp Ile Glu Leu	Leu Thr Leu Thr Glu Val	Asn Ile Ser	
1370		1375	1380	
Asp Val	Lys Lys Val Ala	Cys Gly Asp Asn Tyr Ser	Tyr Ile Ile	
1385		1390	1395	
Lys Thr	Asp Asn Thr Leu	Trp Ser Tyr Gly Lys Asn	Thr Glu Tyr	
1400		1405	1410	
Gln Leu	Gly Val Gly His	Asn Asn Asp Val Arg Glu	Leu Gln Lys	
1415		1420	1425	
Val Thr	Gly Leu Pro Ser	Val Lys Asp Ile Ser Ile	Tyr Asn Ser	
1430		1435	1440	
Met Thr	Leu Val Leu Thr	Asn Glu Gly Glu Leu Tyr	Ala Gln Gly	
1445		1450	1455	
Tyr Asn	Thr Asn Gly Leu	Phe Gly Leu Gly Glu Ser	Glu Lys Asp	
1460		1465	1470	
Lys Ile	Ile Arg Thr Phe	Thr Lys Val Leu Thr Asn	Val Lys Glu	
1475		1480	1485	
Ile Lys	Ser His Asn Asp	Asp His Ile Leu Val Ile	Lys Asn Asp	
1490		1495	1500	
Asn Ser	Leu Trp Ile Thr	Gly Lys Asn Lys Ser Met	Tyr Lys Ile	
1505		1510	1515	
Ser Ile	Ser Ile Thr Asp	Leu Tyr Glu Phe Thr Lys	Ile Pro Ile	
1520		1525	1530	
Pro Glu	His Leu Asn Asp	Ile Leu Asp Ile Glu Leu	Ser Asp Asp	
1535		1540	1545	

ES 2 622 280 T3

Thr Ile Tyr Met Ile Thr Lys Val Asp Thr Ser Lys Ala Ser Ile
 1550 1555 1560

Glu Ile Val Glu Lys Ser Ile Ser Gln Val Arg Val Val Val Gln
 1565 1570 1575

Asp Pro Asn Asn Val Ile Glu Lys Leu Glu Met Phe Ile Asn Asp
 1580 1585 1590

Glu Leu Ile Ser Thr Lys Thr Asn Leu Glu Ile Asn Ser Ile Ile
 1595 1600 1605

Phe Glu Ile Pro Gln Asn Lys Ile Val Leu Gly Glu Asn Lys Ile
 1610 1615 1620

Leu Ile Lys Ala Ser Ser Pro Thr Gly Asp Leu Tyr Ser Ser Met
 1625 1630 1635

Phe Ile Phe Lys Ser Glu Thr Gly Leu Lys Val Lys Lys Asp Ser
 1640 1645 1650

Ile Leu Met Ile Asn Asn Lys Val Tyr Ser Ile Ile Asn Ile Thr
 1655 1660 1665

Glu Asn Asn Thr Asp Leu Ile Val Thr Leu Asn Glu Gly Leu Lys
 1670 1675 1680

Asp Asp Met Met Glu Asn Asn Pro Ile Tyr Gln Leu Ile Asn Lys
 1685 1690 1695

Thr Lys Val Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp
 1700 1705 1710

Met Lys Leu Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu
 1715 1720 1725

Ile Tyr Glu Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys
 1730 1735 1740

Ile Ile Val Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro
 1745 1750 1755

Ser Met Ile Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln
 1760 1765 1770

<210> 86
 <211> 1772
 5 <212> PRT
 <213> Clostridium difficile
 <400> 86

ES 2 622 280 T3

Met Lys Arg Thr Lys Leu Leu Gln Arg Gly Asn Phe Phe Gly Asp Lys
 1 5 10 15

Asn Met Val Val Asp Glu Phe Asp Glu Gly Tyr Asp Asn Tyr Asp Phe
 20 25 30

Ile Asn Phe Phe Thr Gly Cys Cys Asn Tyr Thr Phe Gly Leu Lys Asn
 35 40 45

Asn Asn Ile Leu Tyr Gly Cys Gly Asp Asn Ser Asn Phe Gln Leu Gly
 50 55 60

Leu Gly Glu Asp Asn Thr Thr Arg Lys Leu Phe Thr Lys Ile Pro Asn
 65 70 75 80

Ile Ser Thr Asn Ile Lys Lys Val Ala Cys Gly Glu Ser His Ala Val
 85 90 95

Ile Leu Thr Ser Asp Gly Glu Leu Leu Val Ala Gly Ile Asn Thr Asp
 100 105 110

Gly Gln Met Gly Leu Gly Leu Glu Lys Val Gly Lys Thr Val Ser Thr
 115 120 125

Phe Glu Lys Val Pro Glu Ile Lys Gly Val Lys Asp Ile Ala Cys Gly
 130 135 140

Leu Gln Ser Thr Tyr Leu Leu Tyr Asn Asp Gly Thr Leu Tyr Val Ala
 145 150 155 160

Gly Asn Asn Leu Tyr Gly Gln Leu Gly Leu Gly Thr Asn Gly Ala Ser
 165 170 175

Ala Asn Val Asn Thr Phe Thr Lys Val Asp Val Asp Asn Val Lys Ala
 180 185 190

Val Phe Ser Tyr Asn Lys Ser Ala Phe Ile Ile Lys Asn Asp Asn Lys
 195 200 205

Cys Tyr Ser Thr Gly Phe Asn Asn Gln Gly Gln Leu Gly Leu Gly Asp
 210 215 220

Lys Asn Asn Arg Asp Leu Phe Ser Leu Val Ser Ile Asn Asp Val Lys
 225 230 235 240

Thr Ile Ala Cys Gly Ser Glu His Thr Val Leu Met Thr Tyr Asn Asn
 245 250 255

ES 2 622 280 T3

Asp Ile Tyr Gly Cys Gly Lys Glu Lys Cys Phe Gly Asn Ala Leu Gln
 260 265 270
 Ser Ser Leu Phe Thr Lys Ile Glu Glu Val Asn Ile Lys Thr Ile Ala
 275 280 285
 Cys Gly His Gly Asn Thr Met Leu Ile Asp Asn Lys Gly Thr Leu Lys
 290 295 300
 Val Ala Gly Asn Asn Asp Ile Tyr Gln Leu Gly Ile Ala Asn Tyr Ser
 305 310 315 320
 Glu Asn Ile Asp Asn Ser Phe Ile Asp Leu Lys Asn Ile Val Ala Lys
 325 330 335
 Asn Ile Phe Ile Gly Leu Ser His Ser Ile Leu Ile Asp Ser Asn Asn
 340 345 350
 Asp Ser Tyr Cys Thr Gly Asp Asn Thr Tyr Gly Gln Leu Gly Ser Phe
 355 360 365
 Phe Asp Asp Met His Ile Val Glu Phe Lys Lys Met Asp Ser Glu Lys
 370 375 380
 Tyr Ser Tyr Ser Asn Tyr Ile Asn Leu Ile Lys Ser Glu Asp Lys Leu
 385 390 395 400
 Thr Leu Leu Lys Glu Glu Met Glu Ile Lys Asp Ile Glu Leu Pro Leu
 405 410 415
 Asp Ile His Ser Val Arg Asp Val Val Phe Ser Pro Tyr Cys Thr Leu
 420 425 430
 Val Ile Leu Gly Asn Gly Asp Val Tyr Gly Leu Gly Asn Asn Arg Tyr
 435 440 445
 Lys Gly Met Gly Ser Asp Leu Pro Ser Gln Leu Asn Glu Leu Thr Lys
 450 455 460
 Leu Ser Ile Ser Asn Val Lys Ser Ile Val Ala Ser Lys Asn Ile Ser
 465 470 475 480
 Gly Gly Ile Phe Tyr Ile Lys Asn Asp Asp Thr Cys Tyr Tyr Ser Gly
 485 490 495
 Pro Asn Ser Asn Ser Ile Ala Gly Val Leu Pro Ser Asn Ser Asp Val
 500 505 510

ES 2 622 280 T3

Phe Lys Lys Ile Ser Ile Asp Asn Val Lys Lys Val Val Ile Asn Thr
 515 520 525
 Asp Leu Ser Asn Trp Phe Ser Leu Ile Val Thr Asn Asn Lys Gln Ile
 530 535 540
 Tyr Thr Ser Gly Lys Ser Ser Ser Tyr Val Asn Gly Leu Ser Asn Ala
 545 550 555 560
 Leu Ile Ser Gln Tyr Thr Glu Ile Ser Leu Ser Asn Val Thr Asp Ala
 565 570 575
 Tyr Ser Ser Tyr Asn Ala Thr Phe Ile Val Val Asp Glu Lys Lys Val
 580 585 590
 Tyr Ala Thr Gly Ile Asn Thr Asn Tyr Leu Leu Gly Phe Ser Thr Ser
 595 600 605
 Asp Gly Ser Asn Val Asn Leu Gly Leu Leu Ser Asp Trp Tyr Tyr Ile
 610 615 620
 Asn Ile Ser Gly Ser Ser Tyr Ser Arg Val Ser Cys Thr Asn Asn Ile
 625 630 635 640
 Thr Lys Ile Asn Asn Ile Ile Ile Tyr Glu Tyr Val Thr Val Phe Cys
 645 650 655
 Thr Asn Ile Gly Ser Phe Leu Thr Gly Tyr His Gly Thr Ser Trp Thr
 660 665 670
 Lys Pro Thr Asp Ser Ser Tyr Arg Val Gln Tyr Gln Gly Ile Ser Tyr
 675 680 685
 Ala Gly Tyr Leu Asp Ser Tyr Ile Tyr Asn Tyr Tyr Pro Thr Arg Cys
 690 695 700
 Thr Gln Ser Ser Ser Ser Thr Thr Phe Ala Tyr Leu Tyr Asn Gly Glu
 705 710 715 720
 Ser Ser Ser Asn Leu Lys Asn Val Asn Pro Asp Asn Leu Leu Ile Ser
 725 730 735
 Gly Gly Ser Ser Tyr Ile His Gln Tyr Gly Arg Asn Tyr Leu Asn Asn
 740 745 750
 Gln Ser Ser Asn Asn Ile Ala Ala Ser Asn Ile Asn Ser Gly Pro Ile
 755 760 765

ES 2 622 280 T3

Thr Ser Asp Lys Ala Ile Phe Leu Tyr Lys Ala Leu Leu Tyr Leu Ser
 770 775 780
 Ser Asn Thr Leu Tyr Gly Phe Gly Asn Ile Ser Glu Ser Ala Lys Glu
 785 790 795 800
 Leu Asp Val Ser Asp Thr Gln Asp Gly Tyr Asn Ala Thr Asn Tyr Lys
 805 810 815
 Lys Val Met Lys Asn Ile Lys Asn Ile Phe Ile Pro Pro Tyr Asp Leu
 820 825 830
 Ser Arg Asp Lys Thr Arg Phe Ala Ile Leu Thr Asp Lys Ser Leu Phe
 835 840 845
 Ile Cys Gly Tyr Asn Ser Lys Gly Thr His Gly Ile Ser Val Asn Ser
 850 855 860
 Ser Leu Asn Leu Asn Asn Lys Ile Asn Tyr His Lys Lys Asn Ser Ser
 865 870 875 880
 Ser Glu Ile Ser Ser Asn Ile Gln Glu Ile Tyr Ser His Ser Lys Ser
 885 890 895
 Thr Tyr Leu Leu Thr Asn Asn Asn Met Leu Tyr Ser Val Gly Leu Asn
 900 905 910
 Asp Val Gly Gln Leu Gly Val Gly Asp Glu Ile Asn Arg Lys Val Phe
 915 920 925
 Thr Lys Ile Asn Ile Asp Asn Ile Lys Ser Ile Asn Val Asn Arg Phe
 930 935 940
 Thr Asp Asn Ser Lys His Ala Phe Ala Ile Lys Asn Asp Asn Thr Cys
 945 950 955 960
 Tyr Ala Val Gly Leu Asn Asn Ser Gly Gln Leu Gly Ile Gly Asp Asn
 965 970 975
 Val Asn Arg Asn Ile Phe Thr Lys Ile Asn Val Glu Asn Val Lys Tyr
 980 985 990
 Val Ala Val Tyr Gly Asn Thr Ser Leu Leu Leu Thr Asn Asp Gly Leu
 995 1000 1005
 Leu Tyr Gly Ala Gly Asn Asn Gly Lys Gly Gln Leu Gly Leu Gly
 1010 1015 1020
 Asp Thr Thr Ser Arg Asn Ile Phe Thr Arg Ile Pro Ile Asn Gly

ES 2 622 280 T3

1025		1030		1035
Val Arg Asp Val Tyr Leu Cys	Asn Asp Val Ser Ile	Ile Val Lys		
1040	1045	1050		
Asn Asp Asn Thr Cys Tyr Val	Cys Gly Leu Val Asn	Gly Tyr Phe		
1055	1060	1065		
Gly Phe Thr Glu Gly Ser Ile	Ser Thr Phe Thr Lys	Ile Asn Ile		
1070	1075	1080		
Glu Asn Val Lys Ser Val Val	Thr Ala Gly Ser Glu	Ala Thr Phe		
1085	1090	1095		
Phe Ile Thr Asn Asp Asn Met	Ile Tyr Thr Thr Gly	Lys Lys Glu		
1100	1105	1110		
Arg Val Phe Phe Ser Thr Glu	Thr Asn Asp Ile Lys	Gly Ile Arg		
1115	1120	1125		
Val Ile Asn Asn Ile Ile Asn	Ala Lys Lys Ile Val	Val Asn Gly		
1130	1135	1140		
Tyr Thr Ser Ala Ile Leu Thr	Asn Asp Asn Lys Leu	Phe Val Gly		
1145	1150	1155		
Gly Leu Ser Gly Tyr Gly Ser	Ile Ala Asn Asn Asn	Asn Thr Asn		
1160	1165	1170		
Ser Val Glu Asp Val Lys Asp	Val Phe Val Thr Ala	Asn Asn Thr		
1175	1180	1185		
Leu Tyr Ile Asp Asn Asn Asn	Asn Leu Ile Ser Ser	Gly Arg Asp		
1190	1195	1200		
Thr Tyr Gly Ile Ser Asp Glu	Ser Tyr Arg Asp Met	Ser Val Pro		
1205	1210	1215		
Tyr Tyr Lys Val Ser Ile Lys	Lys Asp Val Asp Thr	Val Phe Ser		
1220	1225	1230		
Ser Tyr Asn Thr Ile Phe Ile	Lys Asp Ile Tyr Gly	Lys Phe Tyr		
1235	1240	1245		
Ser Ser Thr Arg Asp Asn Arg	Tyr Asn His Leu Gly	Ile His His		
1250	1255	1260		
Arg Tyr Asp Asn Asp Lys Asn	Glu Ala Leu Glu Gly	Ser Leu His		
1265	1270	1275		

ES 2 622 280 T3

Ser Tyr Phe Lys Thr Asp Asn Thr Ser Asp Lys Ile Val Phe Asn
1280 1285 1290

Lys Lys Asn Glu Lys Leu Val Met Phe Asn Asp Lys Tyr Ile Lys
1295 1300 1305

Thr Asn Asn Lys Tyr Ile Asn Tyr Lys Asn Ile Phe Lys Asp Asn
1310 1315 1320

Phe Lys Tyr Thr Ser Ile Ile Leu Pro Phe Glu Val Ser Asp Ile
1325 1330 1335

Asp Ile Ser Lys Thr His Ser Leu Ala Val Ala Lys Asp Gly Lys
1340 1345 1350

Leu Tyr Gly Ile Gly Ser Asn Ser Tyr Lys Glu Ile Asn Gln Thr
1355 1360 1365

Leu Glu Asp Ile Glu Leu Leu Thr Leu Thr Glu Val Asn Ile Ser
1370 1375 1380

Asp Val Lys Lys Val Ala Cys Gly Asp Asn Tyr Ser Tyr Ile Ile
1385 1390 1395

Lys Thr Asp Asn Thr Leu Trp Ser Tyr Gly Lys Asn Thr Glu Tyr
1400 1405 1410

Gln Leu Gly Val Gly His Asn Asn Asp Val Arg Glu Leu Gln Lys
1415 1420 1425

Val Thr Gly Leu Pro Ser Val Lys Asp Ile Ser Ile Tyr Asn Ser
1430 1435 1440

Met Thr Leu Val Leu Thr Asn Glu Gly Glu Leu Tyr Ala Gln Gly
1445 1450 1455

Tyr Asn Thr Asn Gly Leu Phe Gly Leu Gly Glu Ser Glu Lys Asp
1460 1465 1470

Lys Ile Ile Arg Thr Phe Thr Lys Val Leu Thr Asn Val Lys Glu
1475 1480 1485

Ile Lys Ser His Asn Asp Asp His Ile Leu Val Ile Lys Asn Asp
1490 1495 1500

Asn Ser Leu Trp Ile Thr Gly Lys Asn Lys Ser Met Tyr Lys Ile
1505 1510 1515

ES 2 622 280 T3

Ser Ile Ser Ile Thr Asp Leu Tyr Glu Phe Thr Lys Ile Pro Ile
 1520 1525 1530

Pro Glu His Leu Asn Asp Ile Leu Asp Ile Glu Leu Ser Asp Asp
 1535 1540 1545

Thr Ile Tyr Met Ile Thr Lys Val Asp Thr Ser Lys Ala Ser Ile
 1550 1555 1560

Glu Ile Val Glu Lys Ser Ile Ser Gln Val Arg Val Val Val Gln
 1565 1570 1575

Asp Pro Asn Asn Val Ile Glu Lys Leu Glu Met Phe Ile Asn Asp
 1580 1585 1590

Glu Leu Ile Ser Thr Lys Thr Asn Leu Glu Ile Asn Ser Ile Ile
 1595 1600 1605

Phe Glu Ile Pro Gln Asn Lys Ile Val Leu Gly Glu Asn Lys Ile
 1610 1615 1620

Leu Ile Lys Ala Ser Ser Pro Thr Gly Asp Leu Tyr Ser Ser Met
 1625 1630 1635

Phe Ile Phe Lys Ser Glu Thr Gly Leu Lys Val Lys Lys Asp Ser
 1640 1645 1650

Ile Leu Met Ile Asn Asn Lys Val Tyr Ser Ile Ile Asn Ile Thr
 1655 1660 1665

Glu Asn Asn Thr Asp Leu Ile Val Thr Leu Asn Glu Gly Leu Lys
 1670 1675 1680

Asp Asp Met Met Glu Asn Asn Pro Ile Tyr Gln Leu Ile Asn Lys
 1685 1690 1695

Thr Lys Val Gln Val Lys Ile Asn Lys Ser Asp Leu Phe Lys Asp
 1700 1705 1710

Met Lys Leu Val Glu Ile Lys Lys Ser Asp Ser Ser Tyr Gln Glu
 1715 1720 1725

Ile Tyr Glu Leu Glu Glu Ala Asn Ile Lys Ser Ala Gln Pro Lys
 1730 1735 1740

Ile Ile Val Glu Lys Gly Asp Lys Trp Thr Ala Ile Lys Arg Pro
 1745 1750 1755

Ser Met Ile Phe Arg Tyr Asp Ala Glu Asn Asn Glu Pro Gln
 1760 1765 1770

REIVINDICACIONES

1. Una bacteriocina de elevado peso molecular (epm) de tipo R aislada que tiene actividad bactericida, en donde la bacteriocina comprende una bacteriocina de tipo R de *Clostridium difficile*, en donde el polipéptido del dominio de unión al receptor (RBD) nativo ha sido reemplazado por un polipéptido RBD heterólogo de otra cepa de *C. difficile* o por un polipéptido RBD heterólogo de un bacteriófago que infecta *C. difficile*, en donde el polipéptido RBD de reemplazo crea una bacteriocina modificada con un espectro bactericida diferente del creado por el RBD nativo, y en donde el polipéptido de la región de anclaje a la placa basal (BPAR) es idéntico en al menos 80% a un polipéptido compuesto por 50 o más aminoácidos contiguos del SEQ ID NO: 16 ó 78.
2. Una molécula de ácido nucleico aislada que codifica la bacteriocina de epm de tipo R de acuerdo con la reivindicación 1, en donde la molécula de ácido nucleico codifica una bacteriocina de epm de tipo R en la que el polipéptido del dominio de unión al receptor (RBD) nativo se ha reemplazado por un polipéptido RBD codificado por una secuencia de polinucleótidos heteróloga de otra cepa de *C. difficile* o por un polipéptido RBD codificado por una secuencia de polinucleótidos heteróloga de un bacteriófago que infecta *C. difficile*, y en donde el polipéptido RBD de reemplazo crea una bacteriocina modificada con un espectro bactericida diferente de la creada por el polipéptido RBD nativo y en donde la bacteriocina de epm de tipo R comprende:
- (a) polipéptidos de los SEQ ID NO: 4-16, 18 y 19; o
 - (b) polipéptidos de los SEQ ID NO: 66-80; o
 - (c) polipéptidos que tienen una identidad de al menos 80% con (a) o (b).
3. Un casete de expresión que comprende una molécula de ácido nucleico de la reivindicación 2.
4. El casete de expresión de la reivindicación 3, en donde la molécula de ácido nucleico está conectada operablemente a un promotor heterólogo.
5. El casete de expresión de la reivindicación 4, en donde el promotor es inducible o reprimible, opcionalmente en donde el promotor es inducido por un inductor o desrepressor de molécula pequeña.
6. El casete de expresión de la reivindicación 4, que comprende adicionalmente un gen *recA* que codifica una proteína RecA constitutivamente activa y bajo el control de un promotor heterólogo que responde a un inductor o desrepressor de molécula pequeña.
7. Una célula capaz de expresar una bacteriocina de epm de tipo R codificada por una molécula de ácido nucleico que comprende el casete de expresión de una cualquiera de las reivindicaciones 3 a 6, o que comprende una molécula de ácido nucleico de la reivindicación 2 dentro de su cromosoma.
8. La célula de la reivindicación 7, en donde la célula es una bacteria anaerobia no obligada y no patógena, opcionalmente en donde la bacteria anaerobia no obligada y no patógena es una especie de un género de bacterias seleccionado del grupo que consiste en *Bacillus*, *Lactobacillus*, y *Listeria*, preferiblemente en donde la especie es *Bacillus subtilis*, opcionalmente en donde *B. subtilis* carece de la agrupación de genes PBSX.
9. Un método para producir una bacteriocina de epm de tipo R, que comprende exponer una célula que comprende una molécula de ácido nucleico de acuerdo con la reivindicación 2, conectada operablemente a un promotor inducible, a un agente inductor a una concentración eficaz para inducir la expresión de la bacteriocina de epm de tipo R, y purificar la bacteriocina de epm de tipo R expresada, opcionalmente en donde la molécula de ácido nucleico que codifica la bacteriocina de epm de tipo R es heteróloga para el genoma de la célula, y en donde la molécula de ácido nucleico está contenida dentro del cromosoma de la célula o está contenida en un vector de expresión extracromosómico dentro de la célula.
10. El método de la reivindicación 9, en donde la célula es una bacteria anaerobia no obligada y no patógena, opcionalmente en donde la bacteria anaerobia no obligada y no patógena es una especie de un género de bacterias seleccionado del grupo que consiste en *Bacillus*, *Lactobacillus*, y *Listeria*, preferiblemente en donde la especie es *Bacillus subtilis*, opcionalmente en donde *B. subtilis* no se lisa cuando es inducida para producir la bacteriocina de tipo R, opcionalmente en donde *B. subtilis* carece de la agrupación de genes PBSX.
11. Un método *in vitro* para la destrucción de *Clostridium difficile*, que comprende poner en contacto la bacteria patógena con la bacteriocina de tipo R de la reivindicación 1, por medio de lo cual la bacteriocina de epm de tipo R se une a, y destruye, la bacteria patógena.
12. Una bacteriocina de epm de tipo R de la reivindicación 1 o la célula de la reivindicación 7 o de la reivindicación 8, para su uso en un método de tratamiento de una infección por *Clostridium difficile* en un animal.

13. La molécula de ácido nucleico aislada de la reivindicación 2, en donde además del polipéptido RBD heterólogo de reemplazo, la bacteriocina de epm de tipo R comprende secuencias de aminoácidos idénticas en al menos 90% a los SEQ ID NO: 4-15.

