



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 278 436**

51 Int. Cl.:

C12N 15/31 (2006.01)

C07K 14/315 (2006.01)

A61K 39/09 (2006.01)

C12N 1/21 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **99904646 .9**

86 Fecha de presentación : **17.02.1999**

87 Número de publicación de la solicitud: **1054971**

87 Fecha de publicación de la solicitud: **29.11.2000**

54 Título: **Antígenos de estreptococos del grupo B.**

30 Prioridad: **20.02.1998 US 75425 P**

45 Fecha de publicación de la mención BOPI:
01.08.2007

45 Fecha de la publicación del folleto de la patente:
01.08.2007

73 Titular/es: **ID Biomedical Corporation**
525 Cartier Boulevard West
Laval, Quebec H7V 3S8, CA

72 Inventor/es: **Brodeur, Bernard, R.;**
Rioux, Clément;
Boyer, Martine;
Charlebois, Isabelle;
Hamel, Josée y
Martin, Denis

74 Agente: **Elzaburu Márquez, Alberto**

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

Antígenos de estreptococos del grupo B.

5 Campo de la invención

La presente invención se refiere a antígenos, más particularmente a antígenos de patógenos bacterianos de estreptococos del grupo B (GBS), que son útiles como componentes de vacunas para la terapia y/o la profilaxis.

10 Antecedentes de la invención

Los estreptococos son bacterias gram (+) que se diferencian por antígenos de carbohidratos específicos de la A a la O en su superficie celular. Los grupos de estreptococos se distinguen además por antígenos de polisacáridos capsulares de tipo específico. Se han identificado varios serotipos para los estreptococos de Grupo B (GBS): Ia, Ib, II, III, IV, V, VI, VII y VIII. Los GBS también contienen proteínas antigénicas conocidas como "proteínas-C" (alfa, beta, gamma y delta), algunas de las cuales han sido clonadas. Aunque el GBS es un componente común de la flora colónica y vaginal humana, dicho patógeno ha sido reconocido durante mucho tiempo como una causa principal de sepsis neonatal y de meningitis, de meningitis de activación tardía en niños, de endometritis post-parto, así como de mastitis en el ganado lechero. Las madres parturientas expuestas a GBS corren el riesgo de infección post-parto y pueden transferir la infección a sus bebés cuando atraviesan el canal de nacimiento. Aunque el organismo es sensible a antibióticos, la elevada velocidad de ataque y el rápido inicio de la sepsis en neonatos, y de la meningitis en niños, da lugar a una elevada tasa de mortalidad.

Para encontrar una vacuna que proteja a individuos frente a la infección por GBS, los investigadores se han fijado en los antígenos de tipo específico. Desafortunadamente estos polisacáridos han demostrado ser pobremente inmunogénicos en humanos y están restringidos al serotipo concreto del cual procede el polisacárido. Además, el polisacárido capsular provoca una respuesta independiente de células T, es decir, sin producción de IgG. Por consiguiente los antígenos de polisacáridos capsulares no son adecuados como componentes de vacunas para la protección frente a la infección por GBS.

Otros investigadores se han centrado en el antígeno beta de proteína-C, que ha demostrado presentar propiedades inmunogénicas en modelos de ratón y de conejo. Se descubrió que esta proteína era inadecuada como vacuna humana debido a la no deseable propiedad de interaccionar con una alta afinidad y de un modo no inmunogénico con la región Fc del IgA humano. El antígeno alfa de proteína-C es poco común en los serotipos de tipo III de GBS, que es el serotipo responsable de la mayoría de afecciones originadas por GBS y que, por tanto, es de poca utilidad como componente de una vacuna.

Por tanto, sigue existiendo una necesidad no satisfecha de antígenos de GBS que puedan ser usados como componentes de vacuna para la profilaxis y/o la terapia de la infección por GBS.

40 Resumen de la invención

De acuerdo con un aspecto, la presente invención proporciona un polinucleótido aislado que codifica un polipéptido que tiene al menos una identidad del 70% con un segundo polipéptido que comprende una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 39 y SEQ ID NO: 44, o fragmentos de las mismas.

En otros aspectos, se proporcionan vectores que comprenden polinucleótidos de la invención ligados operativamente a una región de control de la expresión, así como células hospedantes transfectadas con dichos vectores, y métodos para producción polipéptidos que comprenden cultivar dichas células hospedantes en las condiciones adecuadas para la expresión.

En otro aspecto adicional, se proporcionan nuevos polipéptidos codificados por polinucleótidos de la invención.

Breve descripción de las figuras

La Figura 1a (únicamente a modo de referencia) es la secuencia de ADN del clon 1 (SEQ ID NO: 1) con las correspondientes secuencias de aminoácidos de los marcos de lectura abiertos;

la Figura 1b es la secuencia de aminoácidos SEQ ID NO: 2;

la Figura 1c es la secuencia de aminoácidos SEQ ID NO: 3;

la Figura 1d es la secuencia de aminoácidos SEQ ID NO: 4;

la Figura 1e es la secuencia de aminoácidos SEQ ID NO: 5;

la Figura 1f es la secuencia de aminoácidos SEQ ID NO: 6.

ES 2 278 436 T3

La Figura 2a (únicamente a modo de referencia) es la secuencia de ADN del clon 2 (SEQ ID NO: 7) con las correspondientes secuencias de aminoácidos de los marcos de lectura abiertos;

la Figura 2b es la secuencia de aminoácidos SEQ ID NO: 8;

la Figura 2c es la secuencia de aminoácidos SEQ ID NO: 9;

la Figura 2d es la secuencia de aminoácidos SEQ ID NO: 10;

la Figura 2e es la secuencia de aminoácidos SEQ ID NO: 11;

la Figura 2f es la secuencia de aminoácidos SEQ ID NO: 12.

La Figura 3a (únicamente a modo de referencia) es la secuencia de ADN del clon 3 (SEQ ID NO: 13) con las correspondientes secuencias de aminoácidos de los marcos de lectura abiertos;

la Figura 3b es la secuencia de aminoácidos SEQ ID NO: 14;

la Figura 3c es la secuencia de aminoácidos SEQ ID NO: 15;

la Figura 3d es la secuencia de aminoácidos SEQ ID NO: 16;

la Figura 3e es la secuencia de aminoácidos SEQ ID NO: 17;

la Figura 3f es la secuencia de aminoácidos SEQ ID NO: 18;

la Figura 3g es la secuencia de aminoácidos SEQ ID NO: 19;

la Figura 3h es la secuencia de aminoácidos SEQ ID NO: 20;

la Figura 3i es la secuencia de aminoácidos SEQ ID NO: 21.

La Figura 4a (únicamente a modo de referencia) es la secuencia de ADN del clon 4 (SEQ ID NO: 22) con las correspondientes secuencias de aminoácidos de los marcos de lectura abiertos;

la Figura 4b es la secuencia de aminoácidos SEQ ID NO: 23;

la Figura 4c es la secuencia de aminoácidos SEQ ID NO: 24;

la Figura 4d es la secuencia de aminoácidos SEQ ID NO: 25;

la Figura 4e es la secuencia de aminoácidos SEQ ID NO: 26.

La Figura 5a (únicamente a modo de referencia) es la secuencia de ADN del clon 5 (SEQ ID NO: 27) con las correspondientes secuencias de aminoácidos de los marcos de lectura abiertos;

la Figura 5b es la secuencia de aminoácidos SEQ ID NO: 28;

la Figura 5c es la secuencia de aminoácidos SEQ ID NO: 29;

la Figura 5d es la secuencia de aminoácidos SEQ ID NO: 30;

la Figura 5e es la secuencia de aminoácidos SEQ ID NO: 31.

La Figura 6a (únicamente a modo de referencia) es la secuencia de ADN del clon 6 (SEQ ID NO: 32) con las correspondientes secuencias de aminoácidos de los marcos de lectura abiertos;

la Figura 6b es la secuencia de aminoácidos SEQ ID NO: 33;

la Figura 6c es la secuencia de aminoácidos SEQ ID NO: 34;

la Figura 6d es la secuencia de aminoácidos SEQ ID NO: 35;

la Figura 6e es la secuencia de aminoácidos SEQ ID NO: 36.

La Figura 7a es la secuencia de ADN del clon 7 (SEQ ID NO: 37); (inventiva);

la Figura 7b es la secuencia de aminoácidos SEQ ID NO: 38;

la Figura 7c es la secuencia de aminoácidos SEQ ID NO: 39;

la Figura 7d es la secuencia de aminoácidos SEQ ID NO: 40;

5 la Figura 7e es la secuencia de aminoácidos SEQ ID NO: 41.

La Figura 8 es la secuencia de ADN de una parte del clon 7 que una secuencia señal (SEQ ID NO: 42); (inventiva).

La Figura 9 es la secuencia de ADN de una parte del clon 7 sin secuencia señal (SEQ ID NO: 43); (inventiva).

10

La Figura 9a es la secuencia de aminoácidos (SEQ ID NO: 44); (inventiva).

La Figura 10 representa la distribución de títulos de ELISA anti-GBS en sueros procedentes de ratones CD-1 inmunizados con proteína de GBS recombinante correspondiente a la SEQ ID NO: 39. (inventiva).

15

Descripción detallada de la invención

La presente invención se refiere a nuevos polipéptidos antigénicos de estreptococos del grupo B (GBS), que se caracterizan por la secuencia de aminoácidos seleccionada del grupo que consiste en: SEQ ID NO: 39 y SEQ ID NO: 44.

20

La invención incluye la SEQ ID NO: 39 y la SEQ ID NO: 44.

Una realización preferida adicional de la invención es la SEQ ID NO: 39.

25

Una realización preferida adicional de la invención es la SEQ ID NO: 44.

Tal como se usan en la presente memoria, "fragmentos" de los polipéptidos de la invención incluyen aquellos polipéptidos en los que se sustituye uno o más residuos de aminoácido con un residuo de aminoácido conservado o no conservado (preferiblemente conservado) y que pueden ser naturales o no naturales.

30

El término "fragmentos" de polipéptidos de la presente invención también incluye polipéptidos que se modifican mediante adición, eliminación, sustitución de aminoácidos, siempre que los polipéptidos retengan la capacidad de inducir una respuesta inmune.

35

Por el término "aminoácido conservado" se entiende una sustitución de uno o más aminoácidos por otro en el cual el determinante antigénico (incluyendo su estructura secundaria y la naturaleza hidropática) de un antígeno dado se conserva completa o parcialmente a pesar de la sustitución.

40

Por ejemplo, se puede sustituir uno o más residuos de aminoácido de la secuencia con otro aminoácido de polaridad similar, que actúa como un equivalente funcional, dando como resultado una alteración silenciosa. Los sustitutos de un aminoácido dentro de la secuencia se pueden seleccionar entre otros miembros de la clase a la que pertenece el aminoácido. Por ejemplo, los aminoácidos no polares (hidrófobos) incluyen alanina, leucina, isoleucina, valina, prolina, fenilalanina, triptófano y metionina. Los aminoácidos polares neutros incluyen glicina, serina, treonina, cisteína, tirosina, asparagina y glutamina. Los aminoácidos cargados positivamente (básicos) incluyen arginina, lisina e histidina. Los aminoácidos cargados negativamente (ácidos) incluyen ácido aspártico y ácido glutámico.

45

Preferiblemente, los derivados y análogos de polipéptidos de la invención tendrán aproximadamente un 70% de identidad con respecto a las secuencias ilustradas en las figuras, o a fragmentos de las mismas. Es decir, el 70% de los residuos son los mismos. Más preferiblemente, los polipéptidos tendrán una homología superior al 95%. En otra realización preferida, los derivados y análogos de polipéptidos de la invención tendrán menos de aproximadamente 20 sustituciones, modificaciones o eliminaciones de residuos de aminoácido, y más preferiblemente menos de 10. Las sustituciones preferidas son las que se conocen en la técnica como conservadas, es decir, los residuos sustituidos comparten propiedades físicas o químicas tales como la hidrofobicidad, el tamaño, la carga o grupos funcionales.

55

Además, en las situaciones en las que se descubre que las regiones de aminoácidos son polimórficas, puede ser deseable variar uno o más aminoácidos en particular para imitar de forma más eficaz los diferentes epítomos de las diferentes cepas de GBS.

60

También se incluyen polipéptidos fusionados a otros compuestos que alteran las propiedades biológicas y farmacológicas de los polipéptidos, por ejemplo, polietilenglicol (PEG) para aumentar la vida media; secuencias de aminoácidos líderes o secretoras para facilitar la purificación; prepro- y pro-secuencias; y (poli)sacáridos.

Además, los polipéptidos de la presente invención se pueden modificar mediante acilación terminal de -NH₂ (por ejemplo mediante acetilación, o amidación con ácido tioglicólico, amidación carboxiterminal, por ejemplo con amoníaco o metilamina) para proporcionar estabilidad, y aumentar la hidrofobicidad para la unión a un soporte u a otra molécula.

65

También se contemplan los multímeros de hetero y homo polipéptido de los fragmentos de polipéptido, análogos y derivados. Estas formas poliméricas incluyen, por ejemplo, uno o más polipéptidos que han sido reticulados con agentes reticulantes tales como avidina/biotina, glutaraldehído o dimetil-superimidato. Dichas formas poliméricas también incluyen polipéptidos que contienen dos o más secuencias contiguas invertidas o emparejadas, producidos a partir de mARNs multicistronicos generados mediante tecnología de ADN recombinante.

Preferiblemente, un fragmento, análogo o derivado de un polipéptido de la invención comprenderá al menos una región antigénica, es decir, al menos un epítopo.

Con el fin de alcanzar la formación de polímeros antigénicos (es decir, multímeros sintéticos), se pueden utilizar polipéptidos que tengan grupos bishaloacetilo, nitroarilhaluros, o similares, en donde los reactivos sean específicos para los grupos tiol. Por tanto, el enlace entre dos grupos mercapto de los diferentes péptidos puede ser un enlace sencillo o puede estar compuesto por un grupo de unión de al menos dos átomos de carbono, típicamente de al menos cuatro, y no más de 16, pero normalmente no más de aproximadamente 14.

En una realización concreta, los fragmentos de polipéptido de la invención no contienen un residuo de metionina (Met) de inicio. Preferiblemente, los polipéptidos no incorporarán una secuencia líder o secretora (secuencia señal). La porción señal de un polipéptido de la invención se puede determinar de acuerdo con técnicas biológicas moleculares establecidas. En general, el polipéptido de interés se puede aislar a partir de un cultivo de GBS, y posteriormente puede secuenciarse para determinar el residuo inicial de la proteína madura y, por tanto, la secuencia del polipéptido maduro.

De acuerdo con otro aspecto, se proporcionan composiciones de vacuna que comprenden uno o más polipéptidos de GBS de la invención mezclados con un vehículo, diluyente o adyuvante farmacéuticamente aceptable.

Los adyuvantes adecuados incluyen aceites, por ejemplo adyuvante de Freund completo o incompleto; sales, por ejemplo $\text{AlK}(\text{SO}_4)_2$, $\text{AlNa}(\text{SO}_4)_2$, $\text{AlNH}_4(\text{SO}_4)_2$, $\text{Al}(\text{OH})_3$, AlPO_4 , sílice, caolín; derivado de saponina; polinucleótidos de carbono, por ejemplo poli IC y poli AU y también toxina del cólera detoxificada (CTB), y toxina de *E. coli* térmicamente lábil para la inducción de inmunidad mucosal. Los adyuvantes preferidos incluyen QuilA™, Alhydrogel™ y Adjuphos™. Las vacunas de la invención se pueden administrar parenteralmente mediante inyección, infusión rápida, absorción nasofaríngea, dermoabsorción, o bucal u oral.

Las composiciones de vacuna de la invención se usan para el tratamiento o profilaxis de la infección por estreptococos y/o de enfermedades y síntomas mediados por la infección por estreptococos, en particular de estreptococos del grupo A (*pyogenes*), de estreptococos del grupo B (GBS o *agalactiae*), *dysgalactiae*, *uberis*, *nocardia*, así como *Staphylococcus aureus*. Se dispone de información general sobre los estreptococos en el *Manual of Clinical Microbiology* de P.R. Murria y col. (1995, 6ª Edición, ASM Press, Washington, D.C.). Más concretamente, de los estreptococos del grupo B, *agalactiae*. En una realización particular se administran vacunas a individuos con riesgo de infección de GBS, tales como mujeres embarazadas y niños para sepsis, meningitis y neumonía, así como individuos inmunocomprometidos tales como los que padecen diabetes, enfermedades del hígado o cáncer. Las vacunas también pueden tener aplicaciones veterinarias tales como para el tratamiento de la mastitis en ganado, que está mediada por las anteriores bacterias, así como por *E. coli*.

La vacuna de la presente invención también se puede usar para la fabricación de un medicamento usado para el tratamiento o profilaxis de la infección de estreptococos y/o de enfermedades o síntomas mediados por la infección de estreptococos, en particular de estreptococos del grupo A (*pyogenes*), de estreptococos del grupo B (GBS o *agalactiae*), *dysgalactiae*, *uberis*, *nocardia*, así como *Staphylococcus aureus*. Más particularmente estreptococos del grupo B, *agalactiae*.

Las composiciones de vacuna se encuentran preferiblemente en formas unitarias de dosis de entre aproximadamente 0,001 y 100 $\mu\text{g/Kg}$ (antígeno/peso corporal) y más preferiblemente de 0,01 a 10 $\mu\text{g/Kg}$, y aún más preferiblemente de 0,1 a 1 $\mu\text{g/Kg}$ de 1 a 3 veces con un intervalo de aproximadamente 1 a 12 semanas entre inmunizaciones, y más preferiblemente de 1 a 6 semanas.

De acuerdo con otro aspecto, se proporcionan polinucleótidos que codifican polipéptidos de estreptococos del grupo B (GBS) que se caracterizan por la secuencia de aminoácidos seleccionada del grupo que consiste en: SEQ ID NO: 39, y SEQ ID NO: 44, o fragmentos de las mismas.

Los polinucleótidos preferidos son los ilustrados en las Figuras 7a (SEQ ID NO: 37), 8 (SEQ ID NO: 42) y 9 (SEQ ID NO: 43) que corresponden a marcos de lectura abiertos, que codifican polipéptidos de la invención.

Los polinucleótidos preferidos son los ilustrados en las Figuras 7a (SEQ ID NO: 37), 8 (SEQ ID NO: 42) y 9 (SEQ ID NO: 43), y fragmentos de los mismos.

Los polinucleótidos más preferidos de la invención son los ilustrados en las Figuras 7 (SEQ ID NO: 37), 8 (SEQ ID NO: 42) y 9 (SEQ ID NO: 43).

Los polinucleótidos aún más preferidos de la invención son los ilustrados en las Figuras 8 (SEQ ID NO: 42) y 9 (SEQ ID NO: 43).

Cabe destacar que las secuencias de polinucleótidos ilustradas en las figuras pueden alterarse con codones degenerados y aún así seguir codificando los polipéptidos de la invención.

Debido a la degeneración de las secuencias codificadoras de nucleótidos, en la práctica de la invención pueden usarse otras secuencias de polinucleótidos que codifiquen sustancialmente los mismos polipéptidos de la presente invención. Éstas incluyen, aunque sin limitación, secuencias de nucleótidos que son alteradas por la sustitución de diferentes codones que codifican el mismo residuo de aminoácido dentro de la secuencia, produciendo de este modo un cambio silencioso.

En consecuencia, la presente invención proporciona adicionalmente polinucleótidos que se hibridan con las secuencias de polinucleótidos descritas en la presente memoria (o con las secuencias complementarias de las mismas) que tengan al menos un 70% de identidad entre secuencias. Más preferiblemente, los polinucleótidos son hibridables en condiciones severas, por ejemplo que tengan un 95% de identidad y más preferiblemente más de un 97% de identidad.

Por “capaces de hibridarse en condiciones severas” se entiende templar una molécula de ácido nucleico a al menos una región de una segunda secuencia de ácido nucleico (como cADN, mRNA ó ADN genómico) o a su cadena complementaria en condiciones estándar, por ejemplo alta temperatura y/o bajo contenido salino, que tiende a desfavorecer la hibridación de secuencias de nucleótidos no complementarias. Se describe un protocolo adecuado en Maniatis T. y col., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1982, que se incorpora a la presente memoria a modo de referencia.

En un aspecto adicional, se pueden usar polinucleótidos que codifican polipéptidos de la invención, o fragmentos de los mismos, en un método de inmunización de ADN. Es decir, se pueden incorporar en un vector que es replicable y expresable mediante inyección, produciendo con ello el polipéptido antigénico *in vivo*. Por ejemplo, se pueden incorporar polinucleótidos en un vector plásmido bajo el control del promotor CMV que es funcional en células eucarióticas. Preferiblemente el vector se inyecta intramuscularmente.

De acuerdo con otro aspecto, se proporciona un proceso para la producción de polipéptidos de la invención mediante técnicas recombinantes a través de la expresión de un polinucleótido que codifica dicho polipéptido en una célula hospedante, y recuperando el polipéptido producto. De forma alternativa, los polipéptidos se pueden producir de acuerdo con técnicas químicas sintéticas establecidas, por ejemplo, síntesis de oligopéptidos en fase sólida o en disolución, que son ligados para producir el polipéptido completo (ligadura en bloque).

Para la producción recombinante, se transfectan células hospedantes con vectores que codifican el polipéptido, y a continuación se cultivan en un medio nutriente modificado de forma apropiada para activar promotores, seleccionando transformantes o amplificando los genes. Los vectores adecuados son aquellos que son viables y replicables en el hospedante elegido e incluyen secuencias de ADN cromosomales, no cromosomales y sintéticas, por ejemplo, plásmidos bacterianos, ADN de fago, baculovirus, plásmidos de levadura, vectores derivados de combinaciones de plásmidos y ADN de fago. La secuencia de polipéptidos se puede incorporar al vector en una posición apropiada usando enzimas de restricción, de tal forma que se une operativamente a una región de control de la expresión que comprende un promotor, una posición de unión de ribosomas (región de consenso o secuencia de Shine-Dalgarno), y opcionalmente un operador (elemento de control). Uno puede seleccionar componentes individuales de la región de control de la expresión que sean apropiados para un hospedante y un vector dados de acuerdo con principios establecidos de biología molecular (Sambrook y col., *Molecular Cloning: A Laboratory Manual*, 2ª edición, Cold Spring Harbor, N.Y., 1989, incorporado a la presente memoria a modo de referencia). Los promotores adecuados incluyen, aunque sin limitación, el promotor LTR ó SV40, los promotores lac, tac ó trp de *E. coli* y el promotor P_L lambda de fago. Preferiblemente los vectores incorporarán un origen de replicación así como marcadores de selección, por ejemplo gen de resistencia a ampicilina. Los vectores bacterianos adecuados incluyen pET, pQE70, pQE60, pQE-9, pbs, pD10 phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A, ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5, y los vectores eucarióticos pBlueBacIII, pWL-NEO, pSV2CAT, pOG44, pXT1, pSG, pSVK3, pBPV, pMSG y pSVL. Las células hospedantes pueden ser bacterianas, por ejemplo *E. coli*, *Bacillus subtilis*, *Streptomyces*; fúngicas, por ejemplo *Aspergillus niger*, *Aspergillus nidulans*; de levaduras, por ejemplo *Saccharomyces*; o eucarióticas, por ejemplo CHO, COS.

En la expresión del polipéptido en cultivo, las células se recolectan normalmente mediante centrifugación, a continuación se rompen empleando medios físicos o químicos (si el polipéptido expresado no es segregado al medio) y el extracto crudo resultante se retiene para aislar el polipéptido de interés. La purificación del polipéptido a partir del medio de cultivo o lisato se puede lograr empleando técnicas establecidas que dependen de las propiedades del polipéptido, por ejemplo, usando precipitación en sulfato de amonio o etanol, extracción ácida, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrofóbica, cromatografía de hidroxilapatita y cromatografía de lectina. La purificación final se puede lograr empleando HPLC.

El polipéptido se puede expresar con o sin una secuencia líder o de secreción. En el primer caso el líder se puede eliminar usando procesamiento post-traducción (véanse los documentos US 4.431.739, 4.425.437 y 4.338.397) o pueden ser eliminados químicamente después de purificar el polipéptido expresado.

De acuerdo con un aspecto adicional, los polipéptidos de GBS de la invención se pueden usar en un ensayo diagnóstico para la infección de estreptococos, en concreto la infección de GBS. Es posible utilizar varios métodos

diagnósticos, por ejemplo la detección de un organismo estreptococal en una muestra biológica. Se puede seguir el siguiente procedimiento:

- a) obtener una muestra biológica de un paciente;
- b) incubar un anticuerpo o fragmento del mismo que sea reactivo con un polipéptido de GBS de la invención con la muestra biológica para formar una mezcla; y
- c) detectar anticuerpos unidos específicamente o fragmentos unidos en la mezcla, lo que indica la presencia de estreptococos.

De forma alternativa, se puede llevar a cabo un método para la detección de anticuerpos específicos de un antígeno de estreptococos en una muestra biológica que contiene, o que se sospecha que contiene, dicho anticuerpo de la siguiente manera:

- a) aislar una muestra biológica de un paciente;
- b) incubar uno o más polipéptidos de GBS de la invención, o fragmentos de los mismos, con la muestra biológica para formar una mezcla; y
- c) detectar antígenos unidos específicamente o fragmentos unidos en la mezcla, lo que indica la presencia de anticuerpos específicos de estreptococos.

El especialista en la técnica reconocerá que este ensayo diagnóstico puede presentarse en diversas formas, que incluyen un ensayo inmunológico tal como el ensayo inmunosorbente ligado a enzima (ELISA), un radioinmunoensayo o un ensayo de aglutinación de látex, esencialmente para determinar si los anticuerpos específicos de la proteína están presentes en un organismo.

Las secuencias de ADN que codifican polipéptidos de la invención también pueden usarse para diseñar sondas de ADN para su uso en la detección de la presencia de estreptococos en una muestra biológica que se sospecha contiene dichas bacterias. El método de detección de la invención comprende:

- a) aislar la muestra biológica de un paciente;
- b) incubar una o más sondas de ADN que presenten una secuencia de ADN que codifica un polipéptido de la invención o fragmentos del mismo con la muestra biológica para formar una mezcla; y
- c) detectar la sonda de ADN unida específicamente en la mezcla, lo que indica la presencia de bacterias estreptococales.

Las sondas de ADN de esta invención también pueden usarse para detectar estreptococos circulantes, por ejemplo ácidos nucleicos de GBS en una muestra, por ejemplo usando una reacción en cadena de polimerasa, como un método para diagnosticar infecciones de estreptococos. La sonda se puede sintetizar usando técnicas convencionales y puede inmovilizarse sobre una fase sólida, o puede etiquetarse con una marca detectable. Una sonda de ADN preferida para esta aplicación es un oligómero que tiene una secuencia complementaria a al menos 6 nucleótidos contiguos de los polipéptidos de GBS de la invención.

Otro método diagnóstico para la detección de estreptococos es un paciente comprende:

- a) etiquetar un anticuerpo reactivo con un polipéptido de la invención o un fragmento del mismo con una marca detectable;
- b) administrar el anticuerpo marcado o el fragmento marcado al paciente; y
- c) detectar anticuerpos marcados o fragmentos marcados unidos específicamente en el paciente, lo que indica la presencia de estreptococos.

Un aspecto adicional de la invención es el uso de polipéptidos de GBS de la invención como inmunógenos para la producción de anticuerpos específicos para la diagnosis, y en concreto para el tratamiento de la infección de estreptococos. Los anticuerpos adecuados se pueden determinar usando métodos de escrutinio apropiados, por ejemplo midiendo la capacidad de un anticuerpo concreto para proteger pasivamente frente a infección de estreptococos en un modelo de ensayo. Un ejemplo de un modelo animal es el modelo de ratón descrito en los Ejemplos de la presente memoria. El anticuerpo puede ser un anticuerpo entero o un fragmento de unión a antígeno del mismo, y puede pertenecer en general a cualquier clase de inmunoglobulina. El anticuerpo o fragmento puede ser de origen animal, específicamente de origen mamífero y más específicamente de origen murino, de rata o humano. Puede ser un anticuerpo natural o un fragmento del mismo, o si se desea, un anticuerpo o fragmento de anticuerpo recombinante. El término anticuerpo o fragmento de anticuerpo recombinante significa un anticuerpo o fragmento de anticuerpo que se produjo usando técnicas de biología molecular. El anticuerpo o fragmento de anticuerpo puede ser policlonal, o

preferiblemente monoclonal. Puede ser específico de una serie de epítomos asociados con los polipéptidos de GBS, pero preferiblemente es específico de uno solo.

Ejemplo 1

Modelo de murina de infección de Estreptococo de Grupo B (GBS) letal

El modelo de ratón de infección de GBS se describe con detalle en Lancefield y col. (J Exp Med 142: 165-179, 1975). La cepa de GBS C388/90 (producto clínico aislado obtenido en 1990 a partir de fluido cefalorraquídeo de un paciente que sufría meningitis, Hospital Infantil de Eastern Notario, Ottawa, Canadá) y la cepa NCS246 (Nacional Center for Streptococcus, Provincial Laboratory of Public Health for Northern Alberta, Edmonton, Canadá) fueron serotipadas respectivamente como el tipo Ia/c y el tipo II/R.

Para aumentar su virulencia, se sometió en serie a ratones a las cepas de GBS C388/90 (serotipo Ia/c) y NCS246 (serotipo II/R) tal como se ha descrito previamente (Lancefield y col., J. Exp. Med. 142: 165-179, 1975). En resumen, se siguió la evolución del aumento de la virulencia usando inoculaciones intraperitoneales de diluciones en serie de un subcultivo en caldo de Todd-Hewitt obtenido a partir de sangre o de bazo de ratones infectados. Tras la última exposición, las muestras de sangre infectadas se usaron para inocular caldo de Todd-Hewitt. Tras una incubación de 2 horas a 37°C con un 7% de CO₂, se añadió al cultivo glicerol con una concentración final del 10% (v/v). A continuación se dividió en alícuotas el cultivo y se almacenó a -80°C para su uso en los experimentos de exposición a GBS. Se determinó el número de cfu de GBS presentes en estas muestras congeladas. Se determinó que las concentraciones bacterianas necesarias para matar al 100% (LD100) de los ratones de 18 semanas de edad eran de $3,5 \times 10^5$ y de $1,1 \times 10^5$ respectivamente para las cepas de GBS C388/90 y NCS246, que corresponden a un aumento significativo de la virulencia para ambas cepas. De hecho, las LD100 registradas antes de las exposiciones para estas dos cepas eran superiores a 10^9 cfu.

En una exposición bacteriana, se ajustó una alícuota recién congelada de una cepa virulenta de GBS a la concentración bacteriana apropiada usando caldo de Todd-Hewitt, y se inyectó 1 mL intraperitonealmente a cada ratón CD-1 hembra. Los ratones usados para los experimentos de protección pasiva tenían una edad de 6 a 8 semanas, mientras que los usados para los experimentos de protección activa tenían una edad de aproximadamente 18 semanas en el momento de la exposición. Todos los inóculos fueron verificados mediante conteo de colonias. Se observó a los animales en busca de cualquier signo de infección cuatro veces al día durante las primeras 48 horas después de la exposición, y a continuación a diario durante los siguientes 12 días. Al final de dicho periodo, se obtuvieron muestras de sangre de los supervivientes y se congelaron a -20°C. El bazo obtenido de cada ratón que sobrevivió a la exposición se cultivó con el objetivo de identificar cualquier GBS remanente.

Ejemplo 2

Inmunización y protección en ratones con células de GBS enteras muertas con formaldehído

Se prepararon células enteras de GBS muertas por formaldehído de acuerdo con los procedimientos descritos en Lancefield y col. (J. Exp. Med. 142: 165-179, 1975). En resumen, se lavó un cultivo de una noche de placas de agar de sangre de oveja (Quelab Laboratorios, Montreal, Canadá) de una cepa de GBS dos veces en tampón PBS (disolución salina de fosfato tamponada, pH 7,2), se ajustó a aproximadamente 3×10^9 cfu/mL y se incubó toda la noche en PBS que contenía formaldehído al 0,3% (v/v). La suspensión muerta por formaldehído se lavó con PBS y se mantuvo congelada a -80°C.

Se inyectó subcutáneamente ratones CD-1 hembra, con una edad de 6 a 8 semanas (Charles River, St-Constant, Québec, Canadá), tres veces con un intervalo de dos semanas, con 0,1 mL de células muertas por formaldehído de la cepa de GBS C388/90 (aproximadamente 6×10^7 GBS), o con 0,1 mL de PBS para el grupo de control. El día antes de la inmunización, se añadió a estas preparaciones AlhydrogelTM (Superfos Biosector, Frederikssund, Dinamarca) con una concentración final de 0,14 mg ó de 0,21 mg de Al, y se incubó una noche a 4°C con agitación. Se obtuvieron muestras de suero de cada ratón antes de comenzar el protocolo de inmunización y dos semanas después de la última inyección. Los sueros se congelaron a -20°C.

Ocho ratones de cada grupo de control inyectados con PBS y del grupo inmunizado con células de GBS de la cepa C388/90 (Ia/c) enteras muertas por formaldehído, fueron expuestos a $1,5 \times 10^4$ cfu de la cepa de GBS C388/90 (Ia/c) una semana después de la tercera inyección. Todos los ratones inmunizados con las células enteras de GBS muertas por formaldehído sobrevivieron a la exposición homóloga, mientras que, a los 5 días tras la exposición, sólo 4 de cada 8 ratones inyectados con PBS sobrevivieron a la infección. Con el fin de aumentar la tasa de mortalidad en los grupos de control, la suspensión bacteriana se tuvo que ajustar de acuerdo con la edad de los ratones en el momento de la exposición bacteriana. En los experimentos de exposición posteriores, cuando los ratones tenían más de 15 semanas de edad, el inóculo bacteriano fue incrementado hasta concentraciones entre $3,0 \times 10^5$ y $2,5 \times 10^6$ cfu.

TABLA 1

Inmunización de ratones CD1 con células enteras de GBS muertas por formaldehído, y posterior exposición homóloga [cepa C388/90 (Ia/c)] y exposición heteróloga [cepa NCS246 (II/R)]

Preparaciones antigénicas usadas para la inmunización ¹	Numero de ratones vivos 14 días después de la exposición bacteriana (% de supervivencia)	
	Exposición homóloga: cepa C388/90 (Ia/c)	Exposición heteróloga: cepa NCS246 (II/R)
1ª Infección		
Células de GBS muertas por formaldehído, cepa C388/90 (Ia/c) ²	8/8 (100) ³	n.d. ⁵
Control (PBS)	4/8 (50)	n.d.
2ª Infección		
Células de GBS muertas por formaldehído, cepa C388/90 (Ia/c)	6/6 (100) ⁴	0/6 (0) ⁶
Control (PBS)	2/6 (33)	0/6 (0)
¹ se usó alhydrogel™ a una concentración final de 0,14 mg ó 0,21 mg de Al; ² aproximadamente 6×10^7 cfu; ³ exposición intraperitoneal con 1 mL de medio de cultivo de Todd-Hewitt que contiene suspensión de GBS C388/90 (Ia/c) ajustado a $1,5 \times 10^4$ cfu; ⁴ exposición intraperitoneal con 1 mL de medio de cultivo de Todd-Hewitt que contiene suspensión de GBS C388/90 ⁵ (Ia/c) ajustado a $2,1 \times 10^6$ cfu; ⁵ no realizado; ⁶ exposición intraperitoneal con 1 mL de medio de cultivo de Todd-Hewitt que contiene suspensión de GBS NCS246 (II/R) ajustado a $1,2 \times 10^5$ cfu.		

En otro experimento, se inyectó PBS a un grupo de 12 ratones correspondiente a un grupo de control, mientras que un segundo grupo de 12 ratones fue inmunizado con células de GBS enteras muertas por formaldehído de la cepa C388/90 (Ia/c). Seis ratones de cada uno de estos dos grupos fueron expuestos a $2,1 \times 10^6$ cfu de la cepa de GBS C388/90 (Ia/c) (Tabla 1). Como primer experimento de exposición, todos los ratones inmunizados con la cepa de GBS C388/90 (Ia/c) sobrevivieron a la exposición homóloga. Solamente dos de cada 6 ratones inyectados con PBS sobrevivieron a la infección.

Los restantes 6 ratones de ambos grupos fueron usados a continuación una semana después para verificar si dicha preparación antigénica podía conferir protección cruzada frente a la cepa NCS246 (II/R), que produce una cápsula serológicamente distinta. Ninguno de los ratones infectados con esta segunda cepa de GBS sobrevivió a la infección. Este último resultado sugirió que la mayoría de la respuesta inmune protectora inducida por la cepa C388/90 muerta por formaldehído se dirige contra el polisacárido capsular, y que podría estar restringida a cepas de dicho serotipo en particular. Estos resultados indicaron claramente que este modelo concreto de infección se puede usar de forma eficaz para estudiar la protección conferida a través de vacunación.

Ejemplo 3

Inmunización de conejo con células de GBS enteras muertas por formaldehído y protección pasiva en ratones

Se inmunizó un conejo de Nueva Zelanda (2,5 Kg, Charles River, St-Constant, Québec, Canadá) con células de GBS de cepa C388/90 (Ia/c) muertas por formaldehído para obtener suero hiperinmune. Dicho conejo fue inyectado subcutáneamente tres veces, con un intervalo de tres semanas, con aproximadamente $1,5 \times 10^9$ cfu de células enteras

ES 2 278 436 T3

mueras por formaldehído de GBS cepa C388/90 (Ia/c). Se usó adyuvante completo de Freund (Gibco BRL Life Technologies, Grand Island, Nueva York) como adyuvante para la primera inmunización, y se usó adyuvante incompleto de Freund (Gibco BRL) para las siguientes dos inyecciones. Se obtuvieron muestras de suero antes de comenzar el protocolo de inmunización y dos semanas después de la última inyección. Los sueros se congelaron a -20°C.

También se evaluó la capacidad de este particular suero hiperinmune de conejo para proteger pasivamente ratones frente a infección letal con GBS. La inyección intraperitoneal de los ratones con 15 ó 25 µL de suero hiperinmune de conejo 18 horas antes de la exposición protegió a 4 de cada 5 ratones (80%) frente a la infección. De forma comparativa, para los ratones del grupo de control, inyectados con PBS o con suero obtenido a partir de conejo inmunizado con una preparación de membrana exterior meningococal, se registraron tasas de supervivencia inferiores al 20%. Este resultado indica claramente que la inmunización de otra especie animal con células de GBS muertas puede inducir la producción de anticuerpos que pueden proteger pasivamente a ratones. Este reactivo también se usará para caracterizar clones.

TABLA 2

Protección pasiva de ratones CD-1 conferida mediante suero de conejo obtenido tras inmunización con una preparación antigénica de estreptococos enteros de grupo B muertos por formaldehído (cepa C388/90 (Ia/c))

Grupos	Número de ratones vivos 14 días después de la exposición bacteriana con GBS de cepa C388/90 (Ia/c) ²	% de supervivencia
Suero hiperinmune de conejo ² - 25 µL	4/5	80
Suero hiperinmune de conejo ¹ - 15 µL	4/5	80
Suero de control de conejo - 25 µL	1/5	20
Control con PBS	1/10	10
¹ se usó adyuvante de Freund completo para la primera inmunización, y adyuvante de Freund incompleto para las siguientes dos inyecciones; ² exposición intraperitoneal con 1 mL de medio de cultivo de Todd-Hewitt que contiene una suspensión de GBS C388/90 (Ia/c) ajustada a 2×10^4 cfu.		

Ejemplo 4

Producción recombinante de proteína de fusión His.Tag-GBS

Se amplificó la región codificadora de un gen de GBS mediante PCR (Sistema DNA Thermal Cycler GeneAmp PCR 2400 Perkin Elmer, San Jose, CA) a partir de ADN genómico de la cepa de GBS C388/90 (Ia/c) usando los oligos que contenían las extensiones de base correspondientes a la adición de las posiciones de restricción BgIII (AGATCT) y HindIII (AAGCTT), respectivamente. El producto de PCR se purificó en gel de agarosa usando un kit de extracción de gel Qiaex II de Qiagen (Chatsworth, CA), se digirió con las enzimas de restricción BgIII y HindIII (Pharmacia Canada Inc. Baie d'Urfe, Canadá), y se extrajo con fenol:cloroformo antes de ser precipitado en etanol. El vector pET-32b(+) (Novagen, Madison, WI) que contiene la secuencia tiorredoxin-His.Tag se digirió con las enzimas de restricción BgIII y HindIII, se extrajo con fenol:cloroformo, y a continuación se precipitó en etanol. El fragmento de ADN genómico BgIII-HindIII se ligó al vector BgIII-HindIII pET-32b(+) para crear la secuencia codificadora correspondiente a la proteína de fusión tiorredoxin-His.Tag-GBS, cuyo gen estaba bajo el control del promotor T7. Los productos ligados se transformaron en la cepa de *E. coli* XLI Blue MRF' ($\Delta(mcrA)$ 183 Δ (*mcrCB-hsdSMR-mrr*) 173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacI^qZAM15Tn10* (Tet^r)]^c) (Stratagene, La Jolla, CA) de acuerdo con el método de Simanis (Hanahan, D., *DNA Cloning*, 1985, D.M. Glover (ed.), páginas 109-135). El plásmido pET recombinante se purificó usando un kit de Qiagen (Qiagen, Chatsworth, CA) y la secuencia de nucleótidos del injerto de ADN se verificó mediante secuenciamiento de ADN (kit Taq Dye Deoxy Terminador Cycle Sequencing, ABI, Foster City, CA). El plásmido pET recombinante fue transformado mediante electroporación (aparato Gene Pulser II, BIO-RAD Labs, Mississauga, Canadá) en la cepa de *E. coli* AD494 (DE3) ($\Delta ara^{-} leu7697 \Delta lacX74 \Delta phoA PvuII phoR \Delta malF3$

F' [*lac*⁺ (*lacI*^q) *pro*] *trxB*::Kan (DE3)) (Novagen, Madison, WI). En esta cepa de *E. coli*, el promotor T7 que controla la expresión de la proteína de fusión, es reconocido específicamente por la polimerasa de ARN T7 (presente en el profago λ DE3), cuyo gen se encuentra bajo el control del promotor *lac* que es inducible mediante isopropil- β -D-tiogalactopiranosido (IPTG).

El transformante AD494(DE3)/rpET se cultivó a 37°C con agitación a 250 rpm en caldo LB (peptona 10 g/L, extracto de levadura 5 g/L, NaCl 10 g/L) que contiene 100 μ g de ampicilina (Sigma-Aldrich Canada Ltd, Oakville, Canadá) por mL hasta que la A₆₀₀ alcanzó un valor de 0,6. Con el fin de inducir la producción de la proteína de fusión tiorredoxin-His.Tag-GBS, las células se incubaron otras 2 horas más en presencia de IPTG con una concentración final de 1 mM. Las células bacterianas se recogieron por centrifugación.

La proteína de fusión recombinante producida por AD494(DE3)/rpET32 tras inducción con IPTG durante 2 h se obtuvo parcialmente como cuerpos de inclusión insolubles que fueron purificados a partir de las proteínas de *E. coli* endógenas mediante aislamiento de los agregados insolubles (Gerlach, G.F., y col., 1992, Infect. Immun. 60: 892). Las células inducidas a partir de un cultivo de 500 mL se resuspendieron en 20 mL de tampón Tris-HCl 50 mM con un 25% de sacarosa (pH 8,0) y se congelaron a -70°C. La lisis de las células en suspensión congelada se logró mediante la adición de 5 mL de una disolución de lisozima (10 mg/mL) en tampón Tris-HCl 250 mM (pH 8,0) seguida de una incubación de 10 a 15 minutos en hielo, y de la adición de 150 mL de mezcla detergente (5 partes de tampón Tris-HCl 20 mM (pH 7,4)-NaCl 300 mM-ácido desoxicólico al 2%-Nonidet P-40 al 2% y 4 partes de tampón Tris-HCl 100 mM (pH 8)-EDTA 50 mM-Triton X-100 al 2%) seguida de una incubación de 5 minutos en hielo. Después de ser sometido a sonicación, se recogieron los agregados proteicos mediante centrifugación durante 30 minutos a 35.000 x g, y se mantuvo una muestra de la fracción celular soluble. Las proteínas agregadas fueron solubilizadas en hidrocloreuro de guanidina 6 M. El análisis Western blot demostró la presencia de la proteína de fusión en las fracciones soluble e insoluble a la vez, realizado usando el suero de un ratón inyectado con células de GBS cepa C388/90 (Ia/c) muertas por formaldehído que sobrevivió a la exposición bacteriana con la correspondiente cepa de GBS.

La purificación de la proteína de fusión a partir de la fracción soluble de AD494(DE3)/rpET inducida por IPTG se realizó mediante cromatografía de afinidad basada en las propiedades de la secuencia His.Tag (6 residuos de histidina consecutivos) para ligarse a cationes divalentes (Ni²⁺) inmovilizados sobre resina de quelación metálica His.Bind (Novagen, Madison, WI). Los métodos de purificación usados son los descritos en el Manual del sistema pET, 6ª Edición (Novagen, Madison, WI). En resumen, las células paletizadas obtenidas de un cultivo de 100 mL inducido con IPTG se resuspendieron en 4 mL de tampón de Unión (imidazol 5 mM-NaCl 500 mM-Tris-HCl 20 mM, pH 7,9), se sometieron a sonicación, y se giraron a 39.000 x g durante 20 minutos para eliminar los desechos. Se filtró el sobrenadante (membrana con un tamaño de poro de 0,45 μ m) y se depositó en una columna de resina His.Bind equilibrada en tampón de unión. A continuación se lavó la columna con 10 volúmenes de columna de tampón de unión seguidos de 6 volúmenes de columna de tampón de lavado (imidazol 20 mM-NaCl 500 mM-Tris-HCl 20 mM, pH 7,9). La eliminación de la sal y del imidazol de la muestra se realizó mediante diálisis frente a 3 x 1 litro de PBS a 4°C.

Las cantidades de proteína de fusión obtenidas de las fracciones citoplásmicas soluble e insoluble de *E. coli* se estimaron mediante tinción de Coomassie de un gel de dodecilsulfato sódico (SDS)-poliacrilamida con diluciones en serie de dichas proteínas, y con un patrón de albúmina de suero bovino (Pierce Chemical Co., Rockford, IL).

Ejemplo 5

Producción recombinante de proteína de GBS bajo control de promotor *P_L* λ

Se insertó la región de codificación de ADN de una proteína de GBS por debajo del promotor λ P_L en el vector de traducción pURV22. Este plásmido fue derivado de p629 (George y col., 1987, Bio/Technology 5: 600), del cual se eliminó la región codificadora de una porción de la glicoproteína (gD-1) de virus simple del herpes de tipo I (HSV-I), y se sustituyó el gen de resistencia a la ampicilina por una casete de canamicina obtenida del vector plásmido pUC4K (Pharmacia Biotech Canada Inc., Baie D'Urfe, Canadá). El vector contenía una casete del gen represor sensible a la temperatura λ cI857 bacteriófago, en el cual se había eliminado el promotor funcional P_R. La desactivación del represor cI857 mediante un aumento de la temperatura desde los intervalos 30-37°C hasta 37-42°C dio como resultado la inducción del gen bajo el control de λ P_L. La traducción del gen se controló mediante la posición de unión de ribosomas *cro*, seguida por debajo de una posición de restricción BglIII (AGATCT) y la ATG: ACTAAGGAGGTTAGATCTATG.

Se usaron enzimas de restricción y ligasa de ADN T4 de acuerdo con los suministradores (Pharmacia Biotech Canada Inc., Baie D'Urfe, Canadá; y New England Biolabs Ltd., Mississauga, Canadá). Se llevó a cabo una electroforesis en gel de agarosa de los fragmentos de ADN tal como se describe en Sambrook y col. (*Molecular Cloning: A Laboratory Manual*, 1989, Cold Spring Harbor Laboratory Press, N.Y.). Se preparó el AND de las bacterias GBS de acuerdo con procedimientos descritos en Jayarao y col. (*J. Clin. Microbiol.*, 1991, 29: 2774). Se realizaron reacciones de amplificación de ADN mediante reacción en cadena de polimerasa (PCR) usando el Sistema DNA Thermal Cycler GeneAmp PCR 2400 (Perkin Elmer, San Jose, CA). Los plásmidos usados para el secuenciamiento de ADN se purificaron usando kits de plásmidos de Qiagen (Chatsworth, CA). Los fragmentos de ADN fueron purificados en geles de agarosa usando kits de extracción de gel Qiaex II de Qiagen (Chatsworth, CA). Las transformaciones de plásmidos se llevaron a cabo mediante el método descrito por Hanahan (*DNA Cloning*, Glover (ed.), páginas 109-135, 1985). El secuenciamiento de injertos de ADN genómico en plásmidos se realizó usando oligonucleótidos sintéticos que fueron

5 sintetizados mediante el sintetizador de oligonucleótidos modelo 394 (Perkin-Elmer Corp., Applied Biosystems Div. (ABI), Foster City, CA). Las reacciones de secuenciamiento se llevaron a cabo mediante PCR, usando el kit Taq Dye Deoxy Terminator Cycle Sequencing (ABI, Foster City, CA), y se realizó electroforesis de ADN en un secuenciador de ADN automatizado 373A (ABI, Foster City, CA). El montaje de la secuencia de ADN se realizó usando el programa Sequencer 3.0 (Gene Codes Corporation, Ann Arbor, MI). El análisis de las secuencias de ADN y de sus polipéptidos predichos se realizó con el programa Gene Works versión 2.45 (Intelligenetics, Inc., Mountain View, CA).

10 La región codificadora del gen de GBS se amplificó mediante PCR de ADN genómico de la cepa de GBS C388/90 (Ia/c) usando oligos que contenían extensiones base para la adición de posiciones de restricción BgIII (AGATCT) y XbaI (TCTAGA), respectivamente. El producto de PCR se purificó en gel de agarosa usando un kit de extracción de gel Qiaex II de Qiagen (Chatsworth, CA), se digirió con las enzimas de restricción BgIII y XbaI, y se extrajo con fenol:cloroformo antes de ser precipitado en etanol. El vector pURV22 se digirió con las enzimas de restricción BgIII y XbaI, se extrajo con fenol:cloroformo, y se precipitó en etanol. El fragmento de ADN genómico BgIII-XbaI se ligó al vector BgIII-XbaI pURV22, en el cual el gen de GBS estaba bajo el control del promotor λ P_L. Los productos
15 ligados se transformaron en la cepa de *E. coli* XLI Blue MRF' ($\Delta(mcrA)$ 183 Δ (*mcrCB-hsdSMR-mrr*) 173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F' *proAB lacI^qZAM15Tn10* (Tet^r)]^c) (Stratagene, La Jolla, CA) de acuerdo con los métodos descritos en Hanahan (ver anteriormente). Los transformantes que alojan plásmidos con el injerto fueron identificados mediante el análisis de células lisadas sometido a electroforesis sobre gel de agarosa (Sambrook y col., ver anteriormente). El plásmido pURV22 recombinante fue purificado usando un kit Qiagen (Qiagen, Chatsworth, CA) y la secuencia de nucleótidos del injerto de ADN se verificó mediante secuenciamiento de ADN.
20

El transformante XLI Blue MRF'/pURV22 se cultivó a 34°C con agitación a 250 rpm en caldo LB que contenía 50 μ g de canamicina por mL hasta que la A₆₀₀ alcanzó un valor de 0,6. Con el fin de inducir la producción de la proteína de fusión, las células fueron incubadas durante otras 4 horas más a 39°C. Las células bacterianas se recogieron mediante
25 centrifugación, se resuspendieron en un tampón de muestra, se hirvieron durante 10 minutos y se mantuvieron a -20°C.

Ejemplo 6

Subclonación de gen de proteína de GBS en el plásmido pCMV-GH de CMV

30 Se insertó la región codificadora de ADN de una proteína en fase por debajo del gen de la hormona de crecimiento humana (hGH), que estaba bajo el control transcripcional del promotor de citomegalovirus (CMV) en el vector plásmido pCMV-GH (Tang y col., *Nature*, 1992, 356: 152). El promotor CMV no es funcional en células de *E. coli* pero es activo cuando se administra el plásmido en células eucarióticas. El vector también incorporaba el gen de resistencia a ampicilina.
35

La región codificadora del gen se amplificó mediante PCR a partir de ADN genómico de la cepa C388/90 (Ia/c) de GBS usando los oligos que contenían extensiones base para la adición de las posiciones de restricción BgIII (AGATCT) y HindIII (AAGCTT). El producto de la PCR se purificó en gel de agarosa usando el kit de extracción de gel Qiaex II de Qiagen (Chatsworth, CA), se digirió con las enzimas de restricción BgIII y HindIII, y se extrajo con fenol:cloroformo antes de precipitarse con etanol. El vector pCMV-GH (Laboratorio del Dr. Stephen A. Johnston, Departamento de Bioquímica, Universidad de Texas, Dallas, Texas), que contiene hormona de crecimiento humana para crear proteínas de fusión, se digirió con las enzimas de restricción BamHI y HindIII, se extrajo con fenol:cloroformo, y se precipitó con etanol. El fragmento de ADN genómico de 1,3 Kb BgIII-HindIII se ligó al vector BamHI-HindIII pCMV-GH para
40 crear la proteína de fusión hGH-GBS bajo el control del promotor CMV. Los productos ligados fueron transformados en la cepa de *E. coli* DH5 α [ϕ 80 *lacZ* Δ M15 *endA1 recA1 hsdR17* (^rK^{-m}K⁺) *supE44 thi-1* λ^{-} *gyrA96 relA1* Δ (*lacZYA-argF*) U169] (Gibco BRL, Gaithersburg, MD) de acuerdo con los métodos descritos por Hanahan (ver anteriormente). Los transformantes que albergan plásmidos con el injerto fueron identificados mediante análisis de células lisadas sometidas a electroforesis sobre gel de agarosa (Sambrook, J. y col., ver anterior). El plásmido pCMV recombinante fue purificado usando un kit Qiagen (Qiagen, Chatsworth, CA) y la secuencia de nucleótidos del injerto de ADN se verificó mediante secuenciamiento de ADN.
50

Ejemplo 7

Actividad Inmunológica de proteína de GBS para exposición a GBS

Se inyectó subcutáneamente a cuatro grupos de 12 ratones CD-1 hembra (Charles River, St-Constant, Québec, Canadá), de 6 a 8 semanas de edad, tres veces a intervalos de tres semanas con 0,1 mL de las siguientes preparaciones antigénicas: células de cepa C388/90 de GBS muertas por formaldehído (aproximadamente 6 x 10⁷ cfu), 20 μ g de la
60 proteína de fusión, purificada por afinidad (columna de níquel), a partir de la fracción citoplásmica soluble de *E. coli*, ó 20 μ g de polipéptido de control tiorredoxin-His.Tag purificado por afinidad (columna de níquel). Como adyuvante, se añadieron 20 μ g de QuilATM (Cedarlane Laboratorios Ltd., Hornby, Canadá) a cada preparación antigénica. Se extrajeron muestras de suero de cada ratón antes de la inmunización (PB) y en los días 20 (TB1), 41 (TB2) y 54 (TB3) durante los protocolos de inmunización. Los sueros se congelaron a -20°C.
65

Se registró un aumento en los títulos de ELISA después de cada inyección de la proteína de fusión, lo que indica una buena respuesta primaria y un aumento brusco de la respuesta inmune humoral después de la segunda y de la tercera administración. Al final del periodo de inmunización, el valor medio de los títulos de ELISA recíprocos

fue de 456.145 para el grupo inmunizado con 20 μ g de proteína de fusión obtenida de cuerpos de inclusión, en comparación a 290.133 correspondiente al grupo de ratones inmunizados con la proteína procedente de la fracción soluble de *E. coli*. Este último resultado sugiere que la proteína obtenida a partir de cuerpos de inclusión podría ser más inmunogénica que la proteína soluble. El análisis de sueros de ratones en ELISA usando la tiorredoxin-His.Tag purificada por afinidad para cubrir placas demostró que se obtienen títulos despreciables frente a la porción de tiorredoxin-His.Tag de la proteína de fusión. La reactividad de los sueros de ratones inyectados con la proteína de fusión recombinante también fue evaluada mediante ELISA frente a células enteras muertas por formaldehído de GBS de cepa C388/90. Los anticuerpos inducidos por inmunización con proteína de fusión recombinante también reconocieron sus epítomos específicos sobre células de GBS, lo que indica que su conformación está suficientemente próxima a la de la proteína estreptococal nativa para inducir anticuerpos reactivos.

Para verificar si la respuesta inmune inducida por la inmunización podría proteger frente a una infección de GBS, se expuso a los ratones a $3,5 \times 10^5$ cfu de GBS de la cepa C388/90 (Ia/c) y $1,2 \times 10^5$ cfu de GBS de la cepa NCS246 (II/R), cuyos resultados se muestran en las tablas 3 y 4, respectivamente. Los ratones inmunizados con péptido de tiorredoxin-His.Tag de control no fueron protegidos frente a la exposición a cualquiera de las cepas de GBS, mientras que los inmunizados con células enteras muertas por formaldehído de cepa C388/90 sólo proporcionaron protección frente a la exposición homóloga. La proteína de fusión tiorredoxin-His.Tag-GBS de la invención protegió a los ratones frente a la exposición con ambas cepas de GBS. Los cultivos de sangre y de bazo de estos ratones no revelaron presencia alguna de GBS.

TABLA 3

Supervivencia a exposición a GBS cepa C388/90 (Ia/c)¹

Agente inmunizante	Nº de ratones que sobreviven a la exposición	% de supervivencia
Tiorredoxin-His.Tag ²	1/6	17
Células C388/90 muertas por formaldehído ³	6/6	100
Fusión tiorredoxin-His.Tag-GBS (preparación cuerpo de inclusión) ⁴	6/6	100
Fusión tiorredoxin-His.Tag-GBS (fracción citoplásmica) ⁴	6/6	100

¹ administración intraperitoneal con 1 mL de medio de cultivo de Todd-Hewitt ajustado a $3,5 \times 10^5$ cfu;
² administración de 20 μ g; patas posteriores paralizadas en los ratones supervivientes; GBS detectado en sangre y bazo;
³ administración de 6×10^7 cfu;
⁴ administración de 20 μ g.

TABLA 4

Supervivencia a exposición a GBS cepa NCS246 (II/R)¹

Agente inmunizante	Nº de ratones que sobreviven a la exposición	% de supervivencia
Tiorredoxin-His.Tag ²	0/6	0
Células C388/90 muertas por formaldehído ³	2/6	34
Fusión tiorredoxin-His.Tag-GBS (preparación cuerpo de inclusión) ²	5/5 ⁴	100
Fusión tiorredoxin-His.Tag-GBS (fracción citoplásmica) ²	6/6	100
¹ administración intraperitoneal con 1 mL de medio de cultivo de Todd-Hewitt que contiene GBS NCS246 (II/R) ajustado a $1,2 \times 10^5$ cfu; ² administración de 20 µg; ³ administración de 6×10^7 cfu; ⁴ un ratón murió durante la inmunización.		

Ejemplo 8

La inmunización con proteína de GBS recombinante confiere protección frente a infección de GBS experimental

Este Ejemplo ilustra la protección de ratones contra infección fatal de GBS mediante inmunización con la proteína recombinante correspondiente a la SEQ ID NO: 39.

Se inmunizaron grupos de 10 ratones CD-1 hembra (Charles River) subcutáneamente tres veces con intervalos de tres semanas con 20 µg de proteína recombinante purificada a partir de la cepa de *E. coli* BLR (Novagen), que aloja el vector plásmido pURV22 recombinante que contiene el gen de GBS correspondiente a la SEQ ID NO: 42, en presencia de 20 µg de adyuvante QuilATM (Cedarlane Laboratorios Ltd, Hornby, Canadá) o, a modo de control, solo con adyuvante QuilATM en PBS. Se tomaron muestras de sangre del seno orbital en los días 1, 22 y 43 antes de cada inmunización, y catorce días (día 57) después de la tercera inyección. Una semana después, se expusieron los ratones a aproximadamente entre 10^4 y 10^5 cfu de diversas cepas virulentas de GBS. Las muestras del inóculo de exposición a GBS se colocaron en placas de agar de sangre de oveja al 5%/TSA para determinar la CFU y para verificar la dosis de exposición. Se registraron las muertes durante un periodo de 14 días, y en el día 14 después de la exposición los ratones supervivientes fueron sacrificados y se determinó la presencia de organismos GBS en la sangre y en el bazo. Los datos de supervivencia se muestran en la Tabla 5.

Los sueros pre-exposición fueron analizados en busca de la presencia de anticuerpos reactivos con GBS mediante inmunoensayos estándar. Los análisis Elisa y de inmunotinción indicaron que la inmunización con proteína de GBS recombinante producida en *E. coli* provocó anticuerpos reactivos con proteínas de GBS nativas y recombinantes. Las respuestas de anticuerpos a GBS se describen en el Ejemplo 9.

TABLA 5

Capacidad de la proteína de GBS recombinante correspondiente a la SEQ ID NO: 39 para provocar protección frente a la exposición a 8 cepas diversas de GBS

Inmunógeno	Cepa de Exposición			
	Designación	Tipo	Nº vivos	Nº muertos ¹
Proteína rGBS Ninguno	C388/90	Ia/c	8 : 2 0 : 10	(P<0,0001)
Proteína rGBS Ninguno	NCS 246	II/R	10 : 0 3 : 7	(P=0,0012)
Proteína rGBS Ninguno	ATCC 12401	Ib	10 : 0 3 : 7	(P<=0,001)
Proteína rGBS Ninguno	NCS 535	V	10 : 0 5 : 15	(P=0,01)
Proteína rGBS Ninguno	NCS 9842	VI	10 : 0 0 : 10	(P<0,0001)
Proteína rGBS NCS 915-F ³ Ninguno	NCS 915	III	7 : 3 1 : 9 4 : 6	(P=0,0007) ²
Proteína rGBS NCS 954-F Ninguno	NCS 954	III/R	7 : 3 4 : 6 1 : 9	(P=0,002)
Proteína rGBS COH1-F Ninguno	COH1	III	4 : 6 3 : 7 0 : 10	(P=0,0004)
¹ Se usaron grupos de 10 ratones por grupo, se indica el número de ratones que sobrevivieron a la infección y el número de ratones muertos. Las curvas de supervivencia correspondientes a los animales inmunizados con proteína de GBS recombinante se compararon con las curvas de supervivencia correspondiente a los animales inmunizados por simulacro usando el ensayo log-rank para análisis no paramétrico. ² Análisis de comparación con animales inmunizados con NCS915-F. ³ Los animales fueron inmunizados con GBS muerto por formaldehído en presencia de adyuvante QuilA TM .				

Todos los hemocultivos de ratones supervivientes dieron negativo en el 14 después de la exposición. Los cultivos de bazo de los ratones supervivientes dieron negativo excepto para algunos ratones del experimento MB-11.

Ejemplo 9

La vacunación con la proteína de GBS recombinante provoca una respuesta inmune frente a GBS

Se inmunizó grupos de 10 ratones CD-1 hembra subcutáneamente con proteína de GBS recombinante correspondiente a la SEQ ID NO: 39, tal como se describe en el Ejemplo 8. Con el fin de establecer la respuesta de anticuerpos frente a la proteína de GBS nativa, se analizaron los sueros procedentes de muestras de sangre antes de cada inmunización y a los catorce días de la tercera inmunización en busca de anticuerpos reactivos a células de GBS mediante

ELISA usando placas recubiertas con células de GBS muertas por formaldehído de tipo III de cepa NCS 954, de tipo Ib de cepa ATCC12401, de tipo V de cepa NCS 535 ó de tipo VI de cepa NCS 9842. La especificidad de los anticuerpos activados por la proteína de GBS se confirmó mediante análisis Western blot realizados sobre extractos celulares de GBS y sobre antígenos recombinantes purificados. Los resultados mostrados en la Figura 10 demuestran claramente que los animales responden enérgicamente a la proteína de GBS recombinante usada como inmunógeno, con títulos de anticuerpos recíprocos medios que varían entre 12.000 y 128.000, para sueros recogidos después de la tercera inmunización, dependiendo del antígeno que recubre. Todos los sueros preinmunización resultaron negativos cuando fueron evaluados con una dilución de 1 : 100. Los anticuerpos reactivos a GBS eran detectables en los sueros de cada animal después de una única inyección de proteína de GBS recombinante.

Ejemplo 10

Conservación antigénica de la proteína de GBS de la presente invención

Se usaron anticuerpos monoclonales (MAbs) específicos de la proteína de GBS de la presente invención para demostrar que dicho antígeno superficial se produce en todos los GBS, y también que es antigénicamente altamente conservado.

Se usó una colección de 68 elementos aislados de GBS para evaluar la actividad de MAbs específicos de GBS. Dichas cepas se obtuvieron del National Center for Streptococcus, del Provincial Laboratory of Public Health for Northern Alberta, Canadá; del Centre Hospitalier Universitaire de Québec, Pabellón CHUL, Québec, Canadá; de la American Type Culture Collection, EE.UU.; del Laboratoire de Sante Publique du Québec, Canadá; y del Dept. of Infectious Disease, Children's Hospital and Medical Center, Seattle, EE.UU. Se evaluó los ocho MAbs frente al siguiente panel de cepas: 6 elementos aislados de serotipo Ia ó Ia/c, 3 elementos aislados de serotipo Ib, 4 elementos aislados de serotipo II, 14 elementos aislados de serotipo III, 2 elementos aislados de serotipo IV, 2 elementos aislados de serotipo V, 2 elementos aislados de serotipo VI, 2 elementos aislados de serotipo VII, 1 elemento aislado de serotipo VIII, 10 elementos aislados que no estaban serotipados y 3 cepas bovinas de *S. agalactiae*. También se hizo reaccionar MAb 3A2 con GBS adicional: 9 elementos aislados de serotipo Ia/c y 10 elementos aislados de serotipo V. Se cultivaron las cepas una noche sobre placas de agar de sangre a 37°C en una atmósfera con un 5% en CO₂. Los cultivos fueron almacenados a -70°C en un caldo de infusión de corazón con un 20% (v/v) de glicerol.

Para obtener los MAbs específicos de proteína de GBS, se inmunizó ratones tres veces a intervalos de tres semanas con 20 µg de proteína de GBS recombinante purificada (SEQ ID NO: 44), en presencia de adyuvante QuilA™ al 20%. Se generaron líneas celulares de hibridoma mediante fusión de células de bazo recuperadas de los ratones inmunizados con la línea celular de mieloma SP2/0 no secretora tal como se ha descrito previamente (Hamel, J. y col., 1987, *J. Med. Microbiol.* 23: 163-170). Se evaluaron los sobrenadantes de clones híbridos en busca de una producción de anticuerpos específicos mediante ELISA usando GBS desactivado por formaldehído y proteína de GBS recombinante purificada (SEQ ID NO: 39 ó 44) como antígeno cubriente, tal como se ha descrito previamente (Hamel, J. y col., 1987, *J. Med. Microbiol.* 23: 163-170). Los híbridos específicos fueron clonados mediante diluciones limitantes, fueron expandidos y congelados en nitrógeno líquido. La producción de proteína de GBS recombinante se presentó en los Ejemplos 4 y 5. La proteína de GBS recombinante purificada o el GBS desactivado por formaldehído fueron redisoluidos mediante electroforesis usando el sistema tampón discontinuo de Laemmli, tal como recomienda el fabricante, y a continuación se transfirieron a una membrana de nitrocelulosa para realizar la inmunotinción Western tal como se ha descrito previamente (Martin y col., 1992, *Infect. Immun.* 60: 2718-2725).

Los experimentos de inmunotinción de Western claramente indicaron que los ocho MAbs reconocen una banda de proteína que se corresponde con proteína de GBS recombinante purificada (SEQ ID NO: 39). Dichos MAbs también reaccionaron con una banda de proteína presente en todos los elementos aislados de GBS evaluados. La reactividad de estos MAbs específicos de GBS se presenta en la Tabla 6. Cada MAb reaccionó bien con los 46 GBS. Además, estos MAbs también reconocieron las 3 cepas de *S. agalactiae* de origen bovino que fueron evaluadas. El MAb 3A2 también reconoció diecinueve GBS; 9 elementos aislados de serotipo Ia/c y 10 de serotipo V. Los otros MAbs no fueron evaluados frente a estas cepas adicionales.

Estos resultados demuestran que la proteína de GBS (SEQ ID NO: 39) fue producida por todos los 65 GBS, y por las 3 cepas de *S. agalactiae* de origen bovino que fueron evaluadas. Aún más importante, estos resultados claramente demuestran que los epítomos reconocidos por estos ocho MAbs específicos de GBS estaban distribuidos y conservados ampliamente entre los GBS. Estos resultados también indicaron que dichos epítomos no estaban restringidos a elementos aislados relacionados serológicamente, ya que se evaluaron representantes de todos los serotipos de GBS conocidos, incluyendo los grupos principales causantes de enfermedades.

En conclusión, los datos presentados en este Ejemplo demuestran claramente que la proteína de GBS de la presente invención es producida por todos los GBS y que es antigénicamente altamente conservada.

TABLA 6

Reactividad de ocho MAbs específicos de proteína de GBS con diferentes cepas de *S. agalactiae* evaluada mediante inmunotinciones Western

MAbs	Número de cada serotipo de cepas de <i>S. agalactiae</i> reconocidas por los MAbs											
	Ia ó Ia/c (6)	Ib (3)	II (4)	III (4)	IV (2)	V (2)	VI (2)	VII (2)	VIII (1)	NT (10) ²	TOTAL (26)	Bovino (3)
3A2 ¹	6	3	4	4	2	2	2	2	1	10	46	3
5A12	6	3	4	4	2	2	2	2	1	10	46	3
6G11	6	3	4	4	2	2	2	2	1	10	46	3
8B9	6	3	4	4	2	2	2	2	1	10	46	3
8E11	6	3	4	4	2	2	2	2	1	10	46	3
12B12	6	3	4	4	2	2	2	2	1	10	46	3
18F11	6	3	4	4	2	2	2	2	1	10	46	3
20G2	6	3	4	4	2	2	2	2	1	10	46	3
¹ El MAb 3A2 reconoció nueve cepas adicionales de serotipo Ia/c y 10 cepas de serotipo V.												
² Estas cepas no fueron serotipadas.												

REIVINDICACIONES

1. Un polinucleótido aislado que codifica un polipéptido que tiene al menos un 70% de identidad con un segundo polipéptido que tiene una secuencia seleccionada del grupo que consiste en SEQ ID NO: 39 y SEQ ID NO: 44, o un fragmento de dicho polipéptido, en donde dicho polipéptido o fragmento de dicho polipéptido retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.
2. El polinucleótido aislado de acuerdo con la reivindicación 1, que codifica un polipéptido que tiene al menos un 70% de identidad con un segundo polipéptido que tiene una secuencia seleccionada del grupo que consiste en SEQ ID NO: 39 y SEQ ID NO: 44, en donde dicho polipéptido retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.
3. Un polinucleótido de acuerdo con la reivindicación 1 ó 2, en donde dicho polinucleótido codifica un polipéptido que tiene al menos un 95% de identidad con el segundo polipéptido.
4. Un polinucleótido aislado que codifica un polipéptido que tiene al menos un 70% de identidad con un segundo polipéptido capaz de generar anticuerpos que presentan especificidad de unión por un polipéptido que tiene una secuencia seleccionada del grupo que consiste en SEQ ID NO: 39 y SEQ ID NO: 44, o un fragmento de dicho polipéptido, y en donde dicho polipéptido o fragmento de dicho polipéptido retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.
5. Un polinucleótido aislado que codifica un polipéptido que tiene al menos un 70% de identidad con un segundo polipéptido capaz de generar anticuerpos que presentan especificidad de unión por un polipéptido que tiene una secuencia seleccionada del grupo que consiste en SEQ ID NO: 39 y SEQ ID NO: 44, y en donde dichos polipéptidos retienen la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.
6. Un polinucleótido aislado que es complementario al polinucleótido de la reivindicación 1 ó 2.
7. Un polinucleótido aislado que es complementario al polinucleótido de la reivindicación 4 ó 5.
8. El polinucleótido de la reivindicación 1 ó 2, en donde dicho polinucleótido es ADN.
9. El polinucleótido de la reivindicación 4 ó 5, en donde dicho polinucleótido es ADN.
10. El polinucleótido de la reivindicación 1 ó 2, en donde dicho polinucleótido es ARN.
11. El polinucleótido de la reivindicación 4 ó 5, en donde dicho polinucleótido es ARN.
12. Un polinucleótido aislado que se hibrida en condiciones severas con el complemento de un segundo polinucleótido que tiene una secuencia seleccionada del grupo que consiste en SEQ ID NO: 37, SEQ ID NO: 42 y SEQ ID NO: 43, o un fragmento de dicho polinucleótido, en donde dicho polinucleótido tiene al menos un 70% de identidad con dicho segundo polinucleótido y en donde dicho polinucleótido, o fragmento del mismo, codifica un polipéptido que retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.
13. Un polinucleótido aislado que se hibrida en condiciones severas con un segundo polinucleótido que tiene una secuencia seleccionada del grupo que consiste en SEQ ID NO: 37, SEQ ID NO: 42 y SEQ ID NO: 43, en donde dicho polinucleótido tiene al menos un 70% de identidad con dicho segundo polinucleótido y en donde dicho polinucleótido codifica un polipéptido que retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.
14. Un polinucleótido aislado que se hibrida en condiciones severas con el complemento de un segundo polinucleótido que tiene una secuencia seleccionada del grupo que consiste en SEQ ID NO: 37, SEQ ID NO: 42 y SEQ ID NO: 43, en donde dicho polinucleótido tiene al menos un 95% de identidad con dicho segundo polinucleótido y en donde dicho polinucleótido codifica un polipéptido que retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.
15. Un polinucleótido de acuerdo con la reivindicación 14, que se hibrida en condiciones severas con el complemento de un segundo polinucleótido que tiene la secuencia SEQ ID NO: 37.
16. Un polinucleótido de acuerdo con la reivindicación 14, que se hibrida en condiciones severas con el complemento de un segundo polinucleótido que tiene la secuencia SEQ ID NO: 42.
17. Un polinucleótido de acuerdo con la reivindicación 14, que se hibrida en condiciones severas con el complemento de un segundo polinucleótido que tiene la secuencia SEQ ID NO: 43.
18. Un polinucleótido de acuerdo con la reivindicación 12, en donde dicho polinucleótido tiene al menos un 95% de complementariedad con el segundo polinucleótido.

ES 2 278 436 T3

19. Un vector que comprende el polinucleótido de la reivindicación 1 ó 2, en donde dicho polinucleótido está ligado operativamente a una región de control de expresión.

20. Un vector que comprende el polinucleótido de la reivindicación 4 ó 5, en donde dicho polinucleótido está ligado operativamente a una región de control de expresión.

21. Una célula hospedante transfectada con el vector de la reivindicación 19.

22. Una célula hospedante transfectada con el vector de la reivindicación 20.

23. Un proceso para producir un polipéptido que comprende cultivar una célula hospedante de acuerdo con la reivindicación 21 en las condiciones adecuadas para la expresión de dicho polipéptido.

24. Un proceso para producir un polipéptido que comprende cultivar una célula hospedante de acuerdo con la reivindicación 22 en las condiciones adecuadas para la expresión de dicho polipéptido.

25. Un polipéptido aislado que tiene al menos un 70% de identidad con un segundo polipéptido que tiene una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 39 y SEQ ID NO: 44, o un fragmento de dicho polipéptido, y en donde dicho polipéptido o fragmento de dicho polipéptido retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.

26. Un polipéptido aislado que tiene al menos un 70% de identidad con un segundo polipéptido que tiene una secuencia seleccionada del grupo que consiste en SEQ ID NO: 39 y SEQ ID NO: 44, y en donde dicho polipéptido retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.

27. El polipéptido aislado de la reivindicación 25 ó 26 que tiene una secuencia de acuerdo con la SEQ ID NO: 39.

28. El polipéptido aislado de la reivindicación 25 ó 26 que tiene una secuencia de acuerdo con la SEQ ID NO: 44.

29. Un polipéptido aislado que tiene al menos un 70% de identidad con un segundo polipéptido, y que es capaz de generar anticuerpos que presentan especificidad de unión por un polipéptido que tiene una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 39 y SEQ ID NO: 44, o un fragmento de dicho polipéptido, en donde dicho polipéptido o fragmento de dicho polipéptido retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.

30. Un polipéptido aislado que tiene al menos un 70% de identidad con un segundo polipéptido, y que es capaz de generar anticuerpos que presentan especificidad de unión por un polipéptido que tiene una secuencia seleccionada del grupo que consiste en: SEQ ID NO: 39 y SEQ ID NO: 44, en donde dichos polipéptidos retienen la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.

31. El polipéptido aislado de la reivindicación 29 ó 30 que tiene una secuencia de acuerdo con la SEQ ID NO: 39.

32. El polipéptido aislado de la reivindicación 29 ó 30 que tiene una secuencia de acuerdo con la SEQ ID NO: 44.

33. Un polipéptido aislado que tiene una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO: 39, o fragmentos del mismo, y en donde dicho polipéptido o fragmento de dicho polipéptido retiene la capacidad de inducir una respuesta inmune específica de estreptococos del grupo B.

34. Un polipéptido aislado de acuerdo con la reivindicación 33, en donde se elimina el residuo N-terminal Met.

35. Un polipéptido aislado de acuerdo con la reivindicación 33, en donde se elimina la secuencia de aminoácidos secretora.

36. Una composición de vacuna que comprende un polipéptido de acuerdo con una cualquiera de las reivindicaciones 25 a 33 y un vehículo, diluyente o adyuvante farmacéuticamente aceptable.

37. Una composición de vacuna que comprende un polipéptido de acuerdo con la reivindicación 34 y un vehículo, diluyente o adyuvante farmacéuticamente aceptable.

38. Una composición de vacuna que comprende un polipéptido de acuerdo con la reivindicación 35 y un vehículo, diluyente o adyuvante farmacéuticamente aceptable.

39. El uso de un polipéptido de acuerdo con una cualquiera de las reivindicaciones 25 a 33 para la fabricación de una vacuna para el tratamiento terapéutico o profiláctico de infección bacteriana estreptococal de grupo B en un animal susceptible a la infección estreptococal de grupo B.

ES 2 278 436 T3

40. El uso de un polipéptido de acuerdo con la reivindicación 34 para la fabricación de una vacuna para el tratamiento terapéutico o profiláctico de infección bacteriana estreptococal de grupo B en un animal susceptible a la infección estreptococal de grupo B.

5 41. El uso de un polipéptido de acuerdo con la reivindicación 35 para la fabricación de una vacuna para el tratamiento terapéutico o profiláctico de infección bacteriana estreptococal de grupo B en un animal susceptible a la infección estreptococal de grupo B.

10 42. El uso de acuerdo con una cualquiera de las reivindicaciones 39 a 41, en donde dicho animal es un bovino.

 43. El uso de acuerdo con una cualquiera de las reivindicaciones 39 a 41, en donde dicho animal es un humano.

15

20

25

30

35

40

45

50

55

60

65

ES 2 278 436 T3

TATCTGGCAA AGAGCCAGCT AATCGTTTTA GTTGGGCTAA AAATAAATTA TTAATCAATG 60
S G K E P A N R F S W A K N K L L I N G
---->
GATTCAITGC AACTCTAGCA GCAACTATCT TATTTTTTGC AGTTCAATTC ATAGGTCTTA 120
F I A T L A A T I L F F A V Q F I G L K
AACCAGATTA CCCTGGAAAA ACCTACTTTA TTATCCTATT GACAGCATGG ACTTTGATGG 180
P D Y P G K T Y F I I L L T A W T L M A
CATTAGTAAAC TGCTTTAGTG GGATGGGATA ATAGGTATGG TTCCTTCTTG TCGTTATTAA 240
L V T A L V G W D N R Y G S F L S L L I
TATTATTATT CCAGCTTGGT TCAAGCGCAG GAACCTACCC AATAGAATTG AGTCCTAAGT 300
L L F Q L G S S A G T Y P I E L S P K F
TCTTTCAAAC AATTCAACCA TTTTACCAG TGACTTACTC TGTTTCAGGA TTAAGAGAGA 360
F Q T I Q P F L P M T Y S V S G L R E T
CCATCTCGTT GACGGGAGAC GTTAACCATC AATGGAGAAT GCTAGTAATC TTTTATGAT 420
I S L T G D V N H Q W R M L V I F L V S
CATCGATGAT ACTTGCTCTT CTTATTTATC GTAAACAAGA AGATTAATAG AAAGTATCTA 480
S M I L A L L I Y R K Q E D
GTGATAGACT AACAGTATGA TATGGTATGT CAAAGTATTT AGGAGGAGAA GATATGTCTA 540
M S T
|---->
CTTTAACAAT AATTATTGCA ACATTAACCTG CTTTGGAAACA TTTTATATT ATGTATTTGG 600
L T I I I A T L T A L E H F Y I M Y L E
AGACGTTAGC CACCCAGTCA AATATGACTG GGAAGATTTT TAGTATGTCT AAAGAAGAGT 660
T L A T Q S N M T G K I F S M S K E E L
TGTCATATTT ACCCGTTATT AAACCTTTTA AGAATCAAGG TGTATACAAC GCCTTGATTG 720
S Y L P V I K L F K N Q G V Y N G L I G
GCCTATTCCT CTTTATGCGG TTATATATTT CACAGAATCA AGAAATTGTA GCTGTTTTTT 780
L F L L Y G L Y I S Q N Q E I V A V F L
TAATCAATGT ATTGCTAGTT GCTATTTATG GTGCTTTGAC AGTTGATAAA AAAATCTTAT 840
I N V L L V A I Y G A L T V D K K I L L
TAAAACAGGG TGGTTTACCT ATATTAGCTC TTITAACATT CTTATTTTAA TACTACTTAG 900
K Q G G L P I L A L L T F L F
CCGTTGATTT TAGTTGAACG GCTTTTAGTA ATCATTTTTT TCTCATAATA CAGGTAGTTT 960
AAGTAATTTG TCTTTAAAAA TAGTATAATA TAACTACGAA TTCAAAGAGA GGTGACTTTG 1020
ATTATGACTG AGAACTGGTT ACATACTAAA GATGGTTCAG ATATTTATTA TCGTGTCTGT 1080
M T E N W L H T K D G S D I Y Y R V V
|---->
GGTCAAGGTC AACCGATTGT TTTTACAT GGCAATAGCT TAAGTAGTCG CTATTTTGAT 1140
G Q G Q P I V F L H G N S L S S R Y F D
AAGCAAATAG CATATTTTTC TAAGTATTAC CAAGTTATTG TTATGGATAG TAGAGGGCAT 1200
K Q I A Y F S K Y Y Q V I V M D S R G H
GGCAAAAGTC ATGCAAAGCT AAATACCATT AGTTTCAGGC AAATAGCAGT TGACTTAAAG 1260
G K S H A K L N T I S F R Q I A V D L K

ES 2 278 436 T3

GATATCTTAG TTCATTTAGA GATTGATAAA GTTATATTGG TAGGCCATAG CGATGGTGCC 1320
 D I L V H L E I D K V I L V G H S D G A
 AATTTAGCTT TAGTTTTTCA AACGATGTTT CCAGGTATGG TTAGAGGGCT TTTGCTTAAT 1380
 N L A L V F Q T M F P G M V R G L L L N
 TCAGGGAACC TGACTATTCA TGGTCAGCGA TGGTGGGATA TTCTTTTAGT AAGGATTGCC 1440
 S G N L T I H G Q R W W D I L L V R I A
 TATAAATTCC TTCACTATTT AGGGAACTC TTTCCGTATA TGAGGCAAAA AGCTCAAGTT 1500
 Y K F L H Y L G K L F P Y M R Q K A Q V
 ATTTGCTTA TGTTGGAGGA TTTGAAGATT AGTCCAGCTG ATTTACAGCA TGTGTCAACT 1560
 I S L M L E D L K I S P A D L Q H V S T
 CCTGTAATGG TTTTGGTTGG AAATAAGGAC ATAATTAAGT TAAATCATT TAAGAACTT 1620
 P V M V L V G N K D I I K L N H S K K L
 GCTTCTTATT TTCCAAGGGG GGAGTTTTAT TCTTTAGTTG GCTTTGGGCA TCACATTATT 1680
 A S Y F P R G E F Y S L V G F G H H I I
 AAGCAAGATT CCCATGTTTT TAATATTATT GCAAAAAAGT TTATCAACGA TACGTTGAAA 1740
 K Q D S H V F N I I A K K F I N D T L K
 GGAGAAATTG TTGAAAAAGC TAATTGAAAA AGTCAAATCA CTGACTTCTG TGATTAAAT 1800
 G E I V E K A N
 TGTATTTTTT ATATCTGTTT TAGTGCTTAT TATTGTTGAA ATGATTCATT TGAAACGAAC 1860
 M I H L K R T
 |---->
 TATTTCTGTT GAGCAACTAA AGAGTGTTTT TGGGCAATTA TCTCCAATGA ATCTTTTCTT 1920
 I S V E Q L K S V F G Q L S P M N L F L
 AATTATCCTT GTGGGGGTTA TCGCTGTCTT ACCGACAACC GGATATGACT TTGTACTGAA 1980
 I I L V G V I A V L P T T G Y D F V L N
 TGGACTTTTA CGTACAGATA AAAGCAAAAG GTATATTTTA CAGACTAGTT GGTGTATCAA 2040
 G L L R T D K S K R Y I L Q T S W C I N
 CACTTTTAAT AACTGTGTCAG GATTCCGGTGG CTTAATCGAT ATTGGGTTGC GCATGGCTTT 2100
 T F N N L S G F G G L I D I G L R M A F
 TTATGGTAAA AAAGGTCAAG AGAAGAGTGA CCTAAGAGAA GTGACTCGTT TTTTACCCTA 2160
 Y G K K G Q E K S D L R E V T R F L P Y
 TCTTATTTCT GGTCTGTCAT TTATTAGTGT GATTGCCTTA ATCATGAGCC ATATTTTCA 2220
 L I S G L S F I S V I A L I M S H I F H
 TGCCAAAGCT AGTGTGATT ACTATTATTT GGTATTAATT GGTGCTAGTA TGTATTTTCC 2280
 A K A S V D Y Y Y L V L I G A S M Y F P
 TGTTATTTAT TGGATTTCTG GTCATAAAGG AAGCCATTAT TTCGGAGATA TGCCATCTAG 2340
 V I Y W I S G H K G S H Y F G D M P S S
 TACTCGTATA AAATTAGGTG TTGTTTCTTT TTTTGAATGG GGATGTGCGG CCGCAGCATT 2400
 T R I K L G V V S F F E W G C A A A A F
 TATAATTATC GGTTATTTAA TGGGCATTCA TCTACCAGTT TATAAAATTT TACCACTATT 2460
 I I I G Y L M G I H L P V Y K I L P L F

ES 2 278 436 T3

TTGTATTGGT TGTGCCGTCG GGATTGTATC CCTATTCCCG GGTGGATTAG GAAGTTTTGA	2520
C I G C A V G I V S L I P G G L G S F E	
ATTAGTTCTA TTTACAGGGT TTGCTGCCGA GGGACTACCT AAAGAACTG TGGTTGCATG	2580
L V L F T G F A A E G L P K E T V V A W	
GTTATTACTT TATCGTTTAG CCTACTATAT TATTCCATTC TTTGCAGGTA TCTATTTCTT	2640
L L L Y R L A Y Y I I P F F A G I Y F F	
TATCCATTAT TTAGGTAGTC AAATAAATCA ACGTTATGAA AATGTCCCGA AAGAGTTAGT	2700
I H Y L G S Q I N Q R Y E N V P K E L V	
ATCAACTGTT CTACAAACCA TGGTGAGCCA TTTGATGCGT ATTTTAGGTG CATTCTTAAT	2760
S T V L Q T M V S H L M R I L G A F L I	
---->	
ATTTTCAACA GCATTTTTTG AAAATATTAC TTATATTATG TGGTTGCAGA AGCTAGGCTT	2820
F S T A F F E N I T Y I M W L Q K L G L	
GGACCCATTA CAAGAACAAA TGTTATGGCA GTTTCAGGT TTATTGCTGG GGGTTTGT	2880
D P L Q E Q M L W Q F P G L L L G V C F	
TATTCTCTTA GCTAGAACTA TTGATCAAAA AGTGAAAAAT GCTTTTCCAA TTGCTATTAT	2940
I L L A R T I D Q K V K N A F P I A I I	
CTGGATTACT TTGACATTGT TTTATCTTAA TTTAGGTCAT ATTAGTTGGC GACTATCTTT	3000
W I T L T L F Y L N L G H I S W R L S F	
CTGGTTTATT TTACTATTGT TAGGCTTATT AGTCATTAAG CCAACTCTCT ATAAAAACA	3060
W F I L L L L G L L V I K P T L Y K K Q	
ATTTATTTAT AGCTGGGAAG AGCGTATTAA GGATGGAATC ATTATCGTTA GTTTAATGGG	3120
F I Y S W E E R I K D G I I I V S L M G	
AGTTCTATTT TATATTGCAG GACTACTATT CCCTATCAGG GCTCATATTA CAGGTGGTAG	3180
V L F Y I A G L L F P I R A H I T G G S	
TATTGAACGC CTGCATTATA TCATAGCATG GGAGCCGATA GCATTGGCTA CGTTGATTCT	3240
I E R L H Y I I A W E P I A L A T L I L	
TACTCTCGTT TATTTATGTT TGGTTAAGAT TTTACAAGGA AAATCTTGTC AGATTGGTGA	3300
T L V Y L C L V K I L Q G K S C Q I G D	
TGTGTTCAAT GTGGATCGTT ATAAAAAAT ACTTCAAGCT TACG3TGGTT CTTCGGATAG	3360
V F N V D R Y K K L L Q A Y G G S S D S	
CGGTTTAGCC TTTTAAATG ATAAAAGGCT CTACTGGTAC CAAAAAATG GAGAAGATTG	3420
G L A F L N D K R L Y W Y Q K N G E D C	
CGTTGCGTTC CAATTTGTAA TTGTCAATAA TAAATGTCTT ATTATGGGGG AACCAGCCGG	3480
V A F Q F V I V N N K C L I M G E F A G	
TGATGACACT TATATTCGTG AAGCTATTGA ATCGTTTATT GATGATGCTG ATAAGCTAGA	3540
D D T Y I R E A I E S F I D D A D K L D	
CTATGACCTT GTTTTTTACA GTATTGGACA GAAGTTGACA CTACTTTTAC ATGAGTATGG	3600
Y D L V F Y S I G Q K L T L L L H E Y G	
TTTTGACTTT ATGAAAGTTG GTGAGGATGC TTTAGTTAAT TTAGAAACGT TTACTCTTAA	3660
F D F M K V G E D A L V N L E T F T L K	

ES 2 278 436 T3

```

AGGGAATAAG TACAAACCTT TCAGAAATGC CCTAAATAGA GTTGAAAAGG ATGGTTTCTA 3720
G N K Y K P F R N A L N R V E K D G F Y

TTTCGAAGTT GTACAATCGC CACATAGTCA AGAGCTACTA AATAGTTTGG AAGAGATTTC 3780
F E V V Q S P H S Q E L L N S L E E I S

TAATACTTGG TTAGAAGGAC GTCCTGAAAA AGGTTTCTCA CTAGGATATT TTAATAAAGA 3840
N T W L E G R P E K G F S L G Y F N K D

TTATTTCCAA CAAGCCCCAA TAGCTTTGGT AAAAAATGCT GAACACGAAG TTGTTGCTTT 3900
Y F Q Q A P I A L V K N A E H E V V A F

TGCTAATATT ATGCCAAACT ATGAAAAGAG TATTATCTCT ATTGATTTAA TGCCTCACGA 3960
A N I M P N Y E K S I I S I D L M R H D

TAAACAGAAA ATTCCGAATG GCGTTATGGA TTCTCTCTTT TTATCATTAT TCTCTTATTA 4020
K Q K I P N G V M D F L F L S L F S Y Y

TCAAGAGAAG GGATACCACT ATTTTGATTT GGGGATGGCA CCTTTATCAG GAGTTGGTCG 4080
Q E K G Y H Y F D L G M A P L S G V G R

CGTTGAAACA AGTTTTGCTA AAGAGAGAAT GCGGTATCTT GTCTATCATT TCGGTAGTCA 4140
V E T S F A K E R M A Y L V Y H F G S H

TTTCTACTCA TTTAATGGTT TACACAAGTA TAAGAAGAAG TTTACACCAT TGTGGTCGGA 4200
F Y S F N G L H K Y K K K F T P L W S E

ACGTTATATT TCTTGTTCTC GTTCGTCCTG GTTAATTTGT GCTATTTGTG CCCTATTAAT 4260
R Y I S C S R S S W L I C A I C A L L M

GGAAGATAGT AAAATTAAGA TTGTTAAATA AGCTTTATTT GGCAATTAAA AAGAGCATGT 4320
E D S K I K I V K

CATGCGACAT GCTCTTTTTA AATCATTTAA TACCATTGAT TGCTTGAATC TACTTTATAA 4380

TATGATGTGC TTTTAAATAT TGTTTAGCTA CTGTAGCTGC TGATTTATGC TTTACAGCTA 4440

CTTGGTAGTT CATTTCTTGC ATTTCTTTTT CAGTGATATG ACCAGCAAGT TTATTGAGAG 4500

CTTTTTTTAC TTGA (SEQ ID NO:1) 4514

```

FIG. 1a
[clon1-adn/aa]

SGKEPANRFS WAKNKLLING FIATLAATIL FFAVQFIGLK PDYPGKTYFI 50
 ILLTAWTLMA LVTALVGWDN RYGSFLSLLI LLFQLGSSAG TYPIELSPKF 100
 FQTIQPFELPM TYSVSGLRER ISLTGDVNHQ WRMLVIFLVS SMILALLIYR 150
 KQED (SEQ ID NO:2) 154

FIG. 1b

MSTLTIIIIAT LTALEHFYIM YLETLATQSN MTGKIFSMSK EELSYLPVIK 50
 LEKNQGVYNG LIGLFLLYGL YISQNEIVA VFLINVLLVA IYGALTVDKK 100
 ILLKQGGLPI LALLTFLF (SEQ ID NO:3) 118

FIG. 1c

MTENWLHTKD GSDIYYRVVG QGQPIVELHG NSLSSRYFDK QIAYFSKYYQ 50
 VIVMDSRGHG KSHAKLNTIS FRQIAVDLKD ILVHLEIDKV ILVGHSDGAN 100
 LALVFQTMFP GMVRGLLLNS GNLTIHGQRW WDILLVRIAY KFLHYLGKLF 150
 PYMRQKAQVI SLMLEDLKIS PADLQHVSTP VMVLVGNKDI IKLNHSHKLA 200
 SYFPRGEFYS LVGFGHHIIK QDSHVFNIIA KKFINDTLKG EIVEKAN 247
 (SEQ ID NO:4)

FIG. 1d

ES 2 278 436 T3

MIHLKRTISV	EQLKSVFGQL	SPMNLFLIIL	VGVI AVLPTT	GYDFVLNGLL	50
RTDKSKRYIL	QTSWCINTFN	NLSGFGGLID	IGLRMAFYGK	KGQEKSDLRE	100
VTRFLPYLIS	GLSFISVIAL	IMSHIFHAKA	SVDY YYLVLI	GASMYFPVIY	150
WISGHKGSHY	FGDMPSSTRI	KLGVVSFFEW	GCAAAAFIII	GYLMGIHLPV	200
YKILPLFCIG	CAVGIVSLIP	GGLGSFEIVL	FTGFAAEGLP	KETVVAWLLL	250
YRLAYYIIPF	FAGIYFFIHY	LGSQINQRYE	NVPKELVSTV	LQTMVSHLMR	300
ILGAFLIFST	AFFENITYIM	WLQKLGLDPL	QEOMLWQFPG	LLLGVCFILL	350
ARTIDQKVKN	AFPIAIIWIT	LTLFYLN LGH	ISWRLSFWFI	LLLLGLLVIK	400
PTLYKKQFIY	SWEERIKDGI	IIVSLMGVLF	YIAGLLFPIR	AHITGGSIER	450
LHYIIAWEPI	ALATLILTLV	YLCLVKILQG	KSCQIGDVFN	VDRYKKLLQA	500
YGGSSDSGLA	FLNDKRLYWY	QKNGEDCVAF	QFVIVNNKCL	IMGEPAGDDT	550
YIREAIESFI	DDADKLDYDL	VFYSIGQKLT	LLLHEYGFDF	MKVGEDALVN	600
LETFTLKG NK	YKPFERNALNR	VEKDGIFYFEV	VQSPHSQELL	NSLEEISNTW	650
LEGRPEKGFS	LGYFNKD YFQ	QAPIALVKNA	EHEVVAFANI	MPNYEKSIIS	700
IDLMRHDKQK	IPNGVMDFLF	LSLFSYYQEK	GYHYFDLGMA	PLSGVGRVET	750
SFAKERMAYL	VYHFGSHFYS	FNGLHKYKKK	FTPLWSERYI	SCSRSSWLIC	800
AICALLMEDS	KIKIVK	(SEQ ID NO:5)			816

FIG. 1e

MRILGAFLIF	STAFFENITY	IMWLQKLGLD	PLQEOMLWQF	PGLLLGVCFI	50
LLARTIDQKV	KNAFPIAIIW	ITLTLFYLN L	GHISWRLSFW	FILLLLGLLV	100
IKPTLYKKQF	IYSWEERIKD	GIIIVSLMGV	LFYIAGLLFP	IRAHITGGS I	150
ERLHYIIAWE	PIALATLILT	LVYLCLVKIL	QGKSCQIGDV	FNVDRYKKLL	200
QAYGGSSDSG	LAFLNDKRLY	WYQKNGEDCV	AFQFVIVNNK	CLIMGEPAGD	250
DTYIREAIES	FIDDADKLDY	DLVFYSIGQK	LTLLLHEYGF	DFMKVGEDAL	300
VNLETFTLKG	NKYKPFERNAL	NRVEKDGIFY	EVVQSPHSQE	LLNSLEEISN	350
TWLEGRPEKG	FSLGYFNKDY	FQQAPIALVK	NAEHEVVAFA	NIMPNYEKSI	400
ISIDLMRHDK	QKIPNGVMDF	LFLSLFSYYQ	EKG YHYFDLG	MAPLSGVGRV	450
ETSFAKERMA	YLVYHFGSHF	YSFNGLHKYK	KKFTPLWSE R	YISCSRSSWL	500
ICAICALLME	DSKIKIVK	(SEQ ID NO:6)			518

FIG. 1f

ES 2 278 436 T3

AATTTTGATA TCGAAACAAC AACTTTTGAG GCAATGAAAA AGCACGCGTC ATTATTGGAG 60
 N F D I E T T T F E A M K K H A S L L E
 ---->
 AAAATATCTG TTGAGCGTTC TTTTATTGAA TTTGATAAAC TTCTATTAGC ACCTTATTGG 120
 K I S V E R S F I E F D K L L L A P Y W
 CGTAAAGGAA TGCTGGCACT AATAGATAGT CATGCTTTTA ATTATCTACC ATGCTTAAAA 180
 R K G M L A L I D S H A F N Y L P C L K
 AATAGGGAAT TACAATTAAG CGCCTTTTTG TCCCAGTTAG ATAAAGATTT TTTATTGAG 240
 N R E L Q L S A F L S Q L D K D F L F E
 ACATCAGAAC AAGCTTGGGC ATCACTCATC TTGAGTATGG AAGTTGAACA CACAAAGACT 300
 T S E Q A W A S L I L S M E V E H T K T
 TTTTAAAAA AATGGAAGAC ATCAACTCAC TTTCAAAAAG ATGTTGAGCA TATAGTGGAT 360
 F L K K W K T S T H F Q K D V E H I V D
 GTTTATCGTA TTCGTGAACA AATGGGATTG GCTAAAGAAC ATCTTTATCG TTATGGAAAA 420
 V Y R I R E Q M G L A K E H L Y R Y G K
 ACTATAATAA AACAAGCGGA AGGTATTTCG AAAGCAAGAG GCTTGATGGT TGATTTCGAA 480
 T I I K Q A E G I R K A R G L M V D F E
 AAAATAGAAC AACTAGATAG TGAGTTAGCA ATCCATGATA GGCATGAGAT AGTTGTCAAT 540
 K I E Q L D S E L A I H D R H E I V V N
 GGTGGCACCT TAATCAAGAA ATTAGGAATA AAACCTGGTC CACAGATGGG AGATATTATC 600
 G G T L I K K L G I K P G P Q M G D I I
 TCTCAAATTG AATTAGCCAT TGTTTTAGGA CAACTGATTA ATGAAGAAGA GGCTATTTTA 660
 S Q I E L A I V L G Q L I N E E E A I L
 CATTTTGTTA AGCAGTACTT GATGGATTAG AGAGGATTAT ATGAGCGATT TTTTAGTAGA 720
 H F V K Q Y L M D M S D F L V D
 |----->
 TGGATTGACT AAGTCGGTTG GTGATAAGAC GGTCTTTAGT AATGTTTCAT TTATCATCCA 780
 G L T K S V G D K T V F S N V S F I I H
 TAGTTTAGAC CGTATTGGGA TTATTGGTGT CAATGGAAC TGGAAAGACAA CACTATTAGA 840
 S L D R I G I I G V N G T G K T T L L D
 TGTTATTTTCG GGTGAATTAG GTTTTGATGG TGATCGTTCC CCTTTTTCAT CAGCTAATGA 900
 V I S G E L G F D G D R S P F S S A N D
 TTATAAGATT GCTTATTTAA AACAAGAACC AGACTTTGAT GATCTCAGA CAATTTTGGA 960
 Y K I A Y L K Q E P D F D D S Q T I L D
 CACCGTACTT TCTTCTGACT TAAGAGAGAT GGCTTTAATT AAAGAATATG AATTATTGCT 1020
 T V L S S D L R E M A L I K E Y E L L L
 TAATCACTAC GAAGAAAGTA AGCAATCAG TCTAGAGAAA GTAATGGCAG AAATGGATTC 1080
 N H Y E E S K Q S R L E K V M A E M D S
 TTTAGATGCT TGGTCTATTG AGAGCGAAGT CAAAACAGTA TTATCCAAAT TAGGTATTAC 1140
 L D A W S I E S E V K T V L S K L G I T
 TGATTTGCAG TTGTCGGTTG GTGAATTATC AGGAGGATTA CGAAGACGTG TTCAATTAGC 1200
 D L Q L S V G E L S G G L R R R V Q L A