5

10

Figura 1. SEQ ID NO:1-61

```

SEQ ID NO:1          > locus Dif16
GGCCGCAATACCCACTACACCTTCgTCATCTTTAAATTTAAGAGTTTTTACTATTGAATAAT
AAAGGTATATTCCAGTAAAAATAATCTTTAAATACAAGAAAAATAAACTCTTTGGGTATATT
AAAAAGCTAAAAAGTGTAAATATAAAAAGCAAGTAGAGTACTTATCCTGTAAAAGAAATCTAT
TTGTGTAATGCTTTTATATTTTATCATAAACACCCGAATATAAAATGATGAAAATAATTGCGA
CGATTGCATATATGGTAAAATAACATATTTTCAAGAGTACCATTTGAAATTAATCCACTTA
TACCACATAAATGGCCAAAATAATAGTGCTAAGAACTTAAAATAATTATCAAACAACCTTTTC
TTTATACATTTCATCAAACAACCTTTCTTAAACAAAAGCATATATTTGTTTTTAGAATTTTAAA
TAATATGATATCATTATATATATTAATATTGAATTTATAGAAAACCAAAATTTGTTAAAATA
AATATATAGATTTTACTGTTAAGCCAGTTAAAATTACTACTATTTTATATGAAATTGGAT
CAAATATGTAGAAAATACGGCAAATTAGTTAATATTAATATTTATTTATTTCCAAGTTGATAA
GACTGTTTTTTTTAATGATAAAAAATCTAATCTTTTTTGAAGAAAAGTAATATCCACATTAAG
TATGCTGCCATTTTATAAACGCAAGTGATGCCAGAATTAATTATGTTTATATATCTTCTT
CAGTAATTAAGAACTCACAAGCCCATTTTAAAGGCTTTATTTTCGCACCTATCTATAATAATT
TTTGATAATAATCGTTATAAGAGGATACATAGTATCCAAGGCTAGTGAATGATGTCCAAG
TTCTTCAGCTAAGATGGATGCAATTTTTTTGAGTTTTGTTTTAAATTAAGTGAATGATA
TAATTTTAAATACCATGTTTGTATATATAGCCCTTCTAAATCACCTGCAATATAAGTGGTA
TAATGAATTATTATCTCTCTTGAGAAAGCTAATTCAAAAGCTTATCCAAATTAATCATAAA
AATCCCCCTAAAATAGAATGTATGTTGCCTTTAAATTATATTAAGAGCAGAAAAATAGA
CTGCTCATCATATGGTTTATTTTTTTTATATTTATTTAGTAAAAATCTATATAATCATT
AGTTGTTCTTGTGCTTCTTCAGGTAACCTCTTCATGTGGATTTTTTCTATGTGCAGCTACTGT
ATCAATATTTTCTTAACTAAGGTTCTTCCAAGAAGGTAATCAACTGATACATTAATACAT
CAGCCAATTTGTTTAAAATGTGTTTATCAGGAAATCTGTTTTCTGTTTCATAGTACCCTAAG
ACTCTTTGGGAAACGCCTACTTTTTCTCCAAGTTCTCTTTGAGTCAATCCAAATTCCTTTCT
AAGTTCTCTTAATCTTTTGGCAAACATTATAACACCACCTTATGTATAGATTATAACAAATT
GTTCTAAAAAATAAACTAATAAATATAAAGAATATTTTTTCTAAAACTATTGATAAAG
AACAAATGATTCTATATAATCTAAGTGTGGAAGAACAATAATTTCTTAATGGTAATGGAGGT
ATAAAACAATGTTTAAAATAAAGTGAATATATAGAAAATGCAAAGGTATGACACAAATT
CAACTTCCGAGAAAGGCTGGAAATACAAATGACTATATATCTCAAATAGAAAGAGGTATAAA
AAATCCTGCTCTTCTTATGGCTAAGAGATTTCTAGTATTTTAGAACAAAATATAGAAGAAG
TTTTTTTTTATACAGTTATAGAACAAATATGTTCTTGAAGTTGTGAGATTAGTAAAAAAGTGT
GCCTAAAGAGATTATTGTAATTTGAAGCTAATAATAAGTATATAAAAAAGGAGAAAGTACT
ATGGAAAACAAAAAGATATATTTTAAAGAAACAGATGAAAGATTACATAATTATAAGTA
TTTGGATATAAAGATAAAGAATATTAATTTGGACATAAAAAGATGTGAGAAATGAATACTCTG
GATGTGGAGCAATGGTATATACAGAAAAGACTAGTAACACATATAACATAAGTTCTTCTGTG
GAAAATGAGGTGTTAAAAGAGAGGAAAGATTAAGAAAATTTAAAATGGAAAAAGAGATAT
AGAAAATAGAAAAGAGAAAGATAGAAAATGCTCTAACATGTCTAAATGATATAGAAATGGAA
TTTTAATCTTTTTTATAATAGTAAGACAAAAACAATATGACATATATTTCTATGAAACTA
CACTTAGATAGAACATCTTGCTACAATTTAAAGAAAAAATGATATTTAAATGAGTGAGAT
ATTATAAAAAAATATGACAACCTTTACAACACTTTATATACACTATTGCAACACTAGGCAATAA
AATATGTGAGATAATGTTATTGTGAAAGAAATCCATATTGAAGGAGGTGATAAATGAAAAG
AATAATATTACCTAAAATATAGAAGATATTTGACAGGAATAAATGAGATGTATATTTAAAA
ATGACTTATATCATTTATAGTAAGATTATCAGATTAAGCAAGAATATTTAGTGATAGTGTGG
TGATTATTTGCTTAAATACAAGGAAATATTAGAAACAATTTAGAGATTTCAAAAAAACT
TFACTGAAAGTATTTTTATTGATGATGAAAGTGTGCAAGGCTCTGAAGGGTCTGTTTTTTT
GTAAGTATACTATCAGTTATTTGTACACCTATAATGTTAAATACGAATAATAAAGATATTGT
TATCTCTATAAAATACTTACCAAACCCACAGTCAAAGAGTATTAGAATGTATGAAATTCAG
ATGAATTAATAAGTTATTCACAGAAAATAAAGGTAACAGACAGAAAATTAATATAACA
    
```

AAGCTAGAACAAAGTATTAATAAAGAAAGAGTCAATTTATGTATTGAACTTTACAATTACACT
 AAATTATCTGGATAGTGTATATGAAGAAGATGTAGTATATGAAAATATGGAAGAAATCAATT
 TAAATTTAGGAGAGTGATAGTATGGCTATAGGATTACCAAGTATCAACATATCATTTAAGGA
 GCTAGCTACAACGTAAAGAACGTTACAGCTAGAGGAATAATTGCAATGGTCTTAAAGATG
 CTAAGGCACTAGGTCCTTAAATGAAATACATGAAAAAGAGGATATAACCAGTTGATTTATCTGCT
 GAAAATAAAGAGTATATAAATTTAGCTTTGATGGGAAATGTTAACACTCCAAATAAATTTATT
 AGTTTATGTAATAGAAGGAGAAGCAGATATTCAAACTGCATTAGATTTTTTTAGAGACTAAGG
 AATTTAATTATCTATGTATGCCAAAAGCAGTAGAAGCTGATAAGACTGCTATAAAAAATTTGG
 ATAATTAACCTTAGAGATATAGATAAGGTTAAGGTTAAGCTGTATTAGGAAAAGTTGTAGG
 AAATCATGAAGGGATAAATTTACTACAGAAGATGTGTTAGTTGGAGAAAAGAAATACA
 GTGTTGATGAGTTTTACAAGTAGGGTGGCTGGACTTATAGCAGGAACACCTTTAAGTCAATCA
 GTAACCTTATACTAAGCTTAGTGATGTAGTTGATATACCTAAGATGACGAAAGTTGATGCAGA
 ATCAAGGGTTAATAAAGGAGAGCTTATACTTATTAAGGAAGCAGGGGCTATAAGAATTTGCAA
 GAGGAGTAAATTTCTTAACTGAGTTAACAGAAGAAAAGGAGAAATGTTCCAGAAAATAAAA
 ATAGTTGACACTTTAGATATTATACATAGTGACATAAGAAAGGTGATAATAGATGACTATAT
 AGGAAAGGTTACTAACAGTTATGACAACAAATGTTTATTGATAGTAGCTATAAAAAAGTTATT
 TAGAAGAATTAGAAAAGTCAGCACTTATAGAATCTGATTCTACTGTTGAAATAGATTTTTGAA
 GCACAAAATCGTATTTAAAATCAAAGGAGTAGATTTATCTTATATGACATTCACAAGAAAT
 AAAAGAAGCTAACACAGGTTCTAAAGTATTTTTAAAAGCAAAAATAAAAGTACTTGATGCTA
 TGGAAGATATAGATTTTCAATAGAAAATATAGGAGGATTATTAATATGGCAAATATGGAAGC
 TAGAAATGTAATGAGTGGTACTTTGGGGAGAACTTTGGCTTGATGAAACAAAGTAGCAGAA
 TAAAGAAGTTTTCAAGCAAAGATGGAATTTACAAAAGAGGATATTATAATAGCAGGTCAAATG
 GGTACTGATACAAAGTATATGGGATATAAAGGAAAAGGTTCAATAACTCTATACCATGTTAG
 TTCAAGAATGCACAAGTTAATTGGAGAAAAGATAAAGAGAGGTTCTGAACCTAGATTTGTTG
 CTATATCTAAATTAATGACCCAGATTCTTATGGAGCAGAAAGAATAGCAGTAAAAATATA
 GCATTTGATGATTTAAGCTTTAGCTGATGGGAGGTTGCAGTAAAAGGAGAGATAGAAGCACC
 TTTACATTTACTGAGTATGATTTTCTTGATATAAATTTAGTTTTATATTTGGTTTTTACTG
 ATATTTAGTAGATATATACTTAAATAAATTTAGGTAGTTAATAAGTAAAAAGTTAGTTGATT
 GAATTTGATTGATAAAGGAGCAAATAAATAATGAATGAAAATGGATTATCAAAAATAATAAAC
 ATAGTATTTACTTTTAAATGCAGATACAGAAAACCTAGAAAAGACCAAGTACTATAGTTGA
 ACTTAAGAGATTATCAACTATATTTGGGCAGGAATTTAAAGTAATGTGTAGAGCTTTAACAA
 TAAGTAAAAGATGAAGAGATACAAAATACTTGTCTTAAAATTTGATGAAAATATGAAAACGGAT
 ATAGACTTACCGGAGATGCAGATGCTTACAATTATAGAAGGTGTTTGTGATTTGGATGGAAA
 GCTTTTATTTAAAATAAAGGAACATAATGGATAAATTTAAGGCTCCAACACCAAAGAATTGG
 CAAGAAAATATTTACCAGGTGAAATTTACCAACCTATATAGAATACTTCAAGATGTTATG
 GGTATGGTAAAATGCAGTGATAGAAGAGGTAATAAAGGACGATACCAAGACT
 ACAATAATGTACTATTATTGGAAGAAAAGGTTATAAGACCGTCCCTTTTTTATGCAATGGA
 TAAAGGCGAATTAAGCTTATTGAAGCTTTTTTCGCCTTAGAAATTGAGGAAGAAGTTGAAA
 AAATGAAACATGGATATGGAGTGTGTCCTTTGACAGGAGGTGGTATGTAATGGGAAATGTGA
 GAGAAGAAGGTATAAATATGTATCTTACAGATAATTACACACCAAAAATGAACCAAATTTATA
 TCAGTAACTGATAAATTTAGGAGAGCAACTGTGGCTGTTTCACTTTCCACTAATGTAATGGC
 TAGTAGCATAAAAAATTTCTATTGGAAGTGCAAGTAGTAGAGTAAACAGTTTTAAATTCCTCGT
 TAAGAAAAGTTCAAACACTGCTAGTAGTGAAGTTCAACTATGGCAAATTAAGTTCTAGC
 ATAAATGCTGTTTCAGGAGTTATTGGAAGTTTAAATGGAAGTATTATGAGACTAGCAATAAC
 TATAGCTATGATTATTGATTATTTTAAATAAGTTGATTCAAAGAAAATGAGTTTAAATCAA
 ATATTATGATTATATTAATTTTAAAGCTAAAAGTGATGAAAGTAGAAAAACTAAAAATAAA
 TTACTTTGAAAATTTAAAAAGATTGGTGGCAAGATTTGGAATATCGTAATAAAGCAAAGA
 TATGACTAAGAGAGTGATAAGTAGTATCTTTGGGAAAATTAACAGAGTAGAGAAACGTCCTT
 ATCAAGGAAGTATTAATCTTAAAGATATGGTAAGTAGTGCTATGGCTAGAATTTTGCCTAAG
 TTAATGTTGTTTAAAATACTTTTTGGAGTGGTGAATAGCTATAAAGATATGGCAAGTAG
 CATTATAAGTAAAGTATTTCCCAAATTGAGATTGTTTGCAGGTAAGGTATGGAGTGGTGCAA

TAGCTGTAAAGGATATGGCAAGTGGAACTTGGTTCGATAAAAAGGGAAGATATCTGATTTG
 ACAAATGGTGTACTATAGGTGTGCTGTGAAAAAGGGTGTGACTTACTTGGTCAGGAACA
 AAATCAGAAAAGTTGTTCTAGAAAAGTGAATGAAAAGAAATACTGGAAAACTAGCCAAAAAG
 ATGTTGATAAGTATTATGACAGTTTAGTAAATATGGCAAATGATACGCCTTTTGACCCTGAA
 GATGTTGTTGCAATGGGAAGTAAAGCTAAAATGATTAGTAATATTACTGGTGGCAAAAAAGA
 AAAAGATATAACTCAAGCTATGGTAGATCTTAGAGCTTTAAATATGAATACAAGTAGTGAAC
 AAGATGTATCAGCAGCTTTCTTAAGTGCAGCAAAAGGAAATATGGAATCTCTTAATACTCTG
 GTAGGAGAAAAATTATAAACTTTTGATGAAGCATFGGAAGGCATAAGTGTAAAGCAGATGGG
 GTTAGCTAAAGAAATGAGTAATACAATACCAGGTATAATATCAGGAGCTCAAACAAGCATTA
 ACAATGGTTTGAAGAGTATTGTTAAACCTTTTGATGATATTTTAGGTCAAGGACTAAAGAAA
 ATAAAAACTTTTATAGAAAAGTGGATTAGGGAATTTAGCTGGCTTATCTGAAAAAATGGCTGG
 TAAAAATAGGCAATGTAATGAATGGTAAGATAATTATFGGCAACAAATATGACCAGATGCAAT
 CTAGAAGTGTAAAAAATGGAAAAGAGTTTTCTGATTCTACTCAATATCGAATTTCTAATGAG
 GCTGAAAAGCGTAAAATGATGGTTGAAAATAAGCAAGAACGTTTTGAAAATCATGCAGCAAC
 AATGATAGGGAATGCACCAAAAGCAATTGTTAACGCAGGAAGTACACTATTACAAAATATTG
 ATTTTACAGCATTAATAGATTCACTACTTCCAGTAGTAAACTTAGTAAATAATTTACTAGAT
 AGTATAAAACAATAAATCACCAATTCACAAAGGATTAATAAGTATATTTGGTACAATAGTAAC
 TACAGCATTCACACTAATCGGACCTGTAGTTGAAGCTGTAGTCTTATATCACAGAATTT
 TTAGTTTTTTAGGTGAATATGCACCTCAAATAACAATTTTATAGAGACACTGGGTGTTATT
 TGGAAAACCTGTATGGGAGACCTTAGGACCTCTGTTGGAAAACCTGGATGGAAAATTATAGAGCC
 AATATTTGGGAGCTTTTTTAAACATATTAGATAAAAGTATGTAAAATAGTTAAAGATATATGCA
 AATGGTGGCAAACTATGATTAATAAGATAAAAAATGGGAAGCATCACAGGAACAGTTTTAAAT
 CTAGTGGAAAAGAGTAAAAAATTACAAAGATAATCCATATGCTGGAACAAAGGCTGGTGA
 TTCTGGTAAAGCTTATTCAAGTAAGAAAGGTAATAATGCATTTGGATTGAACTATGTTCCCTT
 ATAATGACTATCAAACCAGACTCCATGAAGGTGAAATGGTTTTAACTAAACAAGAAGCAAAT
 CAATATAGAAGCAGAAAAAATGGTGGAAATATAAACATAGCTAAGTTAGCTGATACAATAGT
 GATTAGAGAAGAAGCTGATATAGAAAAGATAACATCAAATTAGTTGCAAGTATCCAATTTGG
 CACAGTTAGGGGGTGTCTTATAATGGAAATGTGGCTTAGACAAGCAGAAGATAGATTTAGAT
 TTCCAGTATTTCCATCTTCTTTAGTATTAATGGAAAAGCTGCTGTAAACTCTTCTAGTATA
 CTCAAAATAGGTGAAGTAGCAACTTTTGGTGGTGTAGTCTTAAAAGCATTTCAATATCAAG
 TTTTTTTCCAAATAAAGACTACACTTTCTGTGACTATACAGGTTTTCCATCACCATATGATT
 GTGTAATAAGATAGAAAAATGGATGAAGGAAGGTTTTATATTAAGATTTACAATTACGGAA
 ACAAATATAAATATGGAAGTCATAATTGAAGGGTTTAGTTATGAAGAAAGAGATGGGAC
 TCGAGATGTATTTTTACATTAGATTTAAAAGAGTATAAAAAGAATAAAGATACCAAAAGTAA
 CTCAAAACAATAACTATTATAGATAATAAGTTGTAAGTAACTGCTGATAGAATTAATGAA
 AAGGCAGGTGATTTTTTATTATTAAGATTTGGGTACACATAAAAAACGGAAGTATATATGAC
 ATAACCTGACATAGTAGACAAGGTATCATGGTCAGGTGATTATAAATCTCCATCAAGGACACT
 AGAGTTTTCAATAATACAATCATCTTTGATGTAATTTCCAACAAATCGATATACCAATAG
 CTAGTACAGTCTGTTCTATGTAGATGAGAAAGAACTCTTTAGAGGAATGATAATTAATAGG
 TCTAAAGATTTCAAGCAGTAATGAAATTTAGTTTTGTATCTAAAGATATGGGATTTTTACTTAC
 ACAAAGTGAAGTGTACATAAATTTAAAGATAAGTTAGTTGAAGACATAGCAAAGCAAGTAT
 TTGCTGAAAATAGGCTTTCAGTTGGAACAATAGCAAAGACCAATGTCAAGTATACAAAGATG
 TTTATAGGAGTAAATGGTTATGACACAATAATGAGTGCATATACAGAGGCAAGTAAAAAGAC
 AAAGAAAAAGTATATGATAGAGGCTAATTTAGATAAGTTAATGTTATTGAAAAAGGAACTG
 TTACATTAAGTGTATGTTTGAAGAGGGATTTAATATTATAAATACCACCTTTTCGGAGAGC
 ATGAAAAATGTAAAAAATAAAGTAATAGTGGTAGACCAGTATCGAAGCAAGATTAGCGAAAA
 AATAGATAATGAAATTTTAAAGGAAGTAAATGTAATAATGCAAAAAAGTAAATTCAGCAACAAG
 AAAATCAAGATGTAGATATTGATAGCGAGTTAATGGGATAGAAAAAGCTGTTCTCTTAAa
 GGTATGGAGATGTAAGTTGTATAACTGGTAGAGGAGTAAAAGTTAAAGATTTCTTATACAAA
 GCTTGTAGGACTATTTTATATAGATACAGACAAACATACTTGGCAAAATGGAGAATATCAA
 TTGAGCTTGAACCTAATTTTTCAAATCTTATGGATGAAAAGTCAGCAGGACAGGATGAACCT

AAGGAAGAAAAGTAATTTAGGGGGGAGAAGATTATGCAGGAGGAAAAGAGTTTACAGCAGAATT
TACAGCTTACTGICCTAGAAAAGAAGAAGGTGGAGATACAGATTGTAGAAAAGAAAAAAGCTTG
ACCCATCTAAAAAACTTGGCCTGCTCCTATGGTTGGTAAATATGAGCAAAGCTTATTATACAA
AAGAGTTTTTAAATAAACATCCTTTATTGAACTATGGAGATGAAATACAGGTAATTACAGGA
GTTTCTGGTCGTGATGGAGTCTATAAAGTAAATGACGTAGGACCTGCAATAACTATAGAAAA
AAATGGAACATACCATATAGATATTTTATTTGGAAATGTTGAAGAAGCTAGTAAATTTGGAA
GAAGAAAAGGAAAAATTATTATTGGTGGTTATTCTGGTAATGTATCTGATAAAGCTAAAATA
GTAATATCAGAGGCAAAAAAACATCTAGGTAACCTTATAAATGGGGTGGAAATGGACCAAG
TAGTTTTGACTGTTCTGGTTAATGGICTACTGTTTTAAAAAAGTTAATGTTAGTTTGCCAA
GAACGTCAAAATCAACAATCTAAAAAAGGCAAGAAAGTAGAACAAAAAATCTTCAAGCAGGA
GATTTAGTATTTTTTCATAATCCAGTCAGCCATGTTGGATTATATATAGGTAATGGAGAATT
TTTACATGCTCCACAAAAAGGTGATGTAGTTAAAATAAGTAAGTTAAGTAGTAGAAGAGATT
TTAATACAGCTAGGAGATTATAAAAAGGATGGTGATATAATGGCTAATCCAATAAATGAA
TTTATAGGAATAAGAGAGAAGGAAAGTATCATAATCAACCTTCTTTTTTATTGGAAAA
TTAAAAAGTAAATTACCAGATTTAAAAATAGAGACAAATAACATCATATTAGAAAAAGAAGAT
ATTTTGATAGATAGTTGGATGATGATAGACAGCTAGAAACATTTGACACAGAAACAAATCA
AGAACACCAGCATGAAGTAAAAAATCCTTTTATAGATAACTTTGAATCTGGGGATATGGTAA
TAATGTTTAGAATAGGCGAAAAATTTGCTGTTGTAAGTAAGTTGGTGAGCTTATAATGAGTA
CAATATTTCTTTTATAGGTGTCCAGAGGATTATATCTTACCTAAAACAGAAGAATTGCCA
ATCTTTCTGTAAGTGGCATGGGATTTTGAAAAAGATGAACCTATTTTAGAAAAAGGTGACTT
TAAAAATAATTGAAAAAAGAAAGCCTTAAAAGTTTGGATATACAAGTGTATAAAGACAAAT
AGATATGAACATGAGATATACTCTTTAGAATATGGGACAGAGCTTTTCAGAACTAATAGGACA
AAAAATACAAAAAGGTCTTACAGAAAAGTGAAGCTAGTAGATTATAAAAAGAGGCCCTTCTAA
TAAATCCATATATATAGAAAGTAAACGTAAAAAGTGAACCTTTAACAGAGACGTATTGAGT
GCAAAATGTAAGTATCCACTATCTATGGGGAGGTGGAATAAATGTATAGTGACCAGACAT
ATGAAGTAATAAAAAATAGAAGTCTTTGAAAAATTAATCTTTGATATTTATAAAGGAGAAGT
TCTTTCTAAACAACATGGTATCTGGAAATAATCTAGAAGCTTTTCGAAGATATATCTAGAAGT
TTCAAAGATACATAAAATGGCTTTTATACAAGACACATATAACCAGTTTCTTGATAAAAAGAG
TCAATGAATTTGGTGTATATAGAAAAGTTAGGTACAGAGTCAAATGGAGAAGTTGAATTTATT
GGAGAGAAAAGGAACTGTAATAAATAATGGCACAATAATATCATATAGAGATTTACTATTTGT
AGTAATAAAAAGATGTAACATTTGGTAGTGAAGAAGGTGACAATAGCCCAGTTCAAGCTCTGG
AAGTTGGTAAAGAAATATAATTTACCTACAAATTTGTAATTTAAACTAGTTGATAATATATCT
GGAGTAACAAAGATTACTAACACAAGAAGTTTTGAAAGGTGGTACAGATATAGAGACAGATGA
AGAACTAAAAGAAAGATTTTATAAAAATCCAAAGAAATCAAGCTACAAGTGGAAATAAAGCTC
ACTATGAAGAAATGGGCTTTGGAAAGTAGATGGAGTCTATAATGTTAAGGTTTATCCAAGATGG
GATGGTCCAGGAACAGTTAAGGTCTTGATATTTGGGGAAAATAATCAAGCTGTTGATACAGA
AACGATTTGAAAGGTGTCAGCAACATATAGATGAAGAGAAGCCTATTGGACCAACTATAACAG
TTGTGACACCATTACCAATAGAAATAAGTATAAGTGCAGTAATGAACTAGAAGATGGATAT
ACATTAGACAATGTAAGAAGTCTTTCTAGAAAAGTATAAATACATACTTTAGAGATATTAG
AGGAGAGATAATCTATACAAAAGTCATGGGAATACTTATAAATACTACTGGTGTACACGATT
TAAGTAATCTACTTATAAATGGAAAGTACAGATAATATAACTATTAATGAAGATAAAAATACCT
AGTGTAACAAC'TGTTAATTTTAGTGAGGTGGAAAAATCAATGAAGCTAATTTGATAAACTACCA
TCATTTGATAGAAATACATTTGTAGAGGAGATACAAGGTGCATACGATACAGAATTAATAT
TCTTAAAGAAGATATTTGATGATACCTTTAACCAATTTATTTGTTGATACAGCGACATGGGGAT
TAGATATGTGGGAAGACATACTCTGCATTTGAAAAAAGAAGCTTGAATTTTACACAAGACGT
AGCAATATAAAAAGCTAAAATGAGAAGCAGAGGTAAGTACTATTGAAGTTATAAAAAGTAT
ATGTGAGGCATATACAAAATCAGAAACAGATATAAAAAGTTTATAGTGATGAATTTACATTCG
TATTGAGTTTTATAGCAAATAACTGTGACTATAAAAAGCTTTTTAGATTGTAGCGAGATGATT
GAAAGAGTAAAACCTGCTCACTTATTACACTATTTAGAACCAATAATACTAGATAAAAAGTAT
GGTCTATTGTGGTGGAGGTATGGTATGTAGTGAAGAGGTAAAAGTTCATCCATACTTTGAAC
CAATTATAAAAATGTAGTGCTGTTGTAACCTGTGGAGCTGGAATGTTAAGTAGAGAAGAAATA

AAGGTTTATCCTTTAAGCATTAAATGCATTGAAAAATAATTGTAAGATTAATATAGCTATTGC
AAATGATACAGGCGTAGAAAAATGTAGTAGTTTATCCTAAAATCGGAGGTGGTATAAATTGGAAG
AAAAATTTTATATAATATTAACCAAAAATTTGGTAGAGAAAAAATAGCAAATGCAACTGCACTA
GGAGAGCTTGTGGATTAACCAAGTTTCAAGTTGGAGATAGTAATGGAGAATATTTATGAGCC
AACAGAGGAACAACTGCTTTAAAGAATGTAGTTTGGGAAGGAAATATAAATTTCTCTAAGAA
TTGATGAAAAAATCCTAATTTGGATAGTTATAGAGACTATTTTACCAGGAACAGTTGGTGGAA
TTTATGATAAGAGAAGCTGCTGTCTGGATAATGAGAATAATATAATAGCTATWGGTAAGTA
TCCAGAGACGTATAAGCCACGTGCTGAAGATGGCAGTATTAAGATTTGGTTGTAAAAATGA
TTTTACAATTTGTCCAATACTTCAAAATGTTACATTAGAAGTAGACCCGACGTTGGTTTTTGT
ACTCAAAGGATATTTCAAGATTTAGATGATAAGTTTGATAAAAAATATAAAAGAAATAAAAGT
AAAAATTTGGAGATACAGATATATTAACCTACAGATTTCTAAAGATTTATCAGGAGCTATAAATG
AGGTAGTTAAAAAATAGAAAAATATCTTTTGGATGATGTTATAAGTGGTCAAATACAACT
GATATATCAGTATTTAAAAAATAGCTATAACAAAATTTATCTGAAAAAGTCTAGATATATTAAT
ATACCTAGAATTTAGAGTCAGAAGTAACTGTAGATGAGGCTGGTTATTGGTATGATACATTAG
CAAATGGAAATAACATAGTAGCTATAGAAGGGCTTAAGTTAGATTTAAATAGAAAAATGTATA
ACAGGTGAAATTTGGTAATGTGATTTTGTAGAGATGTAGTATTACCATTTAGTGCAAATAGAGT
TAGATATATACATGATATGGATAATAACTTTGTTGAGACAAAATCTAGTAACACTTATTTAA
AAGAACAAAAAGATATAACTCTAAGTAAATATTCATATGAAATAAGATAAAATAAAGGAGGTA
GTACTAATAATGAAGCAAAAATAAATTTTACAGCGTGGTCTTATTTTAAATGATAAGAACAT
ATTTGATTTGATGATTTTGTATAAAGATATAATGATTATGATTTTGTAGAAATTTTACTGGTA
TAAGTAATAGTACCTTTGGTTTTAAATCAGATGGTAATTTATATGCTTGTGGCGATAATACA
GGTTTTCAACTAGGACTTTGGAAAAGATTCGTGAGAGAGAAGGATGTTTAGTAAAGTAAAAAT
TGATAATGTA AAAATATGTATCTTTGTGGTTCAAAAACACAGTGTAGCAGTAACTAAAGATGGAT
TTGCATATGGAGCAGGAACAAGTAATGTAGGTCAATTAGGTGTAATTGAGTCTACAGTATAT
TATGAATTTACTAAGCTACCAATAGATGATGTA AAAACTGTTGCATGTGGTTATGACTTTTAC
ATTTGTGCTTAAAAATGATGGAACATTATATTCAGCAGGTTTAAACTCAAGTGGTCAACTTG
GACTAGGTGATACTAACAATAGAGCTACTTTTACTAAAGTAAATATAGATAGTGTGAAAGAT
GTAGTGACTTATAATCAATCTGTATTTATCATAAAAATGGATGGGACAGCACATGCATGTGG
ATTAATTTCAAATGGGCAGTTGGGAATTAATAGTACTTTAAATAAAAAGTGTATTTAATAAAA
TAGAAGGTATGGATAATGTAAAACAGATAGCGTGTGGTAGTAGTCATACAATTTCTTATTAAG
AATGATGGAACATATGTATACTACAGGCTATAATGGAGTTGGTCAGCTTGGTACAGGAAATAA
TAATAATTTCAATTTGATTTTACTCTTTCTAGTATAAATAATGTTAAGTATGCTTCTTGTGGAA
ATAATCATACTATGATATTTAAAAATACGATAATACACTGTTTAGTACAGGACAAAACAATTTAT
GGTCAACTAGCCAAATGCCAAATAAGATGTAGCATCAAGAAATACTTTTGTCAAGGTTAATGT
AGAAAATATAAAAAGATATTAATTTAAATTTTCAATTTTAAATTTTAAATAAATGGTTCAAAA
AGATATTTGTATCTGGCTGTAATTTAGCAGGTCAACTTGGTTTCAATTTTTCATACAACCTTTT
CTGTATGAGTTTTCAAATGTGCAATCTTCAAATTTAGATAATTATTAGGTTTTATTGGTTAA
TGATGATTATTTATATGTTTACAAAGGACAATAGTGAATTTTAAATGTAAAGTTAAGTGATA
ATTTTCAAGATTATAAGAAGATAGAGTTAACAGATAGCAATATGTTTATTGTTATGAATGAT
GGTACATTGTATGCTTGTGGTTTTAAATAATTATGGACAGTTAGGATTTGGGAGATACTGTTAA
CAGGTCAAGTTATGACTAAGGTGGATATAGATAATGTTTGGATATAAAAAGGAAACGGAAACT
CAACTTTTGTGCTTAAAGAATAATGGAACATTATATTCATGTGGTTTAAATAGTAATGGACAA
TTGGGTTTAAAGAGATGAAGTTAATAGAAAATATATTTACAAAAATAGAAAATAGAGAATGTAAA
GGATTTTGTGTAGGAAGCAATTTATGTCATAGCTTTAAATCACTCAAAGAAGTATATGGAT
GGGGAATAAATCCTTATAATAATATAGAAAAACTTCTAATTATCCATATAAGCAGGGGAATA
AGTAAATTTGAAAAGATAGCAGCATATGATTTCTGTATATATGATAAACAGTGAAGGGAA
ACTATATGTTTCTGGATACAATTATAAATTTATCAATTAGGTAAAAGGAAATAATAGTAACCAAA
GCAAAGCATTAGTATCTCAATGTAGAACAAATTCACATCTTCTACATCAAATGGACTTAGA
ACGTTACCTAAAATAACTAATGTTTTCTTTTATGATGGTTGTGCAATAATTGACGAAGG
AGGTTATGTTTTTAAACAGGATATCATGGATTTTAAAGACATTAATAGCAGTCCAAGTA
TATCTGATTATTTCAAGATATGGAACTTTTATTGAGGCTACAAATTTCAAATCATAATACTTAT