ES 2 278 436 T3

GCAAGTATTA TTAAATGATG CAGATTTATT GCTCTTAGAC GAACCTACTA ACCACTTAGA	1260
Q V L L N D A D L L L L D E P T N H L D	
TATTGACACT ATTGCATGGT TAACGAATTT TTTGAAAAAT AGTAAAAAGA CAGTGCTTTT	1320
I D T I A W L T N F L K N S K K T V L F	
TATAACTCAT GATCGTTATT TTCTAGACAA TGTTGCAACA CGTATTTTTG AATTAGATAA	1380
I T H D R Y F L D N V A T R I F E L D K	
GGCACAGATT ACAGAATATC AAGGCAATTA TCAGGATTAT GTCCGACTTC GTGCAGAACA	1440
A Q I T E Y Q G N Y Q D Y V R L R A E Q	
AGACGAGCGT GATGCTGCTA GTTTACATRA AAAGAAACAG CTTTATAAAC AGGAACTAGC	1500
D E R D A A S L H K K K Q L Y K Q E L A	
TTGGATGCGT ACTCAGCCAC AAGCTCGTGC AACGAAACAA CAGGTCGTA TTAATCGTTT	1560
W M R T Q P Q A R A T K Q Q A R I N R F	
TCAAATCTA AAAAACGATT TACACCAAAC AAGCGATACA AGCGATTGG AAATGACATT	1620
Q N L K N D L H Q T S D T S D L E M T F	
TGAAACAAGT CGAATTGGGA AAAAGGTTAT TAATTTTGAA AATGTCTCTT TTTCTTACCC	1680
E T S R I G K K V I N F E N V S F S Y P	
AGATAAATCT ATCTTGAAAG ACTTTAATTT GTTAATTCAA AATAAGACC GTATTGGCAT	1740
D K S I L K D F N L L I Q N K D R I G I	
CGTTGGAGAT AATGGTGTG GAAAGTCAAC CTTACTTAAT TTAATTGTTT AAGATTTACA	1800
V G D N G V G K S T L L N L I V Q D L Q	
GCCGGATTCT GGTAAATGTCT CTATTGGTGA AACGATACGT GTAGGTTACT TTTCACAACA	1860
P D S G N V S I G E T I R V G Y F S Q Q	
ACTTCATAAT ATGGATGGCT CAAAACGTGT TATTAATTAT TTGCAAGAGG TTGCAGATGA	1920
L H N M D G S K R V I N Y L Q E V A D E	
GGTTAAACT AGTGTCGGTA CAACAAGTGT GACAGAACTA TTGGAACAAT TTCTCTTTCC	1980
V K T S V G T T S V T E L L E Q F L F P	
ACGTTGACA CATGGAACAC AAATTGCAAA ATTATCAGGT GGTGAGAAAA AAAGACTTTA	2040
R S T H G T Q I A K L S G G E K K R L Y	
CCTTTTAAAA ATCCTGATTG AAAAGCCTAA TGTGTTACTA CTTGATGAGC CGACAAATGA	2100
L L K I L I E K P N V L L L D E P T N D	
CTTAGATATT GCTACATTAA CTGTTCTTGA AAATTTTSTA CAAGGCTTTG GTGGTCTGT	2160
L D I A T L T V L E N F L Q G F G G F V	
GATTACAGTT AGTCACGATC GTTACTTTTT AGATAAAGTG GCTAATAAAA TTATTGCGTT	2220
I T V S H D R Y F L D K V A N K I I A F	
TGAAGATAAC GATATCCGTG AATTTTTTGG TAATTATACT GATTATTTAG ATGAAAAAGC	2280
E D N D I R E F F G N Y T D Y L D E K A	
ATTTAATGAG CAAAATAATG AAGTTATCAG TAAAAAGAG AGTACCAAGA CAAGTCGTGA	2340
F N E Q N N E V I S K K E S T K T S R E	
AAAGCAAAGT CGTAAAAGAA TGTCTTACTT TGAAAAACAA GAATGGGCGA CAATTGAAGA	2400
K Q S R K R M S Y F E K Q E W A T I E D	
CGATATTATG ATATTGAAAA ATACTATCAC TCGTATAGAA AATGATATGC AAACATGTGG	2460

ES 2 278 436 T3

D I M I L E N T I T R I E N D M Q T C G
 TAGTGATTTT ACAAGGTTAT CTGATTTACA AAAGGAATTA GATGCAAAAA ATGAAGCACT 2520
 S D F T R L S D L Q K E L D A K N E A L
 TCTAGAAAAG TATGACCGTT ATGAGTACCT TAGTGAGTTA GACACATGAT TATCCGTCCG 2580
 L E K Y D R Y E Y L S E L D T M I I R P
 |---->
 ATTATTAAAA ATGATGACCA AGCAGTTGCA CAATTAATTC GACAAAGTTT ACGCGCCTAT 2640
 I I K N D D Q A V A Q L I R Q S L R A Y
 GATTTAGATA AACCTGATAC AGCATATTCA GACCCTCACT TAGATCATTG GACCTCATAC 2700
 D L D K P D T A Y S D P H L D H L T S Y
 TACGAAAAAA TAGAGAAGTC AGGATTCTTT GTCATTGAGG AGAGAGATGA GATTATTGGC 2760
 Y E K I E K S G F F V I E E R D E I I G
 TGTGGCGGCT TTGGTCCGCT GAAAAATCTA ATTGCAGAGA TGCAGAAGGT GTACATTGCA 2820
 C G G F G P L K N L I A E M Q K V Y I A
 GAACGTTTCC GTGGTAAGGG GCTTGCTACT GATTTAGTGA AAATGATTGA AGTAGAAGCT 2880
 E R F R G K G L A T D L V K M I E V E A
 CGAAAAATG GGTATAGACA ACTTTATTTA GAGACAGCCA GTACTTTGAG TAGGGCAACT 2940
 R K I G Y R Q L Y L E T A S T L S R A T
 GCGGTTTATA AGCATATGGG ATATTGTGCC TTATCGCAAC CAATAGCAAA TGATCAAGGT 3000
 A V Y K H M G Y C A L S Q P I A N D Q G
 CATAAGCTA TGGATATTG GATGATTAAA GATTTATAAG TTGAAAGTGG ATTAGTGAAC 3060
 H T A M D I W M I K D L
 ATGGATTAAT TATTTTGAGA TAAGAGCAAA GAAAAGGAGA CATATATGGC ATATATTTGG 3120
 M A Y I W
 |---->
 TCTTATTTGA AAAGGTACCC CAATTGGTTA TGGCTTGATT TACTAGGAGC TATGCTTTTT 3180
 S Y L K R Y P N W L W L D L L G A M L F
 GTGACGGTTA TCCTAGGAAT GCCCAGAGCC TTAGCGGGTA TGATTGATAA TGGCGTTACA 3240
 V T V I L G M P T A L A G M I D N G V T
 AAAGGTGATC GGAAGTGACT TTATCTGTGG ACGTTCATCA TGTTTATATT TGTTGTACTA 3300
 K G D R T G V Y L W T F I M F I F V V L
 GGTATTATTG GGCATATTAC GATGGCTTAC GCATCTAGTC GCTTAACGAC AACAAATGATT 3360
 G I I G R I T M A Y A S S R L T T T M I
 AGAGATATGC GTAATGATAT GTATGCTAAG CTTCAAGAAT ACTCCCATCA TGAATATGAA 3420
 R D M R N D M Y A K L Q E Y S H H E Y E
 CAGATAGGTG TATCTTCACT AGTGACACGT ATGACAAGCG ATACTTTTGT TTTGATGCAA 3480
 Q I G V S S L V T R M T S D T F V L M Q
 TTTGCTGAAA TGTCTTTACG TTTAGGCCTA GTAACCTCCTA TGGTAATGAT TTTTAGCGTG 3540
 F A E M S L R L G L V T P M V M I F S V
 GTTATGATAC TAATTACGAG TCCATCTTTG GCTTGGCTTG TAGCGGTTGC GATGCCTCTT 3600
 V M I L I T S P S L A W L V A V A M P L
 TTGGTAGGAG TCGTTTTATA TGTAGCTATA AAAACAAAAC CTTTATCTGA AAGACAACAG 3660
 L V G V V L Y V A I K T K P L S E R Q Q

ES 2 278 436 T3

ACTATGCTTG ATAAAAATCAA TCAATATGTT CGTGAAAATT TAACAGGGTT ACGCGTTGTT	3720
T M L D K I N Q Y V R E N L T G L R V V	
AGAGCCTTTG CAAGAGAGAA TTTTCAATCA CAAAAATTC AAGTCGCTAA CCAACGTTAC	3780
R A F A R E N F Q S Q K F Q V A N Q R Y	
ACAGATACTT CAACTGGTCT TTTTAAATTA ACAGGGCTAA CAGAACCACT TTTCGTTCAA	3840
T D T S T G L F K L T G L T E P L F V Q	
ATTATTATTG CAATGATTGT GGCTATCGTT TGGTTTGCTT TGGATCCCTT ACAAAGAGGT	3900
I I I A M I V A I V W F A L D P L Q R G	
GCTATTAAAA TAGGGGATTT AGTTGCTTTT ATCGAATATA GCTTCCATGC TCTCTTTTCA	3960
A I K I G O L V A F I E Y S F H A L F S	
TTTTTGCTAT TTGCCAATCT TTTTACTATG TATCCTCGTA TGGTGGTATC AAGCCATCGT	4020
F L L F A N L F T M Y P R M V V S S H R	
ATTAGAGAGG TGATGGATAT GCCAATCTCT ATCAATCCTA ATGCCGAAGG TGTTACGGAT	4080
I R E V M D M P I S I N P N A E G V T D	
ACGAAACTTA AAGGGCATTT AGAATTTGAT AATGTAACAT TCGCTTATCC AGGAGAAACA	4140
T K L K G H L E F D N V T F A Y P G E T	
GAGAGTCCCG TTTTGCATGA TATTTCTTTT AAAGCTAAGC CTGGAGAAAC AATTGCTTTT	4200
E S P V L H D I S F K A K P G E T I A F	
ATTGGTTCAA CAGGTTCAAG AAAATCTTCT CTTGTTAATT TGATTCCACG TTTTATGAT	4260
I G S T G S G K S S L V N L I P R F Y D	
GTGACACTTG GAAAAATCTT AGTAGATGGA GTTGATGTAA GAGATTATAA CCTTAAATCA	4320
V T L G K I L V D G V D V R D Y N L K S	
CTTCGCCAAA AGATTGGATT TATCCCCCAA AAAGCTCTTT TATTTACAGG GACAATAGGA	4380
L R Q K I G F I P Q K A L L F T G T I G	
GAGAATTTAA AATATGGAAG AGCTGATGCT ACTATTGATG ATCTTAGACA AGCGGTTGAT	4440
E N L K Y G K A D A T I D D L R Q A V D	
ATTTCTCAAG CTAAAGAGTT TATTGAGAGT CACCAAGAAG CCTTTGAAAC GCATTTAGCT	4500
I S Q A K E F I E S H Q E A F E T H L A	
GAAGGTGGGA GCAATCTTTC TGGGGGTCAA AAACAACGGT TATCTATTGC TAGGGCTGTT	4560
E G G S N L S G G Q K Q R L S I A R A V	
GTTAAAGATC CAGATTTATA TATTTTGTAT GATTCATTTT CTGCTCTCGA TTATAAGACA	4620
V K D P D L Y I F D D S F S A L D Y K T	
GACGCTACTT TAAGAGCGCG TCTAAAAGAA GTAACCGGTG ATTCTACAGT TTTGATAGTT	4680
D A T L R A R L K E V T G D S T V L I V	
GCTCAAAGGG TGGGTACGAT TATGGATGCT GATCAGATTA TTGTCCTTGA TGAAGGCGAA	4740
A Q R V G T I M D A D Q I I V L D E G E	
ATGTGCGGTC GTGGTACCCA CGCTCAATTA ATAGAAAATA ATGCTATTTA TCGTGAAATC	4800
I V G R G T H A Q L I E N N A I Y R E I	
GCTGAGTCAC AACTGAAGAA CCAAACTTA TCAGAAGGAG AGTGATTGTA TGAGAAAAAA	4860
A E S Q L K N Q N L S E G E M R K K	
---->	

ATCTGTTTTT TTGAGATTAT GGTCTTACCT AACTCGCTAC AAAGCTACTC TTTTCTTAGC	4920
S V F L R L W S Y L T R Y K A T L F L A	
GATTTTTTTG AAAGTTTTAT CTAGTTTTAT GAGTGTCTG GAGCCTTTTA TTTAGGGTT	4980
I F L K V L S S F M S V L E P F I L G L	
AGCGATAACA GAGTTGACTG CTAACCTTGT TGATATGGCT AAGGGAGTTT CTGGGGCAGA	5040
A I T E L T A N L V D M A K G V S G A E	
ATTGAACGTT CCTTATATTG CTGGTATTTT GATTATTTAT TTTTCAGAG GTGTTTCTA	5100
L N V P Y I A G I L I I Y F F R G V F Y	
TGAATTAGGT TCTTATGGCT CAAATT (SEQ ID NO:7)	5126
E L G S Y G S N	

FIG. 2a

ES 2 278 436 T3

NFDIETTTFE AMKKHASLLE KISVERSFIE FDKLLLAPYW RKGMLALIDS	50
HAFNYLPCLK NRELQLSAFL SQLDKDFLFE TSEQAWASLI LSMEVEHTKT	100
FLKKWKTSTH FQKDVEHIVD VYRIREQMGL AKEHLYRYGK TTIKQAEGR	150
KARGLMVDFE KIEQLDSELA IHDRHEIVVN GGTLIKLLGI KPGPQMGDII	200
SQIELAIVLG QLINEEEAIL HFVKQYLM (SEQ ID NO:8)	229

FIG. 2b

MSDFLVDGLT KSVGDKTVFS NVSFIIHSLD RIGIIGVNGT GKTTLDDVIS	50
GELGFDGDRS PFSSANDYKI AYLKQEPDFD DSQTILDTVL SSDLREMAI	100
KEYELLLNHY EESKQSRLEK VMAEMDSLDA WSIESEVKT VLSKLGITDLQ	150
LSVGELSGGL RRRVQLAQVL LNDADLLLLD EPTNHLIDIT IAWLTNFLKN	200
SKKTVLFITH DRYFLDNVAT RIFELDKAQI TEYQGNQDY VRLRAEQDER	250
DAASLHKKKQ LYKQELAWMR TQPQARATKQ QARINRFQNL KNDLHQTSOT	300
SDLEMTFETS RIGKKVINFE NVSFSYPDKS ILKDFNLLIQ NKDRIGIVGD	350
NGVGKSTLLN LIVQDLQPDS GNVSIGETIR VGYFSQQLHN MDGSKRVINY	400
LQEVADDEVKT SVGTTSVTEL LEQFLFPRST HGTQIAKLSG GEKKRLLYLLK	450
ILIEKPNVLL LDEPTNDLDI ATLTVLENEL QGFGGPVITV SHDRYFLDKV	500
ANKIIAFEDN DIREFFGNYT DYLDEKAFNE QNNEVISKKE STKTSREKQS	550
RKRMSYFEKQ EWATIEDDIM ILENTITRIE NDMQTCGSDF TRLSDLQKEL	600
DAKNEALLEK YDRYEYLSL DT (SEQ ID NO:9)	622

FIG. 2c

MIIRPIIKND DQAVAQLIRQ SLRAYDLKP DTAYSDFHLD HLTSYYEKIE	50
KSGFFVIEER DEIIGCGGFG PLKNLIAEMQ KVIYAERFRG KGLATDLVKM	100
IEVEARKIGY RQLYLETAST LSRATAVYKH MGYCALSQPI ANDQGHAMT	150
IWMIKDL (SEQ ID NO:10)	157

FIG. 2d

MAYIWSYLKR	YPNWLWLDLL	GAMLEFVTVIL	GMPTALAGMI	DNGVTKGDRT	50
GVYLWTFIMF	IFVVLGIIGR	ITMAYASSRL	TTTMIRDMRN	DMYAKLQEYS	100
HHEYEQIGVS	SLVTRMTSDT	FVLMQFAEMS	LRLGLVTPMV	MIFSVVMILI	150
TSPSLAWLVA	VAMPLLVGCV	LYVAIKTKPL	SERQQTMLDK	INQYVRENLT	200
GLRVVRAFAR	ENFQSQKFQV	ANQRYTDTST	GLFKLTGLTE	PLFVQIIIAM	250
IVAIVWFALD	PLQRGAIKIG	DLVAFIEYSF	HALFSFLLFA	NLFTMYPRMV	300
VSSHRIREVM	DMPISINPNA	EGVTDTKLKG	HLEFDNVTFE	YPGETESPVL	350
HDISFKAKPG	ETIAFIGSTG	SGKSSLVNLI	PRFYDVTLGK	ILVDGVDVRD	400
YNLKSLRQKI	GFIPQKALLF	TGTIGENLKY	GKADATIDDL	RQAVDISQAK	450
EFIESHQEAF	ETHLAEGGSN	LSGGQKQRLS	IARAVVKDPD	LYIFDDSFSA	500
LDYKTDATLR	ARLKEVTGDS	TVLIVAQRVG	TIMDADQIIV	LDEGEIVGRG	550
THAQLIENNA	IYREIAESQL	KNQNLSEGE	(SEQ ID NO:11)		579

FIG. 2e

MRKKSFLRL	WSYLTRYKAT	LFLAIFLKV	SSFMSVLEPF	ILGLAITELT	50
ANLVDMAGV	SGAELNVPYI	AGILIIYFFR	GVFYELGSYG	SN	92
(SEQ ID NO:12)					

FIG. 2f

ES 2 278 436 T3

```

AATTTGGAAG TGCTCTATCA ACAGTTGAAG TAAAGGAGAT TATTAGTGAA GAAAACATAT 60
  F G S  A L S  T V E V  K E I  I S E  E N I W
  ---->
GGTTATATCG GCTCAGTTGC TGCCATTTTA CTAGCTACTC ATATTGGAAG TTACCAACTT 120
  L Y R  L S C  C H F T  S Y S  Y W K  L P T W

GGTAAGCATC ATATGGGTCT AGCAACAAAG GACAATCAGA TTGCCTATAT TGATGACAGC 180
      M G L  A T K  D N Q I  A Y I  D D S
      |---->
AAAGGTAAGG CAAAAGCCCC TAAAACAAAC AAAACGATGG ATCAAATCAG TGCTGAAGAA 240
  K G K A  K A P  K T N  K T M D  Q I S  A E E

GGCATCTCTG CTGAACAGAT CGTAGTCAAA ATTACTGACC AAGGCTATGT GACCTCACAC 300
  G I S A  E Q I  V V K  I T D Q  G Y V  T S H

GGTGACCATT ATCATTTTTA CAATGGGAAA GTTCCTTATG ATGCGATTAT TAGTGAAGAG 360
  G D H Y  H F Y  N G K  V P Y D  A I I  S E E

TTGTTGATGA CGGATCCTAA TTACCGTTTT AAACAATCAG ACGTTATCAA TGAAATCTTA 420
  L L M T  D P N  Y R F  K Q S D  V I N  E I L
  |---->
GACGGTTACG TTATTAAAGT CAATGGCAAC TATTATGTTT ACCTCAAGCC AGGTAGTAAG 480
  D G Y V  I K V  N G N  Y Y V Y  L K P  G S K

CGCAAAAACA TTCGAACCAA ACAACAAATT GCTGAGCAAG TAGCCAAAGG AACTAAAGAA 540
  R K N I  R T K  Q Q I  A E Q V  A K G  T K E

GCTAAAGAAA AAGGTTTAGC TCAAGTGGCC CATCTCAGTA AAGAAGAAGT TGCGGCAGTC 600
  A K E K  G L A  Q V A  H L S K  E E V  A A V

AATGAAGCAA AAAGACAAGG ACGCTATACT ACAGACGATG GCTATATTTT TAGTCCGACA 660
  N E A K  R Q G  R Y T  T D D G  Y I F  S P T

GATATCATTG ATGATTTAGG AGATGCTTAT TTAGTACCTC ATGGTAATCA CTATCATTAT 720
  D I I D  D L G  D A Y  L V P H  G N H  Y H Y

ATTCTTAAAA AGGATTTGTC TCCAAGTGAG CTAGCTGCTG CACAAGCCTA CTGGAGTCAA 780
  I P K K  D L S  P S E  L A A A  Q A Y  W S Q

AAACAAGGTC GAGGTGCTAG ACCGTCTGAT TACCGCCCGA CACCAGCCCC AGGTCGTAGG 840
  K Q G R  G A R  P S D  Y R P T  P A P  G R R

AAAGCCCCAA TTCCTGATGT GACGCCTAAC CCTGGACAAG GTCATCAGCC AGATAACGGT 900
  K A P I  P D V  T P N  P G Q G  H Q P  D N G

GGCTATCATC CAGCGCCTCC TAGGCCAAAT GATGCGTCAC AAAACAAACA CCAAAGAGAT 960
  G Y H P  A P P  R P N  D A S Q  N K H  Q R D

GAGTTTAAAG GAAAAACCTT TAAGGAACTT TTAGATCAAC TACACCGTCT TGATTTGAAA 1020
  E F K G  K T F  K E L  L D Q L  H R L  D L K

TACCGTCATG TGGAAGAAGA TGGGTTGATT TTTGAACCGA CTCAAGTGAT CAAATCAAAC 1080
  Y R H V  E E D  G L I  F E P T  Q V I  K S N

GCTTTTGGGT ATGTGGTGCC TCATGGAGAT CATTATCATA TTATCCCAAG AAGTCAGTTA 1140
  A F G Y  V V P  H G D  H Y H I  I P R  S Q L

TCACCTCTTG AAATGGAATT AGCAGATCGA TACTTAGCTG GCCAAACTGA GGACAATGAC 1200
  S P L E  M E L  A D R  Y L A G  Q T E  D N D

TCAGGTTTCA AGCACTCAAA ACCATCAGAT AAAGAAGTGA CACATACCTT TCTTGGTCAT 1260

```

ES 2 278 436 T3

S G S E H S K P S D K E V T H T F L G H
 CGCATCAAAG CTTACGGAAA AGGCTTAGAT GGTAAACCAT ATGATACGAG TGATGCTTAT 1320
 R I K A Y G K G L D G K P Y D T S D A Y
 GTTTTtagTA AAGAATCCAT TCATTcAGTG GATAAATCAG GAGTTACAGC TAAACACGGA 1380
 V F S K E S I H S V D K S G V T A K H G
 GATCATTTCC ACTATATAGG ATTTGGAGAA CTTGAACAAT ATGAGTTGGA TGAGGTCGCT 1440
 D H F H Y I G F G E L E Q Y E L D E V A
 AACTGGGTGA AAGCAAAAGG TCAAGCTGAT GAGCTTGCTG CTGCTTTGGA TCAGGAACAA 1500
 N W V K A K G Q A D E L A A A L D Q E Q
 GGCAAAGAAA AACCCTCTT TGACACTAAA AAAGTGAGTC GCAAAGTAAC AAAAGATGGT 1560
 G K E K P L F D T K K V S R K V T K D G
 AAAGTGGGCT ATATGATGCC AAAAGATGGT AAGGACTATT TCTATGCTCG TGATCAACTT 1620
 K V G Y M M P K D G K D Y F Y A R D Q L
 GATTTGACTC AGATTGCCTT TGCCGAACAA GAACCTAATGC TTAAAGATAA GAAGCATTAC 1680
 D L T Q I A F A E Q E L M L K D K K H Y
 CGTTATGACA TTGTTGACAC AGGTATTGAG CCACGACTTG CTGTAGATGT GTCAAGTCTG 1740
 R Y D I V D T G I E P R L A V D V S S L
 CCGATGCATG CTGGTAATGC TACTTACGAT ACTGGAAGTT CGTTTGTTAT CCCACATATT 1800
 P M H A G N A T Y D T G S S F V I P H I
 GATCATATCC ATGTCGTTCC GTATTCATGG TTGACGCGCG ATCAGATTGC AACAGTCAAG 1860
 D H I H V V P Y S W L T R D Q I A T V K
 TATGTGATGC AACACCCCGA AGTTCGTCCG GATGTATGGT CTAAGCCAGG GCATGAAGAG 1920
 Y V M Q H P E V R P D V W S K P G H E E
 TCAGGTTCCG TCATTCCAAA TGTTACGCCT CTTGATAAAC GTGCTGGTAT GCCAACTGG 1980
 S G S V I P N V T P L D K R A G M P N W
 CAAATTATCC ATTCTGCTGA AGAAGTTCAA AAAGCCCTAG CAGAAGGTCG TTTTGCAACA 2040
 Q I I H S A E E V Q K A L A E G R F A T
 CCAGACGGCT ATATTTTCGA TCCACGAGAT GTTTTGCCA AAGAACTTT TGTATGGAAA 2100
 P D G Y I F D P R D V L A K E T F V W K
 GATGGCTCCT TTAGCATCCC AAGAGCAGAT GGCAGTTCAT TGAGAACCAT TAATAAATCT 2160
 D G S F S I P R A D G S S L R T I N K S
 GATCTATCCC AAGCTGAGTG GCAACAAGCT CAAGAGTTAT TGGCAAAGAA AAATACTGGT 2220
 D L S Q A E W Q Q A Q E L L A K K N T G
 GATGCTACTG ATACGGATAA ACCCAAAGAA AAGCAACAGG CAGATAAGAG CAATGAAAAC 2280
 D A T D T D K P K E K Q Q A D K S N E N
 CAACAGCCAA GTGAAGCCAG TAAAGAAGAA AAAGAATCAG ATGACTTTAT AGACAGTTTA 2340
 Q Q P S E A S K E E K E S D D F I D S L
 CCAGACTATG GTCTAGATAG AGCAACCCTA GAAGATCATA TCAATCAATT AGCACAAAAA 2400
 P D Y G L D R A T L E D H I N Q L A Q K
 GCTAATATCG ATCCTAAGTA TCTCATTTTC CAACCAGAAG GTGTCCAATT TTATAATAAA 2460
 A N I D P K Y L I F Q P E G V Q F Y N K

ES 2 278 436 T3

```

AATGGTGAAT TGGTAACTTA TGATATCAAG ACACTTCAAC AAATAAACCC TTAACCAAAA 2520
N G E L V T Y D I K T L Q Q I N P

GAAGATCTCA TTGTTAAAGC ACTGCTTTGT CAAAGCAAGT TACGGTGATT TTGAAGTCAT 2580

TCTATGTAAC GAGTAGTGAT AAAAGTTGGA TAATAGCGGT TTTCTTTTGC AAAGAAATGG 2640

TATCCATGTT AGAATAGTAA AAAAGAGGA GGATTCTTGG ACTAATGTCA AATAAGTAGA 2700

CAGAAAACCTG TGTTATTTTA TTGCGTTAAA ATAATTTTCT TCTTTCTGAT TAGGGGTTAG 2760
      .K I A N F Y N E E K Q N P T L

TCCTAGATTA GCCGTATGTG GGTTGTAATT GTTATAAAAA TTCTCAATGT ATTCAAAGCA 2820
      G L N A T H P N Y N N Y F N E I Y E F C

GTCTAATTGA ACCTGTTTGA TATTTTGATA ATGTTTTCGG TTGATTGTGC TATGCTTTAA 2880
      D L Q V Q K I N Q Y H K R N I Q R H K L

ATACTTGAAG AATGCTTCAG TTACGGCATT ATCATAAGGA TATCCAGGAT TAGAAAAAGA 2940
      Y K F F A E T V A N D Y P Y G P N S F S

ATGCATGATA TTGGCACTGC ACCCTAATAG TGAGACGCAA GAAAAACACT TTTAGGCAAT 3000
      H M A I
<----|
CAGTTTCTCG TACTGTACAG GCGACTGGTC GTTTAATCTC TGTTGAATTC TAGTTTCATT 3060
      L K R Y Q V P S Q D N L R Q Q I R T E N

ATAAATGTA ATGTAATTTT TAACAATAIT TGTTATACTA TCTTTGTTGT ATTTTCTCCT 3120
      Y F T I Y N K V I N T I S D K N Y K R R

ATTATGGAAA TAAAGGTTT CAGTCTTTAG GACGGTGTGA AACCATTCAA TACAGGCATT 3180
      N H F Y F T E T K L V T H F W E I C A N

ATCTGCAGGT GTTCCTTTTC GAGACATTGA GCGGATAATG TCTTTTCCG TGCAAGCCTG 3240
      D A P T G K R S M S R I I D K E T C A Q

GTAGTAAGCC ATAGAAGTAT ACACTGAGCC TTGGTCACTG TGTAAGATTG CTCCTTTATT 3300
      Y Y A M
      <----|
TAGGCAATTT TAAGTATTA AGGGTGTCTA GTACAAAATC CGTGTCTCTGA CAATCTGAGA 3360
      K P L K L Q N L T D L V F D T D Q C D S

TAGTGTAAGC TATAATTTCT CGGTTATAGA GATTCATAAT TGATGAGAGA TACAATTTAC 3420
      I T Y A I I E R N Y L N M I S S L Y L K

AGTTACCGAA ATATAGGTAG GTAATATCTG TTACGAGCTT TTCCTTAGGC TTATCGGCAT 3480
      C N G F Y L Y T I D T V L K E K P K D A

GGAAATCCCG ACTCAATTTA TTATCTGTGA AATAATAAGC TTTACCCAAA TTGGGAACTT 3540
      H G D R S L K N D T L Y Y A K G L N P V

TCTTGGTACG TGTCGGACAA AGCCAGCCAT TATTTTTCAT GATACGATAG ACTTTCTTTG 3600
      K K T R T R C L W G N N K M I R Y V K K

TATTAACAGT CAATCCGTGG ATTTTGTGTA GCAATCGTGT AATGGTACGA TAGCCATAAA 3660
      T N V T L G H I K K L L R T I T R Y G Y

TAAAGTGATT CTCCATACAG AGCTGTTCAA TTAATTCAT AAGGTCATCT TTTTTTGGCG 3720
      I F H N E M
      <----|

```

ES 2 278 436 T3

CTTCTCATAC TCCTTTTTC AACGGTAATA GGTCGACCGC TTGACCTTAA AACAGTCTAG 3780
 AATGAAACT ATCGGGTAGT TGTTTTTATA GTCTCCACA AGCTTGATAA GACTTACTTT 3840
 ATCGATTTC TTATCAAGCC TCGATACTTT TTTAAGAGGT CAACCTGTAA TTGTAATTGT 3900
 I S K R I L G R Y K K L L D V Q L Q L Q
 TCCACTTCAG ACAGATGTTT CAAGCCTTTA CCGTAGGTAT ATTGCTTGCC AACACCTTGA 3960
 E V E S L H E L G K G Y T Y Q K G V G Q
 TGAACGAT AAAGCTCCTC GTTTTCGTAC CATTTCATCC AAGTATAGAT TTGACTATTA 4020
 H F R Y L E E N E Y W K M W T Y I Q S N
 TTTTGTATGC CTAAAGTCTC CATAATAACT CTGTTAGACT TGCCTGCTTT CTTCATATCG 4080
 N K I G L T E M I V R N S K G A K K M D
 ATGCAAGCCA GCTTAGTTTC CCATGAATAT GCTTTTTTAA CCATAATAAA ACATTCTGT 4140
 I C A L K T E W S Y A K K V M
 <----|
 TTCTAGTTTA CTAAATTCA ACAGGAGTGT TTTCTTTTG TCTCATTTTA GGGATTCAGT 4200
 GCCTATTGTT GTCATCAATT ATTTTCTAA ATTCCCGGA CTAAATTGT GACCCTTGGT 4260
 CGGAATGAAA GAGAAGTGT CTTCAATCT TTCTTTTATT AAGTGAAAAG GCAACACTTT 4320
 TCTGTACAAC ATTTATAAAG TGTTTTTCTA GGCAATTAAT CTTTGTAGTCA TTGGTGTGTTG 4380
 A I L R K T M P T Q
 GTAGTTGAGA CTACCATGAA TGCGGTGGTA ATTCCACCAA TGAACATAGT CTTTAGTCTT 4440
 Y N L S G H I R H Y N W W H V Y D K T K
 AAGAGCTAGT TCTTCCAGCA ATTGAAAGGT TTCTTGATAA ACAAATTCAA TTTTGAAAGC 4500
 L A L E E L L Q F T E Q Y V F E I K F A
 ACGATACGTA CTTTCAGCTA CGGCATTGTC ATAAGGATAA CCAGCCTGAC TAAGCGAACG 4560
 R Y T S E A V A N D Y P Y G A Q S L S R
 TGTGATTCCA AAGGCTTCCA ATATTTTCATC AATTAAGTGA TTATCAAAT CTTTGCCACG 4620
 T I G F A E L I E D I L Q N D F E K G R
 ATCTGAATGG AACATCTTGA CTTTGGTCAG GCGTAAGGG ATGCTTTGTA TGGCTTGCTT 4680
 D S H F M K V K T L A Y P I S Q I A Q K
 AACGAGTTCA GCGGTCTTGT GCCAACCAAG AGACAGGCCG ATGATTTTAC GGTGTATAG 4740
 V L E A T K H W G L S L G I I E R N Y L
 GTCAATGATG AGGCAAACAT AAGCCCAACG ATTGCCTACA CGAACATAGG TTAAGTCAAGT 4800
 D I I L C V Y A W R N G V R V Y T L D T
 GACTAAGGCT TGTAAGTGGT TTTCTTGCTT AAATTGCCTG TCTAAGTGGT TGGGAATAGG 4860
 V L A Q L P R E Q K F Q R D L H N P I P
 GGCTTCATTC TTGCCTCTAG AATGTGGTTT GAAGGTGGCT TTCTGATAAA CAGAAACCAA 4920
 A E N K G R S H P K F T A K Q Y V S V L
 ATGAGTCGC TTCATAATGC GTCGAATCCG ACGACGTGAA AGTGTGATAC CTTCTGTTATT 4980
 N L R K M I R R I R R R S L T I G E N N
 CAAGCATATT TTGATTTTTC TGGATCCGTA TCTAGACTCG CTATCGAGAA AAATCTTTT 5040
 L C I K I K R S G Y R S E S D L F I R K

ES 2 278 436 T3

```

AATAGTTTCT TCAAACCTCCG TTTCAGATAC TGAATCCACG GCTTGATAGT AATAACTTGA 5100
  I T E E F E T E S V S E V A Q Y Y Y S S
GTGTGGCATA TTCAGCCAGC GACACATCTT TGAAATGCTG TATTTATCCT TATTAGCAGT 5160
  H P M N L W R C M K S I S Y K D K N A T
GATTATTTCC CTTTTGTGC CATAATCACC GCTGCTTGCT TTAGGATATC TAATT 5215
  I I E R K T G Y D G S S A K P Y R I
(SEQ ID NO:13)

```

FIG. 3a

```

FGSALSTVEV KEIISEENIW LYRLSCCHFT SYSYWKLPWT 40
(SEQ ID NO:14)

```

FIG. 3b

```

MGLATKDNQI AYIDDSKGKA KAPKTNKTMD QISAEEGISA EQIVVKITDQ 50
GYVTSHGDHY HFYNGKVPYD AIISEELMT DPNYRFKQSD VINEILDGYV 100
IKVNGNYVY LKPGSKRKNI RTKQQIAEQV AKGTKEAKEK GLAQVAHLSK 150
EEVAAVNEAK RQGRYTTDDG YIFSPTDIID DLGDAYLVPK GNHYHYIPKK 200
DLSPSELAAA QAYWSQKQGR GARPSDYRPT PAPGRRKAPI PDVTPNPGQG 250
HQPDNNGGYHP APPRPNDASQ NKHQRDEFKG KTFKELLDQL HRLDLKYRHV 300
EEDGLIFEPT QVIKSNAFGY VVPHGDHYHI IPRSQLSPLE MELADRYLAG 350
QTEDNDSGSE HSKPSDKEVT HTFLGHRIKA YGKGLDGKPY DTSDAYVFSK 400
ESIHSVDKSG VTAKHGDHFI YIGFGELEQY ELDEVANWVK AKGQADELAA 450
ALDQEQGKEK PLFDTKKVSR KVTKDGGKVG YMPKDGKDYF YARDQLDLTQ 500
IAFAEQELML KDKKHYRYDI VDTGIEPRLA VDVSSLPMHA GNATYDTGSS 550
FVIPHDHII VVPYSWLTRD QIATVKYVMQ HPEVRPDVWS KPGHEESGSV 600
IPNVTPLDKR AGMPNWQIIH SAEVQKALA EGRFATPDGY IFDPRDVLAK 650
ETFWWDGSGF SIPRADGSSL RTINKSDLSQ AEWQQAQELL AKKNTGDATD 700
TDKPKEKQQA DKSNNQOPS EASKEEKESD DFIDSLPDYG LDRATLEDHI 750
NQLAQKANID PKYLIFQPEG VQFYNNKNGEL VTYDIKTLQQ INP 793
(SEQ ID NO:15)

```

FIG. 3c

ES 2 278 436 T3

MTDPNYRFBQ	SDVINEILDG	YVIKVNNGYY	VYLKPGSKRK	NIRTKQQIAE	50
QVAKGTKEAK	EKGLAQVAHL	SKEEVAAVNE	AKRQGRTTDD	DGYIFSPTDI	100
IDDLGDAYLV	PHGNHYHYIP	KKDLSPSELA	AAQAYWSQKQ	GRGARPSDYR	150
PTPAPGRRKA	PIPDVTPNPG	QGHQPDNGGY	HPAPPRPND	SDQNKHQDEF	200
KGKTFKELLD	QLHRLDLKYR	HVEEDGLIFE	PTQVIKSNAF	GYVVPBGDHY	250
HIIPRSQLSP	LEMELADRYL	AGQTEDNDSD	SEHSKPSDKE	VTHTFLGHRI	300
KAYGKGLDGK	PYDTSDAYVF	SKESIHSVVK	SGVTAKHGDH	FHYIGFGELE	350
QYELDEVANW	VKAKGQADEL	AAALDQEQGK	EKPLFDTKKV	SRKVTKDGKV	400
GYMMPKDGKD	YFYARDQLDL	TQIAFAEQEL	MLKDKKHRYR	DIVDTGIEPR	450
LAVDVSSLPM	HAGNATYDTG	SSFVIPHIDH	IHVVPYSWLT	RDQIATVKYV	500
MQHPEVRPDV	WSKPGHEESG	SVIPNVTPLD	KRAGMPNWQI	IHSAAEVQKA	550
LAEGRFATPD	GYIFDPRDVL	AKETFEVWKG	SFSIPRADGS	SLRTINKSDL	600
SQAEWQQAQE	LLAKKNTGDA	TDTDKPKEKQ	QADKSNENQQ	PSEASKEEKE	650
SDDFIDSLPD	YGLDRATLED	HINQLAQKAN	IDPKYLIFQP	EGVQFYKNG	700
ELVTYDIKTL	QQINP	(SEQ ID NO:16)			715

FIG. 3d

MHSFNSPGYP	YDNAVTEAFF	KYLKHRQINR	KHYQNIQVQ	LDCFEYIENF	50
YNNYNPHTAN	LGLTPNQKEE	NYFNAIK	(SEQ ID NO:17)		77

FIG. 3e

MAYYQACTEK	DIIRMSRKG	TPADNACIEW	FHTVLKTETF	YFHNRRKYNK	50
DSITNIVKNY	ITFYNETRIQ	QRLNDQSPVQ	YRKLIA	(SEQ ID NO:18)	86

FIG. 3f

MENHFIYGYR	TITRLLKKIH	GLTVNTKKVY	RIMKNNGWLC	RTRTKKVPNL	50
GKAYYLTDNK	LSRDFHADKP	KEKLVDITY	LYFGNCKLYL	SSIMNLYNRE	100
IIAYTISDCQ	DTDFVLDTLN	QLKLPK	(SEQ ID NO:19)		126

FIG. 3g

ES 2 278 436 T3

MVKKAYSWET KLACIDMKKA GKSNRVIMET LGIKNNSQIY TWMKWEENE 50
 LYRFHQGVGK QYTYGKGLEH LSEVEQLQLQ VDLLKKYRGL IRKSIK 96
 (SEQ ID NO:20)

FIG. 3h

IRYPKASSGD YGTKREIITA NKDKYSISKM CRWLNMPHSS YYYQAVESVS 50
 ETEFEETIKR IFLDSESRYG SRKIKICLNN EGITLSRRRI RRIMKRLNLV 100
 SVYQKATFKP HSRGKNEAPI PNHLDRQFKQ ERPLQALVTD LTYVRVGNRW 150
 AYVCLIIDLY NREIIGLSLG WHKTAELVKQ AIQSIPYALT KVKMFHSDRG 200
 KEFDNQLIDE ILEAFGITRS LSQAGYPYDN AVAESTYRAF KIEFVYQETF 250
 QLLEELALKT KDYVHWWNYH RIHGSLNYQT PMTKRLIA (SEQ ID NO:21) 288

FIG. 3i

AATTTGAAAG CAGAATTATC TGTAAGAT GAGCAATATA CAGCAACAGT TTATGGTAAA 60
 N L K A E L S V E D E Q Y T A T V Y G K
 ---->
 TCTGCTCATG GTTCAACACC ACAAGAAGGT GTTAATGGGG CGACTTATTT AGCTCTTTAT 120
 S A H G S T P Q E G V N G A T Y L A L Y
 CTAAGTCAAT TTGATTTTGA AGGTCCTGCT CGTGCTTTCT TAGATGTTAC AGCCAACATT 180
 L S Q F D F E G P A R A F L D V T A N I
 ATTCACGAAG ACTTCTCAGG TGAAAACTT GGAGTAGCTT ATGAAGATGA CTGTATGGGA 240
 I H E D F S G E K L G V A Y E D D C M G
 CCATTGAGCA TGAATGCAGG TGTCTCCAG TTTGATGAAA CTAATGATGA TAATACTATC 300
 P L S M N A G V F Q F D E T N D D N T I
 GCTCTTAATT TCCGTTACCC ACAAGGGACA GATGCTAAAA CTATCCAAAC TAAGCTTGAG 360
 A L N F R Y P Q G T D A K T I Q T K L E
 AAACCTAACG GAGTTGAAAA AGTGACTCTT TCTGACCATG AACACACACC AACTATGTGA 420
 K L N G V E K V T L S D H E H T P H Y V
 CCTATGGACG ATGAATTAGT ATCAACCTTA CTAGCTGTCT ATGAAAAGCA AACTGGTCTT 480
 P M D D E L V S T L L A V Y E K Q T G L
 AAAGGACATG AACAGGTTAT TGGTGGTGGG ACATTTGGTC GCTTACTTGA ACGGGGTGTT 540
 K G H E Q V I G G G T F G R L L E R G V
 GCATACGGTG CCATGTTCCC AGGAGATGAA AACACTATGC ATCAAGCTAA TGAGTACATG 600
 A Y G A M F P G D E N T M H Q A N E Y M
 CCTTTAGAAA ATATTTTCCG TTCGGCTGCT ATCTACGCAG AAGCTATCTA TGAATTAATC 660

ES 2 278 436 T3

P L E N I F R S A A I Y A E A I Y E L I

AAATAAAATA ATCCTTAAAC TAAATATGTG ATCAATGATA AAGGGTGGTG AAGACATGAA 720
K .