TTTATACAAGAGACTGATTTTAGTGGAAATTGAAAAAGTAATAGGGATGTCAAATAATATATT
 ATTTTTTAAAGAAAGGAAGTTCATATATTTACTGGATATCCAAAAACATTTGGCTCAACCATTA
 CTGGACATAGAAGTTATACTAGTATTAATTTCTGAGAGTTCTAATTTAGGAAGTAATTTTATA
 ATATATCATAGTAATTTCCAAGTTATATGGAAAAGGGATTGCTAATAGTGGGCAATTTGGGAA
 TTCAACAAATATAGATGGCACAAGTAACTATGATACAGGATTA AAAAGACATAAAAAGATATAA
 TTGTA AAAAGGAAATACTGTAGTAGTAGATAAAAAATAACAATATATATGTAACAGGAATG
 AATCAGAATAACAACTTTGGGATAGGGGAATATAACAACGAACCAGTAAAAAAATTCACAAA
 TATAACTGAACAATCAAACCTCATTTATATTTATGGATGATATAAAAAGAAATTACAACATCAA
 GAAATACAATGTTTTATAGTAAAAAATGATGGAAACAGCCTATGCCACAGGAAATAATAGTTCT
 GGACAATTAGGATTAGGTGACACAATAAATAGAAAATAAGTTCACTCAGATAAACCTTGATAA
 TATAAAAGAAAATATCAACAAGTATAGATGGTAACACAACATTTGCAATTAGAAATGATGGAA
 CACTATACTCCACAGGATTAATAACCAAAGGACAACCTGGGATTAGGTGATATAGTAAATAGA
 AATACATTTACCAAAGTAAACATCCAAAATGTAAGAGATGTTGTTTTAGGGACTACTCACTC
 GCATGCAATCAAAGATGATAACACATTTATTTATCATGTGGAGaAAACACTCATGGGCAACTGG
 GCTTAGGAAGCGAAAAGCAACCATCCAGACGTATTGACATTTACTGTAAACAATATAACTAAT
 GTAAGAGATGTGTA CTGCTCAGATACAACAACATTTATTGTA AAAAGGACACAAAACATTGCATA
 TTGTTGTGGATACAATAATAATTCACAACCTAGGTATGGGAAAATACTACTGACCAGTATAGTT
 TTATAAAAGTGTATGGAAAATGTAAAAGAAGTTATAACCAAATGAAATAAATACCTATATAATA
 ACAATCTATAACTGCATATAGTACAGGTTTAAATACTGATTATTGCTTAGGTCTAAATAG
 TAATAGCAATCAAAGTTCATTTTCTGAAATTTCCAATTTCAAATGTAGTAAAAGTAGCTCCAA
 ACAGAAATAATGCAGTACTTTTACTTTACAAGTGAAGGGGATGTATATACTGCAGGCAAATGT
 AGTAATGGTTCAGGTACAGGGAAGTGAGACTCCAGAGAAGATTA AAAAAAAATAGCATCAAAGGC
 AAAGGATATTGGAAATGAATTTATAGATGTGGACATTTATGTAAGTGATAATGGAGACCTATATG
 GTACAGGTTTTAATAATAATGGACAATTAGGTGTTGGTGATGTAACAAAAAGAGATACATTT
 ATAAAAACCAATACAAGAGTAAAGAAAATACTTCCTTTAGAATATGCAAAATATAGCAATAAA
 AGATACTAATGATATATATATTTGTGGATTAATAACTATGGACAATTAGGTGTTGGAAAATA
 GATACGATAGTAGAAAATAATGATAATAGAAATATTTAATTATAAGCATATGAATTTTGTAATG
 GGTGATTTTGACATCTATTAAAAACAGACATAA CTTTATACTTCTAAAACAATAAGATAGTGAT
 ACCTACCACAAAAGACATAGATTATGGTTTTAGTATTAGGAAAATTTATACAAAAGGAGACCTTT
 ATACTGAGCTTCCATATGAAGATATAAAAAGAAATATCTATTTCTAAGACTCATATTTATTTATA
 TTACTTAATGATGGAAACAATGTATGGATGTGGTACAACTACCATGGAGAATTTATTTGCAAGA
 CTTGTCTATAAATCAAGTGGATGAATTTGTGCAGATTAATGTATCAGATGTAAGCATGTTT
 CATGTGGAGATAACTTTACTTATTTTATAAAAATCTGATGATAGTCTTTGGTCTATTGGTAAA
 AATTCCGAATATCAATTAGGTATAGGTACAAATAATCCAGTTACTGAATTACAAAAGAAATAC
 AACTATATCTAGCTGTAAAGAAGTACATTTGTGGTAAAAACTATACATTAGTAGTAACCTACAG
 GTAATGAATTTATTTGTACAAGGATATAAATGATAAGGGAGCTTTTAGGATTAGGAAGCGATAGT
 GAAAATACTATAAATTAAGTTCTTTACAAAAGCACTAACAGACATAAGAGAAAATAAATCTTA
 TGGAAAGTGACCATATATAGTACTTAAAAATGATAAATTCAGTATGGGTTACTGGAAAAATA
 GGGATGTATATAAAAAATTGAACAACCAGTAGAATTTTAAAAAGAAATTTACTATAGTACCTATT
 TCTGAAGATGTAAATACAGTAAAAGGATGTACTTGCAACAGACAATACATTTATATATTTATATC
 AGAAGTAGGAACGACAAAATGCTGCTATAGAAAATTA CTGAAAAATCAATTTCAATCAATTAAGA
 TAAAAATACAAGACCCTAATAAAGATATAAGTAGAATAGAAATGCTTATAAATGGTGAAAGT
 GTAAAATCTGTAAGTGATTTAACTACTGAAAAAATATCCTTTGAAGTACCACCAGATAAAAAT
 TAAAATAGGAGAGAATAAGATACTATTTAGAGCTTATTTGTAAGGTGATGATTTATATGCAT
 CTTTTATTTATTTTTTAAAGAGAGTACTGGAAATTTCTATAAATTAAGATTCTTATGTTATGATA
 GGTAATAGAATGTACAAGGTAGTTAATACAACATCTAATGAACAAGATATTACAATTTACT
 AGATAGAGGACTTGAAGAAGATTTAAATCTTTGGAGACCCTATATATCAATTTAATAAATAAAA
 CTAAAGTTCAAGTAAAAATAAATAAATCTGACTTATTTCAAAGACATGAACTAGTTGAAATC
 AAAAAATCAGACTCAAGTTATCAAGAAATCTATGAATTAGAAGAAGCCAACATAAAAAGTGC
 TCAGCCTAAAATCATAGTAGAAAAAGGAGATAAATGGACAGCTATAAAAACGTCCATCTATGA
 TTTTTAGATATGATGCTGAAAACAACGAGCCACAAGCTTAAAATGGAGGTGTAAAAATTTGTT

TAAATTCGATAAAAAATAAAATAGAACAAATCAAACAAGGTAGAAAAAGTAGAAATGCAGTATA
AAGACATTTTCAGACATAAGTATAGGTCAAGCAAAGCAAGATGATGATATAACAAATAATTTT
ATAGCAAATGCAGAAATATATGAGATGTTGTTAAGTCAAAGTTCTGTCAATGAAGCAAGTAA
TATAAGCACTTTTAGTGTAAAGAAAATCTGGAGGTGAGAGTGGAAATGGTAGAAGTATATGTAG
CTTTAATTTTAAAGAGGCAGAAAAACAATAGAAGAAGTACCAGCAGTAATTAGAGAGCAAGTT
AGAATTAGATGTAAAGAATTAGAAATACCAGTTGAATAGTAAATTTAGAATAACTATGTATT
AGTTATTTTTTTTATGTAAAGTACAAGGTCTTAACTTTAATAAGTAAGCCTTGTACTTATTT
TTTGTATATATTAGAAATTTGTATATATATTTTATTATTTTCAATCTATAAATTAACCTACA
ATTTAAAGTACAGAAGATTAATTTGATAATCCTGAAAAATAATATTGCATGATGTAAGAAT
ATAACAAAAATTAAGCTATAAGTATAAAAAATTTAGACAAATAGGAGGCTATAATGGATAAA
TTAATAACCGAATTGAGTAGTCTGGGGCAATAGGTATACTATGTGCTCTATTATTTAAAAA
TACTATGCAGGAGAAAAAAGAAGATAGAGACATGTATAAAAAACTGTAGAAAAATTTATAG
AATTATCTACACAACAAGAATAAACAATAAATACTTGTCAAATGGGAATAATGAAA
ACAGATGTAGAGGAAATTAAGGAAGATGTTACTGATATAAAAAGGTATGTTACAAAAACGGTGT
ATAACATGAAAGTAGCAGTAGCACCAGATTATATATTATTAGGAAAAAGATAAAGTAGTATTG
TAGATAGTGCCCTATTTTATTGAGAAGGATTTTATATTTTAAAAATATTAATTAAAAAAGTA
ATAAAAAATAATATAAAAAATAACATATAAAAAATTCAAAAAGGAGTTAAGCTTAAATTTGAT
TAGAAAAATCAATTTTAAAGACAACCTCTTTTTTTTtATTAAATTATTGTCTATTAACCAAAA
TAGCTATTTTAGCATCTGGATTATAAATTATCTGAACCATTTGATTTTCTTAACATGTTCA
AGGTCTTACCACCATAAGCTATTTGTAACCTTAACTGGTAACTTACCTTGTTTTATAATAGC
AACGTACTCTTTTTTACCTTTTTCTCTAAACTAATCAAAATGCCAACATAAGGTTTAAAGTT
CTGATACTTTTTACTAGAATTTCTTATGTAGAAGAAGCACCACAGCAATAACTAAATTTA
TGCCAAGTGTAAACCAAGAATTGATTTTAAAGCATAGCTCCAGCGATTATTATCACGAACATT
AAAACGATAGGTAATATAGCTTTCTTAAAGAAGCAATTTACCCATTTATTTTATTAGCCTTTTT
CTCAGGGCCACTCATAGTTTTTGTACTAGCAAATGATTGCGCGAATTTGTCTCTTAAAGCCCA
TTTTATCCTCCTAATTTTAAATAAATATTTAGTTATAATAACGAGATATTACTTGAAACTAAA
AATTTACTACATTTATATTATGTTTGACTTTTTGTATAAATAATTACATTCAGTAAAGCAAA
ATATACTAATTTATTTATCATAAAAATTATAAAAAAGAAAATAAATGAAATAAAAAATATTAGA
ACAAAGAAATGATGTAAAATCGTATCAAAAGCAACATAAAAAATTTATTTATCTATTTTCTCAT
CTTTATTTTTGTTATACTCAATTTTTCTTAAATCCTTCTCTTTTTTCAATTCATGAAGTTTT
AATTCATCATACTTCTATTTTGGCTTTATCATAATCATTTAACTTTCTAAAAGTTGTTTAA
AAGCTTTATTTTATTAGAGTTTATGCTATTAAGTGGATAGTTTGAGGAGGAATCGCAAATTA
AATCTGATTTATGTGATAGATTATCTCCATTTAATAGCCAGTCTACTGAAACATTAATATATC
TCAGCTATAGATTTTAAATTTTATAATTTTCAATAATTTGGTTTTCTAATGTTTCTCTCAAATTTACTTAA
GTTGTACAGCCTAACATTTCTTCTAGTTTATATTGTTTAAAGTTTTTTTGTCTTTTCTCAAAT
AAACAATTTCTTCTCCTAAAGTATCCATAAACACTCTCCATTCAATTAATGTCAAAAAAGACT
TTTTAAGATGTAATAAGTTTCAAATTAAGGTCAAATGACATAAAAACCATTGACTTAAAGG
TCAAATGACTTTTATAATTAACCTAATGATACGAATTTACATCCTAATTTTAGCACAAAGTA
ATCAAAAAATCTTATTTAGTATTAATAAATTTATATACTTAATATGTGTACATATTAAAAA
TATATACTAAATAGAGGGGGTGCCTAAGCTAAAGTAATATAAAAGTAAATATAAATCACTTA
GAAAGGAAGTTGATAAATGGATGCTCGAAAAAAATGGATACCTTTTTTTGGGAGTGCAAGTCA
AGCAAAGACTTATTGAATTAATATGACTCAAAGGGAATTAGCCAAGAAAAAGGTGTTAAT
GAAAACATTTTGTGAGCTATTTTAAATGGAAGAAGAAGGTAAAAAATATAAATCATCAAT
TTATCAATTTACTTAAATATAGAATATTCAGAAGATGATTAATAAATAGTATATAAAGTAGGTG
AATATTCTTGTGTCAAATTTGGATTGAGATGGGGTTATAGAGTGTGTAGAGCAATTGATGA
TTTTATTACAGCACTTAGTAATATAAAAAAGCTTAAATATGGAAAGATTAAATACTTTAACTA
AATATTCTAGTACATGTTCAATCCTTCTTAAAGAGGGGAATTATGAAGGATGTACAATTTGTG
TATAGAAAGATGTTGGAAGAATTA AAAACATGAGTAATGCATTTCTTAGGAATATAAATTTAT
ACATAGAAATGTATTATTTTTCAAAGTACTTAACTAAAATATGGATAAGATAATCTAAA
TATTATAAATGTGCTTGAATTTAGACTATACTTGTTTTTTAAATAATCCAATATCCATATTTT
AGTAATATACTACAAAAAAGAAGGTTAATAGATGATGTA AAAATCGTATCAAATTTATGTATG

ES 2 622 280 T3

TTTAAACCATTTTATCTTCATTATTATTAGAGGAATGCTTTTTTAAGTCTTTATATTTCAGAT
ATCTTAAGTTCAAGTATTCCTTCTATTTTTTATTTTATCACGTTTCGTTTAGTTGTCTGTATAG
ATTTAATATCATCATTTTCATCATTAGTAACATGTAAGTAATCTTCTTTATCTTCTTTTACAC
TACTATTGACATTTACCTTCTCTTTACCATAGAGAAGCCAGTCAGTCGTAACATTTAAAATAA
TCAGCTATTGACATTAGTATATCACAAATTAGGTTTTCTATCTCCTGTTTCATACTTGCCTAA
GTTTTCAAATTTTAAAATATCCATAAGTTTTGCGCTGAGTAAGTTTTTTGGAGTTTCTCAAAT
AAGCAATTCCTTTTCCTAAAGTATCCACAAAATACACTCCTTTCTTTTTTATGAGTAATGTCT
AAATGACATTTGAAATTAATAATATAAAATTTATAATATAAAACTACTAAATTAAGTCTA
AATGACATTTTGCTTAAATTAATATGCTCATAATATGATTTTTAACATATTATAGTTGAAAAT
ATATGGTTTATTTTGATTTGTATATATAACAATAGATTTAATTGTTATAAAAATGTAAAGGG
GTGTATGAATAGATTGTATAAATTTATTTTCGATAAACTAAGATTGCTTTTTGATTGTCTGTA
AAAGAGAAAAAGATTAAGATAAAAATAGTATTATATTGTAATTTATATTAATCAATTACAAA
GATTTTATGAATTTATTCTTTAGGGTAAAATATTTAAGAATAAGATAAATTTACAATATAAT
ACTATAACACTCTTTTATCTAGTTTTATTTCTTTTATAGAACAATAATATTATAAATGCTAG
TAGATTTACACAGAATACTGTTATATACATCTGTTTGAATCCTGAGTTTAGAGTAGATTGTA
GTGtGGATCCGG

SEQ ID NO:2 > orf 1360A (100% 630)
MFKNLKYRCKGMTQIQLARKAGITNDYISQIERGIKNPGLLMAKKISSILEQNIIEEVFF
IQL

SEQ ID NO:3 > orf 1361 (100% 630)
MENKKDILFKETDERLHNYKYLDIKIKINLDIKRCENEYSGCGAMVYTEKTSNTYNISSSV
ENEVLKREERLRKLKMEKEDIEIEKEKIENALTCLNDIEMEFFNLFYNSKTKNNMTYISMKL
HLDRTSCYNLKKKMI FKLSEIL

SEQ ID NO:4 > orf 1362
NFTESIFIDDESQVQSEGSCFFVSVILSVICTPIMLNTNNKDIVISIKYLPKPQSKSIRMYEI
SDELNKLFNRNKIVTDRKLNITKLEQSIKKEESIYVLNFTITLNYLDSVYEEDEVVYENMEEI
NLNLGE

SEQ ID NO:5 > orf 1363 (99+% 630)
MAIGLPSINISFKELATTVKERSARGLIAMVLKDAKALGLNEIHEKEDIPVDLSAENKEYIN
LALMGNVNTPNKLLVYVIEGEADIQTALDFLETKEFNYLCMPKAVEADKTAIKNWI I KLRDI
DKVKVAVLGVVGNHEGI INFTTEDVLVGEKKYSVDEFTSRVAGLIAGTPLSQSVTYTKLS
DVVDIPKMTKVDAESRVNKGELILIKEAGAIRIARGVNSLTELTEEKGEMFQKIKIVDTLDI
IHSDIRKVIIDDYIGKVTNSYDNKCLLIVAIKSYLEELEKSALIESDSTVEIDFEAQKSYLK
SKGVDLSYMTLQEIKEANTGSKVFLKAKIKVLDAMEDIDLSEI

SEQ ID NO:6 > orf 1364 (100% 630)
MANMEARNVMSGTWGELWLDGNKVAEVKKFQAKMEFTKEDII IAGQMGTDTKYMGYKKGKSI
TLYHVSSRMHKLIGEKIKRGSEPRFVAISKLNPDPSYGAERIAVKNIAFDDLTLADWEVGVK
GEIEAPFTFTEYDFLDII

SEQ ID NO:7 > orf 1365 (98+% 630)
MNENGLSKNINIVDLLLNADTENLERPSTIVELKRLSTIFGQEFKVMCRALTISKDEEIQNT
CLKIDENMKTDIDLPEMQLTIIEGVCDLDGKLLFKNKELMDKFKAPTPELARKLLLLPGEI
TNLYRILQDVMGYGKNAVIEEVKN

SEQ ID NO:8 > orf 1365A (100% 630)
MYYYWKKKGIRPSLFYAMDKGELKLIIEAFFALEIEEEVEKMKHGYVCPLTGGGM

SEQ ID NO:9 > orf 1366 (98% 630) Tape Measure
 MGNVREEGINMYLTDNYTPKMNQIISVTDNFRATVAVSLSTNVMASSIKNSIGSASSRVNS
 LNSSLRKKVQTTASSVSSTMAKLSSSINAVSGVIGSLNGSIMRLAITIAMIIDYFNKLIQKKN
 EFNSNIMIILIFKAKSDEVEKTKNKLGLNKKIGGKIWNIVIKAKDMTKRVISSILGKLKRV
 EKRPYQGSINLKDMSAMARILPKMLFKNTFWSGVIAIKDMASSIISKVFPKLRRLFAGKV
 WSGAIAVKDMASGILGSIKGISDLTNGATIGVAVKKGVDLLGQEQNQKVVLESVMKRNTGK
 TSQKDVVKYDSDLVNMANTDPFDPEDVAVMGTKAKMISNITGGKKEKDITQAMVDVRLNMM
 TSSEQDVSAAFLSAAKGNMESLNTLVGENYKTFDEALEGISVKQMGGLAKEMSNTIPGIISGA
 QTSINNGLSIVKPFDDILGQGLKIKTFIESGLGNLAGLSEKMGKIGNVMNGKIIIGNKY
 DQMQRSRVKNKGFSDSTQYRISNEAEKRKMMVENKQERFENHAATMIGNAPKAIVNAGSTL
 LQNDIFDALIDSLPVLVNNLLDSINNKSPIAQLISIFGTIVTTAFQLIGPVVEAVSPI
 ITRIFTFLEGEYAPQINNFIETLGVINWKTWETLGPLLETGWKIIIEPILGAFFNILDKVKVCI
 KDICKWQTMINKIKNGSITGTVLNLVEKSKKNYKDNP
 YAGTKAGDSGKAYSSKKGNAFGLNYVPYNDYQTRLHEGEMVLTQKQANQYRSRKNNGNINI
 AKLADTIVIREEADIEKITSKLVASIQLAQLGGVL

SEQ ID NO:10 > orf 1367 (100% 630) Placa basal
 MEMWLRQAEFRFPVFPSSFSINGKAAVNSSSILKIGEVATFGGVALKSISISISFFPNKDY
 TFCDYTGFPSPYDCVNKIEKWMKEGFILRFTITETNINMEVIEGFSYEERDGTDRDVFYFTLD
 LKEYKRKIKIPKVTPKQ

SEQ ID NO:11 > orf 1368 (100% 630) Placa basal
 MIINRSKSSSNEISFVSKDMGFLLTQSEVSYNFKDKLVEDIAKQVFAENRLSVGTIAKTV
 KYTKMFIGVNGYDTIMSAYTEASKKTKKKYMIKANLDFNVIEKGTVTLVSMFEEGFNIINT
 TFSESMENVKNKVIVVDQYGSKISEKIDNEIFKEVNVIMQKVIQQENQDVIDDSEFNGIEK
 SCSLKGYGVDVSCITGRGVKVDKSYTKLVGLFYIDTDKHTWQNGEYQIELELNFQNLMDKSA
 GQDEPKESNLGGEDYAGGKEFTAETAYCPRKEEGDTCRKKKLDPSKKTCAAPVMVGKYE
 QTYTKEFLNKHPLLYGDEIQVITGVSGRDGVYKVNVDVGPATIEKNGTYHIDILFGNVEE
 ASKFGRRKGIIGGYSGNVSDKAKIVISEAKKHLGKPYKGGNGPSSFDCSGLMVYCFKKV
 NVSLPRTSNQQSKKGGKVEQKNLQAGDLVFFHNPVSHVGLYIGNGEFLHAPQKGDVVKISKL
 SSRDFNTARRVL

SEQ ID NO:12 > orf 1369 (100% 630)
 MANPINEFIGIIREEGKYHNQPSFFIGIKSKLPDLKIETNIILEKEDILIDSWMIDRQLE
 TFDTETNQEHEVKNPFIDNFESGDMVIMFRIGEKFAVVSKLVSL

SEQ ID NO:13 > orf 1370 (99+% 630)
 MSTIFPFIVGPEYILPKTEELPIFREVAVDFEKDEPILEKGFKIIEKKEALKVWIYKCIK
 TNRYEHEIYSLEYGTELSELIGQKYTKGLTESEASRFIKEALLINPYILEVNVKSANFNRDV
 LSAVVKVSTIYGEVEINV

SEQ ID NO:14 > orf 1371 (100% 630) Cola
 MYSDQTYEVIKNRLENINLDIYKGEFSFLNNMVSGNNLELSKIYLELSKIHKMAFIQDTYN
 QFLDKRVNEFGVYRKLGTESNGEVEFIGEKGTVINNGTIIISYRDLLFVVIKDVITIGSEEDN
 SPVQALEVGGKYNLPTNCFKLVNDSGVTKITNTRSFEGGTDIETDEELKERFYKIQRNQA
 TSGNKAHYEEWALEVDGVYVNVKVPYPRWDGPGTVKVLIFGENNQAVDTETIERCQQHIDEK
 IGPTITVVTPLPIEISISAVMKLEDGYTLDNVKSFLSINTYFRDIRGEIITYTKVMGILIN
 TTGVHDLNLLINGSTDNITINEDKIPSVTTVNFSEVENQ

SEQ ID NO:15 > orf 1372 (100% 630)
 MKLIDKLPSPDRNYIVEEIQGAYDTELNILKEDIDDTFNQLFVDTATWGLDMWEDILCIEKK
 ELDFDTRRSNIKAKMRSRGTSTIEVIKSICEAYTKSETDIKVYSDEFTFVLSFIANNCDYKT
 LLDCSEMIERVKPAHLLHYLEPIILDKSMVYCGGMVCSEEVKVHPYFEPPIKCSAVVNCGA
 GMLSREEIKVYPLSIKCIENNCKINIAIANDTGVENVVVYPKSEVV

SEQ ID NO:16 > orf 1373 (99% 630) Fibra cola
 MEEKFYIILTKIGREKIANATALGELVGLTKFQVGDSENGEYEPTEEQTALKNVVWEGNINS
 LRIDEKNPNWIVIEITILPGTVGGFMIREAAVLNENNI IAIGKYPETYKPRAE DGSIKDLVV
 KMILQLSNTSNVTLEVDPTLVFVTQKDIQDLDDKFDKNIKEIKVKIGD TDILT TDSKDL SGA
 INEVVKKIENISFDDVISGQIQTDISVLKNSYNKLEKVL DILIYLELESEVTVDEAGYWD
 TLANGNNIVAIEGLKLDLNRKCITGEIGNVIFRDVVL PFSANRVRYIHDMDNNFVETKSSNT
 YLKEQKDITLSKYSYEIR

SEQ ID NO:17 > orf 1374 de unión al receptor núm. 16 - variable
 MKQNKLLQRGAYFNDKNILIDDFDKRYNDYDFVEFFTGISNSTFGLKSDGNLYACGDNTGFQ
 LGLGKDSSERRMFASKVIDNVKYVSCGSKHSVAVT KDGFAYGAGTSNVGQLGVIESTVYYEF
 TKLPIDDVKT VACGYDFTFVLKNDGTLYSAGLNSSGQLGLGDTNRRATFTKVNI DSVKDVVT
 YNQSVFIIKMDGTAHACGLNSNGQLGINSTLNKSVFNKIEGMDNVKQIACGSSHTILIKNDG
 TMYTTGYNGVQGLGTGNNNSIVFTLSSINNPKYASCGNNHTMILKYDNTLFSTGQNNYQQL
 ANANKDVASRNTFAKVNVENIKDIKCGSQFNFLINGSKEIFVSGCNLAGQLGSFFHTTFLYE
 FSNVQSSNLDNYSGLLVNDYLYVTKDNSEFLNVKLSDNFQDYKKIELTDSNMFI VMNDGTL
 YACGLNNYQQLGLGDTVNRSVMTKVDIDNVLDIKGNGNSTFVLKNNGLTLYSCGLNSNGQLGL
 RDEVNRNIFTKIEIENVKDFCVGSNYVIALNHSKEVYGWGNNPYNIEKTSNYPYKQGISNI
 EKIAAYDYSVYMINSEGLYVSGYNYNYQLGKGNNSNQSKALVSQCRTNSTSSSTSNGLRITL
 KITNVFPFYDGCALIDEGGYVYLTGYHG YLRTLNSSPSISDYSRYGT FIEATNSNHNTYFIQ
 ETD FSGIEKVIGMSNNILFFKKGSSYITGYPKTFGSTITGHRSYTSINSESSNLGSNFIIYH
 SNSKLYGKGIANSQGFGNSTNIDGTSNYDTGLKDIKDIIVKGNTVVVV DKNNNIYVTGMNQ
 NKLGI GEYNNP VKKFTNITEQSNSFI FMDDIKEITTSRNTMFIVKNDGTAYATGNNSSGQL
 GLGDTINRNKFTQINLDNIKKISTSIDGNTTFAIRNDGTLYSTGLNTKGQLGLGDIVNRNTF
 TKVNIQNV RDVVLGTT HSHAIKDDNTLYSCGENTHGQLGLGSES NHPDVLTFVNNITNVRD
 VYCSDTTTFIVKDTNIA YCCGYNNNSQLGMCNTTDQYSFIKCMENVKEVIPNEINTYIIT IY
 NTAYSTGLNTDYCLGLNSNSNQSSFSEIPI SNVVKVAPNRNNAVLLTSEG DVYTAGKCSNG
 SGTGSETPEKIKKIAASKAKDIGMNYRCGHYVSDNGDLYGTGFNNNGQLGVGDVTKRDTFIKT
 NTRVKKILPLEYANIAIKDTNDIYICGLNNYGQLGVGNRYDSRNNNDNRI FNYKHMNFVMGDL
 TSIKNRHNFILLNNKIVIPTTKDIDYGLVLGNLYKGDLYTELPYEDIKEVSIKTHIIILLN
 DGTMYGCGTNYHGELLQDLSINQVDFVQINVS DVKHVSCGDNFTYFIKSDDSLWSIGKNSE
 YQLGIGHNNPVTELQRITTISSCKEVHCGKNYTLVVTGNELFVQGYNDK GALGLGSDSENT
 I IKFFT KALTDIREIKSYGSDHILVLKNDNSVWVTGKNRDVYKIEQPVEFLKEFTIVPISED
 VNTVKDVLATDNTLYI ISEVGTNAAIEITEKSISSIKIKIQDPNKDISRIEMLINGESVKS
 VSDLTTEKISFEVFPDKIKIGENKILFRAYCKGDDLYASLFI PKESTGNSI IKDSYVMIGNR
 MYKVVNTTSNEQDITITLDRGLEEDLNLGDPIYQLINKTKVQVKINKSDLFKDMKLV EIKKS
 DSSYQEIYELEBANIKSAQPKIIVEKGDKWT AIKRPSMIFRYDAENNEPQA