AGTGTCTTTG CCTCTTTTCA TAAGGTTAGA TTTGGAGACT TTATGACTGA CTTGGAAAAA 780
M T D L E K
|---->

ATTATTAAG CAATAAAAG TGATTCACAG AATCAAAATT ATACAGAAAA TGGTATTGAT 840
I I K A I K S D S Q N Q N Y T E N G I D

CCTTTGTTTG CTGCTCCTAA AACAGCTAGG ATCAATATTG TTGGCCAAGC ACCTGGTTTA 900
P L F A A P K T A R I N I V G Q A P G L

AAACTCAAG AAGCAAGACT CTATTGGAAA GATAAATCTG GAGATCGTCT ACGCCAGTGG 960
K T Q E A R L Y W K D K S G D R L R Q W

CTTGGAGTTG ATGAAGAGAC ATTTTACCAT TCTGGAAAAT TTGCTGTTTT ACCTTTAGAT 1020
L G V D E E T F Y H S G K F A V L P L D

TTTTATTACC CAGGCAAAGG AAAATCAGGA GATTTACCCC CTAGAAAAGG TTTTGC GGAG 1080
F Y Y P G K G K S G D L P P R K G F A E

AAATGGCACC CTCTTATTTT AAAAGAAATG CCTAATGTTT AATTGACCTT GCTAGTTGGT 1140
K W H P L I L K E M P N V Q L T L L V G

CAGTATGCTC AGAAATATTA TCTTGAAGC TCCGCACATA AAAATCTAAC AGAAACAGTT 1200
Q Y A Q K Y Y L G S S A H K N L T E T V

AAAGCTTACA AAGACTATCT ACCCGATTAT TTACCCCTGG TTCACCCATC ACCGCGAAAT 1260
K A Y K D Y L P D Y L P L V H P S P R N

CAAATTTGGC TAAAGAAGAA TCCATGGTTT GAAAAAGATC TAATCGTTGA TTTACAAAAG 1320
Q I W L K K N P W F E K D L I V D L Q K

ATAGTAGCAG ATATTTTAAA AGATTAAGGA TAGGAGTTGG TATGAGAGAT AATCATCTAC 1380
I V A D I L K D M R D N H L H
|---->

ACACGTATTT TTCCTATGAT TGTCAAACGG CAITTGAGGA CTATATTAAT GGTTTTACAG 1440
T Y F S Y D C Q T A F E D Y I N G F T G

GTGAATTTAT CACGACAGAA CATTTTGATT TATCAAATCC TTACACCGGT CAAGACGATG 1500
E F I T T E H F D L S N P Y T G Q D D V

TTCCTGATTA TAGTGCTTAT TGTCAAAAAA TAGATTATCT TAATCAGAAA TATGGAAATC 1560
P D Y S A Y C Q K I D Y L N Q K Y G N R

GATTTAAAAA AGGAATTGAA ATCGGTTATT TTAAAGATAG GGAATCAGAT ATTTTAGATT 1620
F K K G I E I G Y F K D R E S D I L D Y

ATTTAAAAA TAAAGAATTT GATTTAAAAC TATTGTCAAT CCATCATAAT GGTAGGTATG 1680
L K N K E F D L K L L S I H H N G R Y D

ATTATCTGCA AGAAGAAGCT CTGAAAGTAC CAACAAAGGG AGCTTTTAGC AGATTACTTT 1740
Y L Q E E A L K V P T K G A F S R L L

AATCGTATGG AATTTGCCAT AGGCCGTGTG GAAGCGCACG TTTTAGCTCA CTTTGATTAT 1800

GGTTTTCGTA AGTTAACTT AGATGTAGAA GATTTAAAAC CGTTTGAAAC GCAATTGAAG 1860

CGCATTTTCA TAAAGATGTT ATCTAAGGGG TTAGCTTTTG AACTAAATAC CAAATCCCTT 1920

ES 2 278 436 T3

TATCTATATG GGAATGAAAA ACTTTATCGC TATGCTTTAG AGATACTCAA ACAGCTTGGT 1980
 TGTAACAAT ACTCTATAGG CTCTGACGGT CATATTCCTG AACATTTTGG TTATGAATTT 2040
 GATAGACTTC AAGGTCTGCT AAAGGACTAT CAAATTGATG AAAATCATTT GATATGAGGA 2100
 AATTTTTGAT AAAAAAGCTA GGCAATATTG CTTAGCTTTT TTGTAATGCT ATTGATAGTT 2160
 TTAGTGAAAA TTTCAAAAAA ATAAAGAAAT CATTTACTTG TTGCAAGCGC TTGCGTAAAT 2220
 TGTTATGATT TTATTGGTAA CAATTCATTA AAAAAGGAGA ATGATATGAA AAGAAAAGAC 2280
 M K R K D
 |---->
 TTATTTGGTG ATAAACAAAC TCAATACACG ATTAGAAAGT TAAGTGTTGG AGTAGCTTCA 2340
 L F G D K Q T Q Y T I R K L S V G V A S
 GTTACAACAG GGGTATGTAT TTTTCTTCAT AGTCCACAGG TATTTGCTGA AGAAGTAAGT 2400
 V T T G V C I F L H S P Q V F A E E V S
 GTTTCCTCTG CAACTACAGC GATTGCAGAG TCGAATATTA ATCAGGTTGA CAACCAACAA 2460
 V S P A T T A I A E S N I N Q V D N Q Q
 TCTACTAATT TAAAAGATGA CATAAACTCA AACTCTGAGA CGGTTGTGAC ACCCTCAGAT 2520
 S T N L K D D I N S N S E T V V T P S D
 ATGCCGGATA CCAAGCAATT AGTATCAGAT GAAACTGACA CTCAAAGGG AGTGACAGAG 2580
 M P D T K Q L V S D E T D T Q K G V T E
 CCGGATAAGG CGACAAGCCT GCTTGAAGAA AATAAAGGTC CTGTTTCAGA TAAAAATACC 2640
 P D K A T S L L E E N K G P V S D K N T
 TTAGATTTAA AAGTAGCACC ATCTACATTG CAAAATACTC CCGACAAAAC TTCTCAAGCT 2700
 L D L K V A P S T L Q N T P D K T S Q A
 ATAGGTGCTC CAAGCCCTAC CTTGAAAGTA GCTAATCAAG CTCCACGGAT TGAAAATGGT 2760
 I G A P S P T L K V A N Q A P R I E N G
 TACTTTAGGC TACATCTTAA AGAATTGCCT CAAGGTCATC CTGTAGAAAG CACTGGACTT 2820
 Y F R L H L K E L P Q G H P V E S T G L
 TGGATATGGG GAGATGTTGA TCAACCGTCT AGTAATTGGC CAAATGGTGC TATCCCTATG 2880
 W I W G D V D Q P S S N W P N G A I P M
 ACTGATGCTA AGAAAGATGA TTACGGTTAT TATGTTGATT TTAAATTATC TGAAAAACAA 2940
 T D A K K D D Y G Y Y V D F K L S E K Q
 CGAAAACAAA TATCTTTTTT AATTAATAAC AAAGCAGGGA CAAATTTAAG CGGCGATCAT 3000
 R K Q I S F L I N N K A G T N L S G D H
 CATATTCAT TATTACGACC TGAGATGAAC CAAGTTTGGG TTGATGAAAA GTACGGTATA 3060
 H I P L L R P E M N Q V W I D E K Y G I
 CATACTTATC AACCCCTCAA AGAAGGGTAT GTCCGTATTA ACTATTTGAG TTCCTCTAGT 3120
 H T Y Q P L K E G Y V R I N Y L S S S S
 AACTATGACC ACTTATCAGC ATGGCTCTTT AAAGATGTTG CAACCCCYTC AACAACTTGG 3180
 N Y D H L S A W L F K D V A T P S T T W
 CCAGATGGTA GTAATTTTGT GAATCAAGGA CTATATGGAA GGTATATTGA TGTATCACTA 3240
 P D G S N F V N Q G L Y G R Y I D V S L

ES 2 278 436 T3

AAACTAACG CCAAAGAGAT TGGTTTTCTA ATCTTAGATG AAAGTAAGAC AGGAGATGCA 3300
 K T N A K E I G F L I L D E S K T G D A
 GTGAAAGTTC AACCCAACGA CTATGTTTTT AGAGATTTAG CTAACCATAA CCAAATTTTT 3360
 V K V Q P N D Y V F R D L A N H N Q I F
 GTAAAAGATA AGGATCCAAA GGTTTATAAT AATCCTTATT ACATTGATCA AGTGCAGCTA 3420
 V K D K D P K V Y N N P Y Y I D Q V Q L
 AAGGATGCCC AACAAATTGA TTTAAACAAGT ATTCAAGCAA GTTTTACAAC TCTAGATGGG 3480
 K D A Q Q I D L T S I Q A S F T T L D G
 GTAGATAAAA CTGAAATTTTT AAAAGAATTG AAAGTGACTG ATAAAAATCA AAATGCTATA 3540
 V D K T E I L K E L K V T D K N Q N A I
 CAAATTTCTG ATATCACTCT CGATACTAGT AAATCTCTTT TAATAATCAA AGGCGACTTT 3600
 Q I S D I T L D T S K S L L I I K G D F
 AATCCTAAAC AAGGTCATTT CAACATATCT TATAATGGTA ACAATGTCAT GACAAGGCAA 3660
 N P K Q G H F N I S Y N G N N V M T R Q
 TCTTGGGAAT TTAAAGACCA ACTTTATGCT TATAGTGGAA ATTTAGGTGC AGTTCTCAAT 3720
 S W E F K D Q L Y A Y S G N L G A V L N
 CAAGATGGTT CAAAAGTTGA AGCCAGCCTC TGGTCACCGA GTGCTGATAG TGTCACTATG 3780
 Q D G S K V E A S L W S P S A D S V T M
 ATTATTTATG ACAAAGATAA CCAAAACAGG GTTGTAGCGA CTACCCCCCT TGTGAAAAAT 3840
 I I Y D K D N Q N R V V A T T P L V K N
 AATAAAGGTG TTTGGCAGAC GATACTTGAT ACTAAATTAG GTATTAAAAA CTATACTGGT 3900
 N K G V W Q T I L D T K L G I K N Y T G
 TACTATTATC TTTACGAAAT AAAAGAGGT AAGGATAAGG TTAAGATTTT AGATCCTTAT 3960
 Y Y Y L Y E I K R G K D K V K I L D P Y
 GCAAAGTCAT TAGCAGAGTG GGATAGTAAT ACTGTTAATG ATGATATTAA AACGGCTAAA 4020
 A K S L A E W D S N T V N D D I K T A K
 GCAGCTTTTG TAAATCCAAG TCAACTGGA CCTCAAATT TAAGTTTTGC TAAATTGCT 4080
 A A F V N P S Q L G P Q N L S F A K I A
 AATTTTAAAG GAAGACAAGA TGCTGTTATA TACGAAGCAC ATGTAAGAGA CTTCACTTCT 4140
 N F K G R Q D A V I Y E A H V R D F T S
 GATCGATCTT TGGATGGAAA ATTAAAAAAT CAATTGGTA CCTTTCAGC CTTTTCAGAG 4200
 D R S L D G K L K N Q F G T F A A F S E
 AAAGTAGATT ATTTACAGAA ATTAGGAGTT ACACACATTC AGCTTTTACC GGTATTGAGT 4260
 K L D Y L Q K L G V T H I Q L L P V L S
 TATTTTATG TTAATGAAAT GGATAAGTCA CGCTCAACAG CTTACACTTC CTCAGACAAT 4320
 Y F Y V N E M D K S R S T A Y T S S D N
 AATTACAATT GGGGCTATGA CCCACAGAGC TATTTTGCTC TTTCTGGGAT GTATTCAAG 4380
 N Y N W G Y D P Q S Y F A L S G M Y S E
 AAACCAAAG ATCCATCAGC ACGTATCGCC GAATTAAAC AATTAATACA TGATATTCAT 4440
 K P K D P S A R I A E L K Q L I H D I H

ES 2 278 436 T3

```

AAACGTGGCA TGGGGGTTAT ACTTGATGTC GTCTATAATC AACTGCAAA AACTTATCTC 4500
K R G M G V I L D V V Y N H T A K T Y L

TTTGAGGATA TAGAACCTAA TTATTATCAC TTTATGAATG AAGATGGTTC ACCAAGAGAA 4560
F E D I E P N Y Y H F M N E D G S P R E

AGTTTTGGAG GGGGACGTTT AGGAACCACT CATGCAATGA GTCGTCGTGT TTTGGTTGAT 4620
S F G G G R L G T T H A M S R R V L V D

TCCATTAAAT ATCTTACAAG TGAATTTAAA GTTGATGGTT TCCGTTTTGA TATGATGGGA 4680
S I K Y L T S E F K V D G F R F D M M G

GATCATGATG CGGCTGCGAT TGAATTAGCT TATAAAGAAG CTAAAGCTAT TAATCCTAAT 4740
D H D A A A I E L A Y K E A K A I N P N

ATGATTATGA TTGGTGAGGG CTGGAGAACA TTCCAAGGCG ATCAAGGTCA GCCGGTTAAA 4800
M I M I G E G W R T F Q G D Q G Q P V K

CCAGCTGACC AAGATTGGAT GAAGTCAACC GATACAGTTG GCGTCTTTTC AGATGATATT 4860
P A D Q D W M K S T D T V G V F S D D I

CGTAATAGCT TGAAATCTGG TTTTCCAAAT GAAGGTACTC CAGCTTTCAT CACAGGTGGC 4920
R N S L K S G F P N E G T P A F I T G G

CCACAATCTT TACAAGGTAT TTTTAAAAAT ATCAAAGCAC AACCTGGGAA TTTGAAGCA 4980
P Q S L Q G I F K N I K A Q P G N F E A

GATTGCCCAG GAGATGTGGT GCAGTATATT GCTGCACATG ATAACCTTAC CTTGCATGAT 5040
D S P G D V V Q Y I A A H D N L T L H D

GTGATTGCAA AATCAATT (SEQ ID NO:22) 5058
V I A K S I

```

FIG. 4a

```

NLKAELSVED EQYTATVYGK SAHGSTPQEG VNGATYLALY LSQFDFEGPA 50
RAFLDVTANI IHEDFSGEKL GVAYEDDCMG PLSMNAGVFQ FDETNDNDTI 100
ALNFRYPQGT DAKTIQTKLE KLNGVEKVTL SDHEHTPHYV PMDDELVSTL 150
LAVYEKQTGL KGHEQVIGGG TFGRLLEGRV AYGAMFPGDE NTMHQANEYM 200
PLENIFRSAA IYAEAIYELI K (SEQ ID NO:23) 221

```

FIG. 4b

```

MTDLEKIIKA IKSDSQNQNY TENGIDPLFA APKTARINIV GQAPGLKTQE 50
ARLYWKDKSG DRLRQWLGV D EETFYHSGKF AVLPLDFYYP GKGKSGDLPP 100
RKGFAEKWHP LILKEMPNVQ LTLLVGQY AQ KYYL GSSAHK NLTETVKAYK 150
DYLPDYLPLV HPSPRNQIWL KKNPWFEKDL IVOLQKIVAD ILKD 194
(SEQ ID NO:24)

```

FIG. 4c

ES 2 278 436 T3

MRDNHLHTYF SYDCQTAFED YINGFTGEFI TTEHFDLSNP YTGQDDVPDY	50
SAYCQKIDYL NQKYGNRFKK GIEIGYFKDR ESDILDYLN KEFDLKLISI	100
HHNGRYDYLO EEALKVPTKG AFSRL (SEQ ID NO:25)	126

FIG. 4d

MKRKDLFGDK QTQYTIRKLS VGVASVTTGV CIFLHSPQVF AEEVSVSPAT	50
TAIAESNINQ VDNQQSTNLK DDINSNSETV VTPSDMPDTK QLVSDETDTQ	100
KGVTEPDKAT SLLEENKGPV SDKNTLDLKV APSTLQNTPD KTSQAIGAPS	150
PTLKVANQAP RIENGYFRLH LKELPQGHVP ESTGLWIWGD VDQPSSNWP	200
GAIPMTDAKK DDYGYVDFK LSEKQRKQIS FLINNKAGTN LSGDHHIPLL	250
RPEMNQVWID EKYGIHTYQP LKEGYVRINY LSSSSNYDHL SAWLFKDVAT	300
PSTTWPDGSN FVNQGLYGRY IDVSLKTNK EIGFLILDES KTGDAVKVQP	350
NDYVFRDLN HNQIFVKDKD PKVYNNPYI DQVQLKDAQQ IDLTSIQASF	400
TTLDGVDKTE ILKELKVTOK NQNAIQISDI TLDTSKSLI IKGDFNPKQG	450
HFNISYNGNN VMTRQSWFEK DQLYAYSGNL GAVLNQDGSK VEASLWSPSA	500
DSVTMIIDYK DNQNRVATT PLVKNNKGVW QTILDTKLGI KNYTGYYLY	550
EIKRGKDKVK ILDPYAKSLA EWDSNTVNDD IKTAKAAFVN PSQLGPNLS	600
FAKIANFKGR QDAVIYEAHV RDFTSDRSLD GKLNQFGTF AAFSEKLDYL	650
QKLGVTIQL LPVLSYFYVN EMDKSRSTAY TSSDNNYNWG YDPQSYFALS	700
GMYSEKPKDP SARIAELKQL IHDIHKGGMG VILDVVYNHT AKTYLFEDIE	750
PNNYHFMNED GSPRESFGG RLGTTHAMSR RVLVDSIKYL TSEFKVDGFR	800
FDMMGDHDA AIELAYKEAK AINPNMIMIG EGWRTFQGDQ GQPVKPADQD	850
WMKSTDTGVG FSDDIRNSLK SGFPNEGTPA FITGGPQSLQ GIFKNIKAQP	900
GNFEADSPGD VVQYIAAHDN LTLHDVIAKS I (SEQ ID NO:26)	931

FIG. 4e

ES 2 278 436 T3

```

AATTCAAAGT TTGACAGAAG GTCAACTTCG TTCTGATATC CCTGAGTTCC GTGCTGGTGA 60
  I Q S L T E G Q L R S D I P E F R A G D
---->
TACTGTACGT GTTCACGCTA AAGTTGTTGA AGGTACTCGC GAACGTATTC AGATCTTTGA 120
  T V R V H A K V V E G T R E R I Q I F E

AGGTGTTGTT ATCTCACGTA AAGGTCAAGG AATCTCAGAA ATGTACACAG TACGTAAAAT 180
  G V V I S R K G Q G I S E M Y T V R K I

TTCTGGTGGT ATCGGTGTAG AGCGTACATT CCCAATTCAC ACTCCTCGTG TTGATAAAAT 240
  S G G I G V E R T F P I H T P R V D K I

CGAAGTTGTT CGTTATGGTA AAGTACGTCG TGCTAAACTT TACTACTTAC GCGCATTGCA 300
  E V V R Y G K V R R A K L Y Y L R A L Q

AGGTAAAGCT GCACGTATTA AAGAAATCCG TCGTTAATTT TGATGATCAG ATTTTAAAAA 360

TGCTTG GTT TTTGAGGATA GTAACATATGT TTTAAACTG GACAACCAAG ACGTAAAAAA 420

TCTGCCTGTG GGCAGTTTTT TTAGTAGGTC CCCTTAGTTC AATGGATATA ACAACTCCCT 480
                      H I Y C S G

CCTAAGGAGT AATTGCTGGT TCGATTCCGG CAGGGGACAT ATTCATTGCA TGTAATAGC 540
  G L S Y N S T R N R C P V Y E N C T F L

GGTTTAGAGC TATTTTGCCC CAAATTTCTC TGATTAAGTT TATCGTTCCT ATCTTTTGT 600
  P K S S N Q G L N R Q N L K D N R D K Q

TCTTGTAATT GATGTGCGTA AACTTCTAAA GTGATATTTA AATTCTCGTG ATCTAAACT 660
  E Q L Q H A Y V E L T I N L N E H D L V

TGAGAGATGG AAATTAGATA GCTTGCAAAT GTATGCCTGA GAGAGTGCAC TCGTACCTCG 720
  Q S I S I L Y S A F T H R L S H V R V E

CGACCAGTTA TTTTTCGGAT AGTTTTATTG ACTGCATTAT TTGAAAGTTT GTCGAATAAT 780
  R G T I K R I T K N V A N N S L K D F L

CTGTCGTTTT TATTTTTTGT AAATTCATGC AAAAAAATA ATGTATCATT GTCAATTGGT 840
  R D N K N K T F E H L F F L T D N D I P

ATATTTCTGA TACTACTTTT GTTTTTTGTT GGCAGGTATC TTTGGTTGAA ATGATAATCC 900
  I N R I S S K N K T P L Y R Q N F H Y D

CAAGTTTTAT TAATTGATAA ATATTTGTTA GTGTAATCAA TATCATTAAAC TGTTAAACCT 960
  W T K N I S L Y K N T Y D I D N V T L G

AAACATTGAG CGAAGCGCAT GCCAGTTTTA GCGATGAGGT ATAACGCTGC ATACGATTGA 1020
  L C E A F R M
          <----|
TGTTGTGATT TTTCTTTACA AATTTTATC AAGCGTAAGT ATTCATTGGT TTCAAGAAAT 1080

TTTATCTCTA TTTACGCCCC TTATTTTTTG CTTTAACCTT AGTGAATAAA CAAAAATTTT 1140

TTTCTATATA TCCCTCGTGA ACAGCCATGG ATACGCAGGC TTTTACATGT ATGTTAAAAC 1200

GCTTTACTGT ATCTTGACA TGCGTTTGAC TATAATGATT TATGACTTGT TGATATTTAG 1260

```

ES 2 278 436 T3

TGGAAGTAAT ATTGCAAAGT AATATATTTT CTATTATATG TTTATACGAT ATTCGATATT 1320
 CCCACCCGTT GTCGCGTTTA CGGAAATACG CCATTGATAT ACTCCACATT AGCTAAAGAA 1380
 CAGGGTGTTT AAGGCTACCT TGATGGAAAA GGCTCTCTTA GAGATATTTG TAAATGGTAT 1440
 GATATCTCAA GTCGCTCTGT TCTCCAAAAG TGGATAAAAC GGTATACTAG TGGTGAAGAC 1500
 TTGAAAGCCA CTAGTAGAGG ATATAGCCGT ATGAAACAAG GAAGGCAAGC CACATTTGAA 1560
 GAACGTGTAG AGATTGTTAA CTACACCATT GCCCATGGGA AAGACTATCA AGCAGCTATT 1620
 GAGAAGTTTG GTGTTTCCTA CCAACAAATT TATTCTTGGG TCGTAAGCT TGAGAAGAAT 1680
 GGCTCACAAG GTTTGGTTGA TAGACGTGTG AAAGGGTTGG AGAGTAGGCC TGATTTAACC 1740
 GAGATTGAGC AACTTTAACT CAAGATTAAA CAATTGGAGG AACGTAATCG TCTCTTAGAA 1800
 ATCGAGGTTA GTTTACTAAA AAAGTTAGAA GACATCAAAC GAGGAAACAG ACGSTAAGAC 1860
 TAGGTAAGCA TTTAGCGGAG TTCCAAGTAA TCAAGAATTA TTACGATGAG GAATCTAATG 1920
 TGCCTATTCA GGCCCTATGC CAACTCTTGA AGGGGTCTCG TTCAGGCTAT TACAAGTGGC 1980
 TCAATCGTCA AAAAACAGAT TTTGAGACAA AAAATACAAA GCTAATGGCT AAAATCAAGG 2040
 AACTTCGTAG ACTCTACAAT GGTATCTTAG GTTATCGCCG TATGACAACA TTTATTAATC 2100
 GTCAACTTGG GACAACTTAA AACAAGAAAC GGATTCGTTG ATTGATGAAC ATTCTGGGGA 2160
 TTAGTTCAGT CATTGTCGT GTTAGCCATG CTGTGACAAA AGCTGGTGAC AGATTTTACG 2220
 AAGAAAATAT TCTTAATCGT GAATTACAG CCACAGCTCA TAACCAGAAA TGGTGACAG 2280
 ATGTCACCTA TCTTCAATAC GGTCTGGGAG CTAAAGCTTA TCTCAGTGCG ATTAAGACC 2340
 TGTATAACGG TTCTATTATC GCTTATGAGA TTAGTCACAA CAATGAAATC CACTTGTTAT 2400
 GAAGACCATT AAAAAGGGGC TAGAGCTCAA TCCAGGAGCC ACACCTATCA TCCATAGCGA 2460
 TTGAGGTAGT CAATATACTT CCAAAGAATA CCGTTATATC ATACACAAG CTGGTCTGAC 2520
 CTTATCCATG TCCCGGATTG GCAAATGTAT TCATAATGCA CCAACTGAAA GTTTCTTTGG 2580
 GTTTTCAAG ACTGAGTCTT ACCACCTTAA GAAATACAAC TCTTATGATG AGTTGGTCAA 2640
 TGATGTGGCA CGTTATATCG AATTCTACAA CACACAACGT TATCAATCAA AATTAAACAA 2700
 CCTGACTCCT CTAGAATTCA GGAATCAGGT TGCATAACTT ATCTTTTATT ATTTGACTGT 2760
 CTACTTGACA GGGAGCCGTT CAGATTGCTT AACCTTTCTA AATTGCTAA AATAGCTACA 2820
 AGAAAACGAG CCATTTAATG CTTATTTCTT ATACTGTCTT GCCTCACGCT CTCCTCGACC 2880
 AAAAATTGAG CGTGAGGCTT TTTGTTTCAT TAAACGATGA TATTTCCATA TTCATCAGTT 2940
 TGTTTTCCGA GAGCCATCAA AGCTTCGATA AGGTCGATAA TTCCAGGAAT AAAGGTAATA 3000
 CTAAAAATAA TATATAAAAA AACCTGGCCT ATTTTTCCTG CGTAAATTT ATGCGCTCCA 3060
 ATGCCGCCCA AAAGAACGTT AATAAAACAT AAATACTAT GTTAGCATAA GACTTTATT 3120

ES 2 278 436 T3

```

TTACAACTGA ATTTTCATATA AATGGATTAG AGTAAGGGAT AAAAGAAATT AGCATAGCTC 3180
TTTTGAAAAT AAAAAAATTA ATATAATATG GAAAAAATTT TATTTTCATAA ACGTTTCATA 3240
AAAGGTATGT AATCTAGTAT TTAGGCAACA CTATTTTGTC ACTGGTGTCT AGTAACTTAT 3300
AGATTGATAA TTTTACTAGT AAACGTAATT CTTGCTTTA AGAGTTAAAT GTCTATTTAT 3360
TGTAAGCTAA ATTGGGAGGT GAACTTATGT AAAATTAGAT AGGTACTGTC AAGTACGGGA 3420
TGATTATTGA AACAGCCAGT ATGCATCATA AAATCTGTAT TGCTTAATAA CTATTTCTTT 3480
AACCAGACAT CAGTTCATTG TTTATCATCG CTACCCTAAG TCTAGTTTTT TCAATAGAGC 3540
ATTAGGTAGT TTTTGATAAT AAAACTATAT AAACATGAGA ATTAGATTTC GTATTGCATT 3600
CTTCATAATG AGTTATTTGA GATTTTCCTT TGAATAAATA GATACGAAAT TCAGTAACTT 3660
CATATATAAA CGGCTCTATC ATTGAGATAG TTTGTCAAAT GAAGAAATTT TTAATGGAAA 3720
TAGTTTTTAA AACATTAGTT GTAGGCGATG TAAAAATATT AATCCAGTGG ATGCAATAGT 3780
TGCGGAGTAA AAATAGAGAG GAGTAATTAG GAAGTGATAA AATATGCTAT AGCATATATT 3840
ACCAGAAAAA AAAATAGAAC ACTTATTATA TTTGCTATTT TAACAATTGT TCTTTCTTGC 3900
TTGTATTCAT GTTAAACAAT AATGAAATCA AGTAATGAAA TAGAAAAGGC TTTATATGAA 3960
      M K S S N E I E K A L Y E
      |---->
AGTTCTAATT CTTCAATATC AATTACAAAA AAAGATGGTA AATATTTTAA TATTAATCAA 4020
S S N S S I S I T K K D G K Y F N I N Q
TTTAAGAATA TTGAAAAAAT AAAAGAGGTT GAAGAAAAAA TATTTCAATA TGATGGATTA 4080
F K N I E K I K E V E E K I F Q Y D G L
GCAAAATTGA AAGATCTTAA AGTAGTTAGT GGTGAGCAAA GTATAAATAG AGAAGATTTA 4140
A K L K D L K V V S G E Q S I N R E D L
TCTGACGAAT TTAATAATGT TGTTCCTACTA GAAGCTACAA GTAATACTAA AAGAAATCTT 4200
S D E F K N V V S L E A T S N T K R N L
TTATTTAGTA GTGGAGTATT TAGTTTTTAA GAAGGAAAAA ATATAGAAGA AAATGATAAG 4260
L F S S G V F S F K E G K N I E E N D K
AATCAATTC TTGTTTCATGA AGAATTTGCT AAACAAAACA AACTAAAATT GGGTGATGAA 4320
N S I L V H E E F A K Q N K L K L G D E
ATTGATCTTG AATTACTAGA TACGGAAAAA AGTGGAAAAA TAAAAAGTCA TAAATTTTAA 4380
I D L E L L D T E K S G K I K S H K F K
ATTATAGGAA TCTTTTCTGG TAAAAACAG GAAACATATA CAGGATTATC ATCTGATTTT 4440
I I G I F S G K K Q E T Y T G L S S D F
AGCGAAAAATA TGGTTTTTGT AGATTATTCA ACTAGCCAAG AAATATTAAA TAAATCAGAG 4500
S E N M V F V D Y S T S Q E I L N K S E
AATAATAGAA TTGCAAATAA AATTTTAATG TATTCTGGTA GTTTAGAATC TACAGAGCTT 4560
N N R I A N K I L M Y S G S L E S T E L
GCCTTAAACA AATTGAAAGA CTTTAAATTT GATAAGTCAA AGTATTCTAT TAAGAAAGAT 4620

```


ES 2 278 436 T3

A L N K L K D F K I D K S K Y S I K K D
 AATAAGCAT TCGAAGAGTC TTTAGAGTCA GTGAGTGGAA TAAAACATAT AATTAAAATA 4680
 N K A F E E S L E S V S G I K H I I K I
 ATGACTTATT CGATTATGTT AGGTGGAATA GTTGTTCTTT CATTAACTCT GATTCTATGG 4740
 M T Y S I M L G G I V V L S L I L I L W
 TTAAGAGAAA GAATTTATGA AATAGGTATA TTTTATCTA TTGGAACAAC TAAGATACAA 4800
 L R E R I Y E I G I F L S I G T T K I Q
 ATTATAAGGC AATTTATATT TGAGTTAATA TTCATATCAA TACCAAGTAT AATATCCTCC 4860
 I I R Q F I F E L I F I S I P S I I S S
 TTATTTTATG GGAATCTACT ATTAAAAGTA ATTGTAGAAG GATTTATTAA CTCAGAGAAC 4920
 L F L G N L L L K V I V E G F I N S E N
 TCAATGATTT TCGGTGGAAG TTTAATAAAT AAAAGCAGTT TTATGTTAAA CATAACAACA 4980
 S M I F G G S L I N K S S F M L N I T T
 CTTGCAGAAA GTTATTTAAT ATTAATAAGT ATTATTGTTT TATCAGTTGT AATGGCCTCT 5040
 L A E S Y L I L I S I I V L S V V M A S
 TCATTAATAT TATTTAAGAA ACCACAAGAA ATATTATCAA AAATAAGTTA GGAGCAAATA 5100
 S L I L F K K P Q E I L S K I S
 ATGGATATAT TAGAAATAAA GAATGTAAAT TACAGTTACG CAAATTCTAA AGAAAAAGTT 5160
 M D I L E I K N V N Y S Y A N S K E K V
 |---->
 TTGTCAGGAG TAAATCAAAA ATTTGAACTT GGAAAGTTTT ATGCGATAGT AGGGAAGTCA 5220
 L S G V N Q K F E L G K F Y A I V G K S
 GGAACAGGAA AATCCACACT TCTTTCCTTA CTTGCAGGAC TTGATAAAGT TCAAACAGGA 5280
 G T G K S T L L S L L A G L D K V Q T G
 AAAATCTTGT TTAAGAATGA AGATATAGAA AAGAAAGGAT ATAGTAATCA CAGAAAAAAT 5340
 K I L F K N E D I E K K G Y S N H R K N
 AATATATCTT TGGTATTTCA AAATTATAAT TTAATAGATT ATTTATCGCC GATTGAAAAT 5400
 N I S L V F Q N Y N L I D Y L S P I E N
 ATTAGACTAG TAAATAAATC AGTAGATGAG AGTATCTTGT TCGAATTAGG TTTAGATAAA 5460
 I R L V N K S V D E S I L F E L G L D K
 AAACAAATAA AAAGAAATGT TATGAAATTA TCTGGTGGTC AGCAACAAAG GGTAGCTATT 5520
 K Q I K R N V M K L S G G Q Q Q R V A I
 GCTAGGGCAC TGGTATCAGA TGCCCCAATA ATACTAGCTG ATGAGCCTAC CGGTAACCTA 5580
 A R A L V S D A P I I L A D E P T G N L
 GACAGTGTTA CTGCTGGAGA AATAATT (SEQ ID NO:27) 5607
 D S V T A G E I I

FIG. 5a

ES 2 278 436 T3

IQSLTEGQLR SDIPEFRAGD TVRVHAKVVE GTRERIQIFE GVVISRKGQG	50
ISEMYTVRKI SGGIGVERTF PIHTPRVDKI EVVRYGKVRR AKLYYLRLAQ	100
GKAARIKEIR R (SEQ ID NO:28)	111

FIG. 5b

MRFAECLGLT VNDIDYTNKY LSINKTWDYH FNQRYLPTKN KSSI RNIPID	50
NDTLFFLHEF TKNKNDRLED KLSNNAVNKT IRKITGREVR VHSLRHTFAS	100
YLISISQVLD HENLNITLEV YAHQLQEQKD RNDKLNQRNL GQNSSKPLFT	150
CNEYVPCNRN TSNYSLGGSC YIH (SEQ ID NO:29)	173

FIG. 5c

MKSSNEIEKA LYESSNSSIS ITKKDGKYFN INQFKNIEKI KEVEEKIFQY	50
DGLAKLKD LK VVSGEQSINR EDLSDEFKNV VSLEATSNTK RNLLFSSGVF	100
SFKEGKNIEE NDKNSILVHE EFAKQNK LKL GDEIDLELLD TEKSGKIKSH	150
KFKIIGIFSG KKQETYTGLS SDFSENMFV DYSTSQEILN KSENNRIANK	200
ILMYSGSLES TELALNKLKD FKIDKSKYSI KKD NKAF EES LESVSGIKHI	250
IKIMTYSIML GGIVVLSLIL ILWLRERIYE IGIFLSIGTT KIQIIRQFIF	300
ELIFISIPSI ISSLFLGNLL LKVIVEGFIN SENSMIFGGS LINKSSFMLN	350
ITTLAESYLI LISIIVLSVV MASSLILFKK PQEILSKIS	389
(SEQ ID NO:30)	

FIG. 5d

MDILEIKNVN YSYANSKEKV LSGVNQKFEL GKFYAIVGKS GTGKSTLLSL	50
LAGLDKVQTG KILFKNEDIE KKGYSNHRKN NISLVFQNYN LIDYLSPIEN	100
IRLVNKSVD E SILFELGLDK KQIKRNVML SGGQQQRVAI ARALVSDAPI	150
ILADEPTGNL DSVTAGEII (SEQ ID NO:31)	169

FIG. 5e

ES 2 278 436 T3

CATATGACAA	TATTTTTCAA	AGTCTACATC	ACTTACTCGC	CTGTCGTGGA	AAATCTGGCA	60
ATACATTAAT	CGACCAATTA	GTTGCTGATG	GTTTACTTCA	TGCAGATAAT	CACTACCATT	120
TTTTCAATGG	GAAGTCTCTG	GCCACTTTCA	ATACTAACCA	ATTGATTTCG	GAAGTTGTCT	180
ATGTTGAAAT	ATCCTTAGAT	ACTATGTCTA	GTGGTGAACA	TGATTTAGTA	AAAGTTAACA	240
TTATCAGACC	CACTACCGAG	CATACTATCC	CCACGATGAT	GACAGCTAGC	CCCTATCATC	300
AAGGTATCAA	TGATCCTGCC	GCAGACCAAA	AAACATACCA	AATGGAGGGT	GCGCTAGCAG	360
TTAAACAGCC	TAAACACATA	CAAGTTGACA	CAAAACCATT	TAAAGAAGAA	GTAAAACATC	420
CTTCAAAATT	ACCCATCAGC	CCTGCAACTG	AAAGCTTCAC	ACACATTGAC	AGTTATAGTC	480
TCAATGACTA	TTTTCTTTCT	CGTGGTTTTG	CTAATATATA	CGTTTCAGGT	GTGGGTACTG	540
CTGGCTCTAC	GGGTTTCATG	ACCAGTGGGG	ATTACCAACA	AATACAAAGC	TTTAAAGCAG	600
TCATTGATTG	GTTAAATGGT	AAGGTTACTG	CATTCAACA	TCATAAACGA	GATAAACAA	660
TCAAGGCTGA	TTGGTCAAAC	GGCCTTGTAG	CAACCACAGG	TAAATCTTAT	CTCGGTACCA	720
TGTCAACTGG	TTAGCAACA	ACTGGCGTTG	AGGGGCTGAA	AGTCATTATC	GCTGAAGCCG	780
CAATCTCCAC	ATGGTATGAT	TATTATCGAG	AAAATGGGCT	TGTGTGTAGT	CCAGGCGGCT	840
ACCCCGGTGA	AGATTTAGAC	GTTTTAACAG	AATTAACATA	CTCACGAAAC	CTCTTAGCTG	900
GTGATTACAT	CAAAAACAAC	GATTGCTATC	AAGCATTGTT	AAATGAACAA	TCAAAAGCAA	960
TTGACCGTCA	AAGTGGGGAT	TACAACCAAT	ACTGGCATGA	CCGTAATTAC	CTAACTCACG	1020
TCAATAATGT	CAAAAGTCGA	GTAGTTTACA	CTCATGGACT	ACAGGATTGG	AATGTTAAGC	1080
CAAGACATGT	CTACAAAGTT	TTCAATGCAT	TGCCTCAAAC	CATCAAAAAA	CACCTTTTTT	1140
TACATCAAGG	TCAACATGTG	TATATGCATA	ATTGGCAGTC	GATTGATTTT	CGTGAAAGCA	1200
TGAATGCCTT	ACTAAGCCAA	GAAGTACTTG	GCATTGACAA	TCATTTCCAA	TTAGAAGAGG	1260
TCATTTGGCA	AGATAATACT	ACTGAGCAAA	CTTGGCAAGT	TTTAGATGCT	TTCGGAGGAA	1320
ACCATCAAGA	GCAAATTGGT	TTAGGTGATA	GTAAAAAACT	TATTGATAAC	CATTATGACA	1380
AAGAAGCCTT	TGATACTTAT	TGTAAAGACT	TCAATGTGTT	CAAAAATGAT	CTTTTCAAGG	1440
GAAATAATAA	AACCAATCAA	ATCACTATTA	ATCTTCCTCT	AAAGAAAAAT	TATCTCCTGA	1500
ATGGACAGTG	CAAACTCCAT	CTACGTGTTA	AACTAGTGA	CAAAAAGGCC	ATTTTATCAG	1560
CCCAAATCTT	AGACTATGGT	CCTAAAAAAC	GATTCAAAGA	TACACCAACC	ATCAAATTCT	1620
TAAACAGCCT	TGATAATGGT	AAAAATTTTG	CCAGAGAAGC	TTTACGTGAA	CTCCCGTTTA	1680
CTAAAGATCA	TTATCGTGTC	ATCAGTAAAG	GTGTCTTGAA	CCTTCAAAAT	CGTACAGACT	1740
TACTTACAAT	TGAGGCTATC	GAGCCAGAAC	AATGGTTTGA	TATCGAGTTT	AGCCTCCAAC	1800
CAAGTATATA	TCAATTGAGT	AAAGGTGATA	ATCTAAGGAT	TATCCTTTAT	ACAACTGATT	1860
TTGAACATAC	CATTCGAGAT	AATGCTAGTT	ACTCTATAAC	AGTAGATTTG	AGTCAATCTT	1920
ATTTAACTAT	CCCAACTAAT	CAAGGAAATT	AACCTATGAA	ACTTCTTACT	AAAGAACGGT	1980
TTGATGATTC	TCAACACTTT	TGGTACCAGA	TCAATTTATT	ACAAGAGAGT	AACCTCGGAG	2040
CAGTTTTTGA	CCATGATAAT	AAAAACATTC	CACAGGTTGT	TGCAACTATT	GTTGATGATT	2100
TACAAGGTTT	CGGAAGTTCT	AATCATTTCT	GGTATTTTGG	CAATACTACT	GATACTTCCA	2160
TCCTTATGAT	TGCTCATTTA	AATCGAAAAT	TCTATATTCA	GGTTAATTTA	AAGGACTTTG	2220
ACTTTGCACT	CAATTTAATA	GCTATAAATA	ATTGGAAGAG	TCTCCTCCAA	ACTCAACTTG	2280
AAGCTCTAAA	CGATACCCTA	GCAATATTTT	AATAAATAAG	GTAAGATGGA	GTGACAAAGC	2340
AACGCGAGGG	AGACTGATTA	ATGTCATCTT	ATTGGAATAA	CTATCCTGAA	CTTAAAAAAA	2400

```

ATATTGATGA AACCAATCAA CTAATTCAAG AAAGAATACA GGTCAGAAAT AAAGATATTG 2460
AAGCGGCGCT AAGCCAACTC ACAGCTGCGG GAGGAAAACA GCTCAGACCA GCATTCTTTT 2520
ACCTTTTTTC TCAACTTGGT AATAAGGAGA ATCAAGATAC TCAGCAACTA AAGAAAATCG 2580
CTGCTTCTTT AGAAATCCTT CACGTTGCTA CATTAAATCCA TGATGATGTC ATTGATGACT 2640
CACCACCTAAG ACGTGGAAAT ATGACCATTG AAAGCAAGTT TGGCAAAGAC ATCGCAGTTT 2700
ATACTGGGGA TTTACTTTTC ACAGTCTTTT TCGATCTTAT TTTAGAATCT ATGACTGATA 2760
CACCATTTAT GAGGATTAAT GCAAAATCTA TCGGTAAAAT TCTCATGGGA GAATTGGACC 2820
AGATGCACCT TCGTTACAAT CAACAACAAG GTATCCATCA CTATTTACGT GCGATTTCAG 2880
GTAAGACAGC CGAACTCTTT AAATTAGCTA GCAAAGAAGG AGCTTACTTT GGTGGTGCAG 2940
AGAAGGAGGT TGTTGCTCTA GCAGGCCATA TCGGCTTTAA CATTGGTATG ACATTCCAAA 3000
TTTTGGATGA TATCCTGGAT TATACTGCAG ATAAAAAAC ATTTAATAAG CCTGTCTTAG 3060
AGGATTTAAC ACAAGGCGTT TACAGCCTTC CTCTACTTCT TGCCATTGAA GAAAATCCTG 3120
ATATTTTCAA ACCTATTTTA GATAAAAAAA CAGATATGGC TACTGAAGAC ATGGAAAAAA 3180
TTGCTTATCT CGTCGTTTCC CATAGAGGTG TTGACAAAGC TCGCCATCTA GCTCGTAAAT 3240
TTACTGAGAA AGCTATTAGT GACATAAATA AGCTACCCCA GAACTCTGCA AAAAAACAGT 3300
TGCTACAATT AACTAATTAC CTTTTAAAC GCAAAATTTA AATAATAAAA AAACATTCCA 3360
CAATGCTAGA AAAGCAGTTA GGAATGTTT TTTTATTATC ATTTATTTAT CGCACCTATC 3420
AATCATCATA GATCACCATC ATCAGCGGCT TTCAGCTGAC GGTAACGTTG ACTACTTTGA 3480
GACAATTCCT GAGGAGAACC TTCCAACCTC AATTGCCCAT TTTCTATAAA TAAGATACGA 3540
TCAGCATGTT CAATACCTTT TAAGTGATGT GTAATCCAAA CTAAGGTCTT ACCTTCCAAT 3600
TCTTTCATAA ATACCCCTAG TAAGGCTTGT TCAGTAATAG GATCAAGTCC AACAGTTGGC 3660
TCATCTAAGA TAACAATTGG GACATCTTTT AGTAAGATTC TAGCCAAAGC AATTCTATGC 3720
CTTTCGCCAC CTGAAAACCT AAGTCCAGCT TCATCAACCA TTGTATAGAG ACCATCTGAT 3780
AAATCAGTGA CCATCTCTTT CAATCCAACG CGTTCAAGAA CTTTCCATAC ATCTTCTTCA 3840
CTAGCATCTT GGTTTCCAAT GCGAATGTTA TTTAGCAGGG TTGTATTAAA AAGGTAGGGC 3900
GCTTGTGTGA TCACTCCAAT ATAGTTAGAA ATGCAATCAC CAACTATTGA AACATCAGCA 3960
CCGCCTAGGG TAATCTTCCC TTGACTTGCT TTCAAGTCGC CACGAAGTAG ACTAGCTAAG 4020
GTACTCTTGC CAGAACCACT CCGCCCTAAA ATAGCAATTT TTTCTCCTTC TTTAATATCC 4080
AAATCTAAAT GATGCAAAAC CCATTTCTCT TGTGGCTTAT ACTGGAAACT TAAATTCTTG 4140
ACGGAAAAAT CATATGGCTT ATTAGGCAAT T (SEQ ID NO:32) 4171

```

FIG. 6a

YDNIFQSLHH LLACRGKSGN TLIDQLVADG LLHADNHYHF FNGKSLATFN	50
TNQLIREVVY VEISLDTMSS GEHDLVKVNI IRPTTEHTIP TMMTASPYHQ	100
GINPAADQK TYQMEGALAV KQPKHIQVDT KPFKEEVKHP SKLPISPATE	150
SFTHIDSYSL NDYFLSRGFA NIYVSGVGTA GSTGFMTSGD YQQIQSFKAV	200
IDWLNGKVTA FTSHKRDQV KADWSNGLVA TTGKSYLGTM STGLATTGVE	250
GLKVIIAEAA ISTWYDYYRE NGLVCSPGGY PGEDLDVLTE LTYSRNLLAG	300
DYIKNNDCYQ ALLNEQSKAI DRQSGDYNQY WHDRNYLTHV NNVKSRVVYT	350
HGLQDWNVVP RHVYKVFNAL PQTIKKHLFL HQGQHVYMHN WQSIDFRESM	400
NALLSQELLG IDNHFQLEEV IWQDNTTEQT WQVLDAFGGN HQEQIGLGDS	450
KKLIDNHYDK EAFDTYCKDF NVFKNDLFKG NNKTNQITIN LPLKKNYLLN	500
GQCKLHLRVK TSDKKAILS QILDYGPCKR FKDTPTIKFL NSLDNGKNFA	550
REALRELPFT KDHYRVISKG VLNQNRDLD LTIEAIEPEQ WFDIEFSLQP	600
SIYQLSKGDN LRILYTTDF EHTIRDNASY SITVDLSQSY LTIPTNQGN	649

(SEQ ID NO:33)

FIG. 6b

MKLLTKERFD DSQHFYQIN LLQESNEGAV FDHDNKNIPQ VVATIVDDLQ	50
GSSSNHFWY FGNTTDSIL MIAHLNRKFY IQVNLKDFDF ALNLIAINNW	100
KSLQTQLEA LNDTLAIFQ (SEQ ID NO:34)	119

FIG. 6c

MSSYWNNYPE LKKNIDETNQ LIQERIQVRN KDIEAALSQ LTAAGGQLRP	50
AFFYLFSQLG NKENQDTQQL KKIAASLEIL HVATLIHDDV IDDSPLRRGN	100
MTIQSKFGKD IAVYTGDLLE TVFFDLILES MTDTPFMRIN AKSMRKILMG	150
ELDQMHLRYN QQQGIHHYLR AISGKTAELE KLASKEGAYF GGAEKEVVRL	200
AGHIGFNIGM TFQILDDILD YTADKKTENK PVLEDLTQGV YSLPLLLAIE	250
ENPDIFKPIL DKKTDMATED MEKIAYLVVS HRGVDKARHL ARKFTEKAIS	300
DINKLPQNSA KKQLLQLTNY LLKRKI (SEQ ID NO:35)	326

FIG. 6d

LPNKPYDESV KNLSFQYKPQ EKWVLHHLDL DIKEGEKIAI LGRSGSGKST 50
LASLLRGDLK ASQGKITLGG ADVSIVGDCI SNYIGVIQQA PYLFNTTLLN 100
NIRIGNQDAS EEDVWKVLER VGLKEMVTDL SDGLYTMVDE AGLRFSGGER 150
HRIALARILL KDVPIVILDE PTVGLDPITE QALLRVFMKE LEGKTLVWIT 200
HHLKGIEHAD RILFIENGQL ELEGSPQELS QSSQRYRQLK AADDGDL 247
(SEQ ID NO:36)

FIG. 6e

ES 2 278 436 T3

AATTCTAATT	GGAGGTTTTT	CTTGAATAAA	TGGTTAGTTA	AGGCAAGTTC	CTTAGTTGTT	60
TAGGTGGTA	TGGTTTTATC	TGCGGGTTCC	CGAGTTTTAG	CGGATACTTA	TGTCCGTCCA	120
ATTGATAATG	GTAGAATTAC	AACAGGTTTC	AATGGTTATC	CTGGACATTG	TGGGGTGGAT	180
TATGCTGTTT	CGACTGGAAC	GATTATTAGG	GCAGTGGCAG	ATGGTACTGT	GAAATTTGCA	240
GGAGCTGGAG	CCAACTTTTC	TTGGATGACA	GACTTAGCAG	GAAATTGTGT	CATGATTCAA	300
CATGCGGATG	GAATGCATAG	TGGTTACGCT	CATATGTCAC	GTGTGGTGGC	TAGGACTGGG	360
GAAAAAGTCA	AACAAGGAGA	TATCATCGGT	TACGTAGGAG	CAACTGGTAT	GGCGACGGGA	420
CCTCACCTTC	ATTTTGAATT	TTTACCAGCT	AACCCTAATT	TTCAAAATGG	TTTCCATGGA	480
CGTATCAATC	CAACGTCACT	AATTGCTAAC	GTTGCGACCT	TTAGTGGAAA	AACGCAAGCA	540
TCAGCTCCAA	GCATTAAGCC	ATTACAATCA	GCTCCTGTAC	AGAATCAATC	TAGTAAATTA	600
AAAGTGTATC	GAGTAGATGA	ATTACAAAAG	GTTAATGGTG	TTTGGTTAGT	CAAAAATAAC	660
ACCCTAACGC	CGACTGGGTT	TGATTGGAAC	GATAATGGTA	TACCAGCATC	AGAAATTGAT	720
GAGGTTGATG	CTAATGGTAA	TTTGACAGCT	GACCAGGTTT	TTCAAAAAGG	TGGTTACTTT	780
ATCTTTAATC	CTAAACTCT	TAAGACTGTA	GAAAAACCCA	TCCAAGGAAC	AGCTGGTTTA	840
ACTTGGGCTA	AGACACGCTT	TGCTAATGGT	AGTTCAGTTT	GGCTTCGCGT	TGACAACAGT	900
CAAGAAGTGC	TTTACAAATA	GTTTGAGGTA	TTGATTCAAT	GTTTTAAATG	ACAGTTTTGT	960
TACTAACTAA	GTACAATTTT	TTTAAACCGT	CTGAAAAATA	TTTTATAGTC	CAGTAAAGTG	1020
TGATATTATA	GTCTCGGACT	AATAAAAAGG	AAATAGGAAT	TGAAGCAATG	AAAATGAATA	1080
AAAAGGTACT	ATTGACATCG	ACAATGGCAG	CTTCGCTATT	ATCACTCGCA	AGTGTTCAAG	1140
CACAAGAAAC	AGATACGACG	TGGACAGCAC	GTACTGTTTC	AGAGGTAAAG	GCTGATTTGG	1200
TAAAGCAAGA	CAATAAATCA	TCATATACTG	TGAAATATGG	TGATACACTA	AGCGTTATTT	1260
CAGAAGCAAT	GTCAATTGAT	ATGAATGTCT	TAGCAAAAAAT	TAATAACATT	GCAGATATCA	1320
ATCTTATTTA	TCCTGAGACA	ACACTGACAG	TAACTTACGA	TCAGAAGAGT	CATACTGCCA	1380
CTTCAATGAA	AATAGAAACA	CCAGCAACAA	ATGCTGTCTG	TCAAAACAACA	GCTACTGTGG	1440
ATTTGAAAAC	CAATCAAGTT	TCTGTTGCAG	ACCAAAAAGT	TTCTCTCAAT	ACAATTTCTG	1500
AAGGTATGAC	ACCAGAAGCA	GCAACAACGA	TTGTTTCGCC	AATGAAGACA	TATCTTCTG	1560
CGCCAGCTTT	GAAATCAAAA	GAAGTATTAG	CACAAGAGCA	AGCTGTTAGT	CAAGCAGCAG	1620
CTAATGAACA	GGTATCAACA	GTCCTGTGA	AGTCGATTAC	TTCAAGAGTT	CCAGCAGCTA	1680
AAGAGGAAGT	TAAACCAACT	CAGACGTCAG	TCAGTCAGTC	AACAACAGTA	TCACCAGCTT	1740
CTGTTGCCGC	TGAAACACCA	GCTCCAGTAG	CTAAAGTAGC	ACCGGTAAGA	ACTGTAGCAG	1800
CCCCTAGAGT	GGCAAGTGTT	AAAGTAGTCA	CTCCTAAAGT	AGAACTGGT	GCATCACCAG	1860
AGCATGTATC	AGCTCCAGCA	GTTCTGTGA	CTACGACTTC	AACAGCTACA	GACAGTAAGT	1920
TACAAGCGAC	TGAAGTTAAG	AGCGTTCCGG	TAGCACAAAA	AGCTCCAACA	GCAACACCGG	1980
TAGCACAACC	AGCTTCAACA	ACAAATGCAG	TAGCTGCACA	TCCTGAAAAT	GCAGGGCTCC	2040
AACCTCATGT	TGCAGCTTAT	AAAGAAAAAG	TAGCGTCAAC	TTATGGAGTT	AATGAATTCA	2100
GTACATACCG	TGCAGGTGAT	CCAGGTGATC	ATGGTAAAGG	TTTAGCAGTC	GACTTTATTG	2160
TAGGTAAAAA	CCAAGCACTT	GGTAATGAAG	TTGCACAGTA	CTCTACACAA	AATATGGCAG	2220
CAAATAACAT	TTCATATGTT	ATCTGGCAAC	AAAAGTTTTA	CTCAAATACA	AATAGTATTT	2280
ATGGACCTGC	TAATACTTGG	AATGCAATGC	CAGATCGTGG	TGGCGTTACT	GCCAACCATT	2340
ATGACCATGT	TCACGTATCA	TTTAACAAAT	AATATAAAAA	AGGAAGCTAT	TTGGCTTCTT	2400

```

TTTTATATGC CTTGAATAGA CTTTCAAGGT TCTTATCTAA TTTTATTAA ATTGAGGAGA 2460
TTAAGCTATA AGTCTGAAAC TACTTTCACG TTAACCGTGA CTAAATCAAA ACGTTAAAAC 2520
TAAAATCTAA GTCTGTAAAG ATTATTGAAA ACGCTTTAAA AACAGATATA ATAAGGTTTG 2580
TAGATATCTA AAATTAAAAA AGATAAGGAA GTGAGAATAT GCCACATCTA AGTAAAGAAG 2640
CTTTTAAAAA GCAAATAAAA AATGGCATT TGTGTCTATG TCAAGCTTTG CCTGGGGAGC 2700
CTCTTTATAC TGAAAGTGGA GGTGTTATGC CTCTTTTAGC TTTGGCAGCT CAAGAAGCAG 2760
GAGCGGTTGG TATAAGAGCC AATAGTGTCC GCGACATTAA GGAAATCAA GAAGTTACTA 2820
ATTACCTAT CATCGGCATT ATTAAACGTG AATATCCTCC ACAAGAACCA TTTATCACTG 2880
CTACGATGAC AGAGGTGGAT CAATTAGCTA GTTTAGATAT TGCAGTAATA GCCTTAGATT 2940
GTACACTTAG AGAGCGTCAT GATGGTTTGA GTGTAGCTGA GTTTATTCAA AAGATAAAAG 3000
GGAAATATCC TGAACAGTTG CTAATGGCTG ATATAAGTAC TTTTGAAGAA GGTAAAAATG 3060
CTTTGAAGC AGGAGTTGAT TTTGTGGGTA CAACTCTATC TGGATACACA GATTACAGCC 3120
GCCAAGAAGA AGGACCGGAT ATAGAACTCC TTAATAAGCT TTGTCAAGCC GGTATAGATG 3180
TGATTGCGGA AGGTAAAATT CATACTCCTA AGCAAGCTAA TGAAATTAAT CATATAGGTG 3240
TTGCAGGAAT TGTAGTTGGT GGTGCTATCA CTAGACCAAA AGAAATAGCG GAGCGTTTCA 3300
TCTCAGGACT TAGTTAAAAG TGTTACTCAA AAATCAAAAT CAAAATAAAA AAGGGGAATA 3360
GTTATGAGTA TCAAAAAAAG TGTGATTGGT TTTTGCCTCG GAGCTGCAGC ATTATCAATG 3420
TTTGCTTG TG TAGACAGTAG TCAATCTGTT ATGGCTGCCG AGAAGGATAA AGTCGAAATT 3480
(SEQ ID NO:37)