SEQ ID NO:18 >orf1375
 MQYKDISDISIGQAKQDDITNNFIANA E IYEMLLSQSSVNEASNISTFSVRKSGGESGMVE
 VYVALILRGRKTIEEVPVAVIREQVRIRCKELEIPVE

SEQ ID NO:19 >orf 1376 (100% 630)
 MDKLITELSSLGAIGILCALLFKNTMQEKKEDRDMYKKTVENFIELSTQQQEINKNILVQMG
 IMKTDVEEIKEDVTDIKGMLQNGV

SEQ ID NO:20 >1377 (100% 630)
MGLRDKFAQSFARSKTMSGPEKKANEIMGKLLKAILPIVLMFVIIIAGAMLKINSWVTLG
INLVIAVGAFFYIRNSSKKYQNFKPYVGNLISLEKKGKKEYVAIIKQKLPVKLQIAYGGED
LEHVKKNQMVQISYNPDAKIAILVNRQ

SEQ ID NO:21 >AV1378 (100% 630)
MDTLGERIVYLRKAKNLKQYELEEMLGCDNLSKFERNIRKPNYEILKSIAEIFNVSVDWLLN
GDNLSHKSDLICDSSSNYPLNSINSNEIKLLNFRKLNLDYDKAKIEGMIELKLHEYEKEDL
GKIEYNKNKDEKIDK

SEQ ID NO:22 >1378A (100% 630) regulador
MDARKKWIPFLGVQVKQRLIELNMTQRELAKKIGVNNENYLSAILNGRRTGKKYKSSIIYQLLN
IEYSEDD

SEQ ID NO:23 >AV1379
VDTLGKRIAYLRNSKKLTKRKLMDILKFENLGKYETGDRKPNCIDILMSIADYFNVTTDWLLY
GKEKVNVS SVKEDKEDYLHVTNDEMILNLYRQLNERDKIKIEGILELKISEYKDLKKHSS
NNNEDKMV

SEQ ID NO:24 >AV1419
tttcttGAAGACcAtcgaAgcaccaccaccaccactg

SEQ ID NO:25 >AV1420
ttttttGAAGACAAcgaAggggttcgcccctgtcgctcgac

SEQ ID NO:26 >AV1416
ttccttGAAGACCTaattTggggcaatcccgaaggag

SEQ ID NO:27 >AV1245
ccccttGAAGACCCaattTcgtatggcaatgaaagacgg

SEQ ID NO:28 >AV1372
aattGCGGCCGcagctcGCTAGCggtacCTCGAGgatatcttcgaaGAAGACACATCCg

SEQ ID NO:29 >AV1373
aattccgggatGCATGCctctaGGATCCGGCGCGCC

SEQ ID NO:30 >AV1374
agctGGCGCGCCGGATCctagagGCATGCatcccgaattcGGATGTGTCTTctt

SEQ ID NO:31 >AV1375
cgaagatataCTCGAGgtaccGCTAGCgagctGCGGCCGC

SEQ ID NO:32 fragmento 5'
GGCCGCAATACCCACTACACCTTCgTCATCTTTAAATTTAAGAGTTTTACTATTGAATAAT
AAAGGTATATTCCAGTAAAATAATCTTTAAATACAAGAAAAATAACTCTTTGGGTATATT
AAAAAGCTAAAAGTGTAAATATAAAAGCAAGTAGAGTACTTATCCTGTAAAAGAAATCTAT
TTGTGTAATGTCTTTATATTTTATCATAAACACCGAATATAAAATGATGAAAATAATTGCGA
CGATTGCATATATGGTAAATAACATATTTTCAAGAGTACCATTGAAATFACTATCCACTTA
TACCACATAATTGGCCAAAATAATAGTGCTAAGAACTTAAAATAATTATCAAACAACCTTTTC

TTTATACATTCATCAAAACAACCTTTCTTAACAAAAGCATATATTTGTTTTTAGAATTTTAAA
 TAATATGATATCATTATTATATATTAATATTGAATTTATAGAAACCAAAATTTGTTAAAATA
 AATATATAGATTTTTACTGTTAAGCCAGTTAAAATTACTACTATTTTTATTATGAAATTTGGAT
 CAAATATGTAGAAATACGGCAAATTAGTTAATATTAAATATTTATTATTTCCAAGTTGTA
 GACTGTTTTTTAATGATAAAAATTCTAATCTTTTTTTGAAAGAAAGTAATATCCACATTAAG
 TATGTCTGCCATTTCATAAACGCAAGTGATGCCAGAATTAATTATGTTTATTATATCTTCTT
 CAGTAATTAAGAACTCACAAAGCCCATTTTAAGGCTTTATTTTCGCACCTTATCTATAATAATT
 TTTGTATAATAATCGTTATAAGAGGATACATAGTATCCAAGGCTAGTGAAATGATGTCCAAG
 TTCTTCAGCTAAGATGGATGTCAATTTTTTTGAGTTTTGTTTTAAATTAAGTACTGAGTAATGATA
 TAATTTTAATACCATGTTTGTTTATATATAGCCCTTCTAAATCACCTGCAATATAAGTGGTA
 TAATGAATTATTATCTCTTCTTGAGAAGCTAATTCAAAAGCTTATCCAAATTATTCATAAA
 AATCCCCCTAAAATAGAATGTATGTTTGCCTTTAAATTATATTTAAAAGAGCAGAAAAATAGA
 CTGCTCATCATATGGTTTATTTTTTTTTATATTTTATTAGTAAAATTTCTATATAATCATT
 AGTTGTTCTTGTGCTTCTTCAGGTAACCTTTCATGTGGATTTTTTTCTATGTGCAGCTACTGT
 ATCAATATTTTTCTTAACTAAGGTTCTTCCAAGAAGGTAATCAACTGATACATTAAATACAT
 CAGCCAATTTGTTTTAAAATGTGTTTCATCAGGAAAATCTGTTTTCTGTTTCATAGTACCCTAAG
 ACTCTTTGGGAAACGCCTACTTTTTCTCCAAGTTCTCTTTGAGTCAATCCAAATTCCTTTCT
 AAGTTCTCTTAACTTTTGGCAAACATTATAACACCACCTTATGTATAGATTATAACAAATT
 GTTCTAAAAAATAAACTAATAAAATATAAAAGAATATTTTTCTAAAATCTATTGATAAAG
 AACAAATGATTCTATATAATCTAAGTGTGGAAGAACAAAATATCTTAATGGTAATGGAGGT
 ATAAAACAATGTTTAAAATAAAGTAACTGAAATATATAGAAAATGCAAAGGTATGACACAAATT
 CAACTTGCCAGAAAAGGCTGGAAATACAAATGACTATATATCTCAAAATAGAAAAGAGGTATAAA
 AAATCCTGGTCTTCTTATGGCTAAGAAGATTTCTAGTATTTTAGAACAAAATATAGAAGAAG
 TTTTTTTTATACAGTTATAGAACAATATGTTCTTGAAGTTGTGAGATTAGTAAAAAACTGT
 GCACAAAAGAGATTATTGTAATTTGAAGCTAATAATAAGTATATAAAAAAGGAGAAGTACT
 ATGGAAAACAAAAGATATATTTTAAAGAAACAGATGAAAGATTACATAAATTATAAGTA
 TTTGGATATAAAGATAAAGAATATTAATTTGGACATAAAAAGATGTGAGAATGAATACTCTG
 GATGTGGAGCAATGGTATATACAGAAAAGACTAGTAACACATATAACATAAGTTCTTCTGTG
 GAAAATGAGGTGTTAAAAAGAGAGGAAAAGATTAAGAAAATTAAAAATGGAAAAAGAAATAT
 AGAAATAGAAAAGAGAAGATAGAAAATGCTCTAACATGTCTAAATGATATAGAAAATGGAAT
 TTTTTAATCTTTTTTATAAATAGTAAGACAAAAACAATATGACATATATTTCTATGAACTA
 CACTTAGATAGAACATCTTGCTACAATTTAAAGAAAAAATGATATTTAAATGAGTGAGAT
 ATTATAAAAAATATGACAACTTTACAACACTTTATATACACTATTGCAACACTAGGCAATAA
 AATATGTGAGATAATGTTATTTGTGAAAGAAATCCATATTGAAGGAGGTGATAAATGAAAAAG
 AATAATATTACCTAAAAATATAGAAGATATTTGACAGGAATAAATGAGATGTATATTTAAAA
 ATGACTTATATCATTTTATAGTAAGATTATCAGATTAAGCAAGAAATATTAGTGATAGTGTGG
 TGATTATTTGCTTAAATACAAGGAAATATTAGAAACAATTAATTGAGATTCCTAAAAAACT
 TTACTGAAAGTATTTTTATTGATGATGAAAGTGTGCAAGGCTCTGAAGGGTCTGTTTTTTT
 GTAAGTATACATCAGTTATTTGTACACCTATAATGTTAAATACGAATAATAAAGATATGTT
 TATCTCTATAAAATACCTTACCAAAACCACAGTCAAAGAGTATTAGAATGTATGAAATTTTCTAG
 ATGAATTAATAAGTTATTCAACAGAAATATAAAGGTAACAGACAGAAAATTAATATAACA
 AAGCTAGAACAAGTATTAAAAAGAAGAGTCAATTTATGTATTGAACTTTACAATTACACT
 AAATTATCTGGATAGTGTATATGAAGAAGATGTAGTATATGAAAATATGGAAGAAATCAAT
 TAAATTTAGGAGAGTGATAGTATGGCTATAGGATTTACCAAGTATCAACATATCATTTAAGGA
 GCTAGCTACAACCTGTTAAAGAAGCTTACAGCTAGAGGAATAATTGCAATGGTGTAAAGATG
 CTAAGGCACTAGGTCTTAATGAAATACATGAAAAAGAGGATATACCAGTTGATTTATCTGCT
 GAAAATAAAGAGTATATAAATTTAGCTTTGATGGGAAATGTTAACACTCCAAATAAATTTATT
 AGTTTATGTAATAGAAGGAGAAGCAGATATTCAAACTGCATTAGATTTTTTTAGAGACTAAGG
 AATTTAATTTATCTATGTATGCCAAAAGCAGTAGAAGCTGATAAGACTGCTATAAAAAAATGG
 ATAATTAACCTTAGAGATATAGATAAGGTTAAGGTTAAAGCTGTATTAGGAAAAGTTGTAGG
 AAATCATGAAGGGATAATTAATTTTACTACAGAAGATGTGTTAGTTGGAGAAAAGAAATACA

GTGTTGATGAGTTTACAAGTAGGGTGGCTGGACTTATAGCAGGAACACCTTTAAGTCAATCA
 GTAACCTTATACTAAGCTTAGTGATGTAGTTGATATACCTAAGATGACGAAAGTTGATGCAGA
 ATCAAGGGTTAATAAAGGAGAGCTTATACTTATTAAGGAAGCAGGGGCTATAAGAATTGCAA
 GAGGAGTAAATTCCTTAACTGAGTTAACAGAAGAAAAGGAGAAATGTTCCAGAAAATAAAA
 ATAGTTGACACTTTAGATATTATACATAGTGACATAAGAAAGGTGATAATAGATGACTATAT
 AGGAAAGGTTACTAACAGTTATGACAACAAATGTTTATTGATAGTAGCTATAAAAAGTTATT
 TAGAAGAATTAGAAAAGTCAGCACTTATAGAATCTGATTCTACTGTTGAAATAGATTTTTGAA
 GCACAAAAATCGTATTTAAAATCAAAAGGAGTAGATTTATCITATATGACATTACAAGAAAT
 AAAAGAAGCTAACACAGGTTCTAAAAGTATTTTTAAAAGCAAAAATAAAAAGTACTTGATGCTA
 TGGAAAGATATAGATTTATCAATAGAAAATATAGGAGGATTATTAATATGGCAAATATGGAAGC
 TAGAAATGTAATGAGTGGTACTTTGGGGAGAAGCTTTGGCTTGATGGAAACAAAGTAGCAGAAG
 TAAAGAAGTTTCAAGCAAAGATGGAATTTACAAAAGAGGATATTATAATAGCAGGTCAAATG
 GCTACTGATACAAAATATATGGGATATAAAGGAAAAGGTTCAATAACTCTATACCATGTTAG
 TCAAGAAGATGCACAAGTTAATTGGAGAAAAGATAAAGAGAGGTTCTGAACCTAGATTTGTTG
 CTATATCTAAAATTAATGACCCAGATTCTTATGGAGCAGAAAAGAATAGCAGTAAAAAATATA
 GCATTTGATGATTTAACTTTAGCTGATTGGGAGGTTGGAGTAAAAGGAGAGATAGAAGCACC
 TTTACATTTACTGAGTATGATTTTCTTGATATAAATTTAGTTTTATATTTGGTTTTTACTG
 ATATTTAGTAGATATATACTTAATAAATTTAGGTAGTTAATAAGTAAAAAAGTTAGTTGATT
 GAATTTGATTGATAAAGGAGCAAATAATAATGAATGAAAATGGATTATCAAAAAATATAAAC
 ATAGTAGATTTACTTTTAAATGCAGATACAGAAAAGCTTAGAAAAGACCAAGTACTATAGTTGA
 ACTTAAGAGATTATCAACTATATTTGGGCAGGAATTTAAAGTAATGTGTAGAGCTTTAACAA
 TAAGTAAAGATGAAGAGATACAAAATACTTGTCTTAAAATTTGATGAAAATATGAAAACGGAT
 ATAGACTTACCGGAGATGCAGATGCTTACAATTATAGAAGGTGTTTGTGATTTGGATGGAAA
 GCTTTTATTTAAAAATAAGGAACTAATGGATAAATTTAAGGCTCCAACACCAAAAAGAATTGG
 CAAGAAAACATTTATACCAGGTGAAATTACCAACCTATATAAGATACTTCAAGATGTTATG
 GTTATGGTAAAAATGCAGTGATAGAAGAGGTAAAAAACTAATAGGGACGGATACCAAGACT
 ACAATAATGTAATTTATTTGGAAGAAAAAAGGTATAAGACCGTCCCTTTTTTATGCAATGGA
 TAAAGGCGAATTAAGCTTATTGAAAGCTTTTTTCGCCCTTAGAAAATTGAGGAAGAAGTTGAAA
 AAATGAAACATGGATATGGAGTGTGTCCTTTGACAGGAGGTGGTATGTAATGGGAAATGTGA
 GAGAAGAAGGTATAAATATGTATCTTACAGATAATTACACACCAAAAATGAACCAATTTATA
 TCAGTAACTGATAATTTTAGGAGAGCAACTGTGGCTGTTTCACTTTCCACTAATGTAATGGC
 TAGTAGCATAAAAAATCTATTTGGAAGTGCAAGTAGTAGAGTAAACAGTTTTAAATTCCTCGT
 TAAGAAAAGTTCAAACACTGCTAGTAGTGTAAAGTTCAACTATGGCAAAAATTAAGTTCTAGC
 ATAAATGCTGTTTTCAGGAGTTATTTGGAAGTTTAAATGGAAAGTATTATGAGACTAGCAATAAC
 TATAGCTATGATTATTTGATTATTTTAAATAAGTTGATTCAAAAAGAAAAATGAGTTTAAATCAA
 ATATTTATGATTATATTAATATTTAAAGCTAAAAGTGATGAAAGTAGAAAAAACTAAAAATAAA
 TTACTTTGGAATTTAAAAAAGATTTGGTGGCAAGATTTGGAATATCGTAATAAAAAGCAAAAAGA
 TATGACTAAGAGAGTGATAAGTAGTATCTTGGGAAAATTAAAAACGAGTAGAGAAAACGTCCTT
 ATCAAGGAAGTATTAATCTTAAAGATATGGTAAGTAGTGCTATGGCTAGAATTTTGCCTAAG
 TTAATGTTGTTTTAAAAATACTTTTTGGAGTGGTGAATAGCTATAAAAAGATATGGCAAGTAG
 CATTATAAGTAAAGTATTTCCCAAATTTGAGATTGTTTGCAGGTAAGGTATGGAGTGGTGCAA
 TAGCTGTAAAGGATATGGCAAGTGGAACTACTTGGTTTCGATAAAAAGGGAAGATATCTGATTTG
 ACAAATGGTGTACTATAGGTGTGCTGTGAAAAAGGGTGTGACTTACTTGGTCAGGAACA
 AAATCAGAAAAGTTGTTCTAGAAAAGTGAATGAAAAGAAAATACTGGAAAAACTAGCCAAAAAG
 ATGTTGATAAGTATTATGACAGTTTAGTAAATATGGCAAATGATACGCCCTTTGACCCTGAA
 GATGTTGTTGCAATGGGAACATAAGCTAAAATGATTAGTAAATATTAATGTTGGTGGCAAAAAGA
 AAAAGATATAACTCAAGCTATGGTAGATGTTAGAGCTTTAAATATGAATACAAGTAGTGAAC
 AAGATGTATCAGCAGCTTTCTTAAGTGCAGCAAAAAGGAAATATGGAATCTCTTAATACTCTG
 GTAGGAGAAAATTTAAAACCTTTTGATGAAGCATTGGAAGGCATAAGTGTAAGCAGATGGG
 GTTAGCTAAAGAAATGAGTAATAACAATACCAGGTATAATATCAGGAGCTCAAAACAAGCATTA
 ACAATGGTTTGAAGAGTATTGTTAAACCTTTTGATGATATTTTAGGTCAAGGACTAAAGAAA

ATAAAACTTTTATAGAAAGTGGATTAGGGAATTTAGCTGGCTTATCTGAAAAATGGCTGG
 TAAAATAGGCAATGTAATGAATGGTAAGATAATTATGGCAACAAATATGACCAGATGCAAT
 CTAGAAGTGTAAAAATGGAAGAGATTTTCTGATTCTACTCAATATCGAATTTCTAATGAG
 GCTGAAAAGCGTAAAAATGATGGTTGAAAATAAGCAAGAACGTTTTGAAAATCATGCAGCAAC
 AATGATAGGGAATGCACCAAAAGCAATTGTTAACGCAGGAAGTACACTATTACAAAATATTG
 ATTTTACAGCATTAAATAGATTCACTACTTCCAGTAGTAACTTAGTAAATAATTTACTAGAT
 AGTATAAACAATAAATCACCAATTGCACAAGGATTAATAAGTATATTTGGTACAATAGTAAC
 TACAGCATTCCAACATAATCGGACCTGTAGTTGAAGCTGTAGTCCATTATCACAGAATTT
 TTACTTTTTTAGGTGAATATGCACCTCAAATAAACAATTTTATAGAGACACTGGGTGTATT
 TGGAAAATGTATGGGAGACCTTAGGACCTCTGTTGGAAAATGGATGGAAAATTTATAGAGCC
 AATATTGGGAGCTTTTTTAAACATATTAGATAAAGTATGTAATAAGTTAAAGATATATGCA
 AATGGTGGCAAACATGATTAATAAGATAAAAAATGGAAAGCATCACAGGAACAGTTTTAAT
 CTAGTGGAAAAGAGTAAAAAAATTACAAAGATAATCCATATGCTGGAACAAAGGCTGGTGA
 TTCTGGTAAAGCTTATCAAGTAAGAAAGGTAATAATGCATTTGGATTGAACTATGTTCCTT
 ATAATGACTATCAAACAGACTCCATGAAGGTGAAATGGTTTTAACTAAACAAGAAGCAAT
 CAATATAGAAGCAGAAAAAATGGTGGAAATATAAACATAGCTAAGTTAGCTGATACAATAGT
 GATTAGAGAAGAAGCTGATATAGAAAAGATAACATCAAATTAGTTGCAAGTATCCAATTGG
 CACAGTTAGGGGTCTCTTATAATGGAAATGTGGCTTAGACAAGCAGAAGATAGATTTAGAT
 TTCCAGTATTTCCATCTTCTTTAGTATTAATGGAAAAGCTGCTGTAACTCTTCTAGTATA
 CTCAAAATAGGTGAAGTAGCAACTTTTGGTGGTGTAGCTCTTAAAGCATTTCATATCAAG
 TTTTTTCCAAATAAAGACTACACTTTCTGTGACTATACAGGTTTTCCATCACCATATGATT
 GTGTAATAAGATAGAAAATGGATGAAGGAAGTTTTATATTAAGATTTACAATTACGGAA
 ACAAATATAAATATGGAAGTCATAATTGAAGGTTTTAGTTATGAAGAAAGAGATGGGAC

SEQ ID NO: 33

Fragmento medio

TCGAGATGTATATTTTACATTAGATTTAAAAGAGTATAAAAAGAATAAAGATACCAAAAGTAA
 CTCCAAAACAATAACTATATAGATAATAAGTTGTAAGTAACTGCTGATAGAATTAATGAA
 AAGGCAGGTGATTTTTATTATTAAGATTTGGGTACACATAAAAAACGGAAGTATATATGAC
 ATAACTGACATAGTAGACAAGGTATCATGGTCAGGTGATTATAAATCTCCATCAAGGACACT
 AGAGTTTTCAATAATACAATCATCATTTGATGTAATTTCCAACAAATCGATATACCAATAG
 CTAGTACAGTCTGTTTCTATGTAGATGAGAAAAGACTCTTTAGAGGAATGATAATTAATAGG
 ICTAAAGATTCAGCAGTAATGAAATTAGTTTTGTATCTAAAAGATATGGGATTTTTACTTAC
 ACAAGTGAAGTGTACATCAATTTTAAAGATAAGTTAGTTGAPGACATAGCAAAGCAAGTAT
 TTATAGGAGTAAATGGTTATGACACAATAATGAGTGCATATACAGAGGCAAGTAAAAAGAC
 AAAGAAAAGTATATGATAGAGGCTAATTTAGATAAGTTAATGTTATTGAAAAAGGAACTG
 TTACATTAAGTGTATGTTGAAGAGGGATTTAATATTATAAATACCACCTTTTCGGAGAGC
 ATGGAAAATGTAAAAATAAAGTAATAGTGGTAGACCAGTATGGAAGCAAGATTAGCGAAA
 AATAGATAATGAAATTTTTAAGGAAGTAAATGTAATAATGCAAAAAGTAATTCAGCAACAAG
 AAAATCAAGATGTAGATATGATAGCGAGTTAATGGGATAGAAAAAGCTGTTCTCTTAA
 GGTATGGAGATGTAAGTTGTATAACTGGTAGAGGAGTAAAAGTTAAAGATTCTTATACAAA
 GCTTGTAGGACTATTTTATATAGATACAGACAAACATACTTGGCAAAATGGAGAATATCAA
 TTGAGCTTGAACHTAATTTTCAAATCTTATGGATGAAAAGTCAGCAGGACAGGATGAACCT
 AAGGAAGAAAGTAAATTTAGGGGAGAAAGATTATGCAGGAGGAAAAGAGTTTACAGCAGAATT
 TACAGCTTACTGTCTAGAAAAGAAAGGTGGAGATACAGATTGTAGAAAAGAAAAACTTG
 ACCCATCTAAAAACTTGGCGTCTCTATGGTTGGTAAATATGAGCAAACHTATTATACAA
 AAGAGTTTTTAAATAAACATCTTTTATTGAACTATGGAGATGAAATACAGGTAATTACAGGA
 GTTCTGGTCTGTGATGGAGTCTATAAAGTAAATGACGTAGGACCTGCAATAACTATAGAAA
 AAATGGAACATACCATATAGATATTTTATTGGAAATGTTGAAAGAGCTAGTAAATTTGGAA
 GAAGAAAAGGAAAAATTTATTGGTGGTTATTCTGGTAAATGTATCTGATAAAGCTAAAATA
 GTAATATCAGAGGCAAAAAACATCTAGGTAACCTTATAAATGGGGTGGAAATGGACCAAG

TAGTTTGGACTGTTCTGGTTTAAATGGTCTACTGTTTTAAAAAGTTAATGTTAGTTTGCCAA
 GAACGTCAAATCAACAATCTAAAAAGGCAAGAAAGTAGAACAAAAAATCTTCAAGCAGGA
 GATTTAGTATTTTTTCATAATCCAGTCAGCCATGTTGGATTATATATAGGTAATGGAGAATT
 TTTACATGCTCCACAAAAAGGTGATGTAGTTAAAATAAGTAAGTTAAGTAGTAGAAGAGATT
 TTAATACAGCTAGGAGAGTATTATAAAAGGATGGTGATATAATGGCTAATCCAATAAATGAA
 TTTATAGGAATAATAAGAGAAGAAGGAAAGTATCATAATCAACCTCTTTTTTATTGGAAAA
 TTTAAAGTAAATTACCAGATTTAAAAATAGAGACAAATAACATCATATTAGAAAAAGAAGAT
 ATTTTGATAGATAGTTGGATGATTGATAGACAGCTAGAAACATTTGACACAGAAAACAAATCA
 AGAACACCAGCATGAAGTAAAAAATCCTTTTATAGATAACTTTGAATCTGGGGATATGGTAA
 TAATGTTTAGAATAGGCGAAAAATTTGCTGTTGTAAGTAAGTTGGTGAGCTTATAATGAGTA
 CAATATTTCCCTTTTATAGGTGTCAGGAGATTATATCTTACCATAAACAGAAAGAAATGGCCA
 ATCTTTCCGTGAAGTGGCATGGGATTTTGAAAAAGATGAACCTATTTTAGAAAAAGGTGACTT
 TAAAAAATTGAAAAAAGAAAGCCTTAAAAGTTTGGATATACAAGTGATATAAGACAAAT
 AGATATGAACATGAGATATACTCTTTAGAATATGGGACAGAGCTTTCAGAACTAATAGGACA
 AAAATATACAAAAGGTCTTACAGAAAGTGAAGCTAGTAGATTTCATAAAAGAGGCCCTTCTAA
 TAAATCCATATATATTAGAAGTAAACGTAAAAAGTGCTAACTTTAACAGAGACGTATTGAGT
 GCAAATGTAAAAGTATCCACTATCTATGGGGAGGTGAAATAAATGTATAGTGACCAGACAT
 ATGAAGTATAAAAAATAGAAGTCTTGAATAATTAATCTTGATATTTATAAAGGAGAAGGT
 TCTTTTCTAAACAACATGGTATCTGGAAATAATCTAGAACTTTCGAAGATATATCTAGAAGT
 TTCAAAGATACATAAAATGGCTTTTATACAAGACACATATAACCAGTTTCTTGATAAAAGAG
 TCAATGAAATTTGGTGTATATAGAAAAGTTAGGTACAGAGTCAAATGGAGAAGTTGAATTTATT
 GGAGAGAAAGGAAGTGAATAAATAATGGCACAATAATATCATATAGAGATTTACTATTTGT
 AGTAATAAAAGATGTAAGTATTTGGTAGTGAAGAAGGTGACAATAGCCCAGTTCAAGCTCTGG
 AAGTTGGTAAAGAAATAAATTTACCTACAAATTTGTGAATTTAAACTAGTTGATAATATATCT
 GGAGTAACAAAGATTACTAACACAAGAAGTTTTGAAGGTGGTACAGATATAGAGACAGATGA
 AGAACTAAAAGAAAGATTTTATAAAATCCAAAGAAATCAAGCTACAAGTGGAAATAAAGCTC
 ACTATGAAGAAATGGGCTTTTGAAGTAGATGGAGTCTATAATGTTAAGGTTTTATCCAAGATGG
 GATGGTCCAGGAACAGTTAAGGTCTTGATATTTGGGGAAAATAATCAAGCTGTGATACAGA
 AACGATTTGAAAGGTGTCAGCAACATATAGATGAAGAGAAGCCTATTGGACCAACTATAACAG
 TTGTGACACCAATTACCAATAGAAAATAAGTATAAGTGCAGTAATGAACTAGAAGATGGATAT
 ACATTTAGACAATGTAAAAGAATCTTTCTAGAAAAGTATAAATACATACTTTAGAGATATTAG
 AGGAGAGATAATCTATACAAAAGTCATGGGAATACTTATAAATACTACTGGTGTACACGATTT
 TAAGTAATCTACTTATAAATGGAAGTACAGATAAATAAATACTTAAATGAAGATAAAAATACCT
 AGTGTAACTACTGTTAATTTTAGTGAGGTGGAAAATCAATGAAGCTAATTTGATAAACTACCA
 TCATTTGATAGAAATTACATTTGAGAGGAGATACAAGGTGCATACGATACAGAAATTAATAT
 TCTTAAAGAAGATATTGATGATACCTTTAACCAATTAATTTGTTGATACAGCGACATGGGGAT
 TAGATATGTGGGAAGACATACTCTGCATTGAAAAAAGAAGTGTGATTTTGACACAAGACGT
 AGCAATATAAAAGCTAAAATGAGAAGCAGAGGTACTAGTACTATTGAAGTTATAAAAAGTAT
 ATGTGAGGCATATACAAAATCAGAAACAGATATAAAAGTTTATAGTGATGAATTTACATTCG
 TATTGAGTTTTATAGCAAAATAACTGTGACTATAAAACTCTTTTAGATTTGTAGCGAGATGATT
 GAAAGAGTAAAACCTGCTCACTTATTACACTATTTAGAACCAATAATACTAGATAAAAAGTAT
 GCTCTATTGTGGTGGAGGTATCCTATGTAGTGAAGAGGTTAAAAGTTTCATCCATACTTTGAAC
 CAATTTATAAAATGTAGTGTCTGTTGTAAGTGTGGAGCTGGAATGTTAAGTAGAGAAGAAATA
 AAGTTTTATCCTTTAAGCATTAATGCATTGAAAAATAATTGTAAGATTAATATAGCTATTGC
 AAATGATACAGGCGTAGAAAATGTAGTAGTTTTATCCTAAATCGGAGGTGGTATAAATTGGAA
 AAAATTTTTATATAATATTAACCAAAATTTGGTAGAGAAAAATAGCAAATGCAACTGCCTA
 GGAGAGCTTGTGGATTAACCAAGTTTCAAGTTGGAGATAGTAATGGAGAATATTATGAGCC
 AACAGAGGAACAACTGCTTTAAAGAATGTAGTTTGGGAAGGAAATATAAATTTCTCTAAGAA
 TTGATGAAAAAATCCTAATTTGGATAGTTATAGAGACTATTTTACCAGGAACAGTTGGTGGAA
 TTTATGATAAGAGAAGCTGCTGTTCTGGATAATGAGAATAATATAATAGCTATWGGTAAGTA
 TCCAGAGACGTATAAGCCACGTGCTGAAGATGGCAGTATTAAGATTTGGTTGTAATAATGA