```

FIG. 7a

```

NSIWRFFLNK WLKASSLVV LGGMVLSAGS RVLADTYVRP IDNGRITTGF 50
NGYPGHCGVD YAVPTGTIIR AVADGTVKFA GAGANFSWMT DLAGNCVMIQ 100
HADGMHSGYA HMSRVVARTG EKVQKQDIIG YVGATGMATG PHLHFEFLPA 150
NPNFQNGFHH RINPTSLIAN VATFSGKTQA SAPSIKPLQS APVQNQSSKL 200
KVYRVDELQK VNGVWLKNN TLTPTGFDWN DNGIPASEID EVDANGNLTA 250
DQVLQKGGYF IFNPKTLKTV EKPIQGTAGL TWAKTRFANG SSVWLRVDNS 300
QELLYK (SEQ ID NO:38) 306

```

FIG. 7b

MKMNKKVLLT	STMAASLLSV	ASVQAQETDT	TWTARTVSEV	KADLVKQDNK	50
SSYTVKYGDT	LSVISEAMSI	DMNVLAKINN	IADINLIYPE	TTLTVTYDQK	100
SHTATSMKIE	TPATNAAGQT	TATVDLKTNQ	VSVADQKVSL	NTISEGMTPE	150
AATTIVSPMK	TYSSAPALKS	KEVLAQEQAV	SQAAANEQVS	TAPVKSITSE	200
VPAAKEEVKP	TQTSVSQSTT	VSPASVAAET	PAPVAKVAPV	RTVAAPRVAS	250
VKVVTPKVET	GASPEHVSAP	AVPVTTTSTA	TDSKLQATEV	KSVPVQAQKAP	300
TATPVAQPAS	TTNAVAAHPE	NAGLQPHVAA	YKEKVASTYG	VNEFSTYRAG	350
DPGDHGKGLA	VDFIVGKNQA	LGNEVAQYST	QNMAANNISY	VIWQOKFYSN	400
TNSIYGPANT	WNAMPDRGGV	TANHVDHVHV	SFNK	(SEQ ID NO:39)	434

FIG. 7c

MPHLSKEAFK	KQIKNGIIVS	CQALPGEPLY	TESGGVMPLL	ALAAQEAGAV	50
GIRANSVRDI	KEIQEVTNLP	IIGIIKREYP	PQEPFITATM	TEVDQLASLD	100
IAVIALDCTL	RERHDGLSVA	EFIQKIKGKY	PEQLLMADIS	TFEEGKNAFE	150
AGVDFVGTTL	SGYTDYXRQE	EGPDIELLNK	LCQAGIDVIA	EGKIHTPKQA	200
NEINHIGVAG	IVVGGAITRP	KEIAERFISG	LS	(SEQ ID NO:40)	232

FIG. 7d

MSIKKSVIGF	CLGAAALSMF	ACVDSSQSVM	AAEKDKVEI	39
(SEQ ID NO:41)				

FIG. 7e

ATGAAAATGA	ATAAAAAGGT	ACTATTGACA	TCGACAATGG	CAGCTTCGCT	50
ATTATCAGTC	GCAAGTGTTT	AAGCACAAGA	AACAGATACG	ACGTGGACAG	100
CACGTACTGT	TTCAGAGGTA	AAGGCTGATT	TGGTAAAGCA	AGACAATAAA	150
TCATCATATA	CTGTGAAATA	TGGTGATACA	CTAAGCGTTA	TTTCAGAAGC	200
AATGTCAATT	GATATGAATG	TCTTAGCAAA	AATTAATAAC	ATTGCAGATA	250
TCAATCTTAT	TTATCCTGAG	ACAACACTGA	CAGTAACTTA	CGATCAGAAG	300
AGTCATACTG	CCACTTCAAT	GAAAAATAGAA	ACACCAGCAA	CAAATGCTGC	350
TGGTCAAACA	ACAGCTACTG	TGGATTTGAA	AACCAATCAA	GTTTCTGTTG	400
CAGACCAAAA	AGTTTCTCTC	AATACAATTT	CGGAAGGTAT	GACACCAGAA	450
GCAGCAACAA	CGATTGTTTC	GCCAATGAAG	ACATATTCTT	CTGCGCCAGC	500
TTTGAAATCA	AAAGAAGTAT	TAGCACAAGA	GCAAGCTGTT	AGTCAAGCAG	550
CAGCTAATGA	ACAGGTATCA	ACAGCTCCTG	TGAAGTCGAT	TACTTCAGAA	600
GTTCCAGCAG	CTAAAGAGGA	AGTTAAACCA	ACTCAGACGT	CAGTCAGTCA	650
GTCAACAACA	GTATCACCAG	CTTCTGTTGC	CGCTGAAACA	CCAGCTCCAG	700
TAGCTAAAGT	AGCACCGGTA	AGAACTGTAG	CAGCCCCTAG	AGTGGCAAGT	750
GTTAAAGTAG	TCACTCCTAA	AGTAGAAACT	GGTGCATCAC	CAGAGCATGT	800
ATCAGCTCCA	GCAGTTCCTG	TGACTACGAC	TTCAACAGCT	ACAGACAGTA	850
AGTTACAAGC	GACTGAAGTT	AAGAGCGTTC	CGGTAGCACA	AAAAGCTCCA	900
ACAGCAACAC	CGGTAGCACA	ACCAGCTTCA	ACAACAAATG	CAGTAGCTGC	950
ACATCCTGAA	AATGCAGGGC	TCCAACCTCA	TGTTGCAGCT	TATAAAGAAA	1000
AAGTAGCGTC	AACTTATGGA	GTAAATGAAT	TCAGTACATA	CCGTGCAGGT	1050
GATCCAGGTG	ATCATGGTAA	AGGTTTAGCA	GTCGACTTTA	TTGTAGGTAA	1100
AAACCAAGCA	CTTGGTAATG	AAGTTGCACA	GTA CTCTACA	CAAAATATGG	1150
CAGCAAATAA	CATTTTCATAT	GTTATCTGGC	AACAAAAGTT	TTACTCAAAT	1200
ACAAATAGTA	TTTATGGACC	TGCTAATACT	TGGAATGCAA	TGCCAGATCG	1250
TGGTGGCGTT	ACTGCCAACC	ATTATGACCA	TGTTACGTA	TCATTTAACA	1300
AATAA					1305

(SEQ ID NO:42)

FIG. 8

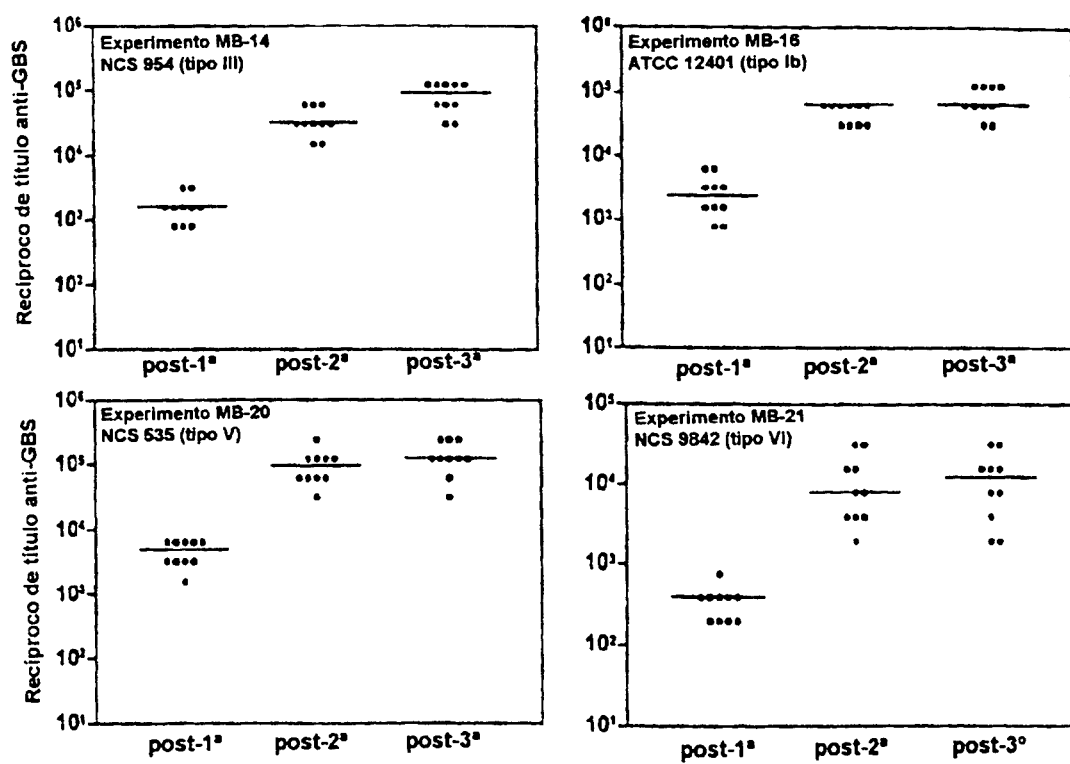
CAAGAAACAG	ATACGACGTG	GACAGCACGT	ACTGTTTCAG	AGGTAAAGGC	50
TGATTTGGTA	AAGCAAGACA	ATAAATCATC	ATATACTGTG	AAATATGGTG	100
ATACACTAAG	CGTTATTTCA	GAAGCAATGT	CAATTGATAT	GAATGTCTTA	150
GCAAAAATTA	ATAACATTGC	AGATATCAAT	CTTATTTATC	CTGAGACAAC	200
ACTGACAGTA	ACTTACGATC	AGAAGAGTCA	TACTGCCACT	TCAATGAAAA	250
TAGAAACACC	AGCAACAAAT	GCTGCTGGTC	AAACAACAGC	TACTGTGGAT	300
TTGAAAACCA	ATCAAGTTTC	TGTTGCAGAC	CAAAAAGTTT	CTCTCAATAC	350
AATTTTCGGAA	GGTATGACAC	CAGAAGCAGC	AACAACGATT	GTTTCGCCAA	400
TGAAGACATA	TTCTTCTGCG	CCAGCTTTGA	AATCAAAAGA	AGTATTAGCA	450
CAAGAGCAAG	CTGTTAGTCA	AGCAGCAGCT	AATGAACAGG	TATCAACAGC	500
TCCTGTGAAG	TCGATTACTT	CAGAAGTTCC	AGCAGCTAAA	GAGGAAGTTA	550
AACCAACTCA	GACGTCAGTC	AGTCAGTCAA	CAACAGTATC	ACCAGCTTCT	600
GTTGCCGCTG	AAACACCAGC	TCCAGTAGCT	AAAGTAGCAC	CGGTAAGAAC	650
TGTAGCAGCC	CCTAGAGTGG	CAAGTGTTAA	AGTAGTCACT	CCTAAAGTAG	700
AAACTGGTGC	ATCACCAGAG	CATGTATCAG	CTCCAGCAGT	TCCTGTGACT	750
ACGACTTCAA	CAGCTACAGA	CAGTAAGTTA	CAAGCGACTG	AAGTTAAGAG	800
CGTTCCGGTA	GCACAAAAAG	CTCCAACAGC	AACACCGGTA	GCACAACCAG	850
CTTCAACAAC	AAATGCAGTA	GCTGCACATC	CTGAAAATGC	AGGGCTCCAA	900
CCTCATGTTG	CAGCTTATAA	AGAAAAAGTA	GCGTCAACTT	ATGGAGTTAA	950
TGAATTCAGT	ACATACCGTG	CAGGTGATCC	AGGTGATCAT	GGTAAAGGTT	1000
TAGCAGTCGA	CTTTATTGTA	GGTAAAAACC	AAGCACTTGG	TAATGAAGTT	1050
GCACAGTACT	CTACACAAAA	TATGGCAGCA	AATAACATTT	CATATGTTAT	1100
CTGGCAACAA	AAGTTTACT	CAAATACAAA	TAGTATTTAT	GGACCTGCTA	1150
ATACTTGGAA	TGCAATGCCA	GATCGTGGTG	GCGTTACTGC	CAACCATTAT	1200
GACCATGTTT	ACGTATCATT	TAACAAATAA	(SEQ ID NO:43)		1230

FIG. 9

QETDTTWTAR	TVSEVKADLV	KQDNKSSYTV	KYGDTLVIS	EAMSIDMNVL	50
AKINNIADIN	LIYPETTLTV	TYDQKSHTAT	SMKIETPATN	AAGQTTATVD	100
LKTNQVSVAD	QKVSINLTISE	GMTPEAATTI	VSPMKTYSSA	PALKSKEVLA	150
QEQAVSQAAA	NEQVSTAPVK	SITSEVPAAK	EEVKPTQTSV	SQSTTVSPAS	200
VAAETPAPVA	KVAPVRTVAA	PRVASVKVVT	PKVETGASPE	HVSAPAVPVT	250
TTSTATDSKL	QATEVKSVPV	AQKAPTATPV	AQPASTTNAV	AAHPENAGLQ	300
PHVAAYKEKV	ASTYGVNEFS	TYRAGDPGDH	GKGLAVDFIV	GKNQALGNEV	350
AQYSTQNMMA	NNISYVIWQQ	KEYSNTNSIY	GPANTWNAMP	DRGGVTANHY	400
DHVHVSFNK	(SEQ ID NO:44)				409

FIG. 9a

Fig. 10



ES 2 278 436 T3

LISTA DE SECUENCIAS

5 <110> BioChem Vaccins
 RIOUX, Clément
 DENIS, Martin
 BRODEUR, Bernard R.
 HAMEL, Josée
 CHARLEBOIS, Isabelle
 BOYER, Martine
 10
 <120> NUEVOS ANTÍGENOS DE ESTREPTOCOCOS DEL GRUPO B
 <130> 12806-9PCT
 15
 <150> 60/075,425
 <151> 1998-02-20
 20
 <160>44
 <170> FastSEQ para Windows Versión 3.0
 25
 <210> 1
 <211> 4514
 <212> ADN
 <213> Estreptococos
 30
 <220>
 <221> CDS
 35
 <222> (3)...(464)
 <221> CDS
 <222> (534)...(887)
 40
 <223>
 <221> CDS
 <222> (1024)...(1767)
 45
 <221> CDS
 <222> (1841)... (4288)
 50
 <221> CDS
 <222> (2735)...(4288)
 <400> 1
 55
 ta tct ggc aaa gag cca gct aat cgt ttt agt tgg gct aaa aat aaa 47
 Ser Gly Lys Glu Pro Ala Asn Arg Phe Ser Trp Ala Lys Asn Lys
 1 5 10 15
 60
 65

ES 2 278 436 T3

	ttt tta atc aat gga ttc att gca act cta gca gca act atc tta ttt	95
	Leu Leu Ile Asn Gly Phe Ile Ala Thr Leu Ala Ala Thr Ile Leu Phe	
	20 25 30	
5	ttt gca gtt caa ttc ata ggt ctt aaa cca gat tac cct gga aaa acc	143
	Phe Ala Val Gln Phe Ile Gly Leu Lys Pro Asp Tyr Pro Gly Lys Thr	
	35 40 45	
10	tac ttt att atc cta ttg aca gca tgg act ttg atg gca tta gta act	191
	Tyr Phe Ile Ile Leu Leu Thr Ala Trp Thr Leu Met Ala Leu Val Thr	
	50 55 60	
15	gct tta gtg gga tgg gat aat agg tat ggt tcc ttc ttg tcg tta tta	239
	Ala Leu Val Gly Trp Asp Asn Arg Tyr Gly Ser Phe Leu Ser Leu Leu	
	65 70 75	
20	ata tta tta ttc cag ctt ggt tca agc gca gga act tac cca ata gaa	287
	Ile Leu Leu Phe Gln Leu Gly Ser Ser Ala Gly Thr Tyr Pro Ile Glu	
	80 85 90 95	
25	ttg agt cct aag ttc ttt caa aca att caa cca ttt tta ccg atg act	335
	Leu Ser Pro Lys Phe Phe Gln Thr Ile Gln Pro Phe Leu Pro Met Thr	
	100 105 110	
30	tac tct gtt tca gga tta aga gag acc atc tcg ttg acg gga gac gtt	383
	Tyr Ser Val Ser Gly Leu Arg Glu Thr Ile Ser Leu Thr Gly Asp Val	
	115 120 125	
35	aac cat caa tgg aga atg cta gta atc ttt tta gta tca tcg atg ata	431
	Asn His Gln Trp Arg Met Leu Val Ile Phe Leu Val Ser Ser Met Ile	
	130 135 140	
40	ctt gct ctt ctt att tat cgt aaa caa gaa gat taatagaaag tatctagtga	484
	Leu Ala Leu Leu Ile Tyr Arg Lys Gln Glu Asp	
	145 150	
45	tagactaaca gcatgatatg gtatgtcaaa gtatttagga ggagaagat atg tct act	542
	Met Ser Thr	
	155	
50	ttt aca ata att att gca aca tta act gct ttg gaa cat ttt tat att	590
	Leu Thr Ile Ile Ile Ala Thr Leu Thr Ala Leu Glu His Phe Tyr Ile	
	160 165 170	
55	atg tat ttg gag acg tta gcc acc cag tca aat atg act ggg aag att	638
	Met Tyr Leu Glu Thr Leu Ala Thr Gln Ser Asn Met Thr Gly Lys Ile	
	175 180 185	
60	ttt agt atg tct aaa gaa gag ttg tca tat tta ccc gtt att aaa ctt	686
	Phe Ser Met Ser Lys Glu Glu Leu Ser Tyr Leu Pro Val Ile Lys Leu	
	190 195 200 205	
65	ttt aag aat caa ggt gta tac aac ggc ttg att ggc cta ttc ctc ctt	734
	Phe Lys Asn Gln Gly Val Tyr Asn Gly Leu Ile Gly Leu Phe Leu Leu	
	210 215 220	

ES 2 278 436 T3

	tat ggg tta tat att tca cag aat caa gaa att gta gct gtt ttt tta	782
	Tyr Gly Leu Tyr Ile Ser Gln Asn Gln Glu Ile Val Ala Val Phe Leu	
	225 230 235	
5	atc aat gta ttg cta gtt gct att tat ggt gct ttg aca gtt gat aaa	830
	Ile Asn Val Leu Leu Val Ala Ile Tyr Gly Ala Leu Thr Val Asp Lys	
	240 245 250	
10	aaa atc tta tta aaa cag ggt ggt tta cct ata tta gct ctt tta aca	878
	Lys Ile Leu Leu Lys Gln Gly Gly Leu Pro Ile Leu Ala Leu Leu Thr	
	255 260 265	
15	ttc tta ttt taatactact tagccgttcg atttagttga acggcttcta	927
	Phe Leu Phe	
	270	
	gtaatcattt tttttctcata atacaggttag tttaagtaat ttgtctttaa aaatagtata	987
20	atataactac gaattcaaag agaggtgact ttgatt atg act gag aac tgg tta	1041
	Met Thr Glu Asn Trp Leu	
	275	
25	cat act aaa gat ggt tca gat att tat tat cgt gtc gtt ggt caa ggt	1089
	His Thr Lys Asp Gly Ser Asp Ile Tyr Tyr Arg Val Val Gly Gln Gly	
	280 285 290	
30	caa ccg att gtt ttt tta cat ggc aat agc tta agt agt cgc tat ttt	1137
	Gln Pro Ile Val Phe Leu His Gly Asn Ser Leu Ser Ser Arg Tyr Phe	
	295 300 305 310	
35	gat aag caa ata gca tat ttt tct aag tat tac caa gtt att gtt atg	1185
	Asp Lys Gln Ile Ala Tyr Phe Ser Lys Tyr Tyr Gln Val Ile Val Met	
	315 320 325	
40	gat agt aga ggg cat ggc aaa agt cat gca aag cta aat acc att agt	1233
	Asp Ser Arg Gly His Gly Lys Ser His Ala Lys Leu Asn Thr Ile Ser	
	330 335 340	
45	ttc agg caa ata gca gtt gac tta aag gat atc tta gtt cat tta gag	1281
	Phe Arg Gln Ile Ala Val Asp Leu Lys Asp Ile Leu Val His Leu Glu	
	345 350 355	
50	att gat aaa gtt ata ttg gta ggc cat agc gat ggt gcc aat tta gct	1329
	Ile Asp Lys Val Ile Leu Val Gly His Ser Asp Gly Ala Asn Leu Ala	
	360 365 370	
55	ttt gtt ttt caa acg atg ttt cca ggt atg gtt aga ggg ctt ttg ctt	1377
	Leu Val Phe Gln Thr Met Phe Pro Gly Met Val Arg Gly Leu Leu Leu	
	375 380 385 390	
60	aat tca ggg aac ctg act att cat ggt cag cga tgg tgg gat att ctt	1425
	Asn Ser Gly Asn Leu Thr Ile His Gly Gln Arg Trp Trp Asp Ile Leu	
	395 400 405	
65	ttt gta agg att gcc tat aaa ttc ctt cac tat tta ggg aaa ctc ttt	1473
	Leu Val Arg Ile Ala Tyr Lys Phe Leu His Tyr Leu Gly Lys Leu Phe	
	410 415 420	

ES 2 278 436 T3

	ccg tat atg agg caa aaa gct caa gtt att tcg ctt atg ttg gag gat	1521
	Pro Tyr Met Arg Gln Lys Ala Gln Val Ile Ser Leu Met Leu Glu Asp	
	425 430 435	
5	ttg aag att agt cca gct gat tta cag cat gtg tca act cct gta atg	1569
	Leu Lys Ile Ser Pro Ala Asp Leu Gln His Val Ser Thr Pro Val Met	
	440 445 450	
10	gtt ttg gtt gga aat aag gac ata att aag tta aat cat tct aag aaa	1617
	Val Leu Val Gly Asn Lys Asp Ile Ile Lys Leu Asn His Ser Lys Lys	
	455 460 465 470	
15	ctt gct tct tat ttt cca agg ggg gag ttt tat tct tta gtt ggc ttt	1665
	Leu Ala Ser Tyr Phe Pro Arg Gly Glu Phe Tyr Ser Leu Val Gly Phe	
	475 480 485	
20	ggg cat cac att att aag caa gat tcc cat gtt ttt aat att att gca	1713
	Gly His His Ile Ile Lys Gln Asp Ser His Val Phe Asn Ile Ile Ala	
	490 495 500	
25	aaa aag ttt atc aac gat acg ttg aaa gga gaa att gtt gaa aaa gct	1761
	Lys Lys Phe Ile Asn Asp Thr Leu Lys Gly Glu Ile Val Glu Lys Ala	
	505 510 515	
30	aat tga aaaagtcaaaa tcaactgactt ctgtgattaa aattgtattt tttatatctg	1817
	Asn *	
35	ttttagtgc tattattggt gaa atg att cat ttg aaa cga act att tct gtt	1870
	Met Ile His Leu Lys Arg Thr Ile Ser Val	
	520 525	
40	gag caa cta aag agt gtt ttt ggg caa tta tct cca atg aat ctt ttc	1918
	Glu Gln Leu Lys Ser Val Phe Gly Gln Leu Ser Pro Met Asn Leu Phe	
	530 535 540 545	
45	ttt att atc ctt gtg ggg gtt atc gct gtc tta ccg aca acc gga tat	1966
	Leu Ile Ile Leu Val Gly Val Ile Ala Val Leu Pro Thr Thr Gly Tyr	
	550 555 560	
50	gac ttt gta ctg aat gga ctt tta cgt aca gat aaa agc aaa agg tat	2014
	Asp Phe Val Leu Asn Gly Leu Leu Arg Thr Asp Lys Ser Lys Arg Tyr	
	565 570 575	
55	att tta cag act agt tgg tgt atc aac act ttt aat aac ttg tca gga	2062
	Ile Leu Gln Thr Ser Trp Cys Ile Asn Thr Phe Asn Asn Leu Ser Gly	
	580 585 590	
60	ttc ggt ggc tta atc gat att ggg ttg cgc atg gct ttt tat ggt aaa	2110
	Phe Gly Gly Leu Ile Asp Ile Gly Leu Arg Met Ala Phe Tyr Gly Lys	
	595 600 605	
65	aaa ggt caa gag aag agt gac cta aga gaa gtg act cgt ttt tta ccc	2158
	Lys Gly Gln Glu Lys Ser Asp Leu Arg Glu Val Thr Arg Phe Leu Pro	
	610 615 620 625	

ES 2 278 436 T3

	tat ctt att tct ggt ctg tca ttt att agt gtg att gcc tta atc atg	2206
	Tyr Leu Ile Ser Gly Leu Ser Phe Ile Ser Val Ile Ala Leu Ile Met	
	630 635 640	
5	agc cat att ttt cat gcc aaa gct agt gtt gat tac tat tat ttg gta	2254
	Ser His Ile Phe His Ala Lys Ala Ser Val Asp Tyr Tyr Tyr Leu Val	
	645 650 655	
10	tta att ggt gct agt atg tat ttt cct gtt att tat tgg att tct ggt	2302
	Leu Ile Gly Ala Ser Met Tyr Phe Pro Val Ile Tyr Trp Ile Ser Gly	
	660 665 670	
15	cat aaa gga agc cat tat ttc gga gat atg cca tct agt act cgt ata	2350
	His Lys Gly Ser His Tyr Phe Gly Asp Met Pro Ser Ser Thr Arg Ile	
	675 680 685	
20	aaa tta ggt gtt gtt tct ttt ttt gaa tgg gga tgt gcg gcc gca gca	2398
	Lys Leu Gly Val Val Ser Phe Phe Glu Trp Gly Cys Ala Ala Ala Ala	
	690 695 700 705	
25	ttt ata att atc ggt tat tta atg ggc att cat cta cca gtt tat aaa	2446
	Phe Ile Ile Ile Gly Tyr Leu Met Gly Ile His Leu Pro Val Tyr Lys	
	710 715 720	
30	att tta cca cta ttt tgt att ggt tgt gcc gtc ggg att gta tcc ctt	2494
	Ile Leu Pro Leu Phe Cys Ile Gly Cys Ala Val Gly Ile Val Ser Leu	
	725 730 735	
35	att ccc ggt gga tta gga agt ttt gaa tta gtt cta ttt aca ggg ttt	2542
	Ile Pro Gly Gly Leu Gly Ser Phe Glu Leu Val Leu Phe Thr Gly Phe	
	740 745 750	
40	gct gcc gag gga cta cct aaa gaa act gtg gtt gca tgg tta tta ctt	2590
	Ala Ala Glu Gly Leu Pro Lys Glu Thr Val Val Ala Trp Leu Leu Leu	
	755 760 765	
45	tat cgt tta gcc tac tat att att cca ttc ttt gca ggt atc tat ttc	2638
	Tyr Arg Leu Ala Tyr Tyr Ile Ile Pro Phe Phe Ala Gly Ile Tyr Phe	
	770 775 780 785	
50	ttt atc cat tat tta ggt agt caa ata aat caa cgt tat gaa aat gtc	2686
	Phe Ile His Tyr Leu Gly Ser Gln Ile Asn Gln Arg Tyr Glu Asn Val	
	790 795 800	
55	ccg aaa gag tta gta tca act gtt cta caa acc atg gtg agc cat ttg	2734
	Pro Lys Glu Leu Val Ser Thr Val Leu Gln Thr Met Val Ser His Leu	
	805 810 815	
60	atg cgt att tta ggt gca ttc tta ata ttt tca aca gca ttt ttt gaa	2782
	Met Arg Ile Leu Gly Ala Phe Leu Ile Phe Ser Thr Ala Phe Phe Glu	
	820 825 830	
65	aat att act tat att atg tgg ttg cag aag cta ggc ttg gac cca tta	2830
	Asn Ile Thr Tyr Ile Met Trp Leu Gln Lys Leu Gly Leu Asp Pro Leu	
	835 840 845	

ES 2 278 436 T3

	caa gaa caa atg tta tgg cag ttt cca ggt tta ttg ctg ggg gtt tgt	2878
	Gln Glu Gln Met Leu Trp Gln Phe Pro Gly Leu Leu Leu Gly Val Cys	
	850 855 860 865	
5	ttt att ctc tta gct aga act att gat caa aaa gtg aaa aat gct ttt	2926
	Phe Ile Leu Leu Ala Arg Thr Ile Asp Gln Lys Val Lys Asn Ala Phe	
	870 875 880	
10	cca att gct att atc tgg att act ttg aca ttg ttt tat ctt aat tta	2974
	Pro Ile Ala Ile Ile Trp Ile Thr Leu Thr Leu Phe Tyr Leu Asn Leu	
	885 890 895	
15	ggt cat att agt tgg cga cta tct ttc tgg ttt att tta cta ttg tta	3022
	Gly His Ile Ser Trp Arg Leu Ser Phe Trp Phe Ile Leu Leu Leu Leu	
	900 905 910	
20	ggc tta tta gtc att aag cca act ctc tat aaa aaa caa ttt att tat	3070
	Gly Leu Leu Val Ile Lys Pro Thr Leu Tyr Lys Lys Gln Phe Ile Tyr	
	915 920 925	
25	agc tgg gaa gag cgt att aag gat gga atc att atc gtt agt tta atg	3118
	Ser Trp Glu Glu Arg Ile Lys Asp Gly Ile Ile Ile Val Ser Leu Met	
	930 935 940 945	
30	gga gtt cta ttt tat att gca gga cta cta ttc cct atc agg gct cat	3166
	Gly Val Leu Phe Tyr Ile Ala Gly Leu Leu Phe Pro Ile Arg Ala His	
	950 955 960	
35	att aca ggt ggt agt att gaa cgc ctg cat tat atc ata gca tgg gag	3214
	Ile Thr Gly Gly Ser Ile Glu Arg Leu His Tyr Ile Ile Ala Trp Glu	
	965 970 975	
40	ccg ata gca ttg gct acg ttg att ctt act ctc gtt tat tta tgt ttg	3262
	Pro Ile Ala Leu Ala Thr Leu Ile Leu Thr Leu Val Tyr Leu Cys Leu	
	980 985 990	
45	gtt aag att tta caa gga aaa tct tgt cag att ggt gat gtg ttc aat	3310
	Val Lys Ile Leu Gln Gly Lys Ser Cys Gln Ile Gly Asp Val Phe Asn	
	995 1000 1005	
50	gtg gat cgt tat aaa aaa cta ctt caa gct tac ggt ggt tct tcg gat	3358
	Val Asp Arg Tyr Lys Lys Leu Leu Gln Ala Tyr Gly Gly Ser Ser Asp	
	1010 1015 1020 1025	
55	agc ggt tta gcc ttt tta aat gat aaa agg ctc tac tgg tac caa aaa	3406
	Ser Gly Leu Ala Phe Leu Asn Asp Lys Arg Leu Tyr Trp Tyr Gln Lys	
	1030 1035 1040	
60	aat gga gaa gat tgc gtt gcg ttc caa ttt gta att gtc aat aat aaa	3454
	Asn Gly Glu Asp Cys Val Ala Phe Gln Phe Val Ile Val Asn Asn Lys	
	1045 1050 1055	
65	tgt ctt att atg ggg gaa cca gcc ggt gat gac act tat att cgt gaa	3502
	Cys Leu Ile Met Gly Glu Pro Ala Gly Asp Asp Thr Tyr Ile Arg Glu	
	1060 1065 1070	

ES 2 278 436 T3

5	gct att gaa tcg ttc att gat gat gct gat aag cta gac tat gac ctt Ala Ile Glu Ser Phe Ile Asp Asp Ala Asp Lys Leu Asp Tyr Asp Leu 1075 1080 1085	3550
10	ggt ttt tac agt att gga cag aag ttg aca cta ctt tta cat gag tat Val Phe Tyr Ser Ile Gly Gln Lys Leu Thr Leu Leu Leu His Glu Tyr 1090 1095 1100 1105	3598
15	ggt ttt gac ttt atg aaa ggt ggt gag gat gct tta gtt aat tta gaa Gly Phe Asp Phe Met Lys Val Gly Glu Asp Ala Leu Val Asn Leu Glu 1110 1115 1120	3646
20	acg ttt act ctt aaa ggg aat aag tac aaa cct ttc aga aat gcc cta Thr Phe Thr Leu Lys Gly Asn Lys Tyr Lys Pro Phe Arg Asn Ala Leu 1125 1130 1135	3694
25	aat aga gtt gaa aag gat ggt ttc tat ttc gaa gtt gta caa tcg cca Asn Arg Val Glu Lys Asp Gly Phe Tyr Phe Glu Val Val Gln Ser Pro 1140 1145 1150	3742
30	cat agt caa gag cta cta aat agt ttg gaa gag att tct aat act tgg His Ser Gln Glu Leu Leu Asn Ser Leu Glu Glu Ile Ser Asn Thr Trp 1155 1160 1165	3790
35	tta gaa gga cgt cct gaa aaa ggt ttc tca cta gga tat ttt aat aaa Leu Glu Gly Arg Pro Glu Lys Gly Phe Ser Leu Gly Tyr Phe Asn Lys 1170 1175 1180 1185	3838
40	gat tat ttc caa caa gcc cca ata gct ttg gta aaa aat gct gaa cac Asp Tyr Phe Gln Gln Ala Pro Ile Ala Leu Val Lys Asn Ala Glu His 1190 1195 1200	3886
45	gaa gtt gtt gct ttt gct aat att atg cca aac tat gaa aag agt att Glu Val Val Ala Phe Ala Asn Ile Met Pro Asn Tyr Glu Lys Ser Ile 1205 1210 1215	3934
50	atc ttt att gat tta atg cgt cac gat aaa cag aaa att ccg aat ggc Ile Ser Ile Asp Leu Met Arg His Asp Lys Gln Lys Ile Pro Asn Gly 1220 1225 1230	3982
55	ggt atg gat ttc ctc ttt tta tca tta ttc tct tat tat caa gag aag Val Met Asp Phe Leu Phe Leu Ser Leu Phe Ser Tyr Tyr Gln Glu Lys 1235 1240 1245	4030
60	gga tac cac tat ttt gat ttg ggg atg gca cct tta tca gga gtt ggt Gly Tyr His Tyr Phe Asp Leu Gly Met Ala Pro Leu Ser Gly Val Gly 1250 1255 1260 1265	4078
65	cgc gtt gaa aca agt ttt gct aaa gag aga atg gcg tat ctt gtc tat Arg Val Glu Thr Ser Phe Ala Lys Glu Arg Met Ala Tyr Leu Val Tyr 1270 1275 1280	4126
	cat ttc ggt agt cat ttc tac tca ttt aat ggt tta cac aag tat aag His Phe Gly Ser His Phe Tyr Ser Phe Asn Gly Leu His Lys Tyr Lys 1285 1290 1295	4174

ES 2 278 436 T3

```

aag aag ttt aca cca ttg tgg tgg gaa cgt tat att tct tgt tct cgt 4222
Lys Lys Phe Thr Pro Leu Trp Ser Glu Arg Tyr Ile Ser Cys Ser Arg
      1300                1305                1310

5   tgg tcc tgg tta att tgt gct att tgt gcc cta tta atg gaa gat agt 4270
    Ser Ser Trp Leu Ile Cys Ala Ile Cys Ala Leu Leu Met Glu Asp Ser
      1315                1320                1325

10  aaa att aag att gtt aaa taagctttat ttggcaatta aaaagagcat 4318
    Lys Ile Lys Ile Val Lys
      1330                1335

15  gtcattgagac atgctctttt taaatcattt aataccattg attgcttgaa tctactttat 4378
    aatatgatgt gcttttaaat attgttttagc tactgtagct gctgatttat gctttacagc 4438
    tacttggtag ttcatttcctt gcatttcctt ttcagtata tgaccagcaa gtttattgag 4498
    agcttttttt acctga 4514

20  <210> 2
    <211> 154
    <212> PRT
    <213> Streptococos

25  <400> 2

30  Ser Gly Lys Glu Pro Ala Asn Arg Phe Ser Trp Ala Lys Asn Lys Leu
     1      5      10      15
    Leu Ile Asn Gly Phe Ile Ala Thr Leu Ala Ala Thr Ile Leu Phe Phe
          20      25      30
    Ala Val Gln Phe Ile Gly Leu Lys Pro Asp Tyr Pro Gly Lys Thr Tyr
          35      40      45
    Phe Ile Ile Leu Leu Thr Ala Trp Thr Leu Met Ala Leu Val Thr Ala
          50      55      60
    Leu Val Gly Trp Asp Asn Arg Tyr Gly Ser Phe Leu Ser Leu Leu Ile
          65      70      75      80
    Leu Leu Phe Gln Leu Gly Ser Ser Ala Gly Thr Tyr Pro Ile Glu Leu
          85      90      95
    Ser Pro Lys Phe Phe Gln Thr Ile Gln Pro Phe Leu Pro Met Thr Tyr
          100     105     110
    Ser Val Ser Gly Leu Arg Glu Thr Ile Ser Leu Thr Gly Asp Val Asn
          115     120     125
    His Gln Trp Arg Met Leu Val Ile Phe Leu Val Ser Ser Met Ile Leu
          130     135     140
    Ala Leu Leu Ile Tyr Arg Lys Gln Glu Asp
          145     150

50  <210> 3
    <211> 118
    <212> PRT
    <213> Streptococos

    <400> 3

60  Met Ser Thr Leu Thr Ile Ile Ile Ala Thr Leu Thr Ala Leu Glu His
     1      5      10      15
    Phe Tyr Ile Met Tyr Leu Glu Thr Leu Ala Thr Gln Ser Asn Met Thr
          20      25      30

```

ES 2 278 436 T3

Gly Lys Ile Phe Ser Met Ser Lys Glu Glu Leu Ser Tyr Leu Pro Val
 35 40 45
 5 Ile Lys Leu Phe Lys Asn Gln Gly Val Tyr Asn Gly Leu Ile Gly Leu
 50 55 60
 Phe Leu Leu Tyr Gly Leu Tyr Ile Ser Gln Asn Gln Glu Ile Val Ala
 65 70 75 80
 Val Phe Leu Ile Asn Val Leu Leu Val Ala Ile Tyr Gly Ala Leu Thr
 85 90 95
 10 Val Asp Lys Lys Ile Leu Leu Lys Gln Gly Gly Leu Pro Ile Leu Ala
 100 105 110
 Leu Leu Thr Phe Leu Phe
 115

15 <210> 4
 <211> 247
 <212> PRT
 20 <213> Estreptococos
 <400> 4

25 Met Thr Glu Asn Trp Leu His Thr Lys Asp Gly Ser Asp Ile Tyr Tyr
 1 5 10 15
 Arg Val Val Gly Gln Gly Gln Pro Ile Val Phe Leu His Gly Asn Ser
 20 25 30
 30 Leu Ser Ser Arg Tyr Phe Asp Lys Gln Ile Ala Tyr Phe Ser Lys Tyr
 35 40 45
 Tyr Gln Val Ile Val Met Asp Ser Arg Gly His Gly Lys Ser His Ala
 50 55 60
 Lys Leu Asn Thr Ile Ser Phe Arg Gln Ile Ala Val Asp Leu Lys Asp
 35 65 70 75 80
 Ile Leu Val His Leu Glu Ile Asp Lys Val Ile Leu Val Gly His Ser
 85 90 95
 Asp Gly Ala Asn Leu Ala Leu Val Phe Gln Thr Met Phe Pro Gly Met
 100 105 110
 40 Val Arg Gly Leu Leu Leu Asn Ser Gly Asn Leu Thr Ile His Gly Gln
 115 120 125
 Arg Trp Trp Asp Ile Leu Leu Val Arg Ile Ala Tyr Lys Phe Leu His
 130 135 140
 45 Tyr Leu Gly Lys Leu Phe Pro Tyr Met Arg Gln Lys Ala Gln Val Ile
 145 150 155 160
 Ser Leu Met Leu Glu Asp Leu Lys Ile Ser Pro Ala Asp Leu Gln His
 165 170 175
 Val Ser Thr Pro Val Met Val Leu Val Gly Asn Lys Asp Ile Ile Lys
 180 185 190
 50 Leu Asn His Ser Lys Lys Leu Ala Ser Tyr Phe Pro Arg Gly Glu Phe
 195 200 205
 Tyr Ser Leu Val Gly Phe Gly His His Ile Ile Lys Gln Asp Ser His
 210 215 220
 55 Val Phe Asn Ile Ile Ala Lys Lys Phe Ile Asn Asp Thr Leu Lys Gly
 225 230 235 240
 Glu Ile Val Glu Lys Ala Asn
 245

60 <210> 5
 <211> 816
 <212> PRT
 65 <213> Estreptococos

ES 2 278 436 T3

<400> 5

5	Met	Ile	His	Leu	Lys	Arg	Thr	Ile	Ser	Val	Glu	Gln	Leu	Lys	Ser	Val
	1				5					10					15	
	Phe	Gly	Gln	Leu	Ser	Pro	Met	Asn	Leu	Phe	Leu	Ile	Ile	Leu	Val	Gly
				20					25					30		
	Val	Ile	Ala	Val	Leu	Pro	Thr	Thr	Gly	Tyr	Asp	Phe	Val	Leu	Asn	Gly
			35					40					45			
10	Leu	Leu	Arg	Thr	Asp	Lys	Ser	Lys	Arg	Tyr	Ile	Leu	Gln	Thr	Ser	Trp
	50						55					60				
	Cys	Ile	Asn	Thr	Phe	Asn	Asn	Leu	Ser	Gly	Phe	Gly	Gly	Leu	Ile	Asp
	65					70					75				80	
15	Ile	Gly	Leu	Arg	Met	Ala	Phe	Tyr	Gly	Lys	Lys	Gly	Gln	Glu	Lys	Ser
				85					90						95	
	Asp	Leu	Arg	Glu	Val	Thr	Arg	Phe	Leu	Pro	Tyr	Leu	Ile	Ser	Gly	Leu
				100					105						110	
20	Ser	Phe	Ile	Ser	Val	Ile	Ala	Leu	Ile	Met	Ser	His	Ile	Phe	His	Ala
				115					120					125		
	Lys	Ala	Ser	Val	Asp	Tyr	Tyr	Tyr	Leu	Val	Leu	Ile	Gly	Ala	Ser	Met
		130					135						140			
	Tyr	Phe	Pro	Val	Ile	Tyr	Trp	Ile	Ser	Gly	His	Lys	Gly	Ser	His	Tyr
	145					150					155				160	
25	Phe	Gly	Asp	Met	Pro	Ser	Ser	Thr	Arg	Ile	Lys	Leu	Gly	Val	Val	Ser
				165						170					175	
	Phe	Phe	Glu	Trp	Gly	Cys	Ala	Ala	Ala	Ala	Phe	Ile	Ile	Ile	Gly	Tyr
				180					185						190	
30	Leu	Met	Gly	Ile	His	Leu	Pro	Val	Tyr	Lys	Ile	Leu	Pro	Leu	Phe	Cys
			195					200					205			
	Ile	Gly	Cys	Ala	Val	Gly	Ile	Val	Ser	Leu	Ile	Pro	Gly	Gly	Leu	Gly
		210					215					220				
35	Ser	Phe	Glu	Leu	Val	Leu	Phe	Thr	Gly	Phe	Ala	Ala	Glu	Gly	Leu	Pro
	225					230					235				240	
	Lys	Glu	Thr	Val	Val	Ala	Trp	Leu	Leu	Leu	Tyr	Arg	Leu	Ala	Tyr	Tyr
				245						250					255	
	Ile	Ile	Pro	Phe	Phe	Ala	Gly	Ile	Tyr	Phe	Phe	Ile	His	Tyr	Leu	Gly
				260					265					270		
40	Ser	Gln	Ile	Asn	Gln	Arg	Tyr	Glu	Asn	Val	Pro	Lys	Glu	Leu	Val	Ser
		275						280					285			
	Thr	Val	Leu	Gln	Thr	Met	Val	Ser	His	Leu	Met	Arg	Ile	Leu	Gly	Ala
		290				295						300				
45	Phe	Leu	Ile	Phe	Ser	Thr	Ala	Phe	Phe	Glu	Asn	Ile	Thr	Tyr	Ile	Met
	305					310					315				320	
	Trp	Leu	Gln	Lys	Leu	Gly	Leu	Asp	Pro	Leu	Gln	Glu	Gln	Met	Leu	Trp
				325						330					335	
50	Gln	Phe	Pro	Gly	Leu	Leu	Leu	Gly	Val	Cys	Phe	Ile	Leu	Leu	Ala	Arg
				340					345					350		
	Thr	Ile	Asp	Gln	Lys	Val	Lys	Asn	Ala	Phe	Pro	Ile	Ala	Ile	Ile	Trp
			355					360					365			
	Ile	Thr	Leu	Thr	Leu	Phe	Tyr	Leu	Asn	Leu	Gly	His	Ile	Ser	Trp	Arg
		370				375						380				
55	Leu	Ser	Phe	Trp	Phe	Ile	Leu	Leu	Leu	Leu	Gly	Leu	Leu	Val	Ile	Lys
	385					390					395				400	
	Pro	Thr	Leu	Tyr	Lys	Lys	Gln	Phe	Ile	Tyr	Ser	Trp	Glu	Glu	Arg	Ile
				405						410					415	
60	Lys	Asp	Gly	Ile	Ile	Ile	Val	Ser	Leu	Met	Gly	Val	Leu	Phe	Tyr	Ile
				420					425					430		

65

ES 2 278 436 T3

Ala Gly Leu Leu Phe Pro Ile Arg Ala His Ile Thr Gly Gly Ser Ile
 435 440 445
 Glu Arg Leu His Tyr Ile Ile Ala Trp Glu Pro Ile Ala Leu Ala Thr
 450 455 460
 Leu Ile Leu Thr Leu Val Tyr Leu Cys Leu Val Lys Ile Leu Gln Gly
 465 470 475 480
 Lys Ser Cys Gln Ile Gly Asp Val Phe Asn Val Asp Arg Tyr Lys Lys
 485 490 495
 Leu Leu Gln Ala Tyr Gly Gly Ser Ser Asp Ser Gly Leu Ala Phe Leu
 500 505 510
 Asn Asp Lys Arg Leu Tyr Trp Tyr Gln Lys Asn Gly Glu Asp Cys Val
 515 520 525
 Ala Phe Gln Phe Val Ile Val Asn Asn Lys Cys Leu Ile Met Gly Glu
 530 535 540
 Pro Ala Gly Asp Asp Thr Tyr Ile Arg Glu Ala Ile Glu Ser Phe Ile
 545 550 555 560
 Asp Asp Ala Asp Lys Leu Asp Tyr Asp Leu Val Phe Tyr Ser Ile Gly
 565 570 575
 Gln Lys Leu Thr Leu Leu Leu His Glu Tyr Gly Phe Asp Phe Met Lys
 580 585 590
 Val Gly Glu Asp Ala Leu Val Asn Leu Glu Thr Phe Thr Leu Lys Gly
 595 600 605
 Asn Lys Tyr Lys Pro Phe Arg Asn Ala Leu Asn Arg Val Glu Lys Asp
 610 615 620
 Gly Phe Tyr Phe Glu Val Val Gln Ser Pro His Ser Gln Glu Leu Leu
 625 630 635 640
 Asn Ser Leu Glu Glu Ile Ser Asn Thr Trp Leu Glu Gly Arg Pro Glu
 645 650 655
 Lys Gly Phe Ser Leu Gly Tyr Phe Asn Lys Asp Tyr Phe Gln Gln Ala
 660 665 670
 Pro Ile Ala Leu Val Lys Asn Ala Glu His Glu Val Val Ala Phe Ala
 675 680 685
 Asn Ile Met Pro Asn Tyr Glu Lys Ser Ile Ile Ser Ile Asp Leu Met
 690 695 700
 Arg His Asp Lys Gln Lys Ile Pro Asn Gly Val Met Asp Phe Leu Phe
 705 710 715 720
 Leu Ser Leu Phe Ser Tyr Tyr Gln Glu Lys Gly Tyr His Tyr Phe Asp
 725 730 735
 Leu Gly Met Ala Pro Leu Ser Gly Val Gly Arg Val Glu Thr Ser Phe
 740 745 750
 Ala Lys Glu Arg Met Ala Tyr Leu Val Tyr His Phe Gly Ser His Phe
 755 760 765
 Tyr Ser Phe Asn Gly Leu His Lys Tyr Lys Lys Lys Phe Thr Pro Leu
 770 775 780
 Trp Ser Glu Arg Tyr Ile Ser Cys Ser Arg Ser Ser Trp Leu Ile Cys
 785 790 795 800
 Ala Ile Cys Ala Leu Leu Met Glu Asp Ser Lys Ile Lys Ile Val Lys
 805 810 815

55 <210> 6
 <211> 518
 <212> PRT
 60 <213> Estreptococos

65

ES 2 278 436 T3

<400> 6

5	Met	Arg	Ile	Leu	Gly	Ala	Phe	Leu	Ile	Phe	Ser	Thr	Ala	Phe	Phe	Glu
	1				5					10					15	
	Asn	Ile	Thr	Tyr	Ile	Met	Trp	Leu	Gln	Lys	Leu	Gly	Leu	Asp	Pro	Leu
				20					25					30		
	Gln	Glu	Gln	Met	Leu	Trp	Gln	Phe	Pro	Gly	Leu	Leu	Leu	Gly	Val	Cys
				35				40					45			
10	Phe	Ile	Leu	Leu	Ala	Arg	Thr	Ile	Asp	Gln	Lys	Val	Lys	Asn	Ala	Phe
	50						55					60				
	Pro	Ile	Ala	Ile	Ile	Trp	Ile	Thr	Leu	Thr	Leu	Phe	Tyr	Leu	Asn	Leu
	65					70					75					80
15	Gly	His	Ile	Ser	Trp	Arg	Leu	Ser	Phe	Trp	Phe	Ile	Leu	Leu	Leu	Leu
					85					90					95	
	Gly	Leu	Leu	Val	Ile	Lys	Pro	Thr	Leu	Tyr	Lys	Lys	Gln	Phe	Ile	Tyr
				100					105					110		
	Ser	Trp	Glu	Glu	Arg	Ile	Lys	Asp	Gly	Ile	Ile	Ile	Val	Ser	Leu	Met
20				115				120					125			
	Gly	Val	Leu	Phe	Tyr	Ile	Ala	Gly	Leu	Leu	Phe	Pro	Ile	Arg	Ala	His
		130					135					140				
	Ile	Thr	Gly	Gly	Ser	Ile	Glu	Arg	Leu	His	Tyr	Ile	Ile	Ala	Trp	Glu
	145					150					155					160
25	Pro	Ile	Ala	Leu	Ala	Thr	Leu	Ile	Leu	Thr	Leu	Val	Tyr	Leu	Cys	Leu
					165					170					175	
	Val	Lys	Ile	Leu	Gln	Gly	Lys	Ser	Cys	Gln	Ile	Gly	Asp	Val	Phe	Asn
				180					185					190		
30	Val	Asp	Arg	Tyr	Lys	Lys	Leu	Leu	Gln	Ala	Tyr	Gly	Gly	Ser	Ser	Asp
		195						200					205			
	Ser	Gly	Leu	Ala	Phe	Leu	Asn	Asp	Lys	Arg	Leu	Tyr	Trp	Tyr	Gln	Lys
		210					215					220				
35	Asn	Gly	Glu	Asp	Cys	Val	Ala	Phe	Gln	Phe	Val	Ile	Val	Asn	Asn	Lys
	225					230					235					240
	Cys	Leu	Ile	Met	Gly	Glu	Pro	Ala	Gly	Asp	Asp	Thr	Tyr	Ile	Arg	Glu
				245						250					255	
	Ala	Ile	Glu	Ser	Phe	Ile	Asp	Asp	Ala	Asp	Lys	Leu	Asp	Tyr	Asp	Leu
				260					265					270		
40	Val	Phe	Tyr	Ser	Ile	Gly	Gln	Lys	Leu	Thr	Leu	Leu	Leu	His	Glu	Tyr
		275						280					285			
	Gly	Phe	Asp	Phe	Met	Lys	Val	Gly	Glu	Asp	Ala	Leu	Val	Asn	Leu	Glu
		290					295					300				
45	Thr	Phe	Thr	Leu	Lys	Gly	Asn	Lys	Tyr	Lys	Pro	Phe	Arg	Asn	Ala	Leu
	305					310					315					320
	Asn	Arg	Val	Glu	Lys	Asp	Gly	Phe	Tyr	Phe	Glu	Val	Val	Gln	Ser	Pro
					325					330					335	
	His	Ser	Gln	Glu	Leu	Leu	Asn	Ser	Leu	Glu	Glu	Ile	Ser	Asn	Thr	Trp
50				340					345					350		
	Leu	Glu	Gly	Arg	Pro	Glu	Lys	Gly	Phe	Ser	Leu	Gly	Tyr	Phe	Asn	Lys
			355					360					365			
	Asp	Tyr	Phe	Gln	Gln	Ala	Pro	Ile	Ala	Leu	Val	Lys	Asn	Ala	Glu	His
		370					375					380				
55	Glu	Val	Val	Ala	Phe	Ala	Asn	Ile	Met	Pro	Asn	Tyr	Glu	Lys	Ser	Ile
	385					390					395					400
	Ile	Ser	Ile	Asp	Leu	Met	Arg	His	Asp	Lys	Gln	Lys	Ile	Pro	Asn	Gly
				405						410					415	
60	Val	Met	Asp	Phe	Leu	Phe	Leu	Ser	Leu	Phe	Ser	Tyr	Tyr	Gln	Glu	Lys
				420					425					430		
	Gly	Tyr	His	Tyr	Phe	Asp	Leu	Gly	Met	Ala	Pro	Leu	Ser	Gly	Val	Gly
			435					440					445			

65

ES 2 278 436 T3

```

      Arg Val Glu Thr Ser Phe Ala Lys Glu Arg Met Ala Tyr Leu Val Tyr
      450          455          460
5     His Phe Gly Ser His Phe Tyr Ser Phe Asn Gly Leu His Lys Tyr Lys
      465          470          475          480
      Lys Lys Phe Thr Pro Leu Trp Ser Glu Arg Tyr Ile Ser Cys Ser Arg
      485          490          495
      Ser Ser Trp Leu Ile Cys Ala Ile Cys Ala Leu Leu Met Glu Asp Ser
      500          505          510
10    Lys Ile Lys Ile Val Lys
      515

```

<210> 7

15 <211> 5126

<212> Adn

<213> Estreptococos

20 <220>

<221> CDS

<222> (1)...(687)

25 <221> CDS

<222> (701)...(2557)

<221> CDS

30 <222> (2566)...(3036)

<221> CDS

35 <222> (3106)...(4842)

<221> CDS

<222> (4850)...(5125)

40 <400> 7

```

      aat ttt gat atc gaa aca aca act ttt gag gca atg aaa aag cac gcg      48
      Asn Phe Asp Ile Glu Thr Thr Thr Phe Glu Ala Met Lys Lys His Ala
45      1          5          10          15

      tca tta ttg gag aaa ata tct gtt gag cgt tct ttt att gaa ttt gat      96
      Ser Leu Leu Glu Lys Ile Ser Val Glu Arg Ser Phe Ile Glu Phe Asp
      20          25          30

50    aaa ctt cta tta gca cct tat tgg cgt aaa gga atg ctg gca cta ata      144
      Lys Leu Leu Leu Ala Pro Tyr Trp Arg Lys Gly Met Leu Ala Leu Ile
      35          40          45

55    gat agt cat gct ttt aat tat cta cca tgc tta aaa aat agg gaa tta      192
      Asp Ser His Ala Phe Asn Tyr Leu Pro Cys Leu Lys Asn Arg Glu Leu
      50          55          60

60    caa tta agc gcc ttt ttg tcc cag tta gat aaa gat ttt tta ttt gag      240
      Gln Leu Ser Ala Phe Leu Ser Gln Leu Asp Lys Asp Phe Leu Phe Glu
      65          70          75          80

65    aca tca gaa caa gct tgg gca tca ctc atc ttg agt atg gaa gtt gaa      288
      Thr Ser Glu Gln Ala Trp Ala Ser Leu Ile Leu Ser Met Glu Val Glu
      85          90          95

```

ES 2 278 436 T3

	cac aca aag act ttt tta aaa aaa tgg aag aca tca act cac ttt caa	336
	His Thr Lys Thr Phe Leu Lys Lys Trp Lys Thr Ser Thr His Phe Gln	
	100 105 110	
5	aaa gat gtt gag cat ata gtg gat gtt tat cgt att cgt gaa caa atg	384
	Lys Asp Val Glu His Ile Val Asp Val Tyr Arg Ile Arg Glu Gln Met	
	115 120 125	
10	gga ttg gct aaa gaa cat ctt tat cgt tat gga aaa act ata ata aaa	432
	Gly Leu Ala Lys Glu His Leu Tyr Arg Tyr Gly Lys Thr Ile Ile Lys	
	130 135 140	
15	caa gcg gaa ggt att cgc aaa gca aga ggc ttg atg gtt gat ttc gaa	480
	Gln Ala Glu Gly Ile Arg Lys Ala Arg Gly Leu Met Val Asp Phe Glu	
	145 150 155 160	
20	aaa ata gaa caa cta gat agt gag tta gca atc cat gat agg cat gag	528
	Lys Ile Glu Gln Leu Asp Ser Glu Leu Ala Ile His Asp Arg His Glu	
	165 170 175	
25	ata gtt gtc aat ggt ggc acc tta atc aag aaa tta gga ata aaa cct	576
	Ile Val Val Asn Gly Gly Thr Leu Ile Lys Lys Leu Gly Ile Lys Pro	
	180 185 190	
30	ggt cca cag atg gga gat att atc tct caa att gaa tta gcc att gtt	624
	Gly Pro Gln Met Gly Asp Ile Ile Ser Gln Ile Glu Leu Ala Ile Val	
	195 200 205	
35	tta gga caa ctg att aat gaa gaa gag gct att tta cat ttt gtt aag	672
	Leu Gly Gln Leu Ile Asn Glu Glu Ala Ile Leu His Phe Val Lys	
	210 215 220	
40	cag tac ttg atg gat tagagaggat tat atg agc gat ttt tta gta gat	721
	Gln Tyr Leu Met Asp Met Ser Asp Phe Leu Val Asp	
	225 230 235	
45	gga ttg act aag tcg gtt ggt gat aag acg gtc ttt agt aat gtt tca	769
	Gly Leu Thr Lys Ser Val Gly Asp Lys Thr Val Phe Ser Asn Val Ser	
	240 245 250	
50	ttt atc atc cat agt tta gac cgt att ggg att att ggt gtc aat gga	817
	Phe Ile Ile His Ser Leu Asp Arg Ile Gly Ile Ile Gly Val Asn Gly	
	255 260 265	
55	act gga aag aca aca cta tta gat gtt att tcg ggt gaa tta ggt ttt	865
	Thr Gly Lys Thr Thr Leu Leu Asp Val Ile Ser Gly Glu Leu Gly Phe	
	270 275 280	
60	gat ggt gat cgt tcc cct ttt tca tca gct aat gat tat aag att gct	913
	Asp Gly Asp Arg Ser Pro Phe Ser Ser Ala Asn Asp Tyr Lys Ile Ala	
	285 290 295 300	
65	tat tta aaa caa gaa cca gac ttt gat gat tct cag aca att ttg gac	961
	Tyr Leu Lys Gln Glu Pro Asp Phe Asp Ser Gln Thr Ile Leu Asp	
	305 310 315	

ES 2 278 436 T3

	acc gta ctt tct tct gac tta aga gag atg gct tta att aaa gaa tat	1009
	Thr Val Leu Ser Ser Asp Leu Arg Glu Met Ala Leu Ile Lys Glu Tyr	
	320 325 330	
5		
	gaa tta ttg ctt aat cac tac gaa gaa agt aag caa tca cgt cta gag	1057
	Glu Leu Leu Leu Asn His Tyr Glu Glu Ser Lys Gln Ser Arg Leu Glu	
	335 340 345	
10		
	aaa gta atg gca gaa atg gat tct tta gat gct tgg tct att gag agc	1105
	Lys Val Met Ala Glu Met Asp Ser Leu Asp Ala Trp Ser Ile Glu Ser	
	350 355 360	
15		
	gaa gtc aaa aca gta tta tcc aaa tta ggt att act gat ttg cag ttg	1153
	Glu Val Lys Thr Val Leu Ser Lys Leu Gly Ile Thr Asp Leu Gln Leu	
	365 370 375 380	
20		
	tcg gtt ggt gaa tta tca gga gga tta cga aga cgt gtt caa tta gcc	1201
	Ser Val Gly Glu Leu Ser Gly Gly Leu Arg Arg Arg Val Gln Leu Ala	
	385 390 395	
25		
	caa gta tta tta aat gat gca gat tta ttg ctc tta gac gaa cct act	1249
	Gln Val Leu Leu Asn Asp Ala Asp Leu Leu Leu Leu Asp Glu Pro Thr	
	400 405 410	
30		
	aac cac tta gat att gac act att gca tgg tta acg aat ttt ttg aaa	1297
	Asn His Leu Asp Ile Asp Thr Ile Ala Trp Leu Thr Asn Phe Leu Lys	
	415 420 425	
35		
	aat agt aaa aag aca gtg ctt ttt ata act cat gat cgt tat ttt cta	1345
	Asn Ser Lys Lys Thr Val Leu Phe Ile Thr His Asp Arg Tyr Phe Leu	
	430 435 440	
40		
	gac aat gtt gca aca cgt att ttt gaa tta gat aag gca cag att aca	1393
	Asp Asn Val Ala Thr Arg Ile Phe Glu Leu Asp Lys Ala Gln Ile Thr	
	445 450 455 460	
45		
	gaa tat caa ggc aat tat cag gat tat gtc cga ctt cgt gca gaa caa	1441
	Glu Tyr Gln Gly Asn Tyr Gln Asp Tyr Val Arg Leu Arg Ala Glu Gln	
	465 470 475	
50		
	gac gag cgt gat gct gct agt tta cat aaa aag aaa cag ctt tat aaa	1489
	Asp Glu Arg Asp Ala Ala Ser Leu His Lys Lys Lys Gln Leu Tyr Lys	
	480 485 490	
55		
	cag gaa cta gct tgg atg cgt act cag cca caa gct cgt gca acg aaa	1537
	Gln Glu Leu Ala Trp Met Arg Thr Gln Pro Gln Ala Arg Ala Thr Lys	
	495 500 505	
60		
	caa cag gct cgt att aat cgt ttt caa aat cta aaa aac gat tta cac	1585
	Gln Gln Ala Arg Ile Asn Arg Phe Gln Asn Leu Lys Asn Asp Leu His	
	510 515 520	
65		
	caa aca agc gat aca agc gat ttg gaa atg aca ttt gaa aca agt cga	1633
	Gln Thr Ser Asp Thr Ser Asp Leu Glu Met Thr Phe Glu Thr Ser Arg	
	525 530 535 540	

ES 2 278 436 T3

	att ggg aaa aag gtt att aat ttt gaa aat gtc tct ttt tct tac cca	1681
	Ile Gly Lys Lys Val Ile Asn Phe Glu Asn Val Ser Phe Ser Tyr Pro	
	545 550 555	
5	gat aaa tct atc ttg aaa gac ttt aat ttg tta att caa aat aaa gac	1729
	Asp Lys Ser Ile Leu Lys Asp Phe Asn Leu Leu Ile Gln Asn Lys Asp	
	560 565 570	
10	cgt att ggc atc gtt gga gat aat ggt gtt gga aag tca acc tta ctt	1777
	Arg Ile Gly Ile Val Gly Asp Asn Gly Val Gly Lys Ser Thr Leu Leu	
	575 580 585	
15	aat tta att gtt caa gat tta cag ccg gat tcg ggt aat gtc tct att	1825
	Asn Leu Ile Val Gln Asp Leu Gln Pro Asp Ser Gly Asn Val Ser Ile	
	590 595 600	
20	ggg gaa acg ata cgt gta ggt tac ttt tca caa caa ctt cat aat atg	1873
	Gly Glu Thr Ile Arg Val Gly Tyr Phe Ser Gln Gln Leu His Asn Met	
	605 610 615 620	
25	gat ggc tca aaa cgt gtt att aat tat ttg caa gag gtt gca gat gag	1921
	Asp Gly Ser Lys Arg Val Ile Asn Tyr Leu Gln Glu Val Ala Asp Glu	
	625 630 635	
30	gtt aaa act agt gtc ggt aca aca agt gtg aca gaa cta ttg gaa caa	1969
	Val Lys Thr Ser Val Gly Thr Thr Ser Val Thr Glu Leu Leu Glu Gln	
	640 645 650	
35	ttt ctc ttt cca cgt tcg aca cat gga aca caa att gca aaa tta tca	2017
	Phe Leu Phe Pro Arg Ser Thr His Gly Thr Gln Ile Ala Lys Leu Ser	
	655 660 665	
40	ggg ggt gag aaa aaa aga ctt tac ctt tta aaa atc ctg att gaa aag	2065
	Gly Gly Glu Lys Lys Arg Leu Tyr Leu Leu Lys Ile Leu Ile Glu Lys	
	670 675 680	
45	cct aat gtg tta cta ctt gat gag ccg aca aat gac tta gat att gct	2113
	Pro Asn Val Leu Leu Leu Asp Glu Pro Thr Asn Asp Leu Asp Ile Ala	
	685 690 695 700	
50	aca tta act gtt ctt gaa aat ttt tta caa ggc ttt ggt ggt cct gtg	2161
	Thr Leu Thr Val Leu Glu Asn Phe Leu Gln Gly Phe Gly Gly Pro Val	
	705 710 715	
55	att aca gtt agt cac gat cgt tac ttt tta gat aaa gtg gct aat aaa	2209
	Ile Thr Val Ser His Asp Arg Tyr Phe Leu Asp Lys Val Ala Asn Lys	
	720 725 730	
60	att att gcg ttt gaa gat aac gat atc cgt gaa ttt ttt ggt aat tat	2257
	Ile Ile Ala Phe Glu Asp Asn Asp Ile Arg Glu Phe Phe Gly Asn Tyr	
	735 740 745	
65	act gat tat tta gat gaa aaa gca ttt aat gag caa aat aat gaa gtt	2305
	Thr Asp Tyr Leu Asp Glu Lys Ala Phe Asn Glu Gln Asn Asn Glu Val	
	750 755 760	

ES 2 278 436 T3

	atc agt aaa aaa gag agt acc aag aca agt cgt gaa aag caa agt cgt	2353
	Ile Ser Lys Lys Glu Ser Thr Lys Thr Ser Arg Glu Lys Gln Ser Arg	
	765 770 775 780	
5	aaa aga atg tct tac ttt gaa aaa caa gaa tgg gcg aca att gaa gac	2401
	Lys Arg Met Ser Tyr Phe Glu Lys Gln Glu Trp Ala Thr Ile Glu Asp	
	785 790 795	
10	gat att atg ata ttg gaa aat act atc act cgt ata gaa aat gat atg	2449
	Asp Ile Met Ile Leu Glu Asn Thr Ile Thr Arg Ile Glu Asn Asp Met	
	800 805 810	
15	caa aca tgt ggt agt gat ttt aca agg tta tct gat tta caa aag gaa	2497
	Gln Thr Cys Gly Ser Asp Phe Thr Arg Leu Ser Asp Leu Gln Lys Glu	
	815 820 825	
20	tta gat gca aaa aat gaa gca ctt cta gaa aag tat gac cgt tat gag	2545
	Leu Asp Ala Lys Asn Glu Ala Leu Leu Glu Lys Tyr Asp Arg Tyr Glu	
	830 835 840	
25	tac ctt agt gag ttagacac atg att atc cgt ccg att att aaa aat gat	2595
	Tyr Leu Ser Glu LeuAspThrMet Ile Ile Arg Pro Ile Ile Lys Asn Asp	
	845 850 855 860	
30	gac caa gca gtt gca caa tta att cga caa agt tta cgc gcc tat gat	2643
	Asp Gln Ala Val Ala Gln Leu Ile Arg Gln Ser Leu Arg Ala Tyr Asp	
	865 870 875	
35	tta gat aaa cct gat aca gca tat tca gac cct cac tta gat cat ttg	2691
	Leu Asp Lys Pro Asp Thr Ala Tyr Ser Asp Pro His Leu Asp His Leu	
	880 885 890	
40	acc tca tac tac gaa aaa ata gag aag tca gga ttc ttt gtc att gag	2739
	Thr Ser Tyr Tyr Glu Lys Ile Glu Lys Ser Gly Phe Phe Val Ile Glu	
	895 900 905	
45	gag aga gat gag att att ggc tgt ggc ggc ttt ggt ccg ctg aaa aat	2787
	Glu Arg Asp Glu Ile Ile Gly Cys Gly Gly Phe Gly Pro Leu Lys Asn	
	910 915 920 925	
50	cta att gca gag atg cag aag gtg tac att gca gaa cgt ttc cgt ggt	2835
	Leu Ile Ala Glu Met Gln Lys Val Tyr Ile Ala Glu Arg Phe Arg Gly	
	930 935 940	
55	aag ggg ctt gct act gat tta gtg aaa atg att gaa gta gaa gct cga	2883
	Lys Gly Leu Ala Thr Asp Leu Val Lys Met Ile Glu Val Glu Ala Arg	
	945 950 955	
60	aaa att ggg tat aga caa ctt tat tta gag aca gcc agt act ttg agt	2931
	Lys Ile Gly Tyr Arg Gln Leu Tyr Leu Glu Thr Ala Ser Thr Leu Ser	
	960 965 970	
65	agg gca act gcg gtt tat aag cat atg gga tat tgt gcc tta tcg caa	2979
	Arg Ala Thr Ala Val Tyr Lys His Met Gly Tyr Cys Ala Leu Ser Gln	
	975 980 985	

ES 2 278 436 T3

	cca ata gca aat gat caa ggt cat aca gct atg gat att tgg atg att	3027
	Pro Ile Ala Asn Asp Gln Gly His Thr Ala Met Asp Ile Trp Met Ile	
	990 995 1000 1005	
5	aaa gat tta taagttgaaa gtggattagt gaacatggat taattatttt	3076
	Lys Asp Leu	
10	gagataagag gaaagaaaag gagacatat atg gca tat att tgg tct tat ttg	3129
	Met Ala Tyr Ile Trp Ser Tyr Leu	
	1010 1015	
15	aaa agg tac ccc aat tgg tta tgg ctt gat tta cta gga gct atg ctt	3177
	Lys Arg Tyr Pro Asn Trp Leu Trp Leu Asp Leu Leu Gly Ala Met Leu	
	1020 1025 1030	
20	ttt gtg acg gtt atc cta gga atg ccc aca gcc tta gcg ggt atg att	3225
	Phe Val Thr Val Ile Leu Gly Met Pro Thr Ala Leu Ala Gly Met Ile	
	1035 1040 1045	
25	gat aat ggc gtt aca aaa ggt gat cgg act gga gtt tat ctg tgg acg	3273
	Asp Asn Gly Val Thr Lys Gly Asp Arg Thr Gly Val Tyr Leu Trp Thr	
	1050 1055 1060	
30	ttc atc atg ttt ata ttt gtt gta cta ggt att att ggg cgt att acg	3321
	Phe Ile Met Phe Ile Phe Val Val Leu Gly Ile Ile Gly Arg Ile Thr	
	065 1070 1075 1080	
35	atg gct tac gca tct agt cgc tta acg aca aca atg att aga gat atg	3369
	Met Ala Tyr Ala Ser Ser Arg Leu Thr Thr Thr Met Ile Arg Asp Met	
	1085 1090 1095	
40	cgt aat gat atg tat gct aag ctt caa gaa tac tcc cat cat gaa tat	3417
	Arg Asn Asp Met Tyr Ala Lys Leu Gln Glu Tyr Ser His His Glu Tyr	
	1100 1105 1110	
45	gaa cag ata ggt gta tct tca cta gtg aca cgt atg aca agc gat act	3465
	Glu Gln Ile Gly Val Ser Ser Leu Val Thr Arg Met Thr Ser Asp Thr	
	1115 1120 1125	
50	ttt gtt ttg atg caa ttt gct gaa atg tct tta cgt tta ggc cta gta	3513
	Phe Val Leu Met Gln Phe Ala Glu Met Ser Leu Arg Leu Gly Leu Val	
	1130 1135 1140	
55	act cct atg gta atg att ttt agc gtg gtt atg ata cta att acg agt	3561
	Thr Pro Met Val Met Ile Phe Ser Val Val Met Ile Leu Ile Thr Ser	
	1145 1150 1155 1160	
60	cca tct ttg gct tgg ctt gta gcg gtt gcg atg cct ctt ttg gta gga	3609
	Pro Ser Leu Ala Trp Leu Val Ala Val Ala Met Pro Leu Leu Val Gly	
	1165 1170 1175	
65	gtc gtt tta tat gta gct ata aaa aca aaa cct tta tct gaa aga caa	3657
	Val Val Leu Tyr Val Ala Ile Lys Thr Lys Pro Leu Ser Glu Arg Gln	
	1180 1185 1190	

ES 2 278 436 T3

	cag act atg ctt gat aaa atc aat caa tat gtt cgt gaa aat tta aca	3705
	Gln Thr Met Leu Asp Lys Ile Asn Gln Tyr Val Arg Glu Asn Leu Thr	
	1195 1200 1205	
5	ggg tta cgc gtt gtt aga gcc ttt gca aga gag aat ttt caa tca caa	3753
	Gly Leu Arg Val Val Arg Ala Phe Ala Arg Glu Asn Phe Gln Ser Gln	
	1210 1215 1220	
10	aaa ttt caa gtc gct aac caa cgt tac aca gat act tca act ggt ctt	3801
	Lys Phe Gln Val Ala Asn Gln Arg Tyr Thr Asp Thr Ser Thr Gly Leu	
	1225 1230 1235 1240	
15	ttt aaa tta aca ggg cta aca gaa cca ctt ttc gtt caa att att att	3849
	Phe Lys Leu Thr Gly Leu Thr Glu Pro Leu Phe Val Gln Ile Ile Ile	
	1245 1250 1255	
20	gca atg att gtg gct atc gtt tgg ttc gct ttg gat ccc tta caa aga	3897
	Ala Met Ile Val Ala Ile Val Trp Phe Ala Leu Asp Pro Leu Gln Arg	
	1260 1265 1270	
25	ggg gct att aaa ata ggg gat tta gtt gct ttt atc gaa tat agc ttc	3945
	Gly Ala Ile Lys Ile Gly Asp Leu Val Ala Phe Ile Glu Tyr Ser Phe	
	1275 1280 1285	
30	cat gct ctc ttt tca ttt ttg cta ttt gcc aat ctt ttt act atg tat	3993
	His Ala Leu Phe Ser Phe Leu Leu Phe Ala Asn Leu Phe Thr Met Tyr	
	1290 1295 1300	
35	cct cgt atg gtg gta tca agc cat cgt att aga gag gtg atg gat atg	4041
	Pro Arg Met Val Val Ser Ser His Arg Ile Arg Glu Val Met Asp Met	
	1305 1310 1315 1320	
40	cca atc tct atc aat cct aat gcc gaa ggt gtt acg gat acg aaa ctt	4089
	Pro Ile Ser Ile Asn Pro Asn Ala Glu Gly Val Thr Asp Thr Lys Leu	
	1325 1330 1335	
45	aaa ggg cat tta gaa ttt gat aat gta aca ttc gct tat cca gga gaa	4137
	Lys Gly His Leu Glu Phe Asp Asn Val Thr Phe Ala Tyr Pro Gly Glu	
	1340 1345 1350	
50	aca gag agt ccc gtt ttg cat gat att tct ttt aaa gct aag cct gga	4185
	Thr Glu Ser Pro Val Leu His Asp Ile Ser Phe Lys Ala Lys Pro Gly	
	1355 1360 1365	
55	gaa aca att gct ttt att ggt tca aca ggt tca gga aaa tct tct ctt	4233
	Glu Thr Ile Ala Phe Ile Gly Ser Thr Gly Ser Gly Lys Ser Ser Leu	
	1370 1375 1380	
60	ggt aat ttg att cca cgt ttt tat gat gtg aca ctt gga aaa atc tta	4281
	Val Asn Leu Ile Pro Arg Phe Tyr Asp Val Thr Leu Gly Lys Ile Leu	
	1385 1390 1395 1400	
65	gta gat gga gtt gat gta aga gat tat aac ctt aaa tca ctt cgc caa	4329
	Val Asp Gly Val Asp Val Arg Asp Tyr Asn Leu Lys Ser Leu Arg Gln	
	1405 1410 1415	

ES 2 278 436 T3

	aag att gga ttt atc ccc caa aaa gct ctt tta ttt aca ggg aca ata	4377
	Lys Ile Gly Phe Ile Pro Gln Lys Ala Leu Leu Phe Thr Gly Thr Ile	
	1420 1425 1430	
5	gga gag aat tta aaa tat gga aaa gct gat gct act att gat gat ctt	4425
	Gly Glu Asn Leu Lys Tyr Gly Lys Ala Asp Ala Thr Ile Asp Asp Leu	
	1435 1440 1445	
10	aga caa gcg gtt gat att tct caa gct aaa gag ttt att gag agt cac	4473
	Arg Gln Ala Val Asp Ile Ser Gln Ala Lys Glu Phe Ile Glu Ser His	
	1450 1455 1460	
15	caa gaa gcc ttt gaa acg cat tta gct gaa ggt ggg agc aat ctt tct	4521
	Gln Glu Ala Phe Glu Thr His Leu Ala Glu Gly Gly Ser Asn Leu Ser	
	1465 1470 1475 1480	
20	ggg ggt caa aaa caa cgg tta tct att gct agg gct gtt gtt aaa gat	4569
	Gly Gly Gln Lys Gln Arg Leu Ser Ile Ala Arg Ala Val Val Lys Asp	
	1485 1490 1495	
25	cca gat tta tat att ttt gat gat tca ttt tct gct ctc gat tat aag	4617
	Pro Asp Leu Tyr Ile Phe Asp Asp Ser Phe Ser Ala Leu Asp Tyr Lys	
	1500 1505 1510	
30	aca gac gct act tta aga gcg cgt cta aaa gaa gta acc ggt gat tct	4665
	Thr Asp Ala Thr Leu Arg Ala Arg Leu Lys Glu Val Thr Gly Asp Ser	
	1515 1520 1525	
35	aca gtt ttg ata gtt gct caa agg gtg ggt acg att atg gat gct gat	4713
	Thr Val Leu Ile Val Ala Gln Arg Val Gly Thr Ile Met Asp Ala Asp	
	1530 1535 1540	
40	cag att att gtc ctt gat gaa ggc gaa att gtc ggt cgt ggt acc cac	4761
	Gln Ile Ile Val Leu Asp Glu Gly Glu Ile Val Gly Arg Gly Thr His	
	1545 1550 1555 1560	
45	gct caa tta ata gaa aat aat gct att tat cgt gaa atc gct gag tca	4809
	Ala Gln Leu Ile Glu Asn Asn Ala Ile Tyr Arg Glu Ile Ala Glu Ser	
	1565 1570 1575	
50	caa ctg aag aac caa aac tta tca gaa gga gag tgattgt atg aga aaa	4858
	Gln Leu Lys Asn Gln Asn Leu Ser Glu Gly Glu Met Arg Lys	
	1580 1585 1590	
55	aaa tct gtt ttt ttg aga tta tgg tct tac cta act cgc tac aaa gct	4906
	Lys Ser Val Phe Leu Arg Leu Trp Ser Tyr Leu Thr Arg Tyr Lys Ala	
	1595 1600 1605	
60	act ctt ttc tta gcg att ttt ttg aaa gtt tta tct agt ttt atg agt	4954
	Thr Leu Phe Leu Ala Ile Phe Leu Lys Val Leu Ser Ser Phe Met Ser	
	1610 1615 1620	
65	ggt ctg gag cct ttt att tta ggg tta gcg ata aca gag ttg act gct	5002
	Val Leu Glu Pro Phe Ile Leu Gly Leu Ala Ile Thr Glu Leu Thr Ala	
	1625 1630 1635	

ES 2 278 436 T3

aac ctt gtt gat atg gct aag gga gtt tct ggg gca gaa ttg aac gtt 5050
 Asn Leu Val Asp Met Ala Lys Gly Val Ser Gly Ala Glu Leu Asn Val
 1640 1645 1650

cct tat att gct ggt att ttg att att tat ttt ttc aga ggt gtt ttc 5098
 Pro Tyr Ile Ala Gly Ile Leu Ile Ile Tyr Phe Phe Arg Gly Val Phe
 1655 1660 1665 1670

tat gaa tta ggt tct tat ggc tca aat t 5126
 Tyr Glu Leu Gly Ser Tyr Gly Ser Asn
 1675

<210> 8

<211> 229

<212> PRT

<213> Estreptococos

<400> 8

Asn Phe Asp Ile Glu Thr Thr Thr Phe Glu Ala Met Lys Lys His Ala
 1 5 10 15
 Ser Leu Leu Glu Lys Ile Ser Val Glu Arg Ser Phe Ile Glu Phe Asp
 20 25 30
 Lys Leu Leu Ala Pro Tyr Trp Arg Lys Gly Met Leu Ala Leu Ile
 35 40 45
 Asp Ser His Ala Phe Asn Tyr Leu Pro Cys Leu Lys Asn Arg Glu Leu
 50 55 60
 Gln Leu Ser Ala Phe Leu Ser Gln Leu Asp Lys Asp Phe Leu Phe Glu
 65 70 75 80
 Thr Ser Glu Gln Ala Trp Ala Ser Leu Ile Leu Ser Met Glu Val Glu
 85 90 95
 His Thr Lys Thr Phe Leu Lys Lys Trp Lys Thr Ser Thr His Phe Gln
 100 105 110
 Lys Asp Val Glu His Ile Val Asp Val Tyr Arg Ile Arg Glu Gln Met
 115 120 125
 Gly Leu Ala Lys Glu His Leu Tyr Arg Tyr Gly Lys Thr Ile Ile Lys
 130 135 140
 Gln Ala Glu Gly Ile Arg Lys Ala Arg Gly Leu Met Val Asp Phe Glu
 145 150 155 160
 Lys Ile Glu Gln Leu Asp Ser Glu Leu Ala Ile His Asp Arg His Glu
 165 170 175
 Ile Val Val Asn Gly Gly Thr Leu Ile Lys Lys Leu Gly Ile Lys Pro
 180 185 190
 Gly Pro Gln Met Gly Asp Ile Ile Ser Gln Ile Glu Leu Ala Ile Val
 195 200 205
 Leu Gly Gln Leu Ile Asn Glu Glu Ala Ile Leu His Phe Val Lys
 210 215 220
 Gln Tyr Leu Met Asp
 225

<210> 9

<211> 622

<212> PRT

<213> Estreptococos

ES 2 278 436 T3

<400> 9

5	Met	Ser	Asp	Phe	Leu	Val	Asp	Gly	Leu	Thr	Lys	Ser	Val	Gly	Asp	Lys
	1				5					10					15	
	Thr	Val	Phe	Ser	Asn	Val	Ser	Phe	Ile	Ile	His	Ser	Leu	Asp	Arg	Ile
				20					25					30		
	Gly	Ile	Ile	Gly	Val	Asn	Gly	Thr	Gly	Lys	Thr	Thr	Leu	Leu	Asp	Val
		35						40					45			
10	Ile	Ser	Gly	Glu	Leu	Gly	Phe	Asp	Gly	Asp	Arg	Ser	Pro	Phe	Ser	Ser
	50					55						60				
	Ala	Asn	Asp	Tyr	Lys	Ile	Ala	Tyr	Leu	Lys	Gln	Glu	Pro	Asp	Phe	Asp
	65				70						75					80
15	Asp	Ser	Gln	Thr	Ile	Leu	Asp	Thr	Val	Leu	Ser	Ser	Asp	Leu	Arg	Glu
				85						90					95	
	Met	Ala	Leu	Ile	Lys	Glu	Tyr	Glu	Leu	Leu	Leu	Asn	His	Tyr	Glu	Glu
				100					105					110		
20	Ser	Lys	Gln	Ser	Arg	Leu	Glu	Lys	Val	Met	Ala	Glu	Met	Asp	Ser	Leu
			115					120					125			
	Asp	Ala	Trp	Ser	Ile	Glu	Ser	Glu	Val	Lys	Thr	Val	Leu	Ser	Lys	Leu
	130					135						140				
25	Gly	Ile	Thr	Asp	Leu	Gln	Leu	Ser	Val	Gly	Glu	Leu	Ser	Gly	Gly	Leu
	145					150					155					160
	Arg	Arg	Arg	Val	Gln	Leu	Ala	Gln	Val	Leu	Leu	Asn	Asp	Ala	Asp	Leu
				165					170					175		
	Leu	Leu	Leu	Asp	Glu	Pro	Thr	Asn	His	Leu	Asp	Ile	Asp	Thr	Ile	Ala
				180				185						190		
30	Trp	Leu	Thr	Asn	Phe	Leu	Lys	Asn	Ser	Lys	Lys	Thr	Val	Leu	Phe	Ile
			195					200					205			
	Thr	His	Asp	Arg	Tyr	Phe	Leu	Asp	Asn	Val	Ala	Thr	Arg	Ile	Phe	Glu
	210					215					220					
35	Leu	Asp	Lys	Ala	Gln	Ile	Thr	Glu	Tyr	Gln	Gly	Asn	Tyr	Gln	Asp	Tyr
	225				230					235					240	
	Val	Arg	Leu	Arg	Ala	Glu	Gln	Asp	Glu	Arg	Asp	Ala	Ala	Ser	Leu	His
				245					250					255		
40	Lys	Lys	Lys	Gln	Leu	Tyr	Lys	Gln	Glu	Leu	Ala	Trp	Met	Arg	Thr	Gln
				260				265						270		
	Pro	Gln	Ala	Arg	Ala	Thr	Lys	Gln	Gln	Ala	Arg	Ile	Asn	Arg	Phe	Gln
		275						280					285			
	Asn	Leu	Lys	Asn	Asp	Leu	His	Gln	Thr	Ser	Asp	Thr	Ser	Asp	Leu	Glu
	290					295						300				
45	Met	Thr	Phe	Glu	Thr	Ser	Arg	Ile	Gly	Lys	Lys	Val	Ile	Asn	Phe	Glu
	305					310					315				320	
	Asn	Val	Ser	Phe	Ser	Tyr	Pro	Asp	Lys	Ser	Ile	Leu	Lys	Asp	Phe	Asn
				325					330					335		
50	Leu	Leu	Ile	Gln	Asn	Lys	Asp	Arg	Ile	Gly	Ile	Val	Gly	Asp	Asn	Gly
				340					345					350		
	Val	Gly	Lys	Ser	Thr	Leu	Leu	Asn	Leu	Ile	Val	Gln	Asp	Leu	Gln	Pro
		355						360				365				
	Asp	Ser	Gly	Asn	Val	Ser	Ile	Gly	Glu	Thr	Ile	Arg	Val	Gly	Tyr	Phe
	370					375						380				
55	Ser	Gln	Gln	Leu	His	Asn	Met	Asp	Gly	Ser	Lys	Arg	Val	Ile	Asn	Tyr
	385					390					395				400	
	Leu	Gln	Glu	Val	Ala	Asp	Glu	Val	Lys	Thr	Ser	Val	Gly	Thr	Thr	Ser
				405					410					415		
60	Val	Thr	Glu	Leu	Leu	Glu	Gln	Phe	Leu	Phe	Pro	Arg	Ser	Thr	His	Gly
				420					425					430		

65

ES 2 278 436 T3

Thr Gln Ile Ala Lys Leu Ser Gly Gly Glu Lys Lys Arg Leu Tyr Leu
 435 440 445
 Leu Lys Ile Leu Ile Glu Lys Pro Asn Val Leu Leu Leu Asp Glu Pro
 450 455 460
 Thr Asn Asp Leu Asp Ile Ala Thr Leu Thr Val Leu Glu Asn Phe Leu
 465 470 475 480
 Gln Gly Phe Gly Gly Pro Val Ile Thr Val Ser His Asp Arg Tyr Phe
 485 490 495
 Leu Asp Lys Val Ala Asn Lys Ile Ile Ala Phe Glu Asp Asn Asp Ile
 500 505 510
 Arg Glu Phe Phe Gly Asn Tyr Thr Asp Tyr Leu Asp Glu Lys Ala Phe
 515 520 525
 Asn Glu Gln Asn Asn Glu Val Ile Ser Lys Lys Glu Ser Thr Lys Thr
 530 535 540
 Ser Arg Glu Lys Gln Ser Arg Lys Arg Met Ser Tyr Phe Glu Lys Gln
 545 550 555 560
 Glu Trp Ala Thr Ile Glu Asp Asp Ile Met Ile Leu Glu Asn Thr Ile
 565 570 575
 Thr Arg Ile Glu Asn Asp Met Gln Thr Cys Gly Ser Asp Phe Thr Arg
 580 585 590
 Leu Ser Asp Leu Gln Lys Glu Leu Asp Ala Lys Asn Glu Ala Leu Leu
 595 600 605
 Glu Lys Tyr Asp Arg Tyr Glu Tyr Leu Ser Glu Leu Asp Thr
 610 615 620

30 <210> 10
 <211> 157
 <212> PRT
 <213> Estreptococos
 35 <400> 10

Met Ile Ile Arg Pro Ile Ile Lys Asn Asp Asp Gln Ala Val Ala Gln
 1 5 10 15
 Leu Ile Arg Gln Ser Leu Arg Ala Tyr Asp Leu Asp Lys Pro Asp Thr
 20 25 30
 Ala Tyr Ser Asp Pro His Leu Asp His Leu Thr Ser Tyr Tyr Glu Lys
 35 40 45
 Ile Glu Lys Ser Gly Phe Phe Val Ile Glu Glu Arg Asp Glu Ile Ile
 50 55 60
 Gly Cys Gly Gly Phe Gly Pro Leu Lys Asn Leu Ile Ala Glu Met Gln
 65 70 75 80
 Lys Val Tyr Ile Ala Glu Arg Phe Arg Gly Lys Gly Leu Ala Thr Asp
 85 90 95
 Leu Val Lys Met Ile Glu Val Glu Ala Arg Lys Ile Gly Tyr Arg Gln
 100 105 110
 Leu Tyr Leu Glu Thr Ala Ser Thr Leu Ser Arg Ala Thr Ala Val Tyr
 115 120 125
 Lys His Met Gly Tyr Cys Ala Leu Ser Gln Pro Ile Ala Asn Asp Gln
 130 135 140
 Gly His Thr Ala Met Asp Ile Trp Met Ile Lys Asp Leu
 145 150 155

65 <210> 11
 <211> 579
 <212> PRT
 <213> Estreptococos

ES 2 278 436 T3

<400> 11

5	Met	Ala	Tyr	Ile	Trp	Ser	Tyr	Leu	Lys	Arg	Tyr	Pro	Asn	Trp	Leu	Trp
	1				5					10					15	
	Leu	Asp	Leu	Leu	Gly	Ala	Met	Leu	Phe	Val	Thr	Val	Ile	Leu	Gly	Met
			20						25					30		
	Pro	Thr	Ala	Leu	Ala	Gly	Met	Ile	Asp	Asn	Gly	Val	Thr	Lys	Gly	Asp
		35						40					45			
10	Arg	Thr	Gly	Val	Tyr	Leu	Trp	Thr	Phe	Ile	Met	Phe	Ile	Phe	Val	Val
	50					55						60				
	Leu	Gly	Ile	Ile	Gly	Arg	Ile	Thr	Met	Ala	Tyr	Ala	Ser	Ser	Arg	Leu
	65				70						75				80	
15	Thr	Thr	Thr	Met	Ile	Arg	Asp	Met	Arg	Asn	Asp	Met	Tyr	Ala	Lys	Leu
				85						90					95	
	Gln	Glu	Tyr	Ser	His	His	Glu	Tyr	Glu	Gln	Ile	Gly	Val	Ser	Ser	Leu
				100					105					110		
	Val	Thr	Arg	Met	Thr	Ser	Asp	Thr	Phe	Val	Leu	Met	Gln	Phe	Ala	Glu
20			115					120					125			
	Met	Ser	Leu	Arg	Leu	Gly	Leu	Val	Thr	Pro	Met	Val	Met	Ile	Phe	Ser
		130					135					140				
	Val	Val	Met	Ile	Leu	Ile	Thr	Ser	Pro	Ser	Leu	Ala	Trp	Leu	Val	Ala
25					150						155				160	
	Val	Ala	Met	Pro	Leu	Leu	Val	Gly	Val	Val	Leu	Tyr	Val	Ala	Ile	Lys
				165						170					175	
	Thr	Lys	Pro	Leu	Ser	Glu	Arg	Gln	Gln	Thr	Met	Leu	Asp	Lys	Ile	Asn
				180					185					190		
30	Gln	Tyr	Val	Arg	Glu	Asn	Leu	Thr	Gly	Leu	Arg	Val	Val	Arg	Ala	Phe
		195						200					205			
	Ala	Arg	Glu	Asn	Phe	Gln	Ser	Gln	Lys	Phe	Gln	Val	Ala	Asn	Gln	Arg
		210					215					220				
35	Tyr	Thr	Asp	Thr	Ser	Thr	Gly	Leu	Phe	Lys	Leu	Thr	Gly	Leu	Thr	Glu
	225				230						235				240	
	Pro	Leu	Phe	Val	Gln	Ile	Ile	Ile	Ala	Met	Ile	Val	Ala	Ile	Val	Trp
				245						250					255	
	Phe	Ala	Leu	Asp	Pro	Leu	Gln	Arg	Gly	Ala	Ile	Lys	Ile	Gly	Asp	Leu
40				260					265					270		
	Val	Ala	Phe	Ile	Glu	Tyr	Ser	Phe	His	Ala	Leu	Phe	Ser	Phe	Leu	Leu
			275					280					285			
	Phe	Ala	Asn	Leu	Phe	Thr	Met	Tyr	Pro	Arg	Met	Val	Val	Ser	Ser	His
		290				295						300				
45	Arg	Ile	Arg	Glu	Val	Met	Asp	Met	Pro	Ile	Ser	Ile	Asn	Pro	Asn	Ala
	305				310						315				320	
	Glu	Gly	Val	Thr	Asp	Thr	Lys	Leu	Lys	Gly	His	Leu	Glu	Phe	Asp	Asn
				325						330					335	
50	Val	Thr	Phe	Ala	Tyr	Pro	Gly	Glu	Thr	Glu	Ser	Pro	Val	Leu	His	Asp
				340					345					350		
	Ile	Ser	Phe	Lys	Ala	Lys	Pro	Gly	Glu	Thr	Ile	Ala	Phe	Ile	Gly	Ser
			355					360					365			
	Thr	Gly	Ser	Gly	Lys	Ser	Ser	Leu	Val	Asn	Leu	Ile	Pro	Arg	Phe	Tyr
55		370				375						380				
	Asp	Val	Thr	Leu	Gly	Lys	Ile	Leu	Val	Asp	Gly	Val	Asp	Val	Arg	Asp
	385				390						395				400	
	Tyr	Asn	Leu	Lys	Ser	Leu	Arg	Gln	Lys	Ile	Gly	Phe	Ile	Pro	Gln	Lys
				405						410					415	
60	Ala	Leu	Leu	Phe	Thr	Gly	Thr	Ile	Gly	Glu	Asn	Leu	Lys	Tyr	Gly	Lys
				420					425						430	

65

ES 2 278 436 T3

Ala Asp Ala Thr Ile Asp Asp Leu Arg Gln Ala Val Asp Ile Ser Gln
 435 440 445
 5 Ala Lys Glu Phe Ile Glu Ser His Gln Glu Ala Phe Glu Thr His Leu
 450 455 460
 Ala Glu Gly Gly Ser Asn Leu Ser Gly Gly Gln Lys Gln Arg Leu Ser
 465 470 475 480
 Ile Ala Arg Ala Val Val Lys Asp Pro Asp Leu Tyr Ile Phe Asp Asp
 485 490 495
 10 Ser Phe Ser Ala Leu Asp Tyr Lys Thr Asp Ala Thr Leu Arg Ala Arg
 500 505 510
 Leu Lys Glu Val Thr Gly Asp Ser Thr Val Leu Ile Val Ala Gln Arg
 515 520 525
 15 Val Gly Thr Ile Met Asp Ala Asp Gln Ile Ile Val Leu Asp Glu Gly
 530 535 540
 Glu Ile Val Gly Arg Gly Thr His Ala Gln Leu Ile Glu Asn Asn Ala
 545 550 555 560
 20 Ile Tyr Arg Glu Ile Ala Glu Ser Gln Leu Lys Asn Gln Asn Leu Ser
 565 570 575
 Glu Gly Glu

<210> 12

<211> 92

<212> PRT

<213> Estreptococos

<400> 12

Met Arg Lys Lys Ser Val Phe Leu Arg Leu Trp Ser Tyr Leu Thr Arg
 1 5 10 15
 35 Tyr Lys Ala Thr Leu Phe Leu Ala Ile Phe Leu Lys Val Leu Ser Ser
 20 25 30
 Phe Met Ser Val Leu Glu Pro Phe Ile Leu Gly Leu Ala Ile Thr Glu
 35 40 45
 40 Leu Thr Ala Asn Leu Val Asp Met Ala Lys Gly Val Ser Gly Ala Glu
 50 55 60
 Leu Asn Val Pro Tyr Ile Ala Gly Ile Leu Ile Ile Tyr Phe Phe Arg
 65 70 75 80
 45 Gly Val Phe Tyr Glu Leu Gly Ser Tyr Gly Ser Asn
 85 90

<210> 13

<211> 5215

<212> ADN

<213> Estreptococos

<220>

<221> CDS

<222> (3)...(122)

<221> CDS

<222> (133)...(2511)

<221> CDS

<222> (367)...(2511)

<221> CDS

<222> (2946)...(2716)

ES 2 278 436 T3

<223> de cadena complementaria

<221> CDS

5 <222> (3252)...(2995)

<223> de cadena complementaria

<221> CDS

10 <222> (3676)...(3299)

<223> de cadena complementaria

<221> CDS

15 <222> (4124)...(3837)

<223> de cadena complementaria

<221> CDS

20 <222> (5214)...(4351)

<223> de cadena complementaria

25 <400> 13

aa ttt gga agt gct cta tca aca gtt gaa gta aag gag att att agt 47
Phe Gly Ser Ala Leu Ser Thr Val Glu Val Lys Glu Ile Ile Ser
1 5 10 15

30 gaa gaa aac ata tgg tta tat cgg ctc agt tgc tgc cat ttt act agc 95
Glu Glu Asn Ile Trp Leu Tyr Arg Leu Ser Cys Cys His Phe Thr Ser
20 25 30

35 tac tca tat tgg aag tta cca act tgg taagcatcat atg ggt cta gca 144
Tyr Ser Tyr Trp Lys Leu Pro Thr Trp Met Gly Leu Ala
35 40

40 aca aag gac aat cag att gcc tat att gat gac agc aaa ggt aag gca 192
Thr Lys Asp Asn Gln Ile Ala Tyr Ile Asp Asp Ser Lys Gly Lys Ala
45 50 55 60

45 aaa gcc cct aaa aca aac aaa acg atg gat caa atc agt gct gaa gaa 240
Lys Ala Pro Lys Thr Asn Lys Thr Met Asp Gln Ile Ser Ala Glu Glu
65 70 75

50 ggc atc tct gct gaa cag atc gta gtc aaa att act gac caa ggc tat 288
Gly Ile Ser Ala Glu Gln Ile Val Val Lys Ile Thr Asp Gln Gly Tyr
80 85 90

55 gtg acc tca cac ggt gac cat tat cat ttt tac aat ggg aaa gtt cct 336
Val Thr Ser His Gly Asp His Tyr His Phe Tyr Asn Gly Lys Val Pro
95 100 105

60 tat gat gcg att att agt gaa gag ttg ttg atg acg gat cct aat tac 384
Tyr Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Thr Asp Pro Asn Tyr
110 115 120

65 cgt ttt aaa caa tca gac gtt atc aat gaa atc tta gac ggt tac gtt 432
Arg Phe Lys Gln Ser Asp Val Ile Asn Glu Ile Leu Asp Gly Tyr Val
125 130 135 140