TTTTACAATTGTCCAATACTTCAAATGTTACATTAGAAGTAGACCCGACGTTGGTTTTTTGTA
 ACTCAAAGGATATTCAAGATTTAGATGATAAGTTTTGATAAAAAATATAAAAGAAATAAAAGT
 AAAAATTGGAGATACAGATATATTAACCTACAGATTCTAAAGATTTATCAGGAGCTATAAATG
 AGGTAGTTAAAAAATAGAAAAATATATCTTTTTGATGATGTTATAAGTGGTCAAATACAACT
 GATATATCAGTATTAATAAATAGCTATAACAAATTATCTGAAAAAGTGGTCTAGATATATTAAT
 ATACCTAGAATTAGAGTCAGAAGTAACTGTAGATGAGGCTGGTTATTGGTATGATACATTAG
 CAAATGGAAATAACATAGTAGCTATAGAAGGGCTTAAGTTAGATTTAAATAGAAAAATGTATA
 ACAGGTGAAATTGGTAAATGTGATTTTTAGAGATGTAGTATTACCATTTAGTGCAAATAGAGT
 TAGATATATACATGATATGGATAATAACTTTGTTGAGACAAAATCTAGTAACACTTATTTAA
 AAGAACAAAAAGATATAACTCTAAGTAAATATTCATATGAAATAAGATAAATAAAGGAGGTA
 GTACTAATAATGAAGCAAAAATAAATTTTTACAGCGTGGTGCTTATTTAATGATAAGAACAT
 ATTGATTGATGATTTTTGATAAAAAGATATAATGATTATGATTTTTGTAGAATTTTTTTACTGGTA
 TAAGTAATAGTACCTTTGGTTTAAATCAGATGGTAATTTATATGCTTGTGGCGATAATACA
 GGTTTTCACTAGGACTTGGAAAAGATTCGTCAGAGAGAAGGATGTTTAGTAAAGTAAAAAT
 TGATAATGTAAATATGTATCTTGTGGTTCAAACACAGTGTAGCAGTAACTAAAGATGGAT
 TTGCATATGGAGCAGGAACAAGTAATGTAGGTCAATTAGGTGTAATGAGTCTACAGTATAT
 TATGAATTTACTAAGCTACCAATAGATGATGTAAAAACTGTTGCATGTGGTTATGACTTTAC
 ATTTGTGCTTAAAAATGATGGAACATTATATTCAGCAGGTTTAACTCAAGTGGTCACTTG
 GACTAGGTGATACTAACAATAGAGCTACTTTTACTAAAGTAAATATAGATAGTGTGAAAGAT
 GTAGTGACTTATAATCAATCTGTATTTATCATAAAAATGGATGGGACAGCACATGCATGTGG
 ATTAATTTCAAATGGGCAGTTGGGAATTAATAGTACTTTAAATAAAAAGTGTATTTAATAAAA
 TAGAAGGTATGGATAATGTAAACAGATAGCGTGTGGTAGTAGTCATACAATTCCTTATTAAG
 AATGATGGAACATGTATACTACAGGCTATAATGGAGTTGGTCAGCTTGGTACAGGAAATAA
 TAATAATTCAAATGTATTTACTCTTTCTAGTATAAATAATGTTAAGTATGCTTCTTGTGGAA
 ATAATCATACTATGATATTAATAACGATAATACACTGTTTAGTACAGGACAAAACAATTAT
 GGTCAACTAGCCAATGCCAATAAAGATGTAGCATCAAGAAATACTTTTGCTAAGGTTAATGT
 AGAAAATATAAAAGATATTAATGTGGTTCTCAATTTAATTTTTtAATAAATGGTTCAAAG
 AGATATTTGTATCTGGCTGTAATTTAGCAGGTCACTTGGTTCATTTTTTTCATACAACCTTTT
 CTGTATGAGTTTTCAAATGTGCAATCTTCAAATTTAGATAATTATTCAGGTTTATTGGTTAA
 TGATGATTATTTATATGTTACAAAGGACAATAGTGAATTTTTAAATGTAAAGTTAAGTGATA
 ATTTTCAAGATTATAAGAAGATAGAGTTAACAGATAGCAATATGTTTATTGTTATGAATGAT
 GGTACATTGTATGCTTGTGGTTTAAATAATTATGGACAGTTAGGATTGGGAGATACTGTTAA
 CAGGTCAGTTATGACTAAGGTGGATATAGATAATGTTTTGGATATAAAAGGAAACGGAACT
 CACTTTTGTGCTTAAAGAATAATGGAACATTATATTCATGTGGTTTAAATAGTAATGGACAA
 TTGGGTTTAAAGAGATGAAGTTAATAGAAATATATTTACAAAAATAGAAATAGAGAATGTAAA
 GGATTTTTGTGTAGGAAGCAATTATGTCATAGCTTTAAATCACTCAAAGAAGTATATGGAT
 GGGGAAATAATCCTTATAATAATAGAAAAAATCTAATTTATCCATATAAGCAGGGAATA
 AGTAATATTGAAAAGATAGCAGCATATGATTATTCTGTATATATGATAAACAGTGAAGGGAA
 ACTATATGTTTCTGGATACAATTATAATTATCAATTAGGTAAGGAAATAATAGTAACCCAAA
 GCAAAGCATTAGTATCTCAATGTAGAACAAATTCACATCTTCTACATCAAATGGACTTAGA
 ACGTTACCTAAAATAACTAATGTTTTtCCTTTTTTATGATGGTTGTGCAATAATTGACGAAGG
 AGGTTATGTTTATTTAACAGGATATCATGGATATTTAAGAACATTAATAGCAGTCCAAGTA
 TATCTGATTATTTCAAGATATGGAACTTTTATTGAGGCTACAAATTCAAATCATAATACTTAT
 TTTATACAAGAGACTGATTTTAGTGGAATTGAAAAAGTAATAGGGATGTCAAATAATATATT
 ATTTTTTAAGAAAGGAAGTTCATATATTTACTGGATATCCAAAAACATTTGGCTCAACCATTA
 CTGGACATAGAAGTTATACTAGTATTAATTTCTGAGAGTTCTAATTTAGGAAGTAATTTTATA
 ATATATCATAGTAATTCCAAGTTATATGGAAAAGGGATTGCTAATAGTGGGCAATTTGGG

SEQ ID NO:34 fragmento 3'

AATTCAACAAATATAGATGGCACAAGTAACTATGATACAGGATTAABAGACATAAAAGATAT
AATTGTAAGAAAGGAAATACIGTAGTAGTAGTAGATAAAAABTAACAATATATATGTAACAGGAA
TGAATCAGAATAACAACTTGGGATAGGGGAATATAACAACGAACCAGTAAAAAAATTCACA
AATATAACTGAACAATCAAACCTCATTATATTTATGGATGATATAAAAAGAAATTACAACATC
AAGAAATACAATGTTTTATAGTAAAAAATGATGGAACAGCCTATGCCACAGGAAATAATAGTT
CTGGACAATTAGGATTAGGTGACACAATAAATAGAAATAAGTTCACTCAGATAAACCTTGAT
AATATAAAGAAAAATATCAACAAGTATAGATGGTAACACAACATTTGCAATTAGAAIGATGG
AACACTATACTCCACAGGATTAATACCAAAGGACAACCTGGGATTAGGTGATATAGTAAATA
GAAATACATTTACCAAAGTAAACATCCAAAATGTAAGAGATGTTGTTTTAGGGACTACTCAC
TCGCATGCAATCAAAGATGATAACACATTATATTCATGTGGAGAAAACACTCATGGGCAACT
GGGCTTAGGAAGCGAAAAGCAACCATCCAGACGTATTGACATTTACTGTAAACAATATAACTA
ATGTAAGAGATGTGTACTGCTCAGATACAACAACATTTATTGTAAGGACACAAACATTGCA
TATTGTTGTGGATACAATAATAATTCAACAACCTAGGTATGGGAAATACTACTGACCAGTATAG
TTTTATAAAGTGTATGGAAAATGTAAAAGAAGTTATACCAAATGAAATAAATACCTATATAA
TAACAATCTATAACTGTCATATAGTACAGGTTTAAATACTGATTATTGCTTAGGTCTAAT
AGTAATAGCAATCAAAGTTCATTTCTGAAATCCAAATTCAAATGTAGTAAAAGTAGCTCC
AAACAGAAATAATGCAGTACTTTACTTACAAGTGAAGGGATGTATATACTGCAGGCAAAAT
GTAGTAATGGTTCAGGTACAGGAAGTGAGACTCCAGAGAAGATTAAAAAAATAGCATCAAAG
GCAAAGGATATTGGAATGAATFATAGATGTGGACATTATGTAAGTATAATGGAGACCTATA
TGGTACAGGTTTTAATAATAATGGACAATTAGGTGTTGGTGATGTAACAAAAAGAGATACAT
TTATAAAAACCAATACAAGGTAAAGAAATACTTCCTTTAGAATATGCAATATAGCAATA
AAGGATACTAATGATATATATATTTGTGGATTAATAACTATGGACAATAGGTGTTGGAAA
TAGATACGATAGTAGAAATAATGATAATAGAATATTTAATTATAAGCATATGAATTTGTAA
TGGGTGATTTGACATCTATTAAAAACAGACATAACTTTATACTTCTAAACAATAAGATAGT
ATACCTACCACAAAAGACATAGATTATGGTTTAGTATTAGGAAATTTATACAAGGAGACCT
TTATACTGAGCTTCCATATGAAGATATAAAAAGAAGTATCTATTTCTAAGACTCATATTATTA
TATTACTTAATGATGGAACAATGTATGGATGTGGTACAACTACCATGGAGAATTATTGCAA
GACTTGTCTATAAATCAAGTGGATGAATTTGTGCAGATTAATGTATCAGATGTAAAGCATGT
TTCATGTGGAGATAACTTTACTTATTTTATAAAATCTGATGATAGTCTTTGGTCTATTGGTA
AAAATCCGAATATCAATTAGGTATAGGTCAACAATAATCCAGTACTGAATTACAAAGAATT
ACAACATATCTAGCTGTAAGAAGTACATTTGTGGTAAAACTATACATTAGTAGTAACCTAC
AGGTAATGAATTATTTGTACAAGGATATAATGATAAGGGAGCTTTAGGATTAGGAAGCGATA
CTGAAAATACTATAAATTAAGTTCTTTACAAAAGCACTAACAGACATAAGAGAAATAAATCT
TATGGAAGTGACCATATATTAGTACTTAAAAATGATAATTCAGTATGGGTTACTGGAAAAAA
TAGGGATGTATATAAAATGAAACAACAGTAGAATTTTTTAAAAGAAATTTACTATAGTACCTA
TTTCTGAAGATGTAATACAGTAAAGGATGACTTGCAACAGACAATACATTATATATTATA
TCAGAAGTAGGAACGACAAATGCTGCTATAGAAATTACTGAAAATCAATTTTCATCAATTAA
GATAAAAATACAAGACCCTAATAAGATATAAGTAGAATAGAAATGCTTATAAATGGTAAA
GTGTAATAATCTGTAAGTGATTTAACTACTGAAAAATATCCTTTGAAAGTACCACCAGATAAA
ATTAATAATAGGAGAGAATAAGATACTATTTAGAGCTTATTGTAAGGTGATGATTTATATGC
ATCTTTATTTATTTTTAAAGAGAGTACTGGAAATCTATAATTAAGATTCTTATGTTATGA
TAGGTAATAGAATGTACAAGGTAGTTAATACAACATCTAATGAACAAGATATTACAATTACA
CTAGATAGAGGACTTGAAGAAGATTTAAATCTTGGAGACCCTATATATCAATTAATAAATAA
AACTAAAAGTTCAAGTAAAAATAAATAAATCTGACTTATTCAAAGACATGAACTAGTTGAAA
TCAAAAAATCAGACTCAAGTTATCAAGAAATCTATGAATTAGAAAGCAACATAAAAAGT
GCTCAGCCTAAAATCATAGTAGAAAAAGGAGATAAATGGACAGCTATAAAACGTCCATCTAT
GATTTTTAGATATGATGCTGAAAACAACGAGCCACAAGCTTAAAATGGAGGTGTAATAAATG
TTTAAATTCGATAAAAATAAATAAGAACAAATCAAACAAGGTAGAAAAGTAGAAATGCAGTA
TAAAGACATTTCAAGACATAAGTATAGGTCAAGCAAAGCAAGATGATGATATAACAAATAATT
TTATAGCAAATGCAGAAATATATGAGATGTTGTTAAGTCAAAGTCTGTCAATGAAGCAAGT

AATATAAGCACTTTTTAGTGTAAAGAAAATCTGGAGGTGAGAGTGGAATGGTAGAAGTATATGT
 AGCTTTAATTTTTAAGAGGCAGAAAAACAATAGAAGAAGTACCAGCAGTAATTAGAGAGCAAG
 TTAGAATTAGATGTAAAGAATTAGAAATACCAGTTGAATAGTAAATTTAGAATAACTATGTA
 TTAGTTATTTTTTTTTATGTAAAGTACAAGGTCTTAACTTTAATAAGTAAGCCTTGTACTTAT
 TTTTTGTTATATTAGAAATTGTATATATATTTTATTATTTTCAATCTATAAATTAACCTA
 CAATTTAAAGTACAGAAGATTAATTTGATAATCCTGAAAATATAATATTGCATGATGTAAGA
 ATATAACAAAAATTAAGCTATAAGTATAAAAAATTTAGACAATAGGAGGCTATAATGGATA
 AATTAATAACCGAATTGAGTAGTCTGGGGCAATAGGTATACTATGTGCTCTATTATTTAAA
 AATACTATGCAGGAGAAAAAGAAGATAGAGACATGTATAAAAAAACTGTAGAAAAATTTTAT
 AGAATTATCTACACAACAACAAGAAATAAACAAAAATATACTTGTTCAAATGGGAATAATGA
 AAACAGATGTAGAGGAAATTAAGGAAGATGTTACTGATATAAAAAGGTATGTTACAAAACGGT
 GTATAACATGAAAGTAGCAGTAGCACCAGATTATATATTTATTAGGAAAAGATAAAGTAGTAT
 GTAGATAGTGCCTATTTTATTGAGAAGGATTTTATATTTTTAAATATTAATTAAAAAAAG
 TAATAAAAAATAATATATAAAAAATAACATATAAAAAATTCAAAAAGGAGTTAAGCTTAAATTTG
 ATTAGAAAAAATCAATTTTTAAGACAACCTCTTTTTTTTTATTAAATTATTGTCTATTAACCAA
 AATAGCTATTTTAGCATCTGGATTATAACTTATCTGAACCATTTGATTTTTCTTAAACATGTT
 CAAGGTCTTCACCACCATAAGCTATTTGTAACCTTAACTGGTAACTTACCTTGTTTTATAATA
 GCAACGTAATCTTTTTTACCTTTTTCTCTAACTAATCAAATTGCCAACATAAGGTTTAAAG
 TTCTGATACTTTTTACTAGAATTTCTTATGTAGAAGAAAGCACCAACAGCAATAACTAAATT
 TATGCCAAGTGTAAACCAAGAATTGATTTTTAAGCATAGCTCCAGCGATTATTATCACGAACA
 TTAACGATAGGTAATATAGCTTTCTAAGAAGCAATTTACCCATTATTTTATTAGCCTTT
 TTCTCAGGGCCACTCATAGTTTTTGTCTAGCAAATGATTGCGCGAATTTGTCTCTTAAGCC
 CATTATCCTCCTAATTTTTAATAAATATTTAGTTATAATAACGAGATATTACTTGAAACTA
 AAAATTTACTACATTTATATTATGTTTGACTTTTTGTATAAATAATTACATTCAGTAAAGCA
 AAATATACTAATTTATTTATCATAAAATTTATAAAAAAGAAAAATAATGAAATAAAAAATATTA
 GAACAAAGAAATGATGATAAATCGTATCAAAAGCAACATAAAAAATTTATTTATCTATTTCTC
 ATCTTTATTTTTGTTATACTCAATTTTTCTAAATCCTTCTCTTTTTTCATATTCATGAAGTT
 TTAATTCATACACTTCTATTTTTGGCTTTATCATAATCATTTAACTTTCTAAAGTTGTTT
 AAAAGCTTTATTTTATTAGAGTTTATGCTATTAAGTGGATAGTTTGGAGGGAATCGCAAT
 TAAATCTGATTTATGTGATAGATTATCTCCATTTAATAGCCAGTCTACTGAAACATTAATA
 TCTCAGCTATAGATTTTTAATATTTTATAAATTTGGTTTTCTAATGTTTCTCTCAAATTTACTT
 AAGTTGTACAGCCTAACATTTCTTCTAGTTTATATTGTTTAAAGGTTTTTTGCTTTTCTCAA
 ATAAACAATTTCTTCTCTAAAGTATCCATAAACACTCTCCATTCAATTAATGTCAAAAAGA
 CTTTTTAAGATGTAATAGTTTCAAATTAAGGTCAAATGACATAAAAACCATTGACTTAA
 GGTCAAATGACTTTATAATTAACCTAATGATACGAATTTACATCCTAATTTTTAGCACAAAG
 TAATCAAAAAATCTTATTTAGTATTAATAAATTTATATACTTAATATGTGTACATATTTAA
 AATATACTAAATAGAGGGGGTGCCTAAGCTAAAGTAATATAAAAAGTAAATATAAATCACT
 TAGAAAGGAAGTTGATAAATGGATGCTCGAAAAAAATGGATACTTTTTTTGGGAGTGAAGT
 CAAGCAAAGACTTATTGAATTAATATGACTCAAAGGGAATTAGCGAAGAAAAATAGGTGTTA
 ATGAAAACATTTGTGAGCTATTTTAAATGGAAGAAGAAGCAGGTAAAAAATATAAATCATCA
 ATTTATCAATTACTTAAATATAGAATATTCAGAAGATCATTAATAAATACTATATAAAGTAGG
 TGAATATTTCTTGTGTGCAAATTTGGATTGAGATGGGGTTATAGAGTGTGTAGAGCAATTTGAT
 GATTTTATTACAGCACTTAGTAATATAAAAAGCTTAAATATGGAAAGATTAAATACTTTAAC
 TAAATATTTCTAGTACATGTTCAATCCTTCTTAAAGAGGGGAATTTATGAAGGATGTACAATTG
 TGTATAGAAAGATGTTGGAAGAATTAAAAACATGAGTAATGCATTTCTTAGGAATATAAATT
 ATACATAGAAATGTATTATATTTTTCAAAGTACTTAACTAAAATATGGATAAGATAATCTA
 AATATTATAAATGTGCTTGAAATTAGACTATACTTGTTTTTAAATAATCCAATATCCATATT
 TTAGTAATATACTACAAAAAAGAAGGTTAATAGATGATGTAAAATCGTATCAAATTTATGTA
 TGTTTAAACCATTTTATCTTATTATTATTAGAGGAATGCTTTTTTAAAGTCTTTATATTCAG
 ATATCTTAAAGTTCAAGTATTCCTTCTATTTTTATTTTATCACGTTGTTTAGTTGTCTGTAT
 AGATTTAATATCATCATTTTATCATTTAGTAACATGTAAGTAATCTTCTTTATCTTCTTTTAC

ES 2 622 280 T3

ACTACTATTGACATTTACCTTCTCTTTACCATAGAGAAGCCAGTCAGTCGTAACATTTAAAAT
AATCAGCTATTGACATTAGTATATCACAATTAGGTTTTCTATCTCCTGTTTCATACTTGCCT
AAGTTTTCAAATTTTAAAATATCCATAAGTTTGGCCTGAGTAAGTTTTTTGGAGTTTCTCAA
ATAAGCAATTCTTTTTTCTAAAGTATCCACAAAATACACTCCTTTCTTTTTATGAGTAATGT
CTAAATGACATTTGAAATTAATAATATATAAATTTATAATATAAACTACTAAATTAAGTC
TAAATGACATTTTGCCTAAATTAATATGCTCATAATATGATTTAACATATTATAGTTGAAA
ATATATGGTTTTATTTGATTTGTATATATAACAATAGATTTAATTGTTATAAAAAATGTAAAG
GGGTGTATGAATAGATTGTATAAATTTATTTTCGATAAACTAAGATTGCTTTTTGATTGTCTG
TAAAAGAGAAAAAGATTAAGATAAAAATAGTATTATATTGTAATTTATATTAATCAATTACA
AAGATTTTATGAATTTATTTCTTTAGGGTAAAATATTTAAGAATAAGATAAATTTACAATATA
ATACTATAACACTCTTTTATCTAGTTTTATTTCTTTATAGAACAATAATATTATAAATGCT
AGTAGATTTACACAGAATACTGTTATATACATCTGTTTGAATCCTGAGTTTAGAGTAGATTG
TAGTGTGGATCCGG

SEQ ID NO:35 >AV1368
ttttttGCGGCCGCaataccactacaccttctgc

SEQ ID NO:36 >AV1289
Tatacatctcgagtcccatctctttc

SEQ ID NO:37 >AV1288
gaagaaagagatgggactcgagatg

SEQ ID NO:38 >AV1366
CTTGTGCCATCTATATTTGTTG

SEQ ID NO:39 >AV1367
GGAAAAGGGATTGCTAATAGTG

SEQ ID NO:40 >AV1300
tccccGGATCCacactacaatctactctaaactcagg

SEQ ID NO:41 >DG1
GGCGCGCCACTAGTACCGGTGCCATGGCGGCCGC

SEQ ID NO:42 >DG2
AGCTGCGGCCCATGGCACCGGTACTAGTGGCGGCCCATG

SEQ ID NO:43 >DG9
ttccttgggtctcAcgcgAACAAAATTTCTCCAGTCTTC

SEQ ID NO:44 >DG10
ttccttgggtctcAggccGTCGCGACTAAGAAAATGCC

SEQ ID NO:45 >DG13
gtgagcggataacaattccc

SEQ ID NO:46 >DG14
AGATTGTAGTGTGGATCCGG

ES 2 622 280 T3

SEQ ID NO:47 >DG15
tccttcggcgcgccTCAAATTTAAGCTTAACTCC

SEQ ID NO:48 >DG16
TTTAGGGACTACTCACTCGC

SEQ ID NO:49 >CD4 Orf1374
MKRTKLLQRGNFFGDKNMVVDEFDEGYDNYDFINFFTGCCNYTFGLKNNNILYGCGDNSNFQ
LGLGEDNTRKLFTKIPNISTNIKKVACGESHAVILTS DGELLVAGINTDGQMGLGLEKVGK
TVSTFEKVPEIKGVKDIACGLQSTYLLYNDGTLVAGNNLYGQLGLGTNGASANVNTFTKVD
VDNVKAVFSYNKSAFIIKNDNKCYSTGFNNQGQLGLGDKNNRDLFSLVSINDVKTIACGSEH
TVLMTYNNDIYGCGKEKCFGNALQSSLFKIEEVNIKTIACGHGNTMLIDNKGTLKVAGNND
IYQLGIANYSENIDNSFIDLKNI VAKNIFIGLSHSILIDSNNDSYCTGDNTY GQLGSFFDDM
HIVEFKKMDSEKYSYSNYINLIKSEDKLTLKEEMEIKDIELPLDIHSVRDVVFSPYCTLVI
LGNGDVYGLGNRRYKMGMSDLPSQLNELTKLSISNVKSIVASKNISGGIFYIKNDDTCYYSG
PNSNSIAGVLPNSDVFKKISIDNVKKVVINTDL SNWFLIVTNNKQIYTS GKSSSYVNGLS
NALISQYTEISLSNVTDAYSSYNATFIVVDEKKVYATGINTNYLLGFSTSDGSNVNLGLLSD
WYYINISGSSYSRVSCTNNITKINNI I IYEYVTVFCTNIGSFLTGYHGT SWTKPTDSSYRVQ
YQGISYAGYLD SYIYNYYPTRCTQSSSSTTFAYLYNGESSNLKVNPNLLISGGSSYIHQ
YGRNYLNNQSSNNAASNINSGPITSDKAI FLYKALLYLSSNTLYGFGNISESAKELDVSDT
QDGYNATNYKKVMKNIKNI FIPPYDL SRDKTRFAILTDKSLFICGYNSKGTHGISVNSSLNL
NNKINYNKKNSSSEISSNIQEIYSHSKSTYLLTNNMMLYSVGLNDV GQLGVGDEINRKVFTK
INIDNIKSINVNRFTDNSKHAFAIKNDNTCYAVGLNNSGQLGIGDNVNRNIFTKINVENVKY
VAVYGNTSLLLLTNDGLLYGAGNNGKQGLGLGDTTSRNI FTRIPINGVRDVYLCNDVSIIVKN
DNTCYVCGLVNGYFGFTEGSI STFTKINIENVKSVVTAGSEATFFITNDNMIYTTGKKERV
FSTETNDIKGIRVINNI INAKKIVVNGYTSAILTNDNKL FVGGLSGYGSIANNNNTNSVEDV
KDVFTANNTLYIDNNNNLISGRD TYGISDES YRDMSPYYKVS IKKDVDTVFSSYNTIFI
KDIYKGFYSSTRDNRYNHLGIHHRYDNDKNEALEGSLHSYFKTDNTSDKIVFNKKNEKLVF
NDKYIKTNNKYINYKNI FKNDFKYTSIILPFEVSDIDISKTHSLAVAKDGKLYGIGSNSYKE
INQTLIEDIELLTLTEVNI SDVKKVACGDNYSYIIKTDNTLWSYGKNTEYQLGVGHNDVREL
QKVTGLPSVKDISIYNSMTLVLTNEGELYAQQYNTNGLFGLGESEKDKI IRTFTKVL TNVKE
IKSHNDDHILVIKNDNSLWITGKNKSMYKISISITDLYEFTKIPIPEHLNDILDIELSDDTI
YMITKVDTSKASIEIVEKSI SQVRVVVQDPNNVIEKLEMF INDELISTKTNLEINSIIFEIP
QNKIVLGENKILIKASSPTGDLYSSMFI FKSETGLKVKKDSILMINNKVYSI INITENNTDL
IVTLNEGLKDDMMENPIYQLINKTKVQVKINKSDLFKDMKLVEIKKSDSSYQEIYELEEAN
IKSAQPKIIVEKGDKWTAIKRPSMI FRYDAENNEPQA