ES 2 278 436 T3

5	att aaa gtc aat ggc aac tat tat gtt tac ctc aag cca ggt agt aag Ile Lys Val Asn Gly Asn Tyr Tyr Val Tyr Leu Lys Pro Gly Ser Lys 145 150 155	480
10	cgc aaa aac att cga acc aaa caa caa att gct gag caa gta gcc aaa Arg Lys Asn Ile Arg Thr Lys Gln Gln Ile Ala Glu Gln Val Ala Lys 160 165 170	528
15	gga act aaa gaa gct aaa gaa aaa ggt tta gct caa gtg gcc cat ctc Gly Thr Lys Glu Ala Lys Glu Lys Gly Leu Ala Gln Val Ala His Leu 175 180 185	576
20	agt aaa gaa gaa gtt gcg gca gtc aat gaa gca aaa aga caa gga cgc Ser Lys Glu Glu Val Ala Ala Val Asn Glu Ala Lys Arg Gln Gly Arg 190 195 200	624
25	tat act aca gac gat ggc tat att ttt agt ccg aca gat atc att gat Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Ser Pro Thr Asp Ile Ile Asp 205 210 215 220	672
30	gat tta gga gat gct tat tta gta cct cat ggt aat cac tat cat tat Asp Leu Gly Asp Ala Tyr Leu Val Pro His Gly Asn His Tyr His Tyr 225 230 235	720
35	att cct aaa aag gat ttg tct cca agt gag cta gct gct gca caa gcc Ile Pro Lys Lys Asp Leu Ser Pro Ser Glu Leu Ala Ala Ala Gln Ala 240 245 250	768
40	tac tgg agt caa aaa caa ggt cga ggt gct aga ccg tct gat tac cgc Tyr Trp Ser Gln Lys Gln Gly Arg Gly Ala Arg Pro Ser Asp Tyr Arg 255 260 265	816
45	ccg aca cca gcc cca ggt cgt agg aaa gcc cca att cct gat gtg acg Pro Thr Pro Ala Pro Gly Arg Arg Lys Ala Pro Ile Pro Asp Val Thr 270 275 280	864
50	cct aac cct gga caa ggt cat cag cca gat aac ggt ggc tat cat cca Pro Asn Pro Gly Gln Gly His Gln Pro Asp Asn Gly Gly Tyr His Pro 285 290 295 300	912
55	gcg cct cct agg cca aat gat gcg tca caa aac aaa cac caa aga gat Ala Pro Pro Arg Pro Asn Asp Ala Ser Gln Asn Lys His Gln Arg Asp 305 310 315	960
60	gag ttt aaa gga aaa acc ttt aag gaa ctt tta gat caa cta cac cgt Glu Phe Lys Gly Lys Thr Phe Lys Glu Leu Leu Asp Gln Leu His Arg 320 325 330	1008
65	ctt gat ttg aaa tac cgt cat gtg gaa gaa gat ggg ttg att ttt gaa Leu Asp Leu Lys Tyr Arg His Val Glu Glu Asp Gly Leu Ile Phe Glu 335 340 345	1056
	ccg act caa gtg atc aaa tca aac gct ttt ggg tat gtg gtg cct cat Pro Thr Gln Val Ile Lys Ser Asn Ala Phe Gly Tyr Val Val Pro His 350 355 360	1104

ES 2 278 436 T3

	gga gat cat tat cat att atc cca aga agt cag tta tca cct ctt gaa	1152
	Gly Asp His Tyr His Ile Ile Pro Arg Ser Gln Leu Ser Pro Leu Glu	
	365 370 375 380	
5	atg gaa tta gca gat cga tac tta gct ggc caa act gag gac aat gac	1200
	Met Glu Leu Ala Asp Arg Tyr Leu Ala Gly Gln Thr Glu Asp Asn Asp	
	385 390 395	
10	tca ggt tca gag cac tca aaa cca tca gat aaa gaa gtg aca cat acc	1248
	Ser Gly Ser Glu His Ser Lys Pro Ser Asp Lys Glu Val Thr His Thr	
	400 405 410	
15	ttt ctt ggt cat cgc atc aaa gct tac gga aaa ggc tta gat ggt aaa	1296
	Phe Leu Gly His Arg Ile Lys Ala Tyr Gly Lys Gly Leu Asp Gly Lys	
	415 420 425	
20	cca tat gat acg agt gat gct tat gtt ttt agt aaa gaa tcc att cat	1344
	Pro Tyr Asp Thr Ser Asp Ala Tyr Val Phe Ser Lys Glu Ser Ile His	
	430 435 440	
25	tca gtg gat aaa tca gga gtt aca gct aaa cac gga gat cat ttc cac	1392
	Ser Val Asp Lys Ser Gly Val Thr Ala Lys His Gly Asp His Phe His	
	445 450 455 460	
30	tat ata gga ttt gga gaa ctt gaa caa tat gag ttg gat gag gtc gct	1440
	Tyr Ile Gly Phe Gly Glu Leu Glu Gln Tyr Glu Leu Asp Glu Val Ala	
	465 470 475	
35	aac tgg gtg aaa gca aaa ggt caa gct gat gag ctt gct gct gct ttg	1488
	Asn Trp Val Lys Ala Lys Gly Gln Ala Asp Glu Leu Ala Ala Ala Leu	
	480 485 490	
40	gat cag gaa caa ggc aaa gaa aaa cca ctc ttt gac act aaa aaa gtg	1536
	Asp Gln Glu Gln Gly Lys Glu Lys Pro Leu Phe Asp Thr Lys Lys Val	
	495 500 505	
45	agt cgc aaa gta aca aaa gat ggt aaa gtg ggc tat atg atg cca aaa	1584
	Ser Arg Lys Val Thr Lys Asp Gly Lys Val Gly Tyr Met Met Pro Lys	
	510 515 520	
50	gat ggt aag gac tat ttc tat gct cgt gat caa ctt gat ttg act cag	1632
	Asp Gly Lys Asp Tyr Phe Tyr Ala Arg Asp Gln Leu Asp Leu Thr Gln	
	525 530 535 540	
55	att gcc ttt gcc gaa caa gaa cta atg ctt aaa gat aag aag cat tac	1680
	Ile Ala Phe Ala Glu Gln Glu Leu Met Leu Lys Asp Lys Lys His Tyr	
	545 550 555	
60	cgt tat gac att gtt gac aca ggt att gag cca cga ctt gct gta gat	1728
	Arg Tyr Asp Ile Val Asp Thr Gly Ile Glu Pro Arg Leu Ala Val Asp	
	560 565 570	
65	gtg tca agt ctg ccg atg cat gct ggt aat gct act tac gat act gga	1776
	Val Ser Ser Leu Pro Met His Ala Gly Asn Ala Thr Tyr Asp Thr Gly	
	575 580 585	

ES 2 278 436 T3

	agt tcg ttt gtt atc cca cat att gat cat atc cat gtc gtt ccg tat	1824
	Ser Ser Phe Val Ile Pro His Ile Asp His Ile His Val Val Pro Tyr	
	590 595 600	
5	tca tgg ttg acg cgc gat cag att gca aca gtc aag tat gtg atg caa	1872
	Ser Trp Leu Thr Arg Asp Gln Ile Ala Thr Val Lys Tyr Val Met Gln	
	605 610 615 620	
10	cac ccc gaa gtt cgt ccg gat gta tgg tct aag cca ggg cat gaa gag	1920
	His Pro Glu Val Arg Pro Asp Val Trp Ser Lys Pro Gly His Glu Glu	
	625 630 635	
15	tca ggt tcg gtc att cca aat gtt acg cct ctt gat aaa cgt gct ggt	1968
	Ser Gly Ser Val Ile Pro Asn Val Thr Pro Leu Asp Lys Arg Ala Gly	
	640 645 650	
20	atg cca aac tgg caa att atc cat tct gct gaa gaa gtt caa aaa gcc	2016
	Met Pro Asn Trp Gln Ile Ile His Ser Ala Glu Glu Val Gln Lys Ala	
	655 660 665	
25	cta gca gaa ggt cgt ttt gca aca cca gac ggc tat att ttc gat cca	2064
	Leu Ala Glu Gly Arg Phe Ala Thr Pro Asp Gly Tyr Ile Phe Asp Pro	
	670 675 680	
30	cga gat gtt ttg gcc aaa gaa act ttt gta tgg aaa gat ggc tcc ttt	2112
	Arg Asp Val Leu Ala Lys Glu Thr Phe Val Trp Lys Asp Gly Ser Phe	
	685 690 695 700	
35	agc atc cca aga gca gat ggc agt tca ttg aga acc att aat aaa tct	2160
	Ser Ile Pro Arg Ala Asp Gly Ser Ser Leu Arg Thr Ile Asn Lys Ser	
	705 710 715	
40	gat cta tcc caa gct gag tgg caa caa gct caa gag tta ttg gca aag	2208
	Asp Leu Ser Gln Ala Glu Trp Gln Gln Ala Gln Glu Leu Leu Ala Lys	
	720 725 730	
45	aaa aat act ggt gat gct act gat acg gat aaa ccc aaa gaa aag caa	2256
	Lys Asn Thr Gly Asp Ala Thr Asp Thr Asp Lys Pro Lys Glu Lys Gln	
	735 740 745	
50	cag gca gat aag agc aat gaa aac caa cag cca agt gaa gcc agt aaa	2304
	Gln Ala Asp Lys Ser Asn Glu Asn Gln Gln Pro Ser Glu Ala Ser Lys	
	750 755 760	
55	gaa gaa aaa gaa tca gat gac ttt ata gac agt tta cca gac tat ggt	2352
	Glu Glu Lys Glu Ser Asp Asp Phe Ile Asp Ser Leu Pro Asp Tyr Gly	
	765 770 775 780	
60	cta gat aga gca acc cta gaa gat cat atc aat caa tta gca caa aaa	2400
	Leu Asp Arg Ala Thr Leu Glu Asp His Ile Asn Gln Leu Ala Gln Lys	
	785 790 795	
65	gct aat atc gat cct aag tat ctc att ttc caa cca gaa ggt gtc caa	2448
	Ala Asn Ile Asp Pro Lys Tyr Leu Ile Phe Gln Pro Glu Gly Val Gln	
	800 805 810	

ES 2 278 436 T3

	ttt tat aat aaa aat ggt gaa ttg gta act tat gat atc aag aca ctt	2496
	Phe Tyr Asn Lys Asn Gly Glu Leu Val Thr Tyr Asp Ile Lys Thr Leu	
	815 820 825	
5	caa caa ata aac cct taacccaaaag aagatctcat tggtaaagca ctgctttgtc	2551
	Gln Gln Ile Asn Pro	
	830	
10	aaagcaagtc acgggtgattt tgaagtcatt ctatgtaacg agtagtgata aaagttggat	2611
	aatagcgggt tctcttttgca aagaaatggt atccatgtta gaatagtaaa aaaagaggag	2671
	gattccttggg ctaatgtcaa ataagtagac agaaaaactgt gttattttattgcgt	2726
	taaaataaatt tctctctttc tgattagggg ttagtcctag attagccgta tgtggggtgt	2786
15	aattgtttata aaaattctca atgtattcaa agcagtcctaa ttgaacctgt ttgatatttt	2846
	gataatgttt tgggttgatt tgtctatgct ttaaataact gaaaaatgct tcagttacgg	2906
	cattatcata aggatatcca ggattagaaa aagaatgcat gatattggca ctgcacccta	2966
	atagtggagac gcaagaaaaa cacttttaggcaatcagtt ttctgtactg tacaggcgac	3025
	tggctgctta atctctgttg aattctagtt tcattataaa atgtaatgta atttttaaca	3085
20	atatttgta tactatcttt gttgtatttt ctctattat ggaaataaaa ggtttcagtc	3145
	tttaggacgg tgtgaaacca ttcaatacag gcattatctg cagggtgttc ttttcgagac	3205
	attgagcggg taatgtcttt ttccgtgcaa gcctggtagt aagccataga agtatacact	3265
	gagccttggt cactgtgtaa gattgtcctt ttatttaggcaatt ttaactgatt	3319
25	aaggggtgctt agtacaaaat ccgtgtcctg acaatctgag atagtgtgag ctataatttc	3379
	tcggttatag agattcataa ttgatgagag atacaattta cagttaccga aatataggta	3439
	ggtaatatct gttacgagct tttccttagg cttatcggca tggaaatccc gactcaattt	3499
	attatctgtt aaataataag ctttacccaa attgggaact ttcttggtac gtgtccgaca	3559
	aagccagcca ttatttttca tgatacgata gactttcttt gtattaacag tcaatccgtg	3619
30	gatttttttg agcaatcgtg taatggtagc atagccataa ataaagtgat tctccataca	3679
	gagctgttca attaatccaa taaggctatc tttttttgcg gcttctcata ctctcttttc	3739
	caacggtaat aggtcgaccg cttgacctta aaacagtcta gaatgaaaac tatcgggtag	3799
	ttgtttttat agtcttccac aagcttgata agacttaactttatcgatt tctttatcaa	3857
	gcctcgatac ttttttaaga ggtcaacctg taattgtaat tgttccactt cagacagatg	3917
35	ttccaagcct ttaccgtagg tatattgctt gccaacacct tgatgaaaac gataaagctc	3977
	ctcgttttcg taccatttca tccaagtata gatttgacta ttatttttga tgcctaaagt	4037
	ctccataata actctgttag acttgccctg tttcttcata tcgatgcaag ccagcttagc	4097
	ttcccatgaa tatgtctttt taaccataat aaaacatttc tgtttctagt ttactaaatt	4157
40	tcaacaggag tgtttttctt ttgtctcatt ttagggatcc agtgccattt gttgtcatca	4217
	attatctttc taaattcccc ggacttaaat tgtgacctt ggtcggaatg aaagagaagt	4277
	gttccctcaa tctttctttt attaagttaa aaggcaacac ttttctgtac aacatttata	4337
	aagtgttttt ctaggcaattaatc ttttagtcat tgggtgttgg tagttgagac	4391
	taccatgaat gcggtggtaa ttccaccaat gaacatagtc tttagtctta agagctagtt	4451
45	cttccagcaa ttgaaagggt tcttgataaa caaattcaat ttgaaagca cgatacgtac	4511
	tttcagctac ggcattgtca taaggataac cagcctgact aagcgaacgt gtgattccaa	4571
	aggtttccaa tatttcatca attaactgat tatcaaatc tttgccacga tctgaatgga	4631
	acatcttgac tttggtcagg gcgtaaggga tgccttgat ggcttgctta acgagttcag	4691
	cggctcttgtg ccaaccaaga gacaggccga tgatttcacg gttgtatagg tcaatgatga	4751
50	ggcaaacata agcccaacga ttgcctacac gaacataggt taagtcagtg actaaggctt	4811
	gtagtggctt tctttgctta aattgcctgt ctaagtgggt gggaataggg gcttcattct	4871
	tgcctctaga atgtgggttg aagggtggctt tctgataaac agaaaccaa ttgagtcgct	4931
	tcataatgcg tcgaatccga cgacgtgaaa gtgtgatacc ttcgttatcc aagcatattt	4991
55	tgatttttct ggatccgtat ctgactcgc tatcgagaaa aattctttta atagtctctt	5051
	caaactccgt ttcagatact gactccacgg cttgatagta ataacttgag tgtggcatac	5111
	tcagccagcg acacatcttt gaaatgcctt atttatctct attagcagtg attatttccc	5171
	tttttgtgcc ataatcccg ctgcttgctt taggatattc aatt	5215
60	<210> 14	
	<211> 40	
	<212> PRT	
65	<213> Estreptococos	

ES 2 278 436 T3

<400> 14

5 Phe Gly Ser Ala Leu Ser Thr Val Glu Val Lys Glu Ile Ile Ser Glu
1 5 10 15
Glu Asn Ile Trp Leu Tyr Arg Leu Ser Cys Cys His Phe Thr Ser Tyr
20 25 30
Ser Tyr Trp Lys Leu Pro Thr Trp
35 40

10

<210> 15

<211> 793

15 <212> PRT

<213> Estreptococos

<400> 15

20

25

30

35

40

45

50

55

60

65

Met Gly Leu Ala Thr Lys Asp Asn Gln Ile Ala Tyr Ile Asp Asp Ser
1 5 10 15
Lys Gly Lys Ala Lys Ala Pro Lys Thr Asn Lys Thr Met Asp Gln Ile
20 25 30
Ser Ala Glu Glu Gly Ile Ser Ala Glu Gln Ile Val Val Lys Ile Thr
35 40 45
Asp Gln Gly Tyr Val Thr Ser His Gly Asp His Tyr His Phe Tyr Asn
50 55 60
Gly Lys Val Pro Tyr Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Thr
65 70 75 80
Asp Pro Asn Tyr Arg Phe Lys Gln Ser Asp Val Ile Asn Glu Ile Leu
85 90 95
Asp Gly Tyr Val Ile Lys Val Asn Gly Asn Tyr Tyr Val Tyr Leu Lys
100 105 110
Pro Gly Ser Lys Arg Lys Asn Ile Arg Thr Lys Gln Gln Ile Ala Glu
115 120 125
Gln Val Ala Lys Gly Thr Lys Glu Ala Lys Glu Lys Gly Leu Ala Gln
130 135 140
Val Ala His Leu Ser Lys Glu Glu Val Ala Ala Val Asn Glu Ala Lys
145 150 155 160
Arg Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Ser Pro Thr
165 170 175
Asp Ile Ile Asp Asp Leu Gly Asp Ala Tyr Leu Val Pro His Gly Asn
180 185 190
His Tyr His Tyr Ile Pro Lys Lys Asp Leu Ser Pro Ser Glu Leu Ala
195 200 205
Ala Ala Gln Ala Tyr Trp Ser Gln Lys Gln Gly Arg Gly Ala Arg Pro
210 215 220
Ser Asp Tyr Arg Pro Thr Pro Ala Pro Gly Arg Arg Lys Ala Pro Ile
225 230 235 240
Pro Asp Val Thr Pro Asn Pro Gly Gln Gly His Gln Pro Asp Asn Gly
245 250 255
Gly Tyr His Pro Ala Pro Pro Arg Pro Asn Asp Ala Ser Gln Asn Lys
260 265 270
His Gln Arg Asp Glu Phe Lys Gly Lys Thr Phe Lys Glu Leu Leu Asp
275 280 285
Gln Leu His Arg Leu Asp Leu Lys Tyr Arg His Val Glu Glu Asp Gly
290 295 300
Leu Ile Phe Glu Pro Thr Gln Val Ile Lys Ser Asn Ala Phe Gly Tyr
305 310 315 320

ES 2 278 436 T3

	Val	Val	Pro	His	Gly	Asp	His	Tyr	His	Ile	Ile	Pro	Arg	Ser	Gln	Leu
					325					330					335	
5	Ser	Pro	Leu	Glu	Met	Glu	Leu	Ala	Asp	Arg	Tyr	Leu	Ala	Gly	Gln	Thr
				340					345					350		
	Glu	Asp	Asn	Asp	Ser	Gly	Ser	Glu	His	Ser	Lys	Pro	Ser	Asp	Lys	Glu
			355					360					365			
	Val	Thr	His	Thr	Phe	Leu	Gly	His	Arg	Ile	Lys	Ala	Tyr	Gly	Lys	Gly
		370					375					380				
10	Leu	Asp	Gly	Lys	Pro	Tyr	Asp	Thr	Ser	Asp	Ala	Tyr	Val	Phe	Ser	Lys
	385					390					395					400
	Glu	Ser	Ile	His	Ser	Val	Asp	Lys	Ser	Gly	Val	Thr	Ala	Lys	His	Gly
				405						410					415	
15	Asp	His	Phe	His	Tyr	Ile	Gly	Phe	Gly	Glu	Leu	Glu	Gln	Tyr	Glu	Leu
				420					425					430		
	Asp	Glu	Val	Ala	Asn	Trp	Val	Lys	Ala	Lys	Gly	Gln	Ala	Asp	Glu	Leu
		435						440					445			
20	Ala	Ala	Ala	Leu	Asp	Gln	Glu	Gln	Gly	Lys	Glu	Lys	Pro	Leu	Phe	Asp
		450					455					460				
	Thr	Lys	Lys	Val	Ser	Arg	Lys	Val	Thr	Lys	Asp	Gly	Lys	Val	Gly	Tyr
	465					470					475					480
	Met	Met	Pro	Lys	Asp	Gly	Lys	Asp	Tyr	Phe	Tyr	Ala	Arg	Asp	Gln	Leu
				485						490					495	
25	Asp	Leu	Thr	Gln	Ile	Ala	Phe	Ala	Glu	Gln	Glu	Leu	Met	Leu	Lys	Asp
				500					505					510		
	Lys	Lys	His	Tyr	Arg	Tyr	Asp	Ile	Val	Asp	Thr	Gly	Ile	Glu	Pro	Arg
			515					520					525			
30	Leu	Ala	Val	Asp	Val	Ser	Ser	Leu	Pro	Met	His	Ala	Gly	Asn	Ala	Thr
		530					535					540				
	Tyr	Asp	Thr	Gly	Ser	Ser	Phe	Val	Ile	Pro	His	Ile	Asp	His	Ile	His
	545					550					555					560
35	Val	Val	Pro	Tyr	Ser	Trp	Leu	Thr	Arg	Asp	Gln	Ile	Ala	Thr	Val	Lys
				565						570					575	
	Tyr	Val	Met	Gln	His	Pro	Glu	Val	Arg	Pro	Asp	Val	Trp	Ser	Lys	Pro
				580					585					590		
	Gly	His	Glu	Glu	Ser	Gly	Ser	Val	Ile	Pro	Asn	Val	Thr	Pro	Leu	Asp
		595					600						605			
40	Lys	Arg	Ala	Gly	Met	Pro	Asn	Trp	Gln	Ile	Ile	His	Ser	Ala	Glu	Glu
		610					615					620				
	Val	Gln	Lys	Ala	Leu	Ala	Glu	Gly	Arg	Phe	Ala	Thr	Pro	Asp	Gly	Tyr
	625					630					635					640
45	Ile	Phe	Asp	Pro	Arg	Asp	Val	Leu	Ala	Lys	Glu	Thr	Phe	Val	Trp	Lys
				645						650					655	
	Asp	Gly	Ser	Phe	Ser	Ile	Pro	Arg	Ala	Asp	Gly	Ser	Ser	Leu	Arg	Thr
				660					665					670		
50	Ile	Asn	Lys	Ser	Asp	Leu	Ser	Gln	Ala	Glu	Trp	Gln	Gln	Ala	Gln	Glu
		675						680					685			
	Leu	Leu	Ala	Lys	Lys	Asn	Thr	Gly	Asp	Ala	Thr	Asp	Thr	Asp	Lys	Pro
		690					695					700				
	Lys	Glu	Lys	Gln	Gln	Ala	Asp	Lys	Ser	Asn	Glu	Asn	Gln	Gln	Pro	Ser
	705					710					715					720
55	Glu	Ala	Ser	Lys	Glu	Glu	Lys	Glu	Ser	Asp	Asp	Phe	Ile	Asp	Ser	Leu
				725						730					735	
	Pro	Asp	Tyr	Gly	Leu	Asp	Arg	Ala	Thr	Leu	Glu	Asp	His	Ile	Asn	Gln
				740					745					750		
60	Leu	Ala	Gln	Lys	Ala	Asn	Ile	Asp	Pro	Lys	Tyr	Leu	Ile	Phe	Gln	Pro
				755				760						765		

ES 2 278 436 T3

Glu Gly Val Gln Phe Tyr Asn Lys Asn Gly Glu Leu Val Thr Tyr Asp
 770 775 780
 Ile Lys Thr Leu Gln Gln Ile Asn Pro
 785 790

<210> 16

<211> 715

<212> PRT

<213> Estreptococos

<400> 16

Met Thr Asp Pro Asn Tyr Arg Phe Lys Gln Ser Asp Val Ile Asn Glu
 1 5 10 15
 Ile Leu Asp Gly Tyr Val Ile Lys Val Asn Gly Asn Tyr Tyr Val Tyr
 20 20 25 30
 Leu Lys Pro Gly Ser Lys Arg Lys Asn Ile Arg Thr Lys Gln Gln Ile
 35 40 45
 Ala Glu Gln Val Ala Lys Gly Thr Lys Glu Ala Lys Glu Lys Gly Leu
 50 55 60
 Ala Gln Val Ala His Leu Ser Lys Glu Glu Val Ala Ala Val Asn Glu
 65 70 75 80
 Ala Lys Arg Gln Gly Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Ser
 85 90 95
 Pro Thr Asp Ile Ile Asp Asp Leu Gly Asp Ala Tyr Leu Val Pro His
 100 105 110
 Gly Asn His Tyr His Tyr Ile Pro Lys Lys Asp Leu Ser Pro Ser Glu
 115 120 125
 Leu Ala Ala Ala Gln Ala Tyr Trp Ser Gln Lys Gln Gly Arg Gly Ala
 130 135 140
 Arg Pro Ser Asp Tyr Arg Pro Thr Pro Ala Pro Gly Arg Arg Lys Ala
 145 150 155 160
 Pro Ile Pro Asp Val Thr Pro Asn Pro Gly Gln Gly His Gln Pro Asp
 165 170 175
 Asn Gly Gly Tyr His Pro Ala Pro Pro Arg Pro Asn Asp Ala Ser Gln
 180 185 190
 Asn Lys His Gln Arg Asp Glu Phe Lys Gly Lys Thr Phe Lys Glu Leu
 195 200 205
 Leu Asp Gln Leu His Arg Leu Asp Leu Lys Tyr Arg His Val Glu Glu
 210 215 220
 Asp Gly Leu Ile Phe Glu Pro Thr Gln Val Ile Lys Ser Asn Ala Phe
 225 230 235 240
 Gly Tyr Val Val Pro His Gly Asp His Tyr His Ile Ile Pro Arg Ser
 245 250 255
 Gln Leu Ser Pro Leu Glu Met Glu Leu Ala Asp Arg Tyr Leu Ala Gly
 260 265 270
 Gln Thr Glu Asp Asn Asp Ser Gly Ser Glu His Ser Lys Pro Ser Asp
 275 280 285
 Lys Glu Val Thr His Thr Phe Leu Gly His Arg Ile Lys Ala Tyr Gly
 290 295 300
 Lys Gly Leu Asp Gly Lys Pro Tyr Asp Thr Ser Asp Ala Tyr Val Phe
 305 310 315 320
 Ser Lys Glu Ser Ile His Ser Val Asp Lys Ser Gly Val Thr Ala Lys
 325 330 335
 His Gly Asp His Phe His Tyr Ile Gly Phe Gly Glu Leu Glu Gln Tyr
 340 345 350

ES 2 278 436 T3

Glu Leu Asp Glu Val Ala Asn Trp Val Lys Ala Lys Gly Gln Ala Asp
 355 360 365
 Glu Leu Ala Ala Ala Leu Asp Gln Glu Gln Gly Lys Glu Lys Pro Leu
 370 375 380
 Phe Asp Thr Lys Lys Val Ser Arg Lys Val Thr Lys Asp Gly Lys Val
 385 390 395 400
 Gly Tyr Met Met Pro Lys Asp Gly Lys Asp Tyr Phe Tyr Ala Arg Asp
 405 410 415
 Gln Leu Asp Leu Thr Gln Ile Ala Phe Ala Glu Gln Glu Leu Met Leu
 420 425 430
 Lys Asp Lys Lys His Tyr Arg Tyr Asp Ile Val Asp Thr Gly Ile Glu
 435 440 445
 Pro Arg Leu Ala Val Asp Val Ser Ser Leu Pro Met His Ala Gly Asn
 450 455 460
 Ala Thr Tyr Asp Thr Gly Ser Ser Phe Val Ile Pro His Ile Asp His
 465 470 475 480
 Ile His Val Val Pro Tyr Ser Trp Leu Thr Arg Asp Gln Ile Ala Thr
 485 490 495
 Val Lys Tyr Val Met Gln His Pro Glu Val Arg Pro Asp Val Trp Ser
 500 505 510
 Lys Pro Gly His Glu Glu Ser Gly Ser Val Ile Pro Asn Val Thr Pro
 515 520 525
 Leu Asp Lys Arg Ala Gly Met Pro Asn Trp Gln Ile Ile His Ser Ala
 530 535 540
 Glu Glu Val Gln Lys Ala Leu Ala Glu Gly Arg Phe Ala Thr Pro Asp
 545 550 555 560
 Gly Tyr Ile Phe Asp Pro Arg Asp Val Leu Ala Lys Glu Thr Phe Val
 565 570 575
 Trp Lys Asp Gly Ser Phe Ser Ile Pro Arg Ala Asp Gly Ser Ser Leu
 580 585 590
 Arg Thr Ile Asn Lys Ser Asp Leu Ser Gln Ala Glu Trp Gln Gln Ala
 595 600 605
 Gln Glu Leu Leu Ala Lys Lys Asn Thr Gly Asp Ala Thr Asp Thr Asp
 610 615 620
 Lys Pro Lys Glu Lys Gln Gln Ala Asp Lys Ser Asn Glu Asn Gln Gln
 625 630 635 640
 Pro Ser Glu Ala Ser Lys Glu Glu Lys Glu Ser Asp Asp Phe Ile Asp
 645 650 655
 Ser Leu Pro Asp Tyr Gly Leu Asp Arg Ala Thr Leu Glu Asp His Ile
 660 665 670
 Asn Gln Leu Ala Gln Lys Ala Asn Ile Asp Pro Lys Tyr Leu Ile Phe
 675 680 685
 Gln Pro Glu Gly Val Gln Phe Tyr Asn Lys Asn Gly Glu Leu Val Thr
 690 695 700
 Tyr Asp Ile Lys Thr Leu Gln Gln Ile Asn Pro
 705 710 715

<210> 17

<211> 77

55 <212> PRT

<213> Streptococcus

<400> 17

Met His Ser Phe Ser Asn Pro Gly Tyr Pro Tyr Asp Asn Ala Val Thr
 1 5 10 15

65

ES 2 278 436 T3

Glu Ala Phe Phe Lys Tyr Leu Lys His Arg Gln Ile Asn Arg Lys His
 20 25 30
 Tyr Gln Asn Ile Lys Gln Val Gln Leu Asp Cys Phe Glu Tyr Ile Glu
 35 40 45
 5 Asn Phe Tyr Asn Asn Tyr Asn Pro His Thr Ala Asn Leu Gly Leu Thr
 50 55 60
 Pro Asn Gln Lys Glu Glu Asn Tyr Phe Asn Ala Ile Lys
 65 70 75

10 <210> 18
 <211> 86
 <212> PRT
 <213> Estreptococos

15
 <400> 18

20 Met Ala Tyr Tyr Gln Ala Cys Thr Glu Lys Asp Ile Ile Arg Ser Met
 1 5 10 15
 Ser Arg Lys Gly Thr Pro Ala Asp Asn Ala Cys Ile Glu Trp Phe His
 20 25 30
 Thr Val Leu Lys Thr Glu Thr Phe Tyr Phe His Asn Arg Arg Lys Tyr
 35 40 45
 25 Asn Lys Asp Ser Ile Thr Asn Ile Val Lys Asn Tyr Ile Thr Phe Tyr
 50 55 60
 Asn Glu Thr Arg Ile Gln Gln Arg Leu Asn Asp Gln Ser Pro Val Gln
 65 70 75 80
 30 Tyr Arg Lys Leu Ile Ala
 85

<210> 19
 <211> 126
 35 <212> PRT
 <213> Estreptococos

40
 <400> 19

Met Glu Asn His Phe Ile Tyr Gly Tyr Arg Thr Ile Thr Arg Leu Leu
 1 5 10 15
 45 Lys Lys Ile His Gly Leu Thr Val Asn Thr Lys Lys Val Tyr Arg Ile
 20 25 30
 Met Lys Asn Asn Gly Trp Leu Cys Arg Thr Arg Thr Lys Lys Val Pro
 35 40 45
 Asn Leu Gly Lys Ala Tyr Tyr Leu Thr Asp Asn Lys Leu Ser Arg Asp
 50 55 60
 50 Phe His Ala Asp Lys Pro Lys Glu Lys Leu Val Thr Asp Ile Thr Tyr
 65 70 75 80
 Leu Tyr Phe Gly Asn Cys Lys Leu Tyr Leu Ser Ser Ile Met Asn Leu
 85 90 95
 55 Tyr Asn Arg Glu Ile Ile Ala Tyr Thr Ile Ser Asp Cys Gln Asp Thr
 100 105 110
 Asp Phe Val Leu Asp Thr Leu Asn Gln Leu Lys Leu Pro Lys
 115 120 125

60 <210> 20
 <211> 96
 <212> PRT
 65 <213> Estreptococos

ES 2 278 436 T3

<400> 20

```

5      Met Val Lys Lys Ala Tyr Ser Trp Glu Thr Lys Leu Ala Cys Ile Asp
      1          5          10          15
      Met Lys Lys Ala Gly Lys Ser Asn Arg Val Ile Met Glu Thr Leu Gly
      20          25          30
      Ile Lys Asn Asn Ser Gln Ile Tyr Thr Trp Met Lys Trp Tyr Glu Asn
      35          40          45
10     Glu Glu Leu Tyr Arg Phe His Gln Gly Val Gly Lys Gln Tyr Thr Tyr
      50          55          60
      Gly Lys Gly Leu Glu His Leu Ser Glu Val Glu Gln Leu Gln Leu Gln
      65          70          75          80
15     Val Asp Leu Leu Lys Lys Tyr Arg Gly Leu Ile Arg Lys Ser Ile Lys
      85          90          95

```

<210> 21

20 <211> 288

<212> PRT

<213> Streptococos

25 <400> 21

```

30     Ile Arg Tyr Pro Lys Ala Ser Ser Gly Asp Tyr Gly Thr Lys Arg Glu
      1          5          10          15
      Ile Ile Thr Ala Asn Lys Asp Lys Tyr Ser Ile Ser Lys Met Cys Arg
      20          25          30
      Trp Leu Asn Met Pro His Ser Ser Tyr Tyr Tyr Gln Ala Val Glu Ser
      35          40          45
35     Val Ser Glu Thr Glu Phe Glu Glu Thr Ile Lys Arg Ile Phe Leu Asp
      50          55          60
      Ser Glu Ser Arg Tyr Gly Ser Arg Lys Ile Lys Ile Cys Leu Asn Asn
      65          70          75          80
40     Glu Gly Ile Thr Leu Ser Arg Arg Arg Ile Arg Arg Ile Met Lys Arg
      85          90          95
      Leu Asn Leu Val Ser Val Tyr Gln Lys Ala Thr Phe Lys Pro His Ser
      100          105          110
      Arg Gly Lys Asn Glu Ala Pro Ile Pro Asn His Leu Asp Arg Gln Phe
      115          120          125
45     Lys Gln Glu Arg Pro Leu Gln Ala Leu Val Thr Asp Leu Thr Tyr Val
      130          135          140
      Arg Val Gly Asn Arg Trp Ala Tyr Val Cys Leu Ile Ile Asp Leu Tyr
      145          150          155          160
50     Asn Arg Glu Ile Ile Gly Leu Ser Leu Gly Trp His Lys Thr Ala Glu
      165          170          175
      Leu Val Lys Gln Ala Ile Gln Ser Ile Pro Tyr Ala Leu Thr Lys Val
      180          185          190
55     Lys Met Phe His Ser Asp Arg Gly Lys Glu Phe Asp Asn Gln Leu Ile
      195          200          205
      Asp Glu Ile Leu Glu Ala Phe Gly Ile Thr Arg Ser Leu Ser Gln Ala
      210          215          220
      Gly Tyr Pro Tyr Asp Asn Ala Val Ala Glu Ser Thr Tyr Arg Ala Phe
      225          230          235          240
60     Lys Ile Glu Phe Val Tyr Gln Glu Thr Phe Gln Leu Leu Glu Glu Leu
      245          250          255
      Ala Leu Lys Thr Lys Asp Tyr Val His Trp Trp Asn Tyr His Arg Ile
      260          265          270
65     His Gly Ser Leu Asn Tyr Gln Thr Pro Met Thr Lys Arg Leu Ile Ala
      275          280          285

```


ES 2 278 436 T3

<210> 22

<211> 5058

<212> ADN

5 <213> Estreptococos

<220>

<221> CDS

10 <222> (1)...(663)

<221> CDS

15 <222> (763)...(1344)

<221> CDS

<222> (1362)...(1739)

20 <221> CDS

<222> (2266)...(5058)

<400> 22

25

aat ttg aaa gca gaa tta tct gta gaa gat gag caa tat aca gca aca	48
Asn Leu Lys Ala Glu Leu Ser Val Glu Asp Glu Gln Tyr Thr Ala Thr	
1 5 10 15	

30

gtt tat ggt aaa tct gct cat ggt tca aca cca caa gaa ggt gtt aat	96
Val Tyr Gly Lys Ser Ala His Gly Ser Thr Pro Gln Glu Gly Val Asn	
20 25 30	

35

ggg gcg act tat tta gct ctt tat cta agt caa ttt gat ttt gaa ggt	144
Gly Ala Thr Tyr Leu Ala Leu Tyr Leu Ser Gln Phe Asp Phe Glu Gly	
35 40 45	

40

cct gct cgt gct ttc tta gat gtt aca gcc aac att att cac gaa gac	192
Pro Ala Arg Ala Phe Leu Asp Val Thr Ala Asn Ile Ile His Glu Asp	
50 55 60	

45

ttc tca ggt gaa aaa ctt gga gta gct tat gaa gat gac tgt atg gga	240
Phe Ser Gly Glu Lys Leu Gly Val Ala Tyr Glu Asp Asp Cys Met Gly	
65 70 75 80	

50

cca ttg agc atg aat gca ggt gtc ttc cag ttt gat gaa act aat gat	298
Pro Leu Ser Met Asn Ala Gly Val Phe Gln Phe Asp Glu Thr Asn Asp	
85 90 95	

55

gat aat act atc gct ctt aat ttc cgt tac cca caa ggg aca gat gct	336
Asp Asn Thr Ile Ala Leu Asn Phe Arg Tyr Pro Gln Gly Thr Asp Ala	
100 105 110	

60

aaa act atc caa act aag ctt gag aaa ctt aac gga gtt gaa aaa gtg	384
Lys Thr Ile Gln Thr Lys Leu Glu Lys Leu Asn Gly Val Glu Lys Val	
115 120 125	

65

act ctt tct gac cat gaa cac aca cca cac tat gta cct atg gac gat	432
Thr Leu Ser Asp His Glu His Thr Pro His Tyr Val Pro Met Asp Asp	
130 135 140	

ES 2 278 436 T3

5	gaa tta gta tca acc tta cta gct gtc tat gaa aag caa act ggt ctt Glu Leu Val Ser Thr Leu Leu Ala Val Tyr Glu Lys Gln Thr Gly Leu 145 150 155 160	480
10	aaa gga cat gaa cag gtt att ggt ggt ggg aca ttt ggt cgc tta ctt Lys Gly His Glu Gln Val Ile Gly Gly Gly Thr Phe Gly Arg Leu Leu 165 170 175	528
15	gaa cgg ggt gtt gca tac ggt gcc atg ttc cca gga gat gaa aac act Glu Arg Gly Val Ala Tyr Gly Ala Met Phe Pro Gly Asp Glu Asn Thr 180 185 190	576
20	atg cat caa gct aat gag tac atg cct tta gaa aat att ttc cgt tcg Met His Gln Ala Asn Glu Tyr Met Pro Leu Glu Asn Ile Phe Arg Ser 195 200 205	624
25	gct gct atc tac gca gaa gct atc tat gaa tta atc aaa taaaataatc Ala Ala Ile Tyr Ala Glu Ala Ile Tyr Glu Leu Ile Lys 210 215 220	673
30	cttaaaactaa atatgtgac aatgataaag ggtggtgaag acatgaaagt gtctttgcct cttttcataa gggttagattt ggagacttt atg act gac ttg gaa aaa att att Met Thr Asp Leu Glu Lys Ile Ile 225	733 786
35	aaa gca ata aaa agt gat tca cag aat caa aat tat aca gaa aat ggt Lys Ala Ile Lys Ser Asp Ser Gln Asn Gln Asn Tyr Thr Glu Asn Gly 230 235 240 245	834
40	att gat cct ttg ttt gct gct cct aaa aca gct agg atc aat att gtt Ile Asp Pro Leu Phe Ala Ala Pro Lys Thr Ala Arg Ile Asn Ile Val 250 255 260	882
45	ggc caa gca cct ggt tta aaa act caa gaa gca aga ctc tat tgg aaa Gly Gln Ala Pro Gly Leu Lys Thr Gln Glu Ala Arg Leu Tyr Trp Lys 265 270 275	930
50	gat aaa cct gga gat cgt cta cgc cag tgg ctt gga gtt gat gaa gag Asp Lys Ser Gly Asp Arg Leu Arg Gln Trp Leu Gly Val Asp Glu Glu 280 285 290	978
55	aca ttt tac cat tct gga aaa ttt gct gtt tta cct tta gat ttt tat Thr Phe Tyr His Ser Gly Lys Phe Ala Val Leu Pro Leu Asp Phe Tyr 295 300 305	1026
60	tac cca ggc aaa gga aaa tca gga gat tta ccc cct aga aaa ggt ttt Tyr Pro Gly Lys Gly Lys Ser Gly Asp Leu Pro Pro Arg Lys Gly Phe 310 315 320 325	1074
65	gcg gag aaa tgg cac cct ctt att tta aaa gaa atg cct aat gtt caa Ala Glu Lys Trp His Pro Leu Ile Leu Lys Glu Met Pro Asn Val Glr. 330 335 340	1122
	ttg acc ttg cta gtt ggt cag tat gct cag aaa tat tat ctt gga agc Leu Thr Leu Leu Val Gly Gln Tyr Ala Gln Lys Tyr Tyr Leu Gly Ser 345 350 355	1170

ES 2 278 436 T3

	ttc gca cat aaa aat cta aca gaa aca gtt aaa gct tac aaa gac tat	1218
	Ser Ala His Lys Asn Leu Thr Glu Thr Val Lys Ala Tyr Lys Asp Tyr	
	360 365 370	
5	cta ccc gat tat tta ccc ctg gtt cac cca tca ccg cga aat caa att	1266
	Leu Pro Asp Tyr Leu Pro Leu Val His Pro Ser Pro Arg Asn Gln Ile	
	375 380 385	
10	tgg cta aag aag aat cca tgg ttt gaa aaa gat cta atc gtt gat tta	1314
	Trp Leu Lys Lys Asn Pro Trp Phe Glu Lys Asp Leu Ile Val Asp Leu	
	390 395 400 405	
15	caa aag ata gta gca gat att tta aaa gat taaggatagg agttggt atg	1364
	Gln Lys Ile Val Ala Asp Ile Leu Lys Asp Met	
	410 415	
20	aga gat aat cat cta cac acg tat ttt tcc tat gat tgt caa acg gca	1412
	Arg Asp Asn His Leu His Thr Tyr Phe Ser Tyr Asp Cys Gln Thr Ala	
	420 425 430	
25	ttt gag gac tat att aat ggt ttt aca ggt gaa ttt atc acg aca gaa	1460
	Phe Glu Asp Tyr Ile Asn Gly Phe Thr Gly Glu Phe Ile Thr Thr Glu	
	435 440 445	
30	cat ttt gat tta tca aat cct tac acc ggt caa gac gat gtt cct gat	1508
	His Phe Asp Leu Ser Asn Pro Tyr Thr Gly Gln Asp Asp Val Pro Asp	
	450 455 460	
35	tat agt gct tat tgt caa aaa ata gat tat ctt aat cag aaa tat gga	1556
	Tyr Ser Ala Tyr Cys Gln Lys Ile Asp Tyr Leu Asn Gln Lys Tyr Gly	
	465 470 475 480	
40	aat cga ttt aaa aaa gga att gaa atc ggt tat ttt aaa gat agg gaa	1604
	Asn Arg Phe Lys Lys Gly Ile Glu Ile Gly Tyr Phe Lys Asp Arg Glu	
	485 490 495	
45	tca gat att tta gat tat tta aaa aat aaa gaa ttt gat tta aaa cta	1652
	Ser Asp Ile Leu Asp Tyr Leu Lys Asn Lys Glu Phe Asp Leu Lys Leu	
	500 505 510	
50	ttg tca atc cat cat aat ggt agg tat gat tat ctg caa gaa gaa gct	1700
	Leu Ser Ile His His Asn Gly Arg Tyr Asp Tyr Leu Gln Glu Glu Ala	
	515 520 525	
55	ctg aaa gta cca aca aag gga gct ttt agc aga tta ctt taatcgtatg	1749
	Leu Lys Val Pro Thr Lys Gly Ala Phe Ser Arg Leu Leu	
	530 535 540	
60	gaatttgcca taggccgtgt ggaagcgcac gtttttagctc acttttgatta tggtttttcgt	1809
	aagtttaact tagatgtaga agatttaaaa cggtttgaaa cgcaattgaa gcgcattttc	1869
	ataaagatgt tatctaagggt gtttagctttt gaactaaata ccaaattccct ttatctatat	1929
	gggaatgaaa aacttttatcg ctatgcttta gagataactca aacagcttgg ttgtaaaaca	1989
	tactctatag gctctgacgg tcatattcct gaacattttt gttatgaatt tgatagactt	2049
	caaggctctgc taaaggacta tcaaattgat gaaaatcatt tgatatgagg aaatttttga	2109
	taaaaaagct aggcaatatt gcttagcttt tttgtaatgc tattgatagt ttttagtgaaa	2169

ES 2 278 436 T3

	atttcaaaaa aataaagaaa tcatttactt gttgcaagcg cttgcgtaaa ttggtatgat	2229
	tttattggta acaattcatt aaaaaaggag aatgat atg aaa aga aaa gac tta	2283
5	Met Lys Arg Lys Asp Leu	545
	ttt ggt gat aaa caa act caa tac acg att aga aag tta agt gtt gga	2331
10	Phe Gly Asp Lys Gln Thr Gln Tyr Thr Ile Arg Lys Leu Ser Val Gly	550 555 560
	gta gct tca gtt aca aca ggg gta tgt att ttt ctt cat agt cca cag	2379
15	Val Ala Ser Val Thr Thr Gly Val Cys Ile Phe Leu His Ser Pro Gln	565 570 575
	gta ttt gct gaa gaa gta agt gtt tct cct gca act aca gcg att gca	2427
20	Val Phe Ala Glu Glu Val Ser Val Ser Pro Ala Thr Thr Ala Ile Ala	580 585 590 595
	gag tcg aat att aat cag gtt gac aac caa caa tct act aat tta aaa	2475
25	Glu Ser Asn Ile Asn Gln Val Asp Asn Gln Gln Ser Thr Asn Leu Lys	600 605 610
	gat gac ata aac tca aac tct gag acg gtt gtg aca ccc tca gat atg	2523
30	Asp Asp Ile Asn Ser Asn Ser Glu Thr Val Val Thr Pro Ser Asp Met	615 620 625
	ccg gat acc aag caa tta gta tca gat gaa act gac act caa aag gga	2571
35	Pro Asp Thr Lys Gln Leu Val Ser Asp Glu Thr Asp Thr Gln Lys Gly	630 635 640
	gtg aca gag ccg gat aag gcg aca agc ctg ctt gaa gaa aat aaa ggt	2619
40	Val Thr Glu Pro Asp Lys Ala Thr Ser Leu Leu Glu Glu Asn Lys Gly	645 650 655
	cct gtt tca gat aaa aat acc tta gat tta aaa gta gca cca tct aca	2667
45	Pro Val Ser Asp Lys Asn Thr Leu Asp Leu Lys Val Ala Pro Ser Thr	660 665 670 675
	ttg caa aat act ccc gac aaa acc tct caa gct ata ggt gct cca agc	2715
50	Leu Gln Asn Thr Pro Asp Lys Thr Ser Gln Ala Ile Gly Ala Pro Ser	680 685 690
	cct acc ttg aaa gta gct aat caa gct cca cgg att gaa aat ggt tac	2763
55	Pro Thr Leu Lys Val Ala Asn Gln Ala Pro Arg Ile Glu Asn Gly Tyr	695 700 705
	ttt agg cta cat ctt aaa gaa ttg cct caa ggt cat cct gta gaa agc	2811
60	Phe Arg Leu His Leu Lys Glu Leu Pro Gln Gly His Pro Val Glu Ser	710 715 720
	act gga ctt tgg ata tgg gga gat gtt gat caa ccg tct agt aat tgg	2859
65	Thr Gly Leu Trp Ile Trp Gly Asp Val Asp Gln Pro Ser Ser Asn Trp	725 730 735
	cca aat ggt gct atc cct atg act gat gct aag aaa gat gat tac ggt	2907
	Pro Asn Gly Ala Ile Pro Met Thr Asp Ala Lys Lys Asp Asp Tyr Gly	740 745 750 755

ES 2 278 436 T3

	tat tat gtt gat ttt aaa tta tct gaa aaa caa cga aaa caa ata tct	2955
	Tyr Tyr Val Asp Phe Lys Leu Ser Glu Lys Gln Arg Lys Gln Ile Ser	
	760 765 770	
5	ttt tta att aat aac aaa gca ggg aca aat tta agc ggc gat cat cat	3003
	Phe Leu Ile Asn Asn Lys Ala Gly Thr Asn Leu Ser Gly Asp His His	
	775 780 785	
10	att cca tta tta cga cct gag atg aac caa gtt tgg att gat gaa aag	3051
	Ile Pro Leu Leu Arg Pro Glu Met Asn Gln Val Trp Ile Asp Glu Lys	
	790 795 800	
15	tac ggt ata cat act tat caa ccc ctc aaa gaa ggg tat gtc cgt att	3099
	Tyr Gly Ile His Thr Tyr Gln Pro Leu Lys Glu Gly Tyr Val Arg Ile	
	805 810 815	
20	aac tat ttg agt tcc tct agt aac tat gac cac tta tca gca tgg ctc	3147
	Asn Tyr Leu Ser Ser Ser Ser Asn Tyr Asp His Leu Ser Ala Trp Leu	
	820 825 830 835	
25	ttt aaa gat gtt gca acc ccy tca aca act tgg cca gat ggt agt aat	3195
	Phe Lys Asp Val Ala Thr Xaa Ser Thr Thr Trp Pro Asp Gly Ser Asn	
	840 845 850	
30	ttt gtg aat caa gga cta tat gga agg tat att gat gta tca cta aaa	3243
	Phe Val Asn Gln Gly Leu Tyr Gly Arg Tyr Ile Asp Val Ser Leu Lys	
	855 860 865	
35	act aac gcc aaa gag att ggt ttt cta atc tta gat gaa agt aag aca	3291
	Thr Asn Ala Lys Glu Ile Gly Phe Leu Ile Leu Asp Glu Ser Lys Thr	
	870 875 880	
40	gga gat gca gtg aaa gtt caa ccc aac gac tat gtt ttt aga gat tta	3339
	Gly Asp Ala Val Lys Val Gln Pro Asn Asp Tyr Val Phe Arg Asp Leu	
	885 890 895	
45	gct aac cat aac caa att ttt gta aaa gat aag gat cca aag gtt tat	3387
	Ala Asn His Asn Gln Ile Phe Val Lys Asp Lys Asp Pro Lys Val Tyr	
	900 905 910 915	
50	aat aat cct tat tac att gat caa gtg cag cta aag gat gcc caa caa	3435
	Asn Asn Pro Tyr Tyr Ile Asp Gln Val Gln Leu Lys Asp Ala Gln Gln	
	920 925 930	
55	att gat tta aca agt att caa gca agt ttt aca act cta gat ggg gta	3483
	Ile Asp Leu Thr Ser Ile Gln Ala Ser Phe Thr Thr Leu Asp Gly Val	
	935 940 945	
60	gat aaa act gaa att tta aaa gaa ttg aaa gtg act gat aaa aat caa	3531
	Asp Lys Thr Glu Ile Leu Lys Glu Leu Lys Val Thr Asp Lys Asn Gln	
	950 955 960	
65	aat gct ata caa att tct gat atc act ctc gat act agt aaa tct ctt	3579
	Asn Ala Ile Gln Ile Ser Asp Ile Thr Leu Asp Thr Ser Lys Ser Leu	
	965 970 975	