ES 2 622 280 T3

SEQ ID NO:50 >CD108 Orf1374
 MKQNKLLQRGAYFNDKNILIDDFDKRYNDYDFVEFFTGISNSTFGLKSDGNLYACGDNTGFQ
 LGLGKDSSERRMFSKVKIDNVKYVSCGSKHSVAVTKDGFAYGAGTSNMGQLGVIESTVYEF
 TKLPIDDVKTVACGYDFTFVLKNDGTLYSAGLNSSGQLGLGDTNRRVTFKVNIDSVKDVVT
 YNQSVFIKMDGTAHACGLNSNGQLGINSTLNKSVFNKIEGMDNVKQIACGSSHTILIKNDG
 TMYTTGSNGYQGLGTGNNNSIVFTLSSINNVKYASCNNHTMILKYDNTLFSTGQNNYGQL
 ANANKDVASRNTFVKVNVENIKDIKCGSQFNFLINGSKEIFVSGCNLAGQLGSFFHTTFLYE
 FSKVQSSNLDNYSGLLVNDDYLVVTKDNSEFLNVKLSDNFQDYKKIELTDNNMFIVMNDGTL
 YACGLNNYQGLGLGDTVNRVSMTKVDIDNVLDIKGNSTFVLKNNGTLYSCGYNSSGILGL
 KDNTNRNIFTKIEIENIKEFCVESNYIVALNHSKELYGWNQSYIVYGDNRNYPYKDRVSN
 VEKIATWSDTLYILDSTGATKTIGYSYNGSGGYPAPSSSSTYREGGYINKNTSYRTLEFYNT
 SKTKLVNLFAYNGCVFVDENGLAYCIGENNINFRGGSTTNENNSLRFINNSGVYYTNTDGT
 DYTQYQWYTKLIRCSI FDS PQNIIIGNSKNILYLSKNNSTFKCTGNCITYGINSQNWYSYFSD
 SSNGAIALGNEFILKNYSGECLLKGYGKATNGEFGNSTNISSISNYDTGLKDIKDIIVKNNT
 VVVVDKNNNIYVTGANQFNKLGIGEYNNQPIRKFTNITEQSNSFI FMDDIKEITTSRNTMFI
 VKNDGTAYATGNSSGQLGLGDTINRNKFTQINLDNIKKISTSIDGNTTFAIRNDGTLYSTG
 LNTKGQLGLGDI VNRNTFTKVN IQNVRDVVLGTTSHAIKDDNTLYSCGENTHGQLGLGSES
 NHPDVLTFVNNITNVRDVYCSDTTTFIVKDTNIAYCCGYNNNSQLGMGNTTDQYSFIKME
 NVKEVIPNEINTYIITTYNTAYSTGLNTDYCLGLNSNSNQSSSEIPI SNVVKVAPNRNNAV
 LLLTSEG DVYTAGKCSNGSGTGSETPEKIKKIASKAKDIGMNYRCGHYVSDNGDLYGTGFND
 CGQLGVGNVTKRDTFIKTNRVKKILPLEYANIAIKDTNDIYICGLNNYQGLGVGNRYDSRN
 NDNRI FNYKHMNFVMDLTSIKNRHNFILLNNKIVIPTTKDIDYGLVLGNLYKGDLYTELPY
 EDIKEVSIKTHI IILLNDGTMYGCGTNYHGELLQDLSINQVDEFVQINVSDVKHVSCGDNF
 TYFIKSDDSLWSIGKNSEYQLGIGHNNPVTELQRITTISSCKEVHCGKNYTLVVTTSNELFV
 QGYNDKALGLGSDSENTI IKFFTKALTDIREIKSYGSDHILVLKNDNSVWVTGKNRDVYKI
 EQPVEFLKEFTIVPISEDVNTVKDVLATDNTLYI ISEVGTNAAEIETEKSISSIKIKIQDP
 NKDISRIEMLINGESVKSVDLITEKISFEVPPDKIKIGENKILFRAYCKGDDLYASLFIK
 ESTGNSIIKDSYVMIGNRMYKVNTTTSNEQDITITLDRGLEEDLNLGDP IYQLINKTKVQVK
 INKSDLFKDMKLV EIKKSDSSYQEIYELEEANIKSAQPKIIVEKGDKWTAIKRPSMIFRYDA
 ENNEPQA

SEQ ID NO:51 >CD123 Orf1374
 MKRKTLLQRGNFFGDKNMVVD EFD EGYDNYDFINFFTGCCNYTFGLKNNN ILYGCGDNSNFQ
 LGLGEDNTTRKLFTKIPNISTNIKKVACGESHAVILTSDGELLVAGINTDGQMGLEKVGK
 TVSTFEKVP EIKGVKDIACGLQSTYLLYNDGTLVAGNNLYGQLGLGTNGASANVNTFTKVD
 VDNVKA VESYNKSAFIKNDNKCYSTGFNNQQLGLGDKNNRDLFSLV SINDVKT IACGSEH
 TVLMTYNNDIYCGKKECFGNALQSSLFTKIEEVNIKT IACGHGNTMLIDNKGTLKVAGNND
 IYQLGIANYSENIDNSFIDLKNIVAKNIFIGLSHSILIDSNND SYCTGDNTYQGLGSFFDDM
 HIVEFKMDSEKYSYSNYINLIKSEDKLTLKEMEIKDIELPLDIHSVRDVVFSPYCTLVI
 LGNGDVYGLGNRYKMGSDLPSQLNELTKLSISNVKSI VASKNISGGIFYIKND DTCYYSG
 PNSNSIAGVLPNSDVFKKISIDNVKKVVINTDLSNWFSLIVTNNKQIYTSKSSSYVNGLS
 NALISQYTEISLSNVTDAYSSYNATFIVVDEKKVYATGINTNYLLGFSTSDGSNVNLGLLSD
 WYYINISGSSYSRVCTNNITKINNIIIEYEVTVFCTNIGSFELTGYHGTSWTKPTDSSYRVQ
 YQGISYAGYLD SYIYNYPTRCTQSSSSTTFAYLYNGESSNLKNVNP DNLLISGGSSYIHQ
 YGRNYLNNQSSNNAASNINSGPITSDKAI FLYKALLYLSSNTLYGFGNISESAKELDVSDT
 QDGYNATNYKKVMKNIKNI FIPPYDL SRDKTRFAILTDKSLFICGYNSKGTHGISVNSSLNL
 NNKINYHKNSSSEISSNIQEIIYSHSKSTYLLTNNNMLYSVGLNDVQGLGVGDEINRKVFTK
 INIDNIKSINVRFTDNSKHAFKNDNTCYAVGLNNSCQLGIGDNVNRNIFTKINVENVKY
 VAVYGNTSLLL TNDGLLYGAGNNGKQGLGLGDTTSRNIFTRIPINGVRDVYLCNDVSIIVKN
 DNTCYVCLVNGYFGFTEGSI STFTKINIENVKSVVTAGSEATFFITNDNMIYTTGKKERVF
 FSTETNDIKGIRVINNI INAKKIVVNGYTSAILTNDNKL FVGGLSGYGSIANNNNTNSVEDV

KDVFVTANNTLYIDN>NNNLISSGRDITYGISDES YRDMSPVYYKVS IKKDVDTVFSSYNTIFI
 KDIYKGFYSSTRDNRYNHLGIHHRVDNDKNEALEGSLHSYFKTDNTSDKIVFNKNEKLVMF
 NDKYIKTNNKYINYKNI FKDNFKYTSIILPFEVSDIDISKTHSLAVAKDGKLYGIGSNYSYKE
 INQTLEDIEELLTTEVNI SDVKKVACGDNYSYIIKT DNTLWSYGKNT EYQLGVGHNDVREL
 QKVTGLPSVKDISIYNSMTLVLTNEGELYAQGYNTNGLFGLGESEKDKIIRTFTKVL TNVKE
 IKSHNDHILV IKN DNSLWITGKNKSMYKISISITDLYEFTKIPIPEHLNDILDIELSDDTI
 YMITKVDTSKASIEIVEKSI SQVRVVVQDPNNVIEKLEMFINDELISTKTNLEINSII FEIP
 QNKIVLGENKILIKASSPTGDLYSSMFI FKSETGLKVKKDSILMINNKVYSIINITENNTDL
 IVTLNEGLKDDMMENPIYQLINKTKVQVKINKSDLFKDMKLVEIKKSDSSYQEIYELEEAN
 IKSAPKIIVEKGDKWTAIKRPSMIFRYDAENNEPQA

SEQ ID NO:52 >CD126 Orf1374
 MKQNKLQRGAYFNDKNILIDDFDKRYNDYDFVEFFFTGISNSTFGLKSDGNLYACGDNTGFP
 LGLGKDSSERRMF SKVKIDNVKYVSCGSKHSVAVTKDGFAYGAGTSNMGQLGVIESTVYYEF
 TKLPIDDVKT VACGYDFTFVLKNDGTYLSAGLNSGQLGLGDTNRRATFTKVNIDSVKDVVT
 YNQSVFI IKMDGTAHACGLNSNGQLGINSTLNKSVFNKIEGMDNVKQIACGSSHTILIKNDG
 TMYTTGYNGVGQLGTGNNMNSIVFTLSSINNVKYASCNNHTMILKYDNTLFSTGQNNYQQL
 ANANKDVASRNTFAKVNVENIKDIKCGSQFNFLINGSKEIFVSGCNLAGQLGSFFHTTFLYE
 FSNVQSSNLDNYSGLLVNDYLYVTKDNSEFLNVKLSDNFQDYKKIELTDSNMFIVMNDGTL
 YACGLNNYQQLGLGDTVNRSVMTKVDIDNVLDIKGNNGSTFVLKNNGTLYSCGLNSNGQLGL
 RDEVNRNIFTKIEIENVKDFCVGSNYVIALNHSKEVYGWGNPNYNNIEKTSNYPYKQGISNI
 EKIAAYDYSVYMINSEGKLYVSGYNYNYQLGKGNNSNQSKALVSQCRNTSSTSNGLRTP
 KITNVFPFYDGCALIDEGGYVYLTGYHGYLRTLNSSPSISDYSRYGT FIEATNSNHNTYFIQ
 ETDFSGIEKVI GMSNNILFFKKGSSYITGYPKTFGSTITGHRYSYTSINSESSNLGNSFI IYH
 SNSKLYGKGIANSQFGNSTNIDGTSNYDTGLKDIKDIIVKGNV VVVVDKNNNIYVTGMNQ
 NKLGI GEYNNPEVKKFTNITEQSNSFI FMDDIKEITTSRNTMFI VKN DGTAYATGNNSSGQL
 GLGDTINRNKFTQINLDNIKKISTSIDGN TTFAIRNDGTYLSTGLNTKGQLGLGDIVNRNTF
 TKVNIQNV RDVVLGTTHSHAIKDDNTLYSCGENTHGOGLGSESNHPDVLTFVNNITNVRD
 VYCSDTTTFIVKDTNIA YCCGYNNNSQLGMGNTTDQYSFIKCMENVKEVIPNEINTYIITIY
 NTAYSTGLNTDYCLGLNSNSNQSSFSEIPI SNVVKVAPNRNNAVLLLTSEGDVYTAGKCSNG
 SGTGSETPEKIKKIASKAKDIGMNYRCGHYVSDNGDLYGTGFNNNGQLGVGDVTKRDTFIKT
 NTRVKKILPLEYANIAIAKDTNDIYICGLNNYQQLGVGNRYDSRNNNDNRI FNYKHMNFVMGDL
 TSIKNRHNFILLNKKI VIPTKDIDYGLVLGNLYKGDLYTELPYEDIKEVSI SKTHI IILLN
 DGTMYGCGTNYHGELLQDLSINQVDEFVQINVS DVKHVSCGDNFTYFIKSDDSLWSIGKNSE
 YQLGIGHNNPVTELRITTISSCKEVHCGKNYTLVVTTGNELEFVQGYNDKGALGLGSDSENT
 I IKFFTKALTDIREIKSYGSDHILVLKNDNSVWVTGKNRDVYKIEQPVEFLKEFTIVPISED
 VNTVKDVLATDNTLYI ISEVGTNAAEITEKSISSIKIKIQDPNKDISRIEMLINGESVKS
 VSDLTTEKISFEVPPDKIKIGENKILFRAYCKGDDLYASLFI FKESTGNSI IKDSYVMIGNR
 MYKVVNTTSNEQDITITLDRGLEEDLNLGDPIYQLINKTKVQVKINKSDLFKDMKLVEIKKS
 DSSYQEIYELEEANIKSAQPKIIVEKGDKWTAIKRPSMIFRYDAENNEPQA

SEQ ID NO:53 >43593 Orf1374
 MKQNKLQRGAYFNDKNILIDDFDKRYNDYDFVEFFFTGISNSTFGLKSDGNLYACGNNTGFP
 LGLGKDSSERRMF SKVKIDNVKYVSCGSKHSVAVTKDGFAYGAGTSNMGQLGVIESTVYYEF
 TKLPIDDVKT VACGYDFTFVLKNDGTYLSAGLNSGQLGLGDTNRRATFTKVNIDSVKDVVT
 YNQSVFI IKMDGTAHACGLNSNGQLGINSTLNKSVFNKIEGMDNVKQIACGSSHTILIKNDG
 TMYTTGYNGVGQLGTGNNMNSIVFTLSSINNVKYASCNNHTMILKYDNTLFSTGQNTYQQL
 ANANKDVASRNTFAKVNVENIKDIKCGSQFNFLINGSKEIFVSGCNLAGQLGSFFHTTFLYE
 FSKVQSSNLDNYSGLLVNDYLYVTKDNSEFLNVKLSDNFQDYKKIELTDNNMFIVMNDGTL
 YACGLNNYQQLGLGDTVNRSVMTKVDIDNVLDIKGNNGSTFVLKNNGTLYSCGYNSSGILGL
 KDNTNRNIFTKIEIENVKAFCVESNYIVVLNHSKELYGWGNESYIVYGNRNPYKDRVSN

VEKIATWSDTLYILDSTGATKIIGYSYNGSGGYAPSSSSTYRDGGYINKNTSYRTLEFYNT
 SKTKLVNLFIFYNGCVFVDENGLAYCIGENNINFRGNSTTNENNSLRFINNSGVYYTNTDGT
 DYTCYQWTYKLRCSIFDSPQNIIGNSKNILYLSKNNSTFKCTGNCITYGINSQNWYSYFSD
 SSNGAIALGNEFILKNYSGECLLKGYGKATNGEFGNSTNISSISNYDTGLKDIKDIIVKNNT
 VVVVDKNNNIYVTGANQFNKLGIGEYNNQPIKFKFTNITEQSNSFIFMDDIKEITTSRNTMFI
 VKNDGTAYATGNSSGQLGLGDTINRNKFTQINLDNIKKISTSIDGNTTFAIRNDGTLYSTG
 LNTKGQLGLGDIVNRNTFTKVNIIQNVRDVVLGTTSHAIKDDNTLYSCGENTHGQLGLGSES
 NHPDVLTFVTNITNVRDVYCSDTTTFIVKDTNIAYCCGYNNNSQLGMGNTTDQYSFIKME
 NVKEVIPNEINTYIITIYNTAYSTGLNTDYCLGLNSNSNQSSFSEIPISNVVKVAPNRNNAV
 LLLTSEGDVYTAGKCSNGSGTGSETPEKIKKIASKAKDIGMNYRCGHYVSDNGDLYGTGFND
 CGQLGVGDVTKRDTFIKTNRVKKILPLEYANIAIKDNTNDIYICGLNNYGQLGVGNRYDSRN
 NDNRIIFYKHMNFVMDLTSIKNRHNFILLNNKIVIPTTKDIDYGLVGLNLYKGDLYTELPY
 EDIKEVSIKTHIILLNDGTMYGCGTNYHGELLQDLSINQVDEFVQINVS DVKHVSCGDNF
 TYFIKSDDSLWSIGKNSEYQLGIGHNNPVTELRITTISSCKEVHCGKNYTLVVTTGNELFV
 QGYNDK GALGLGSDSENTIIKFFTKALTDIREIKSYGSDHILVLKNDNSVWVTGKNRDVYKI
 EQPVEFLKEFTIIPISEDVNTVKDVLATDNTLYIIEVGTNAAEITEKSISSIKIKIQDP
 NKDISRIEMLINGESVKSVDLITEKISFEVPPDKIKIGENKILFRAYCKGDDLYASLFIK
 ESTGNSIIKDSYVMIGNRMVKVNTTSNEQDITITLDRGLEEDLNLGDPYQLINKTKVQVK
 INKSDLFKDMKLVEIKSDSSYQEIYELEEEANIKSAQPKIIVEKGDKWTAIKRPSMIFRYDA
 ENNEPQA

SEQ ID NO:54 >ΦC2 Orf1373
 MAIDKSYTYIITDVGKAKIANASVTGNKVGFKIQLGDDGGSEYTPTESQTALKNVVWEGNI
 GNTTDETAPNCIILESLIPSSVGGFMIREIGYLD DENNLIAISKYKECYKPSIEQGAVVDM
 KVKTVLIVSNVNNIELKIDPTIIFATLKDIQDLETKIGTVNTKIDTTKTELTSNIETTKTEL
 NTRIDTENEKQNIKIDQLIAGGSNVASTQIITIDDWVEDAENGFKATVTHSLLTQRIVVNI
 IDATTKENVVTNFKIIDDNSIEIRSEVKVELNVYVINGNAETHFINATVDDNRVSEMTTYSSK
 KIEDRLVNIIEKVNGLSNIATSVNELITYC

SEQ ID NO:55 >ΦCD119 Orf1373
 MAEQQYFTLVTDIGKAAIANASVTGEKVDFAKIKVGDGGGSSYTPNESQTALKNVVWESTLE
 HAQVDKDNPNWVVIQKFI PGDVGGFEIREVGLFDSKDQLLAVSSYPTTYKPESRFGDCKRTI
 NKSNI SCI

SEQ ID NO:56 >ΦCD27 Orf1373
 MPNELNFNNEIEEYLITTPAHANEFNRRQKLLDNDKYLNKIDTTKTELNTRIDTENEKQNI
 IKIDQLIAGGSNVAYTQRVAIDDWVEDAENGFKATVTHSLLTQRIVVNIIDATTKENVVTNF
 KIIDDNSIEIRSETRSELNVYVINGNAETHFINATVDDNRVSEMTTYSSKKIEDRLVNIIEK
 VNGGLSNIATSVNELITYC

SEQ ID NO:57 >DG211
 ggccgcctcgaggg

SEQ ID NO:58 >DG212
 cgcgcctcgaggc

SEQ ID NO:59 >DG210
 TGAAGTACCATGGTATCCAG

SEQ ID NO:60 >DG209
 ACTGGATACCATGGTACTTC

SEQ ID NO:61 > Locus Dif4 sin orf1 377-1379
GCAATACCCACTACACCTTCGTCATCTTTAAATTTAAGAGTTTTTACTATTGAATAATAAAG
GTATATTTCCAGTAAAAATAATCTTTAAATACAAGAAAAATAAATCTTTGGGTATATTA
AGCTAAAAAGTGTAAATATAAAAGCAAGTAGAGTACTTATCCTGTAAAAGAAATCTATTTGT
GTAATGTCTTTATATTTTATCATAAACACCGAATATAAAATGATGAAAACAATTGCGACGAT
TGCATATATGGTAAATAACATATTTCAAGAGTACCATTTGAAATTACTATCCACTTATACC
ACATAATTGGCCAAAATAATAGTGCTAAGAAGCTAAAATAATTATCAAACAACCTTTTCTTTA
TACATTCATCAAACAACCTTTCTTAACAAAAGCATATATTTGTTTTTAGAATTTTAAATAAT
ATGATATCATTATTATATATTAATATTTGAATTTATAGAAACCAAATTTGTTAAAATAAATA
TATAGATTTTACTGTTAAGCCAGTTAAAATTTACTACTATTTTATTATGAAATTTGGATCAAA
TATGTAGAAAATACGGCAAATTAGTTAATATTAAATATTTATTATTTCCAGTTGTAAAGACT
GTTTTTTTAAATGATAGAAATCTAATCTTTTTTGAAGAAAGTAATATCCACATTAAGTATG
TCTGCCATTTTATAAACCGCAAGTATGCCAGAGTTAATTATGTTTATTATATCTTCTCAGT
AATTAAGAAGCTCACAAAGCCATTTTAAAGCTTTATTTTTCGCACTTATCTATAATAATTTTGT
TATAATAATCGTTATAAGAGGATACATAGTATCCAAGGCTAGTGAAATGATGTTCCAAGTCT
TCAGCTAAGATGGATGTCAATTTTTTTGAGTTTTGTTTTAAATTACTGAGTAATGATATAAT
TTAATACCATGTTTGTATATATAGTCTTCTAAATCACCTGCAATATAAGTGGTATAAT
GAATTTATATCTCTTCTTGAGAAGCTAATTCAAAAGCTTATCCAAATTTATCATAAAAATC
CCCATAAATAGAAATGATGTTTGCCTTTAAATTATATTTAAAAGAGCAGAAAATAGACTGC
TCATCATATGGTTTATTTTTTTTTTATTTTATTTAGTAAAAATCTATATAATCATTAAGTT
GTTCTTGTGCTTCTTCAGGTAACCTCTTCATGTGGATTTTTTCTATGTGCAGCTACTGTATCA
ATATTTTCTTAACTAAGGTTCTTCCAAGGTAATCAACTGATACATTAAATACATCAGC
CAATTTATTTAAAATGTGTTTCATCAGGAAATCTGTTTTCTGTTTCATAGTACCCTAAGACTC
TTTGGGAAACGCCTACTTTTTCTCCAAGTCTCTTTGAGTCAATCCAAATTCCTTTCTAAGT
TCTCTTAATCTTTTGGCAAACATTATAACACCACCTTATGTATAGATTATAACAAAATGTTT
TAAAAATAAAAATAATAAATAATAAAGAAATTTTTTTTCTAAAATCTATTGATAAAGAAC
AAATAATCTATATAATCTAAGTGAGGAAGAACAATAATCTTAAATAGTAATGGAGGTATA
AAACAATGTTTAAAATAAAGTGAATATTATAGAAAATGCAAAGGTATGACACAAATTCAA
CTTGCCGAAAGGCTGGAATTACAAATGATTATATATCTCAAATACAAAGAGGTATAAAAAA
TCTTGGACTTCTTATGGCTAAGAAGATTTCTAGTATTTTAGAACAATAATAGAAAGAGTTT
TTTTTATACAGTTATAGAACAAATATGTTCTTGAAGGTTGTGAGATTAGTAAAAAAGTGTGCA
CTAAGAGATTATTGTAATTTGAAGCTAATAATAAGTATATAAAAAAGGGGAAAGTACTATG
GAAAACAAAAAGATATATTTTAAAGAAACAGATAAAGATTACATAATTATAAGTATTT
GGATATAAAGATAAAGAATATTAACCTGGACATAAAGATGTGAGAATGAATCTCTGGAT
GTGGAGCAATGGTATATACAGAAAAGACTAGTAACACATATAACATAAGCTCTCTGTGGAA
AATGAGGTGTTAAAAGAGAGGAAAGATTAAGAAAATTTAAAATGAAAAGAGATATAGA
AATAGAAAAGAGAAAGATAGAAAATGCTCTAACGTGTCTAATGATATAGAAATGGAATTTT
TTAATCTTTTTTATAAATAGTAAGACAAAAACAATATGACATATATTTCTATGAACTACAC
TTAGATAGAACATCTTGCTACAATTTAAAGAAAAGATGATATTTAAATGAGTGAATATT
ATAAAAAATAGGACAATTTTACAACACTTTATATACACCATTGCAACAATAGGCAATAAAAT
ATGTGAGATAATGTTATTGTGAAAGAAATCCATATTGAAGGAGGTGATAGATTGAAAAGAA
AATATTACCTAAAAATATAGAAGTACTTGACAGGAATAAATGAGATATATTTTAAAATG
ACTTATATCATTTATAGTAAGATTATCAGATTAAGCAAGAATATTTAGTGATAGTGTGGTGA
TTATTTGCTTAAATACAAGGAAATATTAGAAAACAATTTTGGAGATTCTCAAAAAAAGCTTTA
CTGAAAGTATTTTTATTGATGATGAAAGTGTGCAAGGCTCTGAAGGCTCTGTTTTTTTGT
AGTATACTATCAGTTATTTGTACACCTGTAATGTTAAATACGAATAACAAAGATATTGTTAT
CTCTATAAAATACTTACCAAACCACAGTCAAAGAGTATTAGAATGTATGAAATTTTCAAGT
AATTAATAAGCTATTTAACAGAAATATAAAGGTAACAGACAGAAAATTAATATAACAAAG
CTAGAACAAGTATTTAAAAGAAAGAGTCAATTTATGTATTGAACTTTACATTTTACACTAAA
CTATCTGGATAGTGATATGAAGAAGATGTAGTATATGAAAATATGAAAGAAATCAATTTAA

ATTTAGGAGAGTGATAGTATGGCTATAGGATTACCAAGTATCAACATATCATTAAAGGAGCT
 AGCTACAACCTGTTAAAGAACGTTTCAGCTAGAGGAATAATTGCAATGGTACTTAAAGATGCTA
 AGGCACTAGGTTCTTAATGAAATACATGAAAAAGAGGATATACCAGTTGATTTATCTGCTGAA
 AATAAAGAATATATAAATTTAGCTTTGATGGGAAATGTTAACACTCCAAATAAATTATTAGT
 TTATGTAATAGAAGGAGAAGCAGATATTCAAACTGCATTAGATTTTTTTAGAGACTAAGGAAT
 TTAATTATCTATGTATGCCAAAAGCAGTAGAAGCTGATAAGACTGCTATAAAAAAATTGGATA
 ATTAAACTTAGAGATATAGATAAAGGTTAAGGTTAAAGCTGTATTAGGAAAAGTTGTAGGAAA
 TCATGAAGGGATAATTAATTTTACTACAGAAGATGTGTTAGTTGGAGAAAAGAAATACAGTG
 TTGATGAGTTTACAAGCAGGGTGGCTGGACTTATAGCAGGTACACCTTTAAGTCAATCAGTA
 ACTTATACTAAACTTAGTGATGTAGTCGATATACCTAAGATGACGAAAAGTTGATGCAGAATC
 AAGGGTTAATAAAGGAGAGCTTATACTTATTAAGAAGCAGGAGCTATAAGAATTGCTAGAG
 GAGTAAATTCCTTAACTGAATTAACAGCAGAAAAAGGAGAAATGTTCCAGAAAATAAAAATA
 GTTGACACTTTAGATATTATACATAGTGACATAAGAAGGTTGATAATAGATGACTATATAGG
 AAAGGTTACTAACAGTTTATGACAACAAATGTTTATTGATAGTAGCTATAAAAAAGTTATTTAG
 AAGAATTAGAAAAATCAGCACTTATAGAATCTGATTCTACTGTTGAAATAGATTTTGAAGCA
 CAAAAATCGTATTTAAAAATCAAAAAGGAGTAGATTTATCTTATATGACATTACAAGAAATAAA
 AGAAGCTAACACAGGTTCTAAAGTATTTTTAAAAAGCAAAAAATAAAAGTACTTGATGCTATGG
 AAGATATAGATTTATCAATAGAAATATAGGAGGATTATTAATATGGCAAATATGGAAGCTAG
 AAATGTAATGAGTGGTACTTGGGGAGAACTTTGGCTTGATGGAAACAAAGTAGCAGAAGTAA
 AGAAATTTCAAGCAAAGATGGAATTTACAAAAGAAGATATTATAATAGCAGGTCAAATGGGT
 ACTGATACAAAGTATATGGGATATAAAGGAAAAGGCTCAATAACTCTATACCACGTTAGTTC
 AAGAATGCACAAGTTAATTGGAGAAAAGATAAAGAGAGGTTCTGAACCTAGATTTGTTGCTA
 TATCAAAAATTAATGACCCAGATTCCTTATGGAGCAGAAAAGAATAGCAGTAAAAAATATAGCA
 TTTGATGATTTAACTTTAGCTGATTGGGAGGTTGGAGTAAAAGGAGAGATAGAAGCTCCTTT
 CACATTTACTGAGTATGATTTTTCTTGATATAATTTAGTTTTATATTTAGTTTTATACTGATA
 TTTAGTAAGTATATACTTAATAAATTCAGATAGTTAATAAGTAAAAAAGTTAGTTGATTGAA
 TTTGATGATAAAGGAGCAAATAATAATGAGTGAAAATGGATTATCAAAAAATATAAACATA
 GTAGATTTACTTTTTAAATTCAGATACAGAAAACCTTAGAAAAGACCAAGTACTATAGTTGAACT
 TAAGAGATTTACTATATTTGGGCAGGAATTTAAAGTAATGTGTAGAGCTTTAACAATAA
 GTAAAGATGAAGAAATACAAAATACTTGCTTAAATTTGATGAAAATATGAAAACGGATATA
 GACTTACCGGAGATGCAGATGCTTACAATTATAGAAGGTGTTTGTGATTTGGATGGAAAGCT
 TTTATTTAAAAATAAGGAGCTAATGGATAAATTTAAGGCTCCAACACCAAAGAATTAGCAA
 GAAAAATATTATTACCAGGTGAAATTTACCAACTTATATAGAATACTTCAAGATGTTATGGGT
 TATGGTAAAAATGCAGTGATAGAAGAGGTAAAAAACTAATAGGGACGGATACCAGGACTACA
 ATAATGTAATTTATTTGGAAGAAAAAAGGTATAAGACCGTCCCTTTTTTATGCAATGGATAA
 AGGCGAATTAAGCTTATTTGAAGCTTTTTTCGCCTTAGAAAATTGAGGAAGAAGTTGAAAAAA
 TGAAACATGGATATGGAGTGTGTCCTTTGACAGGAGGTGGTATGTAATGGGAAATGTGAGAG
 AAGAAGGTATAAATATGTACCTTACAGATAATTACACACCAAATAATGAATCAAATTATATCA
 GTAACGTATAATTTTAGGAGAGCAACTGTGGCTGTTTCACTTTCCACTAATGTAATGGCTAG
 TAGCATAAAAAATTTCTATTTGGAAGTGCAAGTAATAGAGTAAACAGTTTAAATTCCTCGTTAA
 GAAAAGTTCAAACACTACTGCTAGTAGTGTAAAGTTCAAATGACAAAATTAAGTTCTAGCATA
 AATGCTGTTTTCAGGAGTTATTTGGAAGTTTAAATGGAAGTATTTATGAGACTAGCAATAACTAT
 AGCTATGATTATGATTATTTTAATAAGTTGATTCAAAAAGAAAAATGAGTTTAATTTCAAATA
 TTATGATTATATTAATATTTAAAGCTAAAAGTGATGAAGTAGAAAAAACTAAAAATAAATTA
 CTTGGAAATTTAAAAAAGATTTGGTGGCAAGATTTGGAATATCGTAATAAAAAGCAAAGATAT
 GACTAAGAGAGTGATAAGTAGTATCTTTGGGAAAATTAAAAACAAGTAGAGAAACGTCCTTATC
 AAGGAAGTATTAATCTTAAAGATATGGTGAGTAGTGCTATGGGTAGAATTTTGCCTAAGTTA
 ATGTTGTTTTAAAAATACTTTTTGGAGTGGTGTAAATAGCTATAAAAAGATATGGCAAGTGGCAT
 TATAAGTAAAGTATTTCCCAAATTTGAGATTTATTTGCAGGTAAGGTATGGAGTGGTGCAATAG
 CTGTAAAGGATATGGCAAGTGGAACTTTGGTTCGATAAAAAGGGAAGATATCTGATTTGACA
 AATGGTGTACTATAGGTGTCGCTGTGAAAAAGGGTGTGATTTACTTGGTCAGGAACAAAA

TCAGAAAGTTGTTCTAGAAAGTGTAATGAAAAGAAATACTGGAAAAGTTAATCAAATAGATG
TTGATGATTATTATGGCAGTTTAGTAAGAATGGCAAATGATACGCCTTTTGACCCTGAAGAT
GTTGTTGCAATGGGAACTAAAGCTAAAATGATTAGTAATATTACTGGTGGCAAAAAAGAAAA
AGATATAACTCAAGCTATGGTAGATGTTAGAGCTTTAAATATGAATACAAGTAGTGAACAAG
ATGTATCAGCAGCTTTCTTAAGTGCAGCAAAAGGAAACATGGAATCTCTAATACTCTGGTA
GGAGAAAATTATAAACTTTTGGATGAAGCATTGGAAGGCATAAGTGTAAAGCAGATGGGGTT
AGCTAAAAGAAATGAGTAATACAATACCAGGTATAATATCAGGAGCTCAAACAAGCATTAAACA
ATGGCTTGAAGAGTATTGTTAAACCTTTTGGATGATATTTTAGGTCAAGGACTAAAAGAAAATA
AAAACCTTTTATAGAAAGTGGATTAGGCAATTTAGCTGGCTTATCTGAAAAAATGGCTGGTAA
AATAGGCAATGTAATGAATGGTAAGATAATTATTGGCAACAAATATGACCAGATGCAATCTA
GAAGTGTAAAAAATGGAAAAGGTTTTCTGATTCACATCAATATCGAAATTTCTAATGAGGC
GAAAACGCTAAAATGATGGTTGAAAATAAGCAAGAACGTTTTGAAAATCATGCAGCAAAAT
GATAGGGAATGCACCAAAAGCAATTTGTTAACGCAGGAAGTACACTATTACAAAATATCGATT
TTACAGCATTAAATAGATTCATTACTTCCAGTAGTAACTTAGTAAATAATTTACTAGATAGT
ATAAAACAATAAATCACCAATTCACAAGGATTAATAAGTATATTTGGTACAATAGTAACTAC
AGCATTCCAATAATCGGACCTGTAGTTGAAGCTGTTAGTCCTATTATCACAAGAATTTTTA
CTTTTTTAGGTGAATATGCACCTCAAATAAACAATTTTTATAGAGACACTGGGTGTTATTTGG
AAAACCTGTATGGGAGACCTTAGGACCTCTGTTGGAACTGGATGGAAAATTTATAGAGCCAAT
ATTGGGAGCTTTTTTTAACATATTAGATAAAGTATGTAAAATAGTTAAAGATATATGTAAT
GGTGGCAAACTATGATTAATAAGATAAAAAATGGAAGCATCACAGGAACAGTTTTAAATCTA
GTGAAAAGAGTAAAAAAAATTACAAAGATAATCCATATGCTGGAACAAAGGCTGGTGATTC
TGGTAAAGCTTATTAGGTAAGAAAGGTAATAATGCATTTGGATTGAACTATGTTCTTATA
ATGACTATCAAACCAGACTCCATGAAGGTGAAATGGTTTTAACTAAACAAGAAGCAAATCAA
TATAGAAGCAGAAAAAATGGTGGAAATATAAACATAGCTAAGTTAGCTGATACAATAGTGAT
TAGAGAAGAAGCTGATATAGAAAAGATAACATCAAATTAGTTGCAAGTATCCAATGGCAC
AGTTAGGGGTGTCTTATAATGGAAATGTGGCTTAGACAAGCTGAAGATAGATTTAGATTC
CAGTATTTCCATCTTCTTTAGTATTAATGGAAAAGCTGCTGTAACTCTTCTAGTATACTC
AAAATAGGTGAAATAGCAACTTTTTGGTGGTGTAGCTCTTAAAAGCATTTCATATCAAGTTT
TTTTCCAAATAAAGACTATACTTTCTGTGACTATACAGTTTTTCCATCACCATATGATTGTG
TAAATAAGATAGAAAAATGGATGAAGGAAGGTTTTATATTAAGATTTACAATTACGGAAAACA
AATATAAATATGGAAGTCATAATTGAAGGGTTTAGTTATGAAGAAAGAGATGGGACTCGAGA
TGTATATTTTACATTAGATTTAAAAGAGTATAAAAGAATAAAGATACCAAAAGTAACTCCAA
AACAACTATTATAGATAAATAAGTTATAAATAACTGCTGATAGAATTAATGAAAAGGCA
GGTGATTTTTTATTATTAAGATTTGGGTACACATAAAAAACGGAAGTATATATGACATAACT
GACATAGTAGACAAGGTATCATGGTCAGGTGATTATAAATCTCCATCAAGGACACTAGAGTT
TTCAATAATACAATCATCATTTTGTATGTAATTTCCAACAAATCGATATACCAATAGCTAGTA
CAGTCTGTTTCTATGTAGATGAGAAAAGACTCTTTAGAGGAATGATAATTAATAGGTCTAAA
GATTC AAGCAGTAATGAAATTAGTTTTGTATCTAAAGATATGGGATTTTTACTTACACAAAG
TGAAGTGTACATAAATTTAAAGATAAGTTAGTTGAAGACATAGCAAAGCAAGTATTTGCTG
AAAATAGGCTTTTCAGTTGGAATAATAGCAAAGACCAATGTCAAGTATACAAAGATGTTTATA
GGAGTAAATGGTTATGACACAATAATGAGTGCATATACAGAAGCAAGTAAAAAGACAAAGAA
AAAGTATATGATAGAGGCCAATTTAGATAAGTTTAAATGTTATTGAAAAAGGAACTGTTACAT
TAAGTGTATGTTTGAAGAGGGATTTAATATTATAAATACCACCTTTTCGGAGAGCATGGAA
AATGTAAAAAATAAAGTAATAGTGGTAGACCAGTATGGAAGCAAGATTAGCGAAAAAATAGA
TAATGAAATTTTTAAGGAAGTAAATGTAATAATGCAAAAAGTAATTCAGCAACAAGAAAATC
AAGATGTAGATATTGATAGCGAGTTTTAATGGGATAGAAAAAGCTGTTCTCTTAAAGGTTAT
GGAGATGTAAGTTGTATAACTGGTAGAGGAGTAAAAGTTAAAGATTTCTTATACAAAGCTTGT
AGGACTATTTTATATAGATACAGACAAACATACTTGGCAAAATGGAGAATATCAAATGAGC
TTGAACTAATTTTTCAAATCTTATGGATGAAAAGTCAGCAGGACAGGATGAACCTAAGGAA
GAAAGTAAATTTAGGGGGAGAAGATTATGCAGGAGGAAAAGAGTTTACAGCAGAATTTACAGC
TTACTGTCTAGAAAAGAAGAAGGTGGAGATACAGATTGTAGAAAAGAAAAAATTTGACCCAT

CTAAAAAACTTGCCTGCTCCTATGGTTGGTAAATATGAGCAAACCTTATTATACAAAAGAG
TTTTTAAATAAACATCCTTTATTAACTATGGAGATGAAATACAGGTAATTACAGGAGTTTC
TGGTCGTGATGGAGTCTATAAGTAAATGACGTAGGACCTGCAATAACTATAGAAAAGAATG
GAACATACCATATAGATATTTTATTTGGAAATGTTGAAGAAGCTAGTAAATTTGGAAGAAGA
AAAGGAAAAATTTATTATTGGTGGTTATTCTGGTAATGTATCTGATAAAGCTAAAATAGTAAT
ATCAGAAGCAAAAAAACATCTAGGTAAACCTTATAAATGGGGTGGAAATGGACCAAGTAGTT
TTGACTGTTCTGGTTTAAATGGTCTATTGTTTTAAAAAAGTTAATGTTAGTTTGCCAAGAACG
TCAAATCAACAATCTAAAAAAGGCAAGAAAGTAGAACAAAAAAATCTTCAAGCAGGAGATTT
AGTATTTTTTTCATAATCCAGTCAGCCATGTTGGATTATATATAGGTAATGGAGAATTTTTTAC
ATGCTCCACAAAAAGGTGATGTAGTTAAAAAAGTAAGTTAAGTAGTAGAAGAGATTTTAAAC
ACAGCTAGGAGAGTATTATAAAGGATGGTGATATAATGGCTAATCCAATAAATGAATTTAT
AGGAATAATAAGAGAAGAAGGAAAGTATCATAATCAACCTTCTTTTTTTATTGGAAAAATTA
AAAGTAAATACCAGTTTAAAAATAGAGACAAATAACATCATATTAGAAAAAGAAGATATT
TTGATAGATTGTTGGATTGATAGACAGCTAGAAAACATTTGACACAGAAAACAAATCAAGA
ACACCAGCATGAAGTAAAAAATCCTTTTATAGATAAAGTTGAATCTGGGGATATGGTAATAA
TGTTTAGAATAGGCGAAAAAATTTGCTGTTGTAAGTAAGTTGGTGAGCTTATAATGAGTACAA
TATTTCTTTTATAGGTGTCCAGAGGATTATATCTTACCTAAAACAGAAGAATTGCCAATC
TTTCGTGAAGTGGCATGGGATTTTGAAAAAGATGAACCTATTTTAGAAAAAGGTGACTTTAA
AATAATTTGAAAAAAGAAAGCCTTAAAAAGTTTGGATATACAAGTGTATAAAGACAAATAGAT
ATGAACATGAGATATACTCTTTAGAATATGGGACAGAGCTTTCAGAATAATAGGACAAAAA
TATACAAAAGGTCTTACAGAAAGTGAAGCTAGTAGATTTCATAAAGAGGCCCTTCTAATAAA
TCCATATATATTAGAAGTAAACGTAAAAAGTGCTAACTTTAACAGAGACATATTGAGTGCAA
ATGTAAGTATCCACTATCTATGGGGAGGTGGAAATAAATGTATAGTGACCAGACATATGA
AGTAATAAAAAATAGAAGTCTTGAAAAATATTAATCTTGATATTTATAAAGGAGAAGGTTCTT
TTCTAAACAACATGGTATCTGGAAATAATCTAGAAGTTCGAAGATATATCTAGAAGTTTCA
AAGATGCATAAAATGGCTTTTATACAAGACACATATAACCAGTTTCTTGATAAAGAGTCAA
TGAATTTGGTGTATATAGAAAAGTTAGGTACAGAGTCAAATGGAGAAGTTGAATTTATTGGAG
AGAAAGGTACTGTAATAAATAATGGCACAATAATATCATATAGAGATTTACTATTTGTAGTA
ATAAAAAGATGTAAGTATTGGTAGTGAAGAAGGTGACAATAGCCCAGTTCAAGCTCTGGAAGT
TGGTAAAGAAATATAATTTACCTACAAATTTGTGAATTTAAACTAGTTGATAATATATCTGGAG
TAACAAAGATTAACAACAAGAAGTTTTGAAGGTGGTACAGATATAGAGACAGATGAAGAA
CTAAAAGAAAGATTTTATAAAAATCCAAAGAAATCAAGCTACAAGTGGAAATAAAGCTCACTA
TGAAGAATGGGCTTTGGAAGTAGATGGAGTCTATAATGTAAAGGTTTATCCAAGATGGGATG
GTCCGGGAACAGTTAAGGTCTTGATATTTGGGAAAAATAATCAAGCTGTTGATACAGAAACA
ATTGAAAGGTGTCAGCAACATATAGATGAAGAGAAAGCCTATTGGACCAACTATAACAGTTGT
GACACCATTACCAATAGAAAATAAGTATAAGTGCAGTAATGAAACTAGAAGATGGATATACAT
TAGACAATGTAAGAAGTCTTTCTAGAAAAGTATAAATACTACTTTAGAGATATTAGAGGA
GAGATAATCTATACAAAAGTAATGGGAATACTTATAAATACTACTGGTGTACACGATTTAAG
TAACCTACTTATAAATGGAAGTACAGATAAATAACTATTAATGAAGATAAAAATACCTAGTG
TAACAACGTAAATTTTAGTGAGGTGGAAAAATCAATGAAGCTAATTGATAAACTACCATCAT
TTGATAGAAATFACATTTGTAGAGGAGATACAAGGTGCATACGATACAGAATTAATATTTCTT
AAAGAAGATATTGATGATACCTTTAACCAATTATTTGTTGACACTGCAACATGGGGATTAGA
TATGTGGCAAGACATACTCTGCATTTGAAAAAAGAACTTTGATTTTACACAAAGACGTAGCA
ATATAAAGCTAAAATGAGAAGCAGAGGTACTAGTACTATTGAAGTTATAAAGTATATGT
GAGGCATATACAAAATCAGAAACAGATATAAAGTTTATAGTGATGAATTTACATTCGTATT
GAGTTTATAGCAAATAACTGTGACTATAAACTCTTTTAGATTGTAGCGATATGATTGAAA
GAGTAAAACCTGCTCACTTATTACACTATTTAGAACCAATAATACTAGATAAAAAGTATGGTC
TATTGTGGTGGAGGTATGGTATGTAGTGAAGAGGTAAAAAGTTCATCCATACTTTGAACCAAT
TATAAATGTAGTGCTGTTGTAAGTGTGGAGCTGGAATGATAAGTAGAGAAGAAATAAAGG
TTTATCCTTTAAGCATTAAATGCATTTGAAAAATAATGTAAGATTAATATAGCTATTGCAAAT
GATACAGGTGTAGAAAATGTAGTAGTTTATCCTAAATCGGAGGTGGTATAATTGGAAGAAAA

ATTTTATATAATATTAACCAAAATTGGTAGAGAAAAATAGCAAATGCAACTGCAC TAGGAG
 AGCTTGGTTGGATTAAACCAAGTTTCAAGTTGGAGATAGTAATGGAGAATATTATGAGCCAACA
 GAGGAACAAACTGCTTTAAAGAATGTAGTTTGGGAAGGAAATATAAATTCICTAAGAAATTGA
 TGAAAAAAATCCTAATTGGATAGTTATAGAGACTATTTTACCAGGAACAGTTGGTGGATTTA
 TGATAAGAGAAGCTGCTGTTCTGGATAATGAGAATAATATAATAGCTATAGGTAAGTATCCA
 GAGACGTATAAGCCACGTGCTGAAGATGGCAGTATTAAGATTGGTTGTAAAAATGATTTT
 ACAATTGTCCAATACTTCAAATGTTACATTAGAAGTAGACCCGACGTTGGTTTTTGTAACTC
 AAAAGGATATTCAAGATTTAGATGATAAGTTTGATAAAAATATAAAAAGAAATAAAAAGTAAAA
 ATTGGCGAAGAACTCTTATCTACAGAAGCTAAAAACTTATCAGGAGCTATAAATGAGGTAGT
 AGAAAAAATTA AAAATATATCTATTGATGATGTAATAGGAGGTCAAATACAACTGAAC TAT
 CTGTATTA AAAAATAGTTACAATAAAT TATCTGAAAAAGTATTAGATATCTTAATATACTTA
 GAATTAGAGTCAGAAATAGATGTAGATGAAGCTGGATATTGGTATGATACCTTAAC TAATGC
 TAAAAACATAATAGCTATAGAAGGCCTTAAGTTAGATTTAAATAGAAAGTGTATAACTGGAG
 AACTTGGTAGTGTACATTTAAGAATGTGGTGTACCATTTAATGCAAATAGAGTTAGATAT
 ATACATGAAATGGATAATAA ACTTTGTTGAAACAAAATCTAATAGGGCATATTCAATTGGTCA
 GACAGATATAACTTTAAATAAAATATTTCGTATGAAATAAGATAATTAGGAGGTTTTTATAATG
 AAAAGA ACTAACTACTTCAAAGAGGTAATTTCTTTGGCGATAAAAAATATGGTAGTTGATGA
 ATTTGATGAAGGGTATGATAATTATGACTTTATTAATTTTTTTACTGGATGTTGTA ACTATA
 CATTGGTCTAAAAATAATAATATCTTGTATGGATGTGGAGATAATAGTAACTTTCAACTT
 GGATTGGGAGAAGACAATACAACAAGAAAATTTATTTACGAAAATACCAAATATATCTACCAA
 TATTA AAAAAGTTGCATGTGGAGAATCTCATGCAGTTATACTTACTTCAGATGGAGAATTAC
 TTGTCGCAGGTATAAATACAGATGGTCAAATGGGATTGGGATTAGAAAAAGTAGGGAAAACA
 GTTTTCTACATTTGAGAAGGTTCCAGAAATAAAAAGGCCTAAAGGATATTGCATGTGGACTTCA
 ATCAACATATCTTTTATACAATGATGGAAC TTTATATGTTGCTGGAAATAATTTGTATGGTC
 AATTAGGTCTAGGAACTAATGGAGCATCTGCAAATGTAAATACATTTACAAAAGTAGATGTT
 GACAATGTAAAGGCTGTATTTTCATATAATAAATCAGCTTTTATAATAAAGAATGACAATAA
 ATGCTATTCTACTGGTTTTAATAATCAAGGTC AACTAGGTTTAGGAGATAAGAATAATAGAG
 ATTTATTAGTTAGTTTTCTATTAATGATGTTAAGACTATAGCTTGTGGTTCTGAACACACT
 GTGTTAATGACGTATAAATAATGATATATATGTTTGTGGAAAGGAAAAATGTTTTGGAAATGC
 ACTTCAATCATCACTAATTTACTAAGATAGAAGAAGTAAATATAAAAACTATTGCATGTGGTC
 ATGGTAACACTATGCTTATAGATAACAAAGG TACTTTAAAGGTTGCTGGAAATAATGATATA
 TATCAGTTAGGTATAGCAAAT TACTCTGAGAATATAGATAATTCATTTATAGATTTAAAAAA
 TATTGTAGCTAAGAATATTTTCATTGGTTTATC CATATAGCATACTAATTGATTCAAATAATG
 ATTCATATTGTACAGGAGATAATACTTATGGACAATTAGGTTTCGTTTTTTGATGATATGCAC
 ATTTGTAGAATTTAAGAAAATGGATAGTGAAAAATATAGTTATAGTAATTATATAAATTTAAT
 TAAATCTGAGGATAAATTA ACTTTATTA AAAAGAAGAAATGGAAATAAAGGATATTGAACTTC
 CACTAGATATACATTCTGTAAGAGATGTCGTTTTTAGTCCTTATTGTACTCTGGTTATTTTA
 GGAATGGAGATGTATATGGTCTAGGAAATAATAGATACAAAGGAATGGGTTCTGACTTACC
 AAGTCAATTAATGAGTTGACAAAATTAAGTATCTCTAATGTAAAGTCTATAGTAGCATCAA
 AAAATATTTCTGGAGGAATATTCTACATTA AAAAATGATGATACTTGTATTATTCTGGACCA
 AATAGTAAC TCAATAGCAGGTGTTCTTCCTTCTAATTCAGATGTAITTAAGAAAATATCTAT
 AGATAATGTAAAAAAGTTGTTATAAATACTGATTTATCAA ACTGGTTTTTCATTAATTGTAA
 CTAATAATAAGCAAATATACACTTCTGGAAAGAGTTCAAGTTATGTTAATGGACTTAGTAAT
 GCATTAATAAGTCAATATACTGAGATTAGCCTTAGTAATGTA ACTGATGCTTATAGTTCATA
 TAATGCAACATTTATTGTAGTTGATGAAAAAAGGTATATGCAACTGGTATAAATACAAATT
 ACCTGTTAGGTTTTAGTACTTCTGATGGATCTAATGTAAATCTAGGTTTTATTAAGTGATTGG
 TATTATATAAATATATCAGGCTCAAGTTATAGTAGAGTTTCATGCACGAATAATATTACTAA
 AATTAATAATATTATCATATATGAGTATGTA ACTGTATTTTGTACAAACATTGGATCTTTTC
 TAACTGGATACCATGGTACTTCATGGACAAAACCAACTGATTCAAGCTATAGAGTTCAATAT
 CAGGGAATTTCATATGCAGGATATCTTGATTCTTATATATATAATTATTATCCTACAAGATG
 TACACAATCATCATCTTCTACA ACTTTTGCTTATTTATATAATGGGGAATCGTCAAGTAATT

TAAAAAATGTCAATCCAGATAAATTTACTTATTTCTGGAGGTTTCATCTTATATACATCAATAT
 GGAAGGAATTATCTTAAACAATCAATCATCTAATAATATTGCAGCATCTAATATAAATTCAGG
 TCCTATTACCTCTGATAAAGCCATATTTTTATATAAAGCTCTATTGTATTTATCTTCTAACA
 CGCTATATGGTTTTGGGAATATATCTGAAAGTGCAAAAGAACTAGATGTTTCAGATACACAA
 GATGGATATAATGCCACTAATTATAAAAAGGTAATGAAAAATATAAAAAATATATTTTATACC
 TCCTTATGATTTAAGTAGAGATAAAACTAGATTTGCAATATTAAGTATAAGAGCTTATTTA
 TATGTGGATATAACTCTAAGGGTACGCATGGTATATCAGTTAATAGTAGTTTAAATTTAAAT
 AATAAGATAAATTACAATAAAAAGAATAGCAGTAGTGAAATAICTTCTAATATACAAGAAAT
 ATATAGCCATTTCAAAGTCTACATATTTATTAACCTAATAATAATATGCTTTACAGTGTGGTT
 TAAATGATGTAGGTCAATTAGGAGTGGGAGATGAGATAAATAGAAAGGTATTTACTAAAATA
 AATATTGATAATATAAATCTATAAATGTAAATAGATTTACTGACAATAGTAAACATGCATT
 TGGGATAAAAAATGATAATACCTGTTATGCTGTTGGTTTAAATAAATCTGGTCAGTTAGGAA
 TAGGAGATAAATGTAAGTAAATATATTTACTAAAATAAATGTTGAAAATGTAATAATGTA
 GCTGTATATGGAAACACATCTCTATTATTAACCTAATGATGGTCTTTTATATGGAGCAGGTAA
 TAATGGAAAAGGACAGTTAGGATTGGGTGATACTACAAGTAGGAATATATTTACACGTATAC
 CTATAAATGGTGTAGAGATGTATATCTATGTAATGATGTATCAATCATTGTTAAAAATGAT
 AATACATGCTATGTATGTGGACTTGTAATGGCTATTTTGGGTTTACTGAAGGAAGTATAAG
 TACATTTACAAAAATAAATATTGAGAATGTAAATCTGTTGTGACAGCAGGAAGTGAAGCTA
 CATTTTTATAACAAATGACAATATGATTTATACTACAGGGAAAAAAGAGAGGGTATTCTTT
 TCAACAGAGACTAATGATATAAAGGGGATACGAGTAATTAATAATATTATAAATGCAAAAA
 AATAGTAGTTAATGGATATACTTCAGCCATTTTAAACAAATGACAATAAACTATTTGTTGGAG
 GTCCTTAGTGGATATGGAAGTATAGCAAAATAATAATAACAAATAGTGTGGAAGATGTTAAA
 GATGTTTTTGTAAACAGCTAATAATACACTTTATATAGATAATAATAACAATTTGATATCATC
 AGGTAGAGATACTTATGGTATATCTGATGAATCTTATAGGGATATGTCAGTTCCATATTATA
 AAGTATCTATAAAGAAAGATGTTGATACTGTATTTCTAGTTACAATACTATATTTATTA
 GATATATATGGAAAATTTTATTCTTCAACAAGAGATAATAGATATAATCATTTAGGTATTCA
 CCATAGATATGATAAATGATAAAAAATGAAGCTCTTGAAGGTTCCCTACATTCATATTTAAAA
 CAGATAACACATCAGATAAAAAATAGTTTTTAATAAGAAAAATGAAAAGCTAGTAATGTTAAT
 GATAAGTATATAAAAAACAAATAAAGTATATAAATTTATAAAAAACATATTTAAAGATAATTT
 TAAGTATACTTCAATAATATTGCCATTTGAGGTATCTGATATTGATATATCAAAAACACATT
 CATTTGGCTGTTGCTAAGGATGGCAAGTTATATGGAATAGGAAGTAATTCATATAAAGAAAT
 AATCAAACCTTGAAGATATAGAATTATTAACCTTACTGAAAGTAAATATATCAGATGTCAA
 AAAAGTTGCTTGTGAGATAACTACTCCTATATTATTAAGACAGATAAATACTCTATGGTCAT
 ATGGAAAGAATACTGAGTACCAATTGGGAGTTGGCCACAATAATGATGTAAGAGAGTTACAA
 AAGGTTACTGGATTACCTTCTGTTAAAGATATAAGTATATATAACTCAATGACACTTGTTTT
 AACTAATGAGGGAGAGTTGTACGCTCAAGGGTACAATACAAATGGATTATTTGGACTAGGAG
 AAAGTGAAAAAGATAAGATAATAAGAACTTTTACTAAAGTATTAACCTAATGTTAAAGAAAT
 AAGTCACATAATGATGACCACATACTAGTAATTAATAAATGATAATAGTCTATGGATAACTGG
 TAAAAATAAATCTATGTATAAAAAATCTATATCAATTACTGATTTATATGAATTTACTAAAA
 TACCAATTCCTGAACATCTAAATGATATTTTAGATATAGAGCTTTCAGATGATACAATATAC
 ATGATAACAAAAGTAGATACAAGTAAAGCATCTATAGAAAATAGTTGAAAAATCAATATCTCA
 AGTGAGAGTTGTAGTACAAGACCCATAAATGTTATAGAAAACTTGAATGTTTATAAATG
 ATGAATTAATATCTACTAAGACTAATTTGGAAATAAATAGCATTATATTTGAGATACCACAA
 AATAAATAAGTATTAGGAGAAAAATAAGATACTGATTAAGCCAGTAGTCCTACAGGCGATTT
 ATATTCAGTATGTTTATATTTAAATCAGAAACAGGGCTTAAAGTAAAAAAGGATTCTATTT
 TAATGATAAACAATAAAGTATATTCAATCATAAACATTACTGAAAATAACACTGACTTAATA
 GTAACATTAATGAGGGATTAAAGGATGATATGATGGAAAACAATCCTATATATCAATTAAT
 AAATAAACTAAAGTTCAAGTAAAAATAAATAAATCTGACTTATTCAAAGACATGAAACTAG
 TTGAAATCAAAAAATCAGACTCAAGTTACCAAGAAATCTATGAATTAGAAGAAGCCAACATA
 AAAAGTGCTCAGCCTAAAAATCATAGTAGAAAAAGGAGATAAATGGACAGCTATAAAACGTCC
 ATCTATGATTTTTAGATATGATGCTGAAAACAACGAGCCACAAGCTTAAATGGAGGTGTGA

ES 2 622 280 T3

AAATTGTTTAAATTCGATAAAAAATAAAATAGAACAAATCAAACAAGGTAGAAAAGTAGAAAT
GCAGTATAAAGACATTTTCAGACATAAGTATAGGTCAAGTAAAGCAAGATGATGATATAACAA
ATAATTTTATAGCAAATGTAGAAATATATGAGATGTTGTTAAATCAAAGTTCTGTCAATGAA
GCAAGTAATATAAGCACTTTTAGTGTAAGAAAATCTGGAGGTGAGAGTGGAAATGGTAGAAGT
ATATGTAGCTTTAATTTTAAAGAGGCAAAAAACAATAGAAGAAGTACCAGCAGTAATTAGAG
AGCAAGTTAGAATTAGATGTAAGAATTAGAAATACCAGTTGAATAGTAAATTTAGAATAAC
TATGTATTAGTTATTTTTTTTTATGTAAAGTACAAGGTCTTAACTTTAATAAGTAAGCCTTGT
ACTTATTTTTTTGTTACATTAGAACTTGTATATATATTTATTATTATTCAATCTATAAATTA
CACCTACAATTTAAAGTACAGAAGATTAATTTGATAATCCTGAAAATATAATATTGCATGAT
GTAAGAATACAACAAAAATTAAGCTATAAGTATAAAAAATTTAGACAATAGGAGGCTATAA
TGGATAAATTAATAACCGAATTGAGTAGTCTAGGGGCAATAGGTATACTATGTGCTCTATTA
TTTTAAAAATACTATGCAGGAGAAAAAGAAGATAGAGACATGTATAAAAAAACTGTAGAAAA
TTTTATAGAATTATCTACACAACAAGAATAAACAAAAATATACTTGTTCAAATGGGTA
TAATGAAAACAGATGTAGAGGAAATTAAGGAAGATGTTACTGATATAAAAGGTATGTTACAA
AATGGTGTATAACATGAAAGAGTAGCACCAGATTATATATTGTTAGGAAAAGATAAAGTAGT
ATTGTAGATAGTTCACTATTTTATTGAGAAGGATTTAATATTTAAAAATTAATTAAAAAAA
GTAATAAAAAATAACATATAAAAAATTAAAAAAGGAGTTAAGCTTAAATTTGAGGCGCG

SEQ ID NO:62 >CD4-1359
MNNLDKLFELASQEEIIHYTTYIAGDLEGLYINKHGIKIISLLSNLKQNSKKLTSILAEEL
GHHFTSLGYVSSYNDYYTKI IIDKCNKALKWACEFLITEEDIINI INSGITCVYEMADIL
NVDITFFQKRLEFLSLKKQSLQLGNKYLILTNLPHYFYIFDPI S

SEQ ID NO:63 >CD4-1360
MFAKRLRELKREKFLTQRELGEKVGVSQRVLGYYETENRFPDEHILNKLADVFNVSVDYLLG
RTLKVENIDTVAHRKPNPHEELPEEAQEQLNDYIEFLLNKYKKK

SEQ ID NO:64 >CD4-1360A
MFKNNLKYRCKGKMTQIQ LARKAGITNDYISQIERGIKNPGLLMAKKISSILEQNIIEVFF
IQL

SEQ ID NO:65 >CD4-1361
MENKDKILFKETDKRLHNYKYLDIKIKNINLIDIKRCENEYSGCGAMVYTEKTSNTYNISSSV
ENEVLKREERLRKLMKEDIEIEKEKIENALTCLNDIEMEFFNLFYNSKTKNMITYISMKL
HLDRTSCYNLKKKMI FKLSEIL

SEQ ID NO:66 >CD4-1362
LLKYKEILETIIIEILKKNFTESIFIDDES VQGSEGS CFFVSILSVICTPVMLNTNNKDIVIS
IKYLPKPQSKSIRMYEISDELNKLFNRNKIKVTDKRLNITKLEQSIKKEESIYVLNFTFTLNY
LDSVYEEDVVYENMKEINLNLGE

SEQ ID NO:67 >CD4-1363
MAIGLPSINISFKELATTVKERSARGIIAMVLKDAKALGLNEIHEKEDI PVDLSAENKEYIN
LALMGNVNTPNKLLVYIEGEADIQTALDFLETKEFNLCMPKAVEADKTAIKNWIIKLRDI
DKVKVAVLGKVVGNHEGI INFTTEDVLVGEKKYSVDEFTSRVAGLIAGTPLSQSVTYTKLS
DVVDIPKMTKVDAESRVNKGELILIKEAGAIRIARGVNSLTTELTAEKGEMFQKIKIVDTLDI
IHSDIRKVI IDDYIGKVTNSYDNKCLLIVAIKSYLEELEKSALIESDSTVEIDFEAQKSYLK
SKGVLDLSYMTLQEIKEANTGSKVFLKAKIKVLDAMEDIDLSEI

ES 2 622 280 T3

SEQ ID NO:68 >CD4-1364
MANMEARNVMSGTWGELWLDGNKVAEVKKFQAKMEFTKEDII IAGQMGTDTKYMGYKKGKSI
TLYHVSSRMHKLIGEKIKRGSEPRFVAISKLNPDPSYGAERIAVKNIAFDDLTLDLADWEVGVK
GEIEAPFTFTEYDFLDII

SEQ ID NO:69 >CD4-1365
MSENGLSKNINIVDLLLNSDTENLERPSTIVELKRLSTIFGQEFKVMCRALTISKDEEIQNT
CLKIDENMKTDIDLPEMQMLTIEGVCDLDGKLLFKNKELMDKFKAPTPELARKLLLPGEI
TNLYRILQDVMGYGKNAVIEEVKN

SEQ ID NO:70 >CD4-1365A
MYYYWKKKGIRPSLFYAMDKGELKLIEAFFALEIEEEVEKMKKHGYGVCPLTGGGM

SEQ ID NO:71 >CD4-1366
MGNVREEGINMYLTDNYTPKMNQIISVTDNFRRATVAVSLSTNVMASSIKNSIGSASNRVNS
LNSSLRKVQTTASSVSSTMTKLSSSINAVSGVIGSLNGSIMRLAITIAMIIDYFNKLIQKKN
EFNSNIMIILIFKAKSDEVEKTKNKLGNLKKIGGKIWNIVIKAKDMTKRVISSILGKLKQV
EKRPYQGSINLKMVSSAMGRILPKLMLFKNTFWSGVIAIKDMASGII SKVFPKLRLFAGKV
WSGAIAVKDMASGILGSIKGIKISDLTNGATIGVAVKKGVDLLGQEQNQKVVLESVMKRNTGK
VNQIDVDDYYGSLVRMANDT PFDPEDVVAMGTAKMISNITGGKKEKDITQAMVDVRLNMN
TSSEQDVSA AFLSAAKGNMESLNTLVGENYKTFDEALEGISVKQMGLAKEMSNTIPGII SGA
QTSINGLKSIKVPFDDILGQGLKKIKTFIESGLGNLAGLSEKMAGKIGNVMNGKIIIGNKY
DQMQRSVKNGKEFSDSTQYRISNEAEKRKMMVENKQERFENHAATMIGNAPKAIVNAGSTL
LQNI DFTALIDSLLPVNVNLVNNLLDSINNKSPIAQGLISIFGTIVTTAFQLIGPVVEAVSPI
ITRI FTFLGEYAPQINNFIETLGVWKTWVWETLGPLLETGWKIEPILGAFFNILDVKVCKIV
KDICKWWQTMINKIKNGSITGTVLNLVEKSKKNYKDNPYAGTKAGDSGKAYS GKKGNNAFGL
NYVPYNDYQTRLHEGEMVLTQEQEANQYRSRKNNGNINI AKLADTIVIREEADIEKITSKLVA
SIQLAQLGGVL

SEQ ID NO:72 >CD4-1367
MEMWLRQAEDRFRFPVFPSSFSINGKAAVNSSSILKIGEIATFGGVALKSISISISSFFPNKDY
TFCDYTGFPSPYDCVNKIEKWMKEGFILRFTITETNINMEVIEGFSYEERDGTDRDVYFTLD
LKEYKRKIKPKVTPKQ

SEQ ID NO:73 >CD4-1368
MIINRSKSSSNEISFVSKDMGFLLTQSEVSYNFKDKLVEDIAKQVFAENRLSVGIIAKTNV
KYTKMFIGVNGYDTIMSAYTEASKKTKKKYMI EANLDKFNVIEKGTVTL SVMFEEGFNIINT
TFSESMENVKNKVIIVVDQYGSKISEKIDNEIFKEVNVIMQKVIQQQENQDVIDIDSEFNIEK
SCSLKGYGDVSCITGRGVKVKDSYTKLVGLFYIDTDKHTWQNGEYQIELELNFQNLMDKESA
GQDEPKESNLGGEDYAGGKEFTA EFTAYCPRKEEGGDTDCRKKKLDPSKKTCAAPMVGKYE
QTYYTKEFLNKHPLLNYGDEIQVITGVSGRDGVYKVNVDVGPAITIEKNGTYHIDILFGNVEE
ASKFGRRKGKIIIGGYSGNVSDKAKIVISEAKKHLGKPYKWGGNGPSSFDCSGLMVYCFKKV
NVSLPRTSNQQSKKGGKVEQKNLQAGDLVFFHNPVSHVGLYIGNGEFLHAPQKGDVVKISKL
SSRRDFNTARRVL

SEQ ID NO:74 >CD4-1369
MANPINEFIGIIREEGKYHNQPSFFIGKIKSKLPDLKIETNNIILEKEDILIDSWMIDRQLE
TFDTETNQEHQHEVKNPFIDNFESGDMVIMFRIGEKFAVVSCLVSL

SEQ ID NO:9 > orf 1366 (98% 630) Tape Measure
 MGNVREEGINMYLTDNYTPKMNQIISVTDNFRRATVAVSLSTNVMASSIKNSIGSASSRVNS
 LNSSLRKVQTTASSVSSTMAKLSSSINAVSGVIGSLNGSIMRLAITIAMIIDYFNKLIQKKN
 EFNSNIMIILIFKAKSDEVEKTKNKLGLNKKIGGKIWNIVIKAKDMTKRVISSILGKLKRV
 EKRPYQGSINLKDMSAMARILPKMLLFKNTFWSGVIAIKDMASSIISKVFPKLRLFAGKV
 WSGAIAVKDMASGILGSIKGISDLTNGATIGVAVKKGVDLLGQEQNQKVVLESVMKRNTGK
 TSQKDVVKYDSDLVNMANDTFFDPEDVAVMGTKAKMISNITGGKKEKDITQAMVDVRLNMN
 TSSEQDVSAAFLSAAKGNMESLNTLVGENYKTFDEALEGISVKQMGLAKEMSNTIPGIISGA
 QTSINNGLSIVKPFDDILGQGLKIKTFIESGLGNLAGLSEKMAGKIGNVMNGKIIIGNKY
 DQMQRSRVKNKGFSDSTQYRISNEAEKRKMMVENKQERFENHAATMIGNAPKAIVNAGSTL
 LQNDIFDALIDSLPVVNLVNNLLDSINNKSPIAQLISIFGTIVTTAFQLIGPVVEAVSPI
 ITRIFTFLEGEYAPQINNFIETLGVINWKTWETLGPLLETGWKIIIEPILGAFFNILDKVKVIV
 KDICKWWTMINKIKNGSITGTVLNLVEKSKKNYKDNP
 YAGTKAGDSGKAYSSKKGNAFGLNYVPYNDYQTRLHEGEMVLTQKEANQYRSRKNGGNINI
 AKLADTIVIREEADIEKITSKLVASIQLAQLGGVL

SEQ ID NO:10 > orf 1367 (100% 630) Placa basal
 MEMWLRQAEFRFPVFPSSFSINGKAAVNSSSILKIGEVATFGGVALKSISISISSFFPNKDY
 TFCDYTGFPSPYDCVNKIEKWMKEGFILRFTITETNINMEVIEGFSYEERDGTDRDVYFTLD
 LKEYKRKIKIPKVTPKQ

SEQ ID NO:11 > orf 1368 (100% 630) Placa basal
 MIINRSKSSSNEISFVSKDMGFLLTQSEVSYNFKDKLVEDIAKQVFAENRLSVGTIAKTNV
 KYTKMFIGVNGYDTIMSAYTEASKKTKKKYMIKANLDFNVIEKGTVTLVSMFEEGFNIINT
 TFSESMENVKNKVIVVDQYGSKISEKIDNEIFKEVNVIMQKVIQQENQDVIDDSEFNGIEK
 SCSLKGYGVDVSCITGRGVKVKDSYTKLVGLFYIDTDKHTWQNGEYQIELELNFQNLMDKSA
 GQDEPKESNLGGEDYAGGKEFTAETAYCPRKEEGDTCRKKKLDPSKKTCAAPMVGKYE
 QTYTKEFLNKHPLLYGDEIQVITGVSGRDGVYKVNVDVGPAITIEKNGTYHIDILFGNVEE
 ASKFGRRKGIIGGYSGNVSDKAKIVISEAKKHLGKPYKGGNGPSSFDCSGLMVYCFKKV
 NVSLPRTSNQQSKKGGKVEQKNLQAGDLVFFHNPVSHVGLYIGNGEFLHAPQKGDVVKISKL
 SSRDFNTARRVL

SEQ ID NO:12 > orf 1369 (100% 630)
 MANPINEFIGIIREEGKYHNQPSFFIGIKSKLPDLKIETNNIILEKEDILIDSWMIDRQLE
 TFDTETNQEHEVKNPFIDNFESGDMVIMFRIGEKFAVVSKLVSL

SEQ ID NO:13 > orf 1370 (99+% 630)
 MSTIFPFIVPEDIYLPKTEELPIFREVAVDFEKDEPILEKGFKIIEKKEALKVWIYKCIK
 TNRYEHEIYSLEYGTELSELIGQKYTKGLTESEASRFIKEALLINPYILEVNVKSANFNDRV
 LSANVKVSTIYGEVEINV

SEQ ID NO:14 > orf 1371 (100% 630) Cola
 MYSDQTYEVIKNRLENINLDIYKGEFSFLNNMVSGNNLELSKIYLELSKIHKMAFIQDTYN
 QFLDKRVNEFGVYRKLGTESNGEVEFIGEKGTVINNGTIIISYRDLLFVVIKDVITIGSEEDN
 SPVQALEVGGKYNLPTNCEFKLVDNISGVTKITNTRSFEGGDIETDEELKERFYKIQRNQA
 TSGNKAHYEEWALEVDGVYVNVKVPYPRWDGPGTVKVLIFGENNQAVDTETIERCQQHIDEKPK
 IGPTITVVTPLPIEISISAVMKLEDGYTLDNVKESFLESINTYFRDIRGEIITYTKVMGILIN
 TTGVHDLNLLINGSTDNITINEDKIPSVTTVNFSEVENQ

Figura 2.

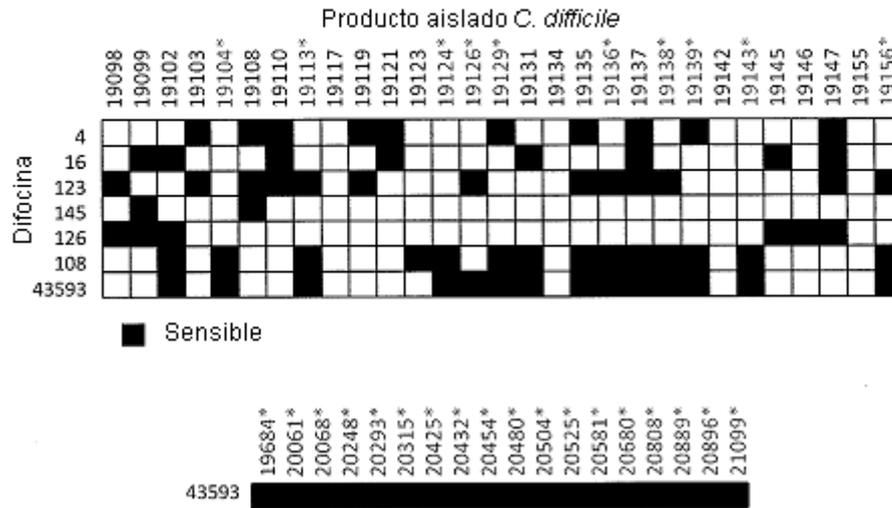
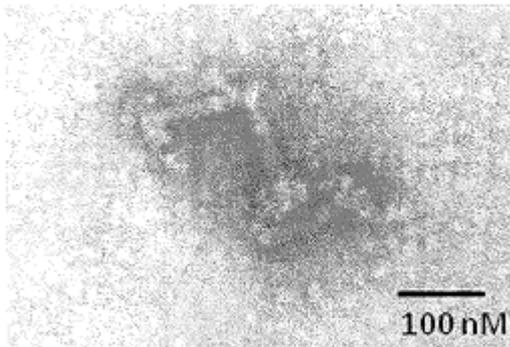
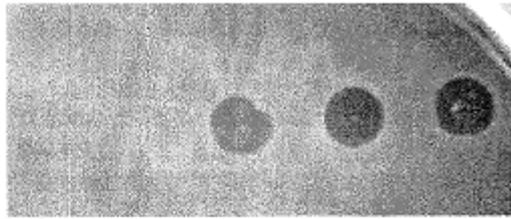


Figura 3.



Difocina 16

Figura 4.



← Diluciones 1/10

Figura 5.

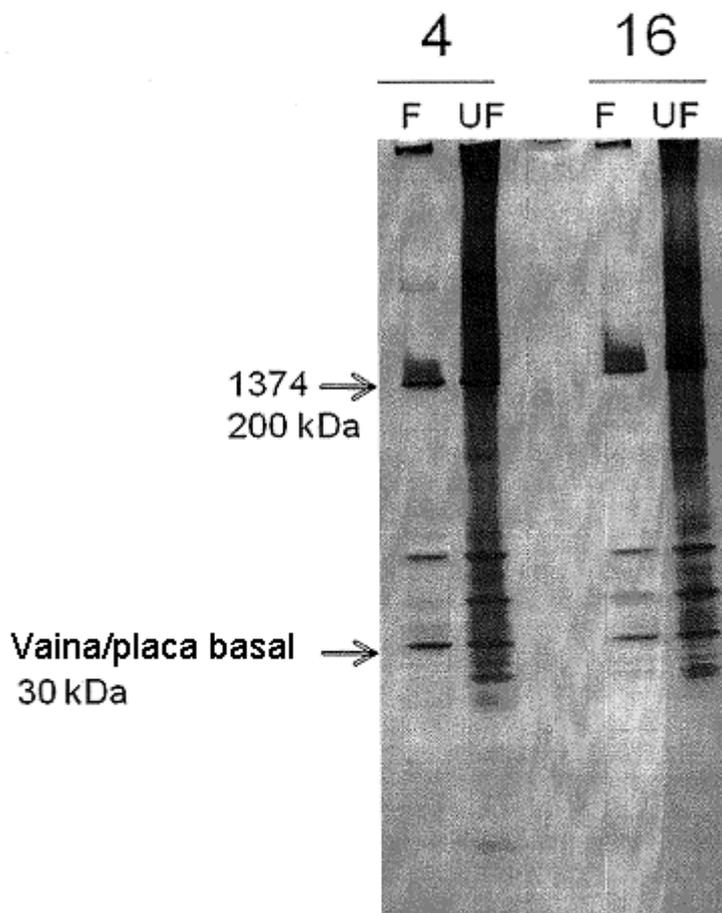


Figura 6.

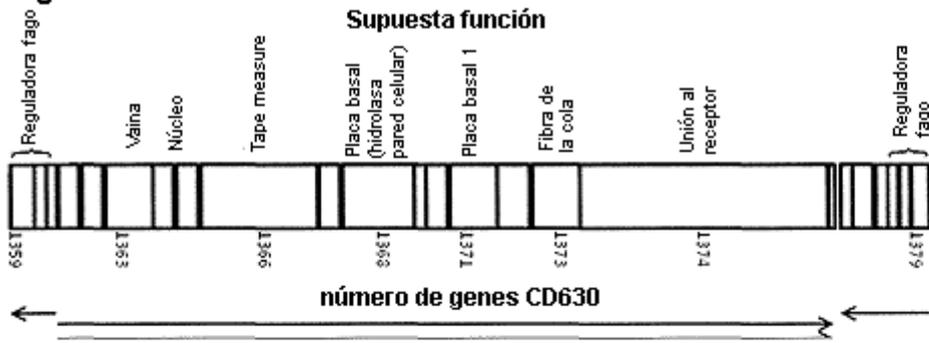
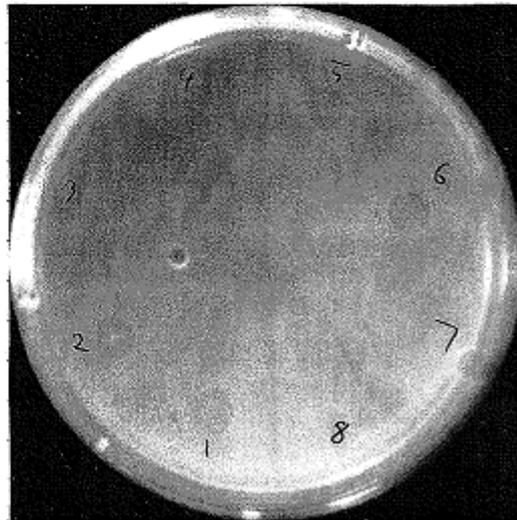


Figura 7.



ES 2 622 280 T3

CD108 NGNST---FVLKNNGLTLYSCGYNSSGILGLKDNTNRNIFTKIEIENIKFCVESN----Y 522
43593 NGNST---FVLKNNGLTLYSCGYNSSGILGLKDNTNRNIFTKIEIENVKAFCVESN----Y 522
CD16 NGNST---FVLKNNGLTLYSCGLNSNGQLGLRDEVNRNIFTKIEIENVKDFCVGSN----Y 522
CD126 NGNST---FVLKNNGLTLYSCGLNSNGQLGLRDEVNRNIFTKIEIENVKDFCVGSN----Y 522
CD4 SKNISGGIFYIKNDDTCYYSGPNSNSIAGVLP-SNSDVFKKISIDNVKKVVINTDLSNWF 534
CD123 SKNISGGIFYIKNDDTCYYSGPNSNSIAGVLP-SNSDVFKKISIDNVKKVVINTDLSNWF 534
. * : * :*:.* * . * *.. * : * :*.**.*:* * . : : :

CD108 IVALNHSKELYGWNQS--YIVYGDNRNYPYKDRVSNVEKIATWSDTLYILDSTGATKT 580
43593 IVVLNHSKELYGWNES--YIVYGNRNYPYKDRVSNVEKIATWSDTLYILDSTGATKT 580
CD16 VIALNHSKEVYGWGNP--YNNIEKTSNYPYKQG-ISNIEKIAAYDYSVYMINSEGKLYV 579
CD126 VIALNHSKEVYGWGNP--YNNIEKTSNYPYKQG-ISNIEKIAAYDYSVYMINSEGKLYV 579
CD4 SLIVTNNKQIYTSKGSSSYVNLGNLISQYTEISLSNVTDAYSSYNATFIVVDEKKVYA 594
CD123 SLIVTNNKQIYTSKGSSSYVNLGNLISQYTEISLSNVTDAYSSYNATFIVVDEKKVYA 594
: :.:.**:* *.. . *.: **: . : : : : .

CD108 IGYSYNGSGGYPAPSSSSTYR---EGGYINKNTSYRTLEFYNTSKTKLVNLFAYNGCV 636
43593 IGYSYNGSGGYPAPSSSSTYR---DGGYINKNTSYRTLEFYNTSKTKLVNLFAYNGCV 636
CD16 SGYNYNQLGKGNNSNQSKAL---VSQCRNNTSSTSNGLR--TLPKITNVFPFYDGCA 633
CD126 SGYNYNQLGKGNNSNQSKAL---VSQCRNNTSSTSNGLR--TLPKITNVFPFYDGCA 633
CD4 TGINTNYLLGFSTSDGSNVNLGLLSDWYYINISGSSYSRVSCTNNITKINNIIIEYEVTV 654
CD123 TGINTNYLLGFSTSDGSNVNLGLLSDWYYINISGSSYSRVSCTNNITKINNIIIEYEVTV 654
* . * * * : . . * : * : : .

CD108 FVDENGLAYCIGENNINFRGGSTTNNENSLRFINNSGVYYT-----NTDGTDYTCYQ 688
43593 FVDENGLAYCIGENNINFRGNSTTNNENSLRFINNSGVYYT-----NTDGTDYTCYQ 688
CD16 IDEGGYVYLTG-----YHGYLRT--LNSSPSISDYSRYGT-----FIEATNSNHNT 678
CD126 IDEGGYVYLTG-----YHGYLRT--LNSSPSISDYSRYGT-----FIEATNSNHNT 678
CD4 FCTNIG-SFLTG-----YHGTSWTKPTDSSYRVQYQGYSYAGYLDYSIYNYPTRCTQSS 708
CD123 FCTNIG-SFLTG-----YHGTSWTKPTDSSYRVQYQGYSYAGYLDYSIYNYPTRCTQSS 708
: : * : * :*: * * :* :. . : * .

CD108 WTYKLIRCSIFDSPQNIIGNSKNILYLSKNNSTFKCTGNCITYGINSQNWYSYFS----- 743
43593 WTYKLIRCSIFDSPQNIIGNSKNILYLSKNNSTFKCTGNCITYGINSQNWYSYFS----- 743
CD16 Y---FIQETDFSGIEKVIKMSNNILFFKKGSS--YITGYPKTFGSTITGHRSYTS----- 728
CD126 Y---FIQETDFSGIEKVIKMSNNILFFKKGSS--YITGYPKTFGSTITGHRSYTS----- 728
CD4 SSTTFAYLYNGESSNLKNVNPDLNLLISGGSS--YIHQYGRNYLNNQSSNIAASNINSG 766
CD123 SSTTFAYLYNGESSNLKNVNPDLNLLISGGSS--YIHQYGRNYLNNQSSNIAASNINSG 766
: . . : : . : * :. . . * :. . . *

CD108 -DSSNGAIALGNEFILKNYSGECLLKGYG---KATNGEFGNSTNISSISNYDTGLKDIAK 798
43593 -DSSNGAIALGNEFILKNYSGECLLKGYG---KATNGEFGNSTNISSISNYDTGLKDIAK 798
CD16 -INSE-SSNLGSNFIIYHSN--SKLYGKG---IANSQFGNSTNIDGTSNYDTGLKDIAK 780
CD126 -INSE-SSNLGSNFIIYHSN--SKLYGKG---IANSQFGNSTNIDGTSNYDTGLKDIAK 780
CD4 PITSDKAI FLYKALLYLSSN---TLYGFGNISESAKELDVSDTQDGYNATNYKKVMKNIAK 823
CD123 PITSDKAI FLYKALLYLSSN---TLYGFGNISESAKELDVSDTQDGYNATNYKKVMKNIAK 823
.*: : * . : : . * * * * . * . : : : : . : **.. :*:**

Figura 9.