ES 2 278 436 T3

	tta ata atc aaa ggc gac ttt aat cct aaa caa ggt cat ttc aac ata	3627
	Leu Ile Ile Lys Gly Asp Phe Asn Pro Lys Gln Gly His Phe Asn Ile	
	980 985 990 995	
5	tct tat aat ggt aac aat gtc atg aca agg caa tct tgg gaa ttt aaa	3675
	Ser Tyr Asn Gly Asn Asn Val Met Thr Arg Gln Ser Trp Glu Phe Lys	
	1000 1005 1010	
10	gac caa ctt tat gct tat agt gga aat tta ggt gca gtt ctc aat caa	3723
	Asp Gln Leu Tyr Ala Tyr Ser Gly Asn Leu Gly Ala Val Leu Asn Gln	
	1015 1020 1025	
15	gat ggt tca aaa gtt gaa gcc agc ctc tgg tca ccg agt gct gat agt	3771
	Asp Gly Ser Lys Val Glu Ala Ser Leu Trp Ser Pro Ser Ala Asp Ser	
	1030 1035 1040	
20	gtc act atg att att tat gac aaa gat aac caa aac agg gtt gta gcg	3819
	Val Thr Met Ile Ile Tyr Asp Lys Asp Asn Gln Asn Arg Val Val Ala	
	1045 1050 1055	
25	act acc ccc ctt gtg aaa aat aat aaa ggt gtt tgg cag acg ata ctt	3867
	Thr Thr Pro Leu Val Lys Asn Asn Lys Gly Val Trp Gln Thr Ile Leu	
	1060 1065 1070 1075	
30	gat act aaa tta ggt att aaa aac tat act ggt tac tat tat ctt tac	3915
	Asp Thr Lys Leu Gly Ile Lys Asn Tyr Thr Gly Tyr Tyr Tyr Leu Tyr	
	1080 1085 1090	
35	gaa ata aaa aga ggt aag gat aag gtt aag att tta gat cct tat gca	3963
	Glu Ile Lys Arg Gly Lys Asp Lys Val Lys Ile Leu Asp Pro Tyr Ala	
	1095 1100 1105	
40	aag tca tta gca gag tgg gat agt aat act gtt aat gat gat att aaa	4011
	Lys Ser Leu Ala Glu Trp Asp Ser Asn Thr Val Asn Asp Asp Ile Lys	
	1110 1115 1120	
45	acg gct aaa gca gct ttt gta aat cca agt caa ctt gga cct caa aat	4059
	Thr Ala Lys Ala Ala Phe Val Asn Pro Ser Gln Leu Gly Pro Gln Asn	
	1125 1130 1135	
50	tta agt ttt gct aaa att gct aat ttt aaa gga aga caa gat gct gtt	4107
	Leu Ser Phe Ala Lys Ile Ala Asn Phe Lys Gly Arg Gln Asp Ala Val	
	1140 1145 1150 1155	
55	ata tac gaa gca cat gta aga gac ttc act tct gat cga tct ttg gat	4155
	Ile Tyr Glu Ala His Val Arg Asp Phe Thr Ser Asp Arg Ser Leu Asp	
	1160 1165 1170	
60	gga aaa tta aaa aat caa ttt ggt acc ttt gca gcc ttt tca gag aaa	4203
	Gly Lys Leu Lys Asn Gln Phe Gly Thr Phe Ala Ala Phe Ser Glu Lys	
	1175 1180 1185	
65	cta gat tat tta cag aaa tta gga gtt aca cac att cag ctt tta ccg	4251
	Leu Asp Tyr Leu Gln Lys Leu Gly Val Thr His Ile Gln Leu Leu Pro	
	1190 1195 1200	

ES 2 278 436 T3

	gta ttg agt tat ttt tat gtt aat gaa atg gat aag tca cgc tca aca	4299
	Val Leu Ser Tyr Phe Tyr Val Asn Glu Met Asp Lys Ser Arg Ser Thr	
	1205 1210 1215	
5		
	gct tac acc tcc tca gac aat aat tac aat tgg ggc tat gac cca cag	4347
	Ala Tyr Thr Ser Ser Asp Asn Asn Tyr Asn Trp Gly Tyr Asp Pro Gln	
	1220 1225 1230 1235	
10		
	agc tat ttt gct ctt tct ggg atg tat tca gag aaa cca aaa gat cca	4395
	Ser Tyr Phe Ala Leu Ser Gly Met Tyr Ser Glu Lys Pro Lys Asp Pro	
	1240 1245 1250	
15		
	tca gca cgt atc gcc gaa tta aaa caa tta ata cat gat att cat aaa	4443
	Ser Ala Arg Ile Ala Glu Leu Lys Gln Leu Ile His Asp Ile His Lys	
	1255 1260 1265	
20		
	cgt ggc atg ggg gtt ata ctt gat gtc gtc tat aat cac act gca aaa	4491
	Arg Gly Met Gly Val Ile Leu Asp Val Val Tyr Asn His Thr Ala Lys	
	1270 1275 1280	
25		
	act tat ctc ttt gag gat ata gaa cct aat tat tat cac ttt atg aat	4539
	Thr Tyr Leu Phe Glu Asp Ile Glu Pro Asn Tyr Tyr His Phe Met Asn	
	1285 1290 1295	
30		
	gaa gat ggt tca cca aga gaa agt ttt gga ggg gga cgt tta gga acc	4587
	Glu Asp Gly Ser Pro Arg Glu Ser Phe Gly Gly Gly Arg Leu Gly Thr	
	1300 1305 1310 1315	
35		
	act cat gca atg agt cgt cgt gtt ttg gtt gat tcc att aaa tat ctt	4635
	Thr His Ala Met Ser Arg Arg Val Leu Val Asp Ser Ile Lys Tyr Leu	
	1320 1325 1330	
40		
	aca agt gaa ttt aaa gtt gat ggt ttc cgt ttt gat atg atg gga gat	4683
	Thr Ser Glu Phe Lys Val Asp Gly Phe Arg Phe Asp Met Met Gly Asp	
	1335 1340 1345	
45		
	cat gat gcg gct gcg att gaa tta gct tat aaa gaa gct aaa gct att	4731
	His Asp Ala Ala Ala Ile Glu Leu Ala Tyr Lys Glu Ala Lys Ala Ile	
	1350 1355 1360	
50		
	aat cct aat atg att atg att ggt gag ggc tgg aga aca ttc caa ggc	4779
	Asn Pro Asn Met Ile Met Ile Gly Glu Gly Trp Arg Thr Phe Gln Gly	
	1365 1370 1375	
55		
	gat caa ggt cag ccg gtt aaa cca gct gac caa gat tgg atg aag tca	4827
	Asp Gln Gly Gln Pro Val Lys Pro Ala Asp Gln Asp Trp Met Lys Ser	
	1380 1385 1390 1395	
60		
	acc gat aca gtt ggc gtc ttt tca gat gat att cgt aat agc ttg aaa	4875
	Thr Asp Thr Val Gly Val Phe Ser Asp Asp Ile Arg Asn Ser Leu Lys	
	1400 1405 1410	
65		
	tct ggt ttt cca aat gaa ggt act cca gct ttc atc aca ggt ggc cca	4923
	Ser Gly Phe Pro Asn Glu Gly Thr Pro Ala Phe Ile Thr Gly Gly Pro	
	1415 1420 1425	

ES 2 278 436 T3

caa tct tta caa ggt att ttt aaa aat atc aaa gca caa cct ggg aat 4971
Gln Ser Leu Gln Gly Ile Phe Lys Asn Ile Lys Ala Gln Pro Gly Asn
1430 1435 1440

5 ttt gaa gca gat tcg cca gga gat gtg gtg cag tat att gct gca cat 5019
Phe Glu Ala Asp Ser Pro Gly Asp Val Val Gln Tyr Ile Ala Ala His
1445 1450 1455

10 gat aac ctt acc ttg cat gat gtg att gca aaa tca att 5058
Asp Asn Leu Thr Leu His Asp Val Ile Ala Lys Ser Ile
1460 1465 1470

15 <210> 23
<211> 221
<212> PRT
<213> Streptococos
20
<400> 23

25 Asn Leu Lys Ala Glu Leu Ser Val Glu Asp Glu Gln Tyr Thr Ala Thr
1 5 10 15
Val Tyr Gly Lys Ser Ala His Gly Ser Thr Pro Gln Glu Gly Val Asn
20 25 30
Gly Ala Thr Tyr Leu Ala Leu Tyr Leu Ser Gln Phe Asp Phe Glu Gly
35 40 45
Pro Ala Arg Ala Phe Leu Asp Val Thr Ala Asn Ile Ile His Glu Asp
50 55 60
Phe Ser Gly Glu Lys Leu Gly Val Ala Tyr Glu Asp Asp Cys Met Gly
65 70 75 80
35 Pro Leu Ser Met Asn Ala Gly Val Phe Gln Phe Asp Glu Thr Asn Asp
85 90 95
Asp Asn Thr Ile Ala Leu Asn Phe Arg Tyr Pro Gln Gly Thr Asp Ala
100 105 110
40 Lys Thr Ile Gln Thr Lys Leu Glu Lys Leu Asn Gly Val Glu Lys Val
115 120 125
Thr Leu Ser Asp His Glu His Thr Pro His Tyr Val Pro Met Asp Asp
130 135 140
Glu Leu Val Ser Thr Leu Leu Ala Val Tyr Glu Lys Gln Thr Gly Leu
45 145 150 155 160
Lys Gly His Glu Gln Val Ile Gly Gly Gly Thr Phe Gly Arg Leu Leu
165 170 175
Glu Arg Gly Val Ala Tyr Gly Ala Met Phe Pro Gly Asp Glu Asn Thr
180 185 190
50 Met His Gln Ala Asn Glu Tyr Met Pro Leu Glu Asn Ile Phe Arg Ser
195 200 205
Ala Ala Ile Tyr Ala Glu Ala Ile Tyr Glu Leu Ile Lys
210 215 220

55
<210> 24
<211> 194
60 <212> PRT
<213> Streptococos

65

ES 2 278 436 T3

<400> 24

```

5      Met Thr Asp Leu Glu Lys Ile Ile Lys Ala Ile Lys Ser Asp Ser Gln
      1          5          10          15
      Asn Gln Asn Tyr Thr Glu Asn Gly Ile Asp Pro Leu Phe Ala Ala Pro
      20          25          30
      Lys Thr Ala Arg Ile Asn Ile Val Gly Gln Ala Pro Gly Leu Lys Thr
      35          40          45
10     Gln Glu Ala Arg Leu Tyr Trp Lys Asp Lys Ser Gly Asp Arg Leu Arg
      50          55          60
      Gln Trp Leu Gly Val Asp Glu Glu Thr Phe Tyr His Ser Gly Lys Phe
      65          70          75          80
15     Ala Val Leu Pro Leu Asp Phe Tyr Tyr Pro Gly Lys Gly Lys Ser Gly
      85          90          95
      Asp Leu Pro Pro Arg Lys Gly Phe Ala Glu Lys Trp His Pro Leu Ile
      100         105         110
20     Leu Lys Glu Met Pro Asn Val Gln Leu Thr Leu Leu Val Gly Gln Tyr
      115         120         125
      Ala Gln Lys Tyr Tyr Leu Gly Ser Ser Ala His Lys Asn Leu Thr Glu
      130         135         140
25     Thr Val Lys Ala Tyr Lys Asp Tyr Leu Pro Asp Tyr Leu Pro Leu Val
      145         150         155         160
      His Pro Ser Pro Arg Asn Gln Ile Trp Leu Lys Lys Asn Pro Trp Phe
      165         170         175
      Glu Lys Asp Leu Ile Val Asp Leu Gln Lys Ile Val Ala Asp Ile Leu
      180         185         190
30     Lys Asp

```

<210> 25

35 <211> 126

<212> PRT

<213> Estreptococos

40 <400> 25

```

45     Met Arg Asp Asn His Leu His Thr Tyr Phe Ser Tyr Asp Cys Gln Thr
      1          5          10          15
      Ala Phe Glu Asp Tyr Ile Asn Gly Phe Thr Gly Glu Phe Ile Thr Thr
      20          25          30
      Glu His Phe Asp Leu Ser Asn Pro Tyr Thr Gly Gln Asp Asp Val Pro
      35          40          45
50     Asp Tyr Ser Ala Tyr Cys Gln Lys Ile Asp Tyr Leu Asn Gln Lys Tyr
      50          55          60
      Gly Asn Arg Phe Lys Lys Gly Ile Glu Ile Gly Tyr Phe Lys Asp Arg
      65          70          75          80
      Glu Ser Asp Ile Leu Asp Tyr Leu Lys Asn Lys Glu Phe Asp Leu Lys
      85          90          95
55     Leu Leu Ser Ile His His Asn Gly Arg Tyr Asp Tyr Leu Gln Glu Glu
      100         105         110
      Ala Leu Lys Val Pro Thr Lys Gly Ala Phe Ser Arg Leu Leu
      115         120         125
60

```

<210> 26

<211> 931

<212> PRT

65 <213> Estreptococos

ES 2 278 436 T3

<400> 26

```

5      Met Lys Arg Lys Asp Leu Phe Gly Asp Lys Gln Thr Gln Tyr Thr Ile
      1          5          10          15
      Arg Lys Leu Ser Val Gly Val Ala Ser Val Thr Thr Gly Val Cys Ile
      20          25          30
10     Phe Leu His Ser Pro Gln Val Phe Ala Glu Glu Val Ser Val Ser Pro
      35          40          45
      Ala Thr Thr Ala Ile Ala Glu Ser Asn Ile Asn Gln Val Asp Asn Gln
      50          55          60
      Gln Ser Thr Asn Leu Lys Asp Asp Ile Asn Ser Asn Ser Glu Thr Val
15     65          70          75          80
      Val Thr Pro Ser Asp Met Pro Asp Thr Lys Gln Leu Val Ser Asp Glu
      85          90          95
      Thr Asp Thr Gln Lys Gly Val Thr Glu Pro Asp Lys Ala Thr Ser Leu
      100          105          110
20     Leu Glu Glu Asn Lys Gly Pro Val Ser Asp Lys Asn Thr Leu Asp Leu
      115          120          125
      Lys Val Ala Pro Ser Thr Leu Gln Asn Thr Pro Asp Lys Thr Ser Gln
      130          135          140
      Ala Ile Gly Ala Pro Ser Pro Thr Leu Lys Val Ala Asn Gln Ala Pro
25     145          150          155          160
      Arg Ile Glu Asn Gly Tyr Phe Arg Leu His Leu Lys Glu Leu Pro Gln
      165          170          175
      Gly His Pro Val Glu Ser Thr Gly Leu Trp Ile Trp Gly Asp Val Asp
      180          185          190
30     Gln Pro Ser Ser Asn Trp Pro Asn Gly Ala Ile Pro Met Thr Asp Ala
      195          200          205
      Lys Lys Asp Asp Tyr Gly Tyr Tyr Val Asp Phe Lys Leu Ser Glu Lys
      210          215          220
35     Gln Arg Lys Gln Ile Ser Phe Leu Ile Asn Asn Lys Ala Gly Thr Asn
      225          230          235          240
      Leu Ser Gly Asp His His Ile Pro Leu Leu Arg Pro Glu Met Asn Gln
      245          250          255
      Val Trp Ile Asp Glu Lys Tyr Gly Ile His Thr Tyr Gln Pro Leu Lys
40     260          265          270
      Glu Gly Tyr Val Arg Ile Asn Tyr Leu Ser Ser Ser Ser Asn Tyr Asp
      275          280          285
      His Leu Ser Ala Trp Leu Phe Lys Asp Val Ala Thr Xaa Ser Thr Thr
45     290          295          300
      Trp Pro Asp Gly Ser Asn Phe Val Asn Gln Gly Leu Tyr Gly Arg Tyr
      305          310          315          320
      Ile Asp Val Ser Leu Lys Thr Asn Ala Lys Glu Ile Gly Phe Leu Ile
      325          330          335
50     Leu Asp Glu Ser Lys Thr Gly Asp Ala Val Lys Val Gln Pro Asn Asp
      340          345          350
      Tyr Val Phe Arg Asp Leu Ala Asn His Asn Gln Ile Phe Val Lys Asp
      355          360          365
      Lys Asp Pro Lys Val Tyr Asn Asn Pro Tyr Tyr Ile Asp Gln Val Gln
55     370          375          380
      Leu Lys Asp Ala Gln Gln Ile Asp Leu Thr Ser Ile Gln Ala Ser Phe
      385          390          395          400
      Thr Thr Leu Asp Gly Val Asp Lys Thr Glu Ile Leu Lys Glu Leu Lys
      405          410          415
60     Val Thr Asp Lys Asn Gln Asn Ala Ile Gln Ile Ser Asp Ile Thr Leu
      420          425          430
      Asp Thr Ser Lys Ser Leu Leu Ile Ile Lys Gly Asp Phe Asn Pro Lys
      435          440          445

```

65

ES 2 278 436 T3

	Gln	Gly	His	Phe	Asn	Ile	Ser	Tyr	Asn	Gly	Asn	Asn	Val	Met	Thr	Arg
	450						455				460					
5	Gln	Ser	Trp	Glu	Phe	Lys	Asp	Gln	Leu	Tyr	Ala	Tyr	Ser	Gly	Asn	Leu
	465					470					475					480
	Gly	Ala	Val	Leu	Asn	Gln	Asp	Gly	Ser	Lys	Val	Glu	Ala	Ser	Leu	Trp
					485					490					495	
	Ser	Pro	Ser	Ala	Asp	Ser	Val	Thr	Met	Ile	Ile	Tyr	Asp	Lys	Asp	Asn
				500					505					510		
10	Gln	Asn	Arg	Val	Val	Ala	Thr	Thr	Pro	Leu	Val	Lys	Asn	Asn	Lys	Gly
			515					520					525			
	Val	Trp	Gln	Thr	Ile	Leu	Asp	Thr	Lys	Leu	Gly	Ile	Lys	Asn	Tyr	Thr
	530					535						540				
15	Gly	Tyr	Tyr	Tyr	Leu	Tyr	Glu	Ile	Lys	Arg	Gly	Lys	Asp	Lys	Val	Lys
	545					550					555					560
	Ile	Leu	Asp	Pro	Tyr	Ala	Lys	Ser	Leu	Ala	Glu	Trp	Asp	Ser	Asn	Thr
					565					570					575	
	Val	Asn	Asp	Asp	Ile	Lys	Thr	Ala	Lys	Ala	Ala	Phe	Val	Asn	Pro	Ser
20				580					585					590		
	Gln	Leu	Gly	Pro	Gln	Asn	Leu	Ser	Phe	Ala	Lys	Ile	Ala	Asn	Phe	Lys
	595						600						605			
	Gly	Arg	Gln	Asp	Ala	Val	Ile	Tyr	Glu	Ala	His	Val	Arg	Asp	Phe	Thr
	610					615						620				
25	Ser	Asp	Arg	Ser	Leu	Asp	Gly	Lys	Leu	Lys	Asn	Gln	Phe	Gly	Thr	Phe
	625					630					635					640
	Ala	Ala	Phe	Ser	Glu	Lys	Leu	Asp	Tyr	Leu	Gln	Lys	Leu	Gly	Val	Thr
					645					650					655	
30	His	Ile	Gln	Leu	Leu	Pro	Val	Leu	Ser	Tyr	Phe	Tyr	Val	Asn	Glu	Met
				660					665					670		
	Asp	Lys	Ser	Arg	Ser	Thr	Ala	Tyr	Thr	Ser	Ser	Asp	Asn	Asn	Tyr	Asn
			675				680						685			
	Trp	Gly	Tyr	Asp	Pro	Gln	Ser	Tyr	Phe	Ala	Leu	Ser	Gly	Met	Tyr	Ser
35	690					695						700				
	Glu	Lys	Pro	Lys	Asp	Pro	Ser	Ala	Arg	Ile	Ala	Glu	Leu	Lys	Gln	Leu
	705					710					715					720
	Ile	His	Asp	Ile	His	Lys	Arg	Gly	Met	Gly	Val	Ile	Leu	Asp	Val	Val
					725					730					735	
40	Tyr	Asn	His	Thr	Ala	Lys	Thr	Tyr	Leu	Phe	Glu	Asp	Ile	Glu	Pro	Asn
				740					745					750		
	Tyr	Tyr	His	Phe	Met	Asn	Glu	Asp	Gly	Ser	Pro	Arg	Glu	Ser	Phe	Gly
			755				760						765			
45	Gly	Gly	Arg	Leu	Gly	Thr	Thr	His	Ala	Met	Ser	Arg	Arg	Val	Leu	Val
	770						775					780				
	Asp	Ser	Ile	Lys	Tyr	Leu	Thr	Ser	Glu	Phe	Lys	Val	Asp	Gly	Phe	Arg
	785					790					795					800
	Phe	Asp	Met	Met	Gly	Asp	His	Asp	Ala	Ala	Ala	Ile	Glu	Leu	Ala	Tyr
					805					810					815	
50	Lys	Glu	Ala	Lys	Ala	Ile	Asn	Pro	Asn	Met	Ile	Met	Ile	Gly	Glu	Gly
				820					825					830		
	Trp	Arg	Thr	Phe	Gln	Gly	Asp	Gln	Gly	Gln	Pro	Val	Lys	Pro	Ala	Asp
				835				840					845			
55	Gln	Asp	Trp	Met	Lys	Ser	Thr	Asp	Thr	Val	Gly	Val	Phe	Ser	Asp	Asp
	850					855						860				
	Ile	Arg	Asn	Ser	Leu	Lys	Ser	Gly	Phe	Pro	Asn	Glu	Gly	Thr	Pro	Ala
	865					870					875					880
60	Phe	Ile	Thr	Gly	Gly	Pro	Gln	Ser	Leu	Gln	Gly	Ile	Phe	Lys	Asn	Ile
					885					890					895	

ES 2 278 436 T3

Lys Ala Gln Pro Gly Asn Phe Glu Ala Asp Ser Pro Gly Asp Val Val
 900 905 910
 Gln Tyr Ile Ala Ala His Asp Asn Leu Thr Leu His Asp Val Ile Ala
 915 920 925
 5 Lys Ser Ile
 930

<210> 27

10 <211> 5607

<212> ADN

<213> Streptococcus

15 <220>

<221> CDS

<222> (2)...(301)

20 <400> 27

a att caa agt ttg aca gaa ggt caa ctt cgt tct gat atc cct gag ttc 49
 Ile Gln Ser Leu Thr Glu Gly Gln Leu Arg Ser Asp Ile Pro Glu Phe
 25 1 5 10 15
 cgt gct ggt gat act gta cgt gtt cac gct aaa gtt gtt gaa ggt act 97
 Arg Ala Gly Asp Thr Val Arg Val His Ala Lys Val Val Glu Gly Thr
 30 20 25 30
 cgc gaa cgt att cag atc ttt gaa ggt gtt gtt atc tca cgt aaa ggt 145
 Arg Glu Arg Ile Gln Ile Phe Glu Gly Val Val Ile Ser Arg Lys Gly
 35 35 40 45
 caa gga atc tca gaa atg tac aca gta cgt aaa att tct ggt ggt atc 193
 Gln Gly Ile Ser Glu Met Tyr Thr Val Arg Lys Ile Ser Gly Gly Ile
 50 55 60
 ggt gta gag cgt aca ttc cca att cac act cct cgt gtt gat aaa atc 241
 Gly Val Glu Arg Thr Phe Pro Ile His Thr Pro Arg Val Asp Lys Ile
 65 70 75 80
 gaa gtt gtt cgt tat ggt aaa gta cgt cgt gct aaa ctt tac tac tta 289
 Glu Val Val Arg Tyr Gly Lys Val Arg Arg Ala Lys Leu Tyr Tyr Leu
 85 90 95
 cgc gca ttg caa ggttaaagctg cacgtattaa agaaatccgt cgtaaatttt 341
 Arg Ala Leu Gln
 100
 gatgatcaga ttttaaaaaat gcttggttgt ttgaggatag taactatggt ttaaaaactgg 401
 acaaccaaga cgtaaaaaat ctgcctgtgg gcagtttttt tactagggtcc ccttagttca 461
 55 atggatataa caactccctc ctaaggagta attgctggtt cgattccggc aggggacata 521
 ttcatgtcat gtaaatagcg gtttagagct attttgcccc aaatttctct gattaagtct 581
 atcgttccca tcttttttgtt cttgtaattg atgtgcgtaa acttctaaaag tgatatttaa 641
 attctcgtga tctaaaactt gagagatgga aattagatag cttgcaaatg tatgcctgag 701
 agagtgcact cgtacctcgc gaccagttat ttttcggata gttttattga ctgcattact 761
 60 tgaaagtctg tccaataatc tctcgttttt attttttgta aattcatgca aaaaaataa 821
 cgtatcattg tcaattggta taattctgat actacttttg ttttttgttg gcaggatatct 881
 ttggttgaaa tgataatccc aagtttttatt aattgataaa tatttgttag tgtaatcaat 941
 atcattaact gttaaaccta aacattcagc gaagcgcgtg ccagtttttag cgatgaggtta 1001
 65

ES 2 278 436 T3

	taacgctgca	tacgattgat	gttgtgattt	ttctttacaa	atttttatca	agcgtaaagta	1061
	ttcattgggt	tcaagaaatt	ttatctctat	ttacgccctt	tattttttgc	tttaacctta	1121
	gtgaataaac	aaaaattttt	ttctatatat	ccctcgtgaa	cagccatgga	tacgcaggct	1181
5	tttacatgta	cgttaaaacg	ctttactgta	tcttgcacat	gcgtttgact	ataatgattt	1241
	atgacttggt	gatatattagt	ggaagtaata	ttgcaaagta	atatatttcc	tatttatatg	1301
	ttatacgata	ttcgatattc	ccaccggttg	tgcggtttac	ggaaatacgc	cattgatata	1361
	ctccacatta	gctaaagaac	agggtgttca	aggctacctt	gatggaaaag	gctctcttag	1421
	agatatttgt	aaatggtatg	atatctcaag	tgcgtctgtt	ctccaaaagt	ggataaaacg	1481
10	gtatactagt	ggtgaagact	tgaaagccac	tagtagagga	tatagccgta	tgaaacaagg	1541
	aaggcaagcc	acatttgaag	aacgtgtaga	gattgttaac	tacaccattg	cccatgggaa	1601
	agactatcaa	gcagctattg	agaagtttgg	tgtttcctac	caacaaattt	attcttgggt	1661
	gcgtaagctt	gagaagaatg	gctcacaaag	tttggttgat	agacgtgtga	aagggttgga	1721
15	gagtaggcct	gatttaaccg	agattgagca	actttaactc	aagattaaac	aattggagga	1781
	acgtaatcgt	ctcttagaaa	tgcaggttag	tttactaaaa	aagttagaag	acatcaaacy	1841
	aggaaacaga	cggtaagact	aggtaagcat	ttagcggagt	tccaagtaat	caagaattrat	1901
	tacgatgagg	aactctaagt	gcctattcag	gccttatgcc	aactcttgaa	ggggtctcgt	1961
	tcaggctatt	acaagtggct	caactcgtca	aaaacagatt	ttgagacaaa	aaatacaaa	2021
20	ctaattggct	aaatcaagga	acttcgtaga	ctctacaatg	gtatcttagg	ttatcgccgt	2081
	atgacaacat	ttattaatcg	tcaacttggg	acaacttaaa	acaagaaacy	gattcgttga	2141
	ttgatgaaca	ttctggggat	tagttcagtc	attcgtcgtg	ttagccatgc	ttgtacaaaa	2201
	gctgggtgaca	gattttacga	agaaaatatt	cttaatcgtg	aattttacagc	cacagctcat	2261
	aaccagaaat	ggtgcacaga	tgtcacctat	cttcaatacg	gtctggggagc	taaagcttat	2321
25	ctcagtgcga	ttaaagacct	gtataacggt	tctatttatcg	cttatgagat	tagtcacaac	2381
	aatgaaatcc	acttggttatg	aagaccatta	aaaaggggct	agagctcaat	ccaggagcca	2441
	cacctatcat	ccatagcgat	tgaggtagtc	aatatacttc	caaagaatac	cgttatatca	2501
	tacaacaagc	tggctctgacc	ttatccatgt	cccggattgg	caaatgtatt	gataatgcac	2561
30	caactgaaag	tttctttggg	tttttcaaga	ctgagtcctta	ccaccttaag	aaatacaact	2621
	ctttatgatga	ggtgggtcaat	gatgtggcac	gttatatcga	attctacaac	acacaacggt	2681
	atcaatcaaaa	attaaaacaac	ctgactcctc	tagaattcag	gaatcagggt	gcataactta	2741
	tctttttatta	tttgactgtc	tacttgacag	ggagccggtt	agattgctta	acctttctaa	2801
	atttgctaaa	atagctacaa	gaaaacgagc	catttaatgc	ttattttctta	tactgtcttg	2861
35	cctcacgctc	tcctcgacca	aaaattgagc	gtgaggcttt	ttgtttcatt	aaacgatgat	2921
	atttccatat	tcactcagttt	gttttccgag	agccatcaaa	gcttcgataa	ggtcgataat	2981
	tccagggaata	aaggtaatac	taaaaataat	atataaaaaa	acctggccta	tttttcctgc	3041
	gtaaaattta	tgcgctccaa	tgccgcccaa	aagaacgtta	ataaaacata	aactactatg	3101
	ttagcataag	actttatttt	tacaactgaa	tttcatataa	atggattaga	gtaagggata	3161
40	aaagaaatta	gcatagctct	tttgaaaata	aaaaaattaa	tataatatgg	aaaaaatttt	3221
	atttcataaaa	cgtttcataa	aagggtatgta	atctagtatt	taggcaacac	tattttgtca	3281
	ctgggtgtcta	gtaacttata	gattgataat	tttactagta	aacgtaattc	ttcgctttaa	3341
	gagttaaaatg	tctattttatt	gtaagctaaa	ttgggagggtg	aacttatgta	aaattagata	3401
45	gggtactgtca	agtacgggat	gattattgaa	acagccagta	tgcatacata	aatctgtatt	3461
	gcttaataaac	tatttcctta	accagacatc	agttcattgt	ttatcatcgc	taccctaagt	3521
	ctagtttttt	caatagagca	ttaggtagtt	tttgataata	aaactatata	aacatgagaa	3581
	ttagattttcg	tattgcattc	ttcataatga	gttatttgag	attttccttt	gaataaatag	3641
	atacgaaatt	cagtaacttc	atatataaac	ggctctatca	ttgagatagt	ttgtc aaatg	3701
50	aagaaattttt	taatggaaat	agtttttaaaa	acattagttg	taggcgatgt	aaaaatatta	3761
	atccagtgga	tgcaatagtt	gcggagtaaa	aatagagagg	agtaattagg	aagtgtataa	3821
	aaatgctata	gcatatatta	ccagaaaaaa	aatagaaca	cttattatat	ttgctatttt	3881
	aacaattgtt	ctttctttgct	tgtattcatg	tttaacaata	atgaaatcaa	gtaatgaaat	3941
	agaaaaggct	ttatatgaaa	gttcttaattc	ttcaatatca	attacaaaaa	aagatgggtaa	4001
55	atatttttaat	attaatcaat	ttaagaatat	tgaaaaaata	aaagagggtg	aagaaaaaat	4061
	attttcaatat	gatggattag	caaaaattgaa	agattcttaa	gtagttagtg	gtgagcaaa	4121
	tataaataga	gaagattttat	ctgacgaatt	taaaaatggt	gtttcactag	aagctacaag	4181
	taataactaaa	agaaatcttt	tatttagtag	tggagtagtt	agtttttaag	aaggaaaaaa	4241
60	tatagaagaa	aatgataaga	attcaattct	tgttcatgaa	gaatttgcta	aacaaaacaa	4301
	actaaaattg	ggcgatgaaa	ttgatcttga	attactagat	acggaaaaaa	gtggaaaaat	4361

ES 2 278 436 T3

```

aaaaagtcac aaatttataa ttataggaat cttttctggt aaaaaacagg aaacatatat 4421
aggattatca tctgatttta gcgaaaatat ggtttttgta gattattcaa cttagccaaga 4481
aatattaaat aaatcagaga ataatagaat tgcaataaaa attttaatgt attctggtag 4541
5 tttagaatct acagagcttg ccttaaacia attgaaagac tttaaaattg ataagtcaaa 4601
gtattctatt aagaaagata ataaagcatt cgaagagctt ttagagtcag tgagtggaa 4661
aaaacatata attaaaataa tgacttattc gattatgta ggtggaatag ttgttctttc 4721
attaatcttg attctatggt taagagaaag aatttatgaa ataggtatat ttttatctat 4781
tggaacaact aagatacaaa ttataaggca atttatattt gagttaatat tcatatcaat 4841
10 accaagtata atatcctcct tttttttagg gaatctacta ttaaaagtaa ttgtagaagg 4901
atattattaac tcagagaact caatgatttt cggcggaagt ttaataaata aaagcagttt 4961
tatgttaaac ataacaacac ttgcagaaag ttatttaata ttaataagta ttattgtttt 5021
atcagttgta atggcctctt cattaatatt atttaagaaa ccacaagaaa tattatcaaa 5081
15 aataagttag gagcaataaa tggatatatt agaaataaag aatgtaaatt acagttacgc 5141
aaattctaaa gaaaaagttt tgtcaggagt aaatcaaaaa tttgaacttg gaaagtttta 5201
tgcatagta ggggaagtcag gaacaggaaa atccacactt ctttccttac ttgcaggact 5261
tgataaagt tcaaacaggaa aaatcttggt taagaatgaa gatatagaaa agaaaggata 5321
tagtaatcac agaaaaata atatatcttt ggtatttcaa aattataatt taatagatta 5381
20 tttatcgccg attgaaaaata ttagactagt aaataaatca gtagatgaga gtatcttggt 5441
cgaattaggt ttagataaaa aacaaataaa aagaaatggt atgaaattat ctggtggtca 5501
gcaacaaagg gtagctattg ctagggcact ggtatcagat gccccaataa tactagctga 5561
tgagcctacc ggtaacctag acagtgttac tgctggagaa ataatt 5607

```

```

25 <210> 28
    <211> 111
    <212> PRT
30 <213> Streptococos

    <400> 28

```

```

35 Ile Gln Ser Leu Thr Glu Gly Gln Leu Arg Ser Asp Ile Pro Glu Phe
    1 5 10 15
Arg Ala Gly Asp Thr Val Arg Val His Ala Lys Val Val Glu Gly Thr
    20 25 30
Arg Glu Arg Ile Gln Ile Phe Glu Gly Val Val Ile Ser Arg Lys Gly
    35 40 45
40 Gln Gly Ile Ser Glu Met Tyr Thr Val Arg Lys Ile Ser Gly Gly Ile
    50 55 60
Gly Val Glu Arg Thr Phe Pro Ile His Thr Pro Arg Val Asp Lys Ile
    65 70 75 80
45 Glu Val Val Arg Tyr Gly Lys Val Arg Arg Ala Lys Leu Tyr Tyr Leu
    85 90 95
Arg Ala Leu Gln Gly Lys Ala Ala Arg Ile Lys Glu Ile Arg Arg
    100 105 110

```

```

50 <210> 29
    <211> 173
    <212> PRT
55 <213> Streptococos

    <400> 29

```

```

60 Met Arg Phe Ala Glu Cys Leu Gly Leu Thr Val Asn Asp Ile Asp Tyr
    1 5 10 15
Thr Asn Lys Tyr Leu Ser Ile Asn Lys Thr Trp Asp Tyr His Phe Asn
    20 25 30
Gln Arg Tyr Leu Pro Thr Lys Asn Lys Ser Ser Ile Arg Asn Ile Pro
    35 40 45
65 Ile Asp Asn Asp Thr Leu Phe Phe Leu His Glu Phe Thr Lys Asn Lys

```

ES 2 278 436 T3

```

      50      55      60
Asn Asp Arg Leu Phe Asp Lys Leu Ser Asn Asn Ala Val Asn Lys Thr
65      70      75      80
5  Ile Arg Lys Ile Thr Gly Arg Glu Val Arg Val His Ser Leu Arg His
      85      90      95
Thr Phe Ala Ser Tyr Leu Ile Ser Ile Ser Gln Val Leu Asp His Glu
      100      105      110
10 Asn Leu Asn Ile Thr Leu Glu Val Tyr Ala His Gln Leu Gln Glu Gln
      115      120      125
Lys Asp Arg Asn Asp Lys Leu Asn Gln Arg Asn Leu Gly Gln Asn Ser
      130      135      140
15 Ser Lys Pro Leu Phe Thr Cys Asn Glu Tyr Val Pro Cys Arg Asn Arg
      145      150      155      160
Thr Ser Asn Tyr Ser Leu Gly Gly Ser Cys Tyr Ile His
      165      170

```

<210> 30

20 <211> 389

<212> PRT

<213> Estreptococos

25 <400> 30

```

Met Lys Ser Ser Asn Glu Ile Glu Lys Ala Leu Tyr Glu Ser Ser Asn
1      5      10      15
30 Ser Ser Ile Ser Ile Thr Lys Lys Asp Gly Lys Tyr Phe Asn Ile Asn
      20      25      30
Gln Phe Lys Asn Ile Glu Lys Ile Lys Glu Val Glu Glu Lys Ile Phe
      35      40      45
35 Gln Tyr Asp Gly Leu Ala Lys Leu Lys Asp Leu Lys Val Val Ser Gly
      50      55      60
Glu Gln Ser Ile Asn Arg Glu Asp Leu Ser Asp Glu Phe Lys Asn Val
65      70      75      80
40 Val Ser Leu Glu Ala Thr Ser Asn Thr Lys Arg Asn Leu Leu Phe Ser
      85      90      95
Ser Gly Val Phe Ser Phe Lys Glu Gly Lys Asn Ile Glu Glu Asn Asp
      100      105      110
Lys Asn Ser Ile Leu Val His Glu Glu Phe Ala Lys Gln Asn Lys Leu
      115      120      125
45 Lys Leu Gly Asp Glu Ile Asp Leu Glu Leu Leu Asp Thr Glu Lys Ser
      130      135      140
Gly Lys Ile Lys Ser His Lys Phe Lys Ile Ile Gly Ile Phe Ser Gly
145      150      155      160
50 Lys Lys Gln Glu Thr Tyr Thr Gly Leu Ser Ser Asp Phe Ser Glu Asn
      165      170      175
Met Val Phe Val Asp Tyr Ser Thr Ser Gln Glu Ile Leu Asn Lys Ser
      180      185      190
55 Glu Asn Asn Arg Ile Ala Asn Lys Ile Leu Met Tyr Ser Gly Ser Leu
      195      200      205
Glu Ser Thr Glu Leu Ala Leu Asn Lys Leu Lys Asp Phe Lys Ile Asp
      210      215      220
Lys Ser Lys Tyr Ser Ile Lys Lys Asp Asn Lys Ala Phe Glu Glu Ser
225      230      235      240
60 Leu Glu Ser Val Ser Gly Ile Lys His Ile Ile Lys Ile Met Thr Tyr
      245      250      255
Ser Ile Met Leu Gly Gly Ile Val Val Leu Ser Leu Ile Leu Ile Leu
      260      265      270

```

65

ES 2 278 436 T3

Trp Leu Arg Glu Arg Ile Tyr Glu Ile Gly Ile Phe Leu Ser Ile Gly
 275 280 285
 Thr Thr Lys Ile Gln Ile Ile Arg Gln Phe Ile Phe Glu Leu Ile Phe
 290 295 300
 Ile Ser Ile Pro Ser Ile Ile Ser Ser Leu Phe Leu Gly Asn Leu Leu
 305 310 315 320
 Leu Lys Val Ile Val Glu Gly Phe Ile Asn Ser Glu Asn Ser Met Ile
 325 330 335
 Phe Gly Gly Ser Leu Ile Asn Lys Ser Ser Phe Met Leu Asn Ile Thr
 340 345 350
 Thr Leu Ala Glu Ser Tyr Leu Ile Leu Ile Ser Ile Ile Val Leu Ser
 355 360 365
 Val Val Met Ala Ser Ser Leu Ile Leu Phe Lys Lys Pro Gln Glu Ile
 370 375 380
 Leu Ser Lys Ile Ser
 385

20 <210> 31
 <211> 169
 <212> PRT
 25 <213> Estreptococos
 <400> 31

Met Asp Ile Leu Glu Ile Lys Asn Val Asn Tyr Ser Tyr Ala Asn Ser
 1 5 10 15
 Lys Glu Lys Val Leu Ser Gly Val Asn Gln Lys Phe Glu Leu Gly Lys
 20 25 30
 Phe Tyr Ala Ile Val Gly Lys Ser Gly Thr Gly Lys Ser Thr Leu Leu
 35 35 40 45
 Ser Leu Leu Ala Gly Leu Asp Lys Val Gln Thr Gly Lys Ile Leu Phe
 50 55 60
 Lys Asn Glu Asp Ile Glu Lys Lys Gly Tyr Ser Asn His Arg Lys Asn
 65 70 75 80
 Asn Ile Ser Leu Val Phe Gln Asn Tyr Asn Leu Ile Asp Tyr Leu Ser
 85 90 95
 Pro Ile Glu Asn Ile Arg Leu Val Asn Lys Ser Val Asp Glu Ser Ile
 100 105 110
 Leu Phe Glu Leu Gly Leu Asp Lys Lys Gln Ile Lys Arg Asn Val Met
 45 115 120 125
 Lys Leu Ser Gly Gly Gln Gln Gln Arg Val Ala Ile Ala Arg Ala Leu
 130 135 140
 Val Ser Asp Ala Pro Ile Ile Leu Ala Asp Glu Pro Thr Gly Asn Leu
 145 150 155 160
 Asp Ser Val Thr Ala Gly Glu Ile Ile
 165

55 <210> 32
 <211> 4171
 <212> ADN
 <213> Estreptococos
 60 <400> 32

catatgacaa tatttttcaa agtctacatc acttactcgc ctgtcgtgga aaatctggca 60
 atacattaat cgaccaatta gttgctgatg gtttacttca tgcagataat cactaccatt 120
 ttttcaatgg gaagtctctg gccactttca atactaacca attgattcgc gaagttgtct 180
 atgttgaaat atccttagat actatgtcta gtggtgaaca tgatttagta aaagttaaca 240

ES 2 278 436 T3

	ttatcagacc	cactaccgag	catactatcc	ccacgatgat	gacagctagc	ccctatcacc	300
	aaggatatcaa	tgatcctgcc	gcagacccaa	aaacatacca	aatggagggt	gcgctagcag	360
	ttaaacagcc	taaacacata	caagttgaca	caaaaccatt	taaagaagaa	gtaaaacatc	420
5	cttcaaaatt	acccatcagc	cctgcaactg	aaagcttcac	acacattgac	agtatatagtc	480
	tcaatgacta	ttttctttct	cgtgggtttg	ctaatatata	cgtttcagg	gtgggtactg	540
	ctggctctac	gggtttcatg	accagtgggg	attaccaaca	aatacaaaag	tttaaagcag	600
	tcattgattg	gttaaattgt	aaggttactg	cattcacaa	tcataaacga	gataaacaag	660
	tcaaggctga	ttgggtcaaac	ggccttgtag	caaccacagg	taaattcttat	ctcggtagca	720
10	tgtcaactgg	tttagcaaca	actggcggtg	aggggctgaa	agtcattatc	gctgaagccg	780
	caatctccac	atgggtatgat	tattatcgag	aaaatgggct	tgtgtgtagt	ccaggcggtc	840
	accccggtga	agattttagac	gttttaacag	aattaacata	ctcacgaaac	ctcttagctg	900
	gtgattacat	caaaaacaac	gatttgctatc	aagcattggt	aaatgaacaa	tcaaaagcaa	960
15	ttgaccgtca	aagtggggat	tacaaccaat	actggcatga	ccgtaattac	ctaactcacg	1020
	tcaataatgt	caaaagtcga	gtagtttaca	ctcatggact	acaggattgg	aatgttaagc	1080
	caagacatgt	ctacaaagtt	ttcaatgcat	tgccctcaaac	catcaaaaaa	cacctttttt	1140
	tacatcaagg	tcaacatgtg	tatatgcata	attggcagtc	gattgatttt	cgtgaaagca	1200
	tgaatgcctt	actaagccaa	gaactacttg	gcattgacaa	tcattttccaa	ttagaagagg	1260
20	tcattttggca	agataaatac	actgagcaaa	cttggcaagt	tttagatgct	ttcggaggaa	1320
	accatcaaga	gcaaatgggt	ttaggtgata	gtaaaaaact	tattgataac	cattatgaca	1380
	aagaagcctt	tgatacttat	tgtaaagact	tcaatgtgtt	caaaaatgat	cttttcaagg	1440
	gaaataataa	aaccaatcaa	atsactatta	atcttctctc	aaagaaaaat	tatctcctga	1500
25	atggacagtg	caaactccat	ctacgtgtta	aaactagtga	caaaaaggcc	attttatcag	1560
	cccaaattctt	agactatggt	cctaaaaaac	gatttcaaaga	tacaccaacc	atcaaattct	1620
	taaacagcct	tgataatggt	aaaaattttg	ccagagaagc	tttacgtgaa	ctcccgttta	1680
	ctaaagatca	ttatcgtgtc	atcagtaaa	gtgtcttgaa	ccttcaaaa	cgtacagact	1740
	tacttacaat	tgaggctatc	gagccagaac	aatggtttga	tatcgagttt	agcctccaac	1800
30	caagtatata	tcaattgagt	aaagggtgata	atctaaggat	tatcctttat	acaactgatt	1860
	ttgaacatac	cattcgagat	aatgctagtt	actctataac	agtagatttg	agtcaatctt	1920
	attttaactat	cccaactaat	caaggaaatt	aacttatgaa	acttcttact	aaagaacggt	1980
	ttgatgatcc	tcaacacttt	tggtaccaga	tcaatttatt	acaagagagt	aacttcggag	2040
	cagtttttga	ccatgataat	aaaaacattc	cacagggtgt	tgcaactatt	gttgatgatt	2100
35	tacaagggtc	cggaggttcg	aatcatcttc	gggtattttg	caatactact	gatacttcca	2160
	tccttatgat	tgctcattta	aatcgaaaat	tctatatcca	gggttaattta	aaggactttg	2220
	actttgcact	caattttaata	gctataaata	attggaagag	tctcctccaa	actcaacttg	2280
	aagctctaaa	cgatacccta	gcaatatctt	aataaataag	gtagaatgga	gtgacaaagc	2340
40	aacgcggagg	agactgatta	atgtcatctt	attggaataa	ctatcctgaa	cttaaaaaaa	2400
	atatgtgatga	aaccaatcaa	ctaattcaag	aaagaataca	ggtcagaaat	aaagatatgt	2460
	aagcggcgct	aagccaactc	acagctgcgg	gaggaaaaaa	gctcagacca	gcattctttt	2520
	accttttttc	tcaacttggt	aataaggaga	atcaagatac	tcagcaacta	aaagaaaatcg	2580
	ctgctctctt	agaaatcctt	caggttgcta	cattaatcca	tgatgatgtc	attgatgact	2640
45	caccactaag	acgtggaaat	atgaccattc	aaagcaagtt	tggaacagac	atcgaggttt	2700
	atactgggga	tttacttttc	acagctcttt	tcgatcttat	tttagaatct	atgactgata	2760
	caccatttat	gaggattaat	gcaaaatcta	tgcgtaaaat	tctcatggga	gaattggacc	2820
	agatgcacct	tcgtttacaat	caacaacaag	gtatccatca	ctattttacgt	gcgatttcag	2880
	gtaagacagc	cgaactcttt	aaattagcta	gcaaaagaag	agcttacttt	gggtgtgcag	2940
50	agaaggagggt	tgttcgtcta	gcaggccata	tcggctttta	cattgggtatg	acattccaaa	3000
	ttttggatga	tatcctggat	tatactgcag	ataaaaaaac	atttaataag	cctgtcttag	3060
	aggattttaac	acaaggcggt	tacagccttc	ctctacttct	tgccattgaa	gaaaatcctg	3120
	atattttcaa	acctatttta	gataaaaaaa	cagatattggc	tactgaagac	atgggaaaaa	3180
	ttgcttatct	cgctcgtttc	catagagggt	ttgacaaagc	tcgccatcta	gctcgtaaat	3240
55	ttactgagaa	agctattagt	gacataaata	agctacccca	gaactctgca	aaaaaacagt	3300
	tgctacaatt	aactaattac	ctttttaaac	gcaaaattta	aataataaaa	aaacattcca	3360
	caatgctaga	aaagcagtta	gggaatgttt	ttttattatc	attttatttat	cgcacctatc	3420
	aatcatcata	gatcaccatc	atcagcggtc	ttcagctgac	ggtaacgttg	actactttga	3480
60	gacaattctt	gaggagaacc	ttccaactct	aattgcccac	tttctataaa	taagatacga	3540
	tcagcatggt	caataccttt	taagtgtagt	gtaatccaaa	ctaagggtctt	accttccaat	3600

ES 2 278 436 T3

```

tctttcataa atacccttag taaggcttgt tcagtaatag gatcaagtcc aacagttggc 3660
tcâtctaaga taacaattgg gacatctttt agtaagattc tagccaaagc aattctatgc 3720
ctttcgccac ctgaaaacct aagtcagct tcarcaacca ttgtatagag accatctgat 3780
5 aaatcagtga ccattctctt caatccaact cgttcaagaa ctttccatac atcttcttca 3840
ctagcatctt ggtttccaat gcgaatgta tttagcaggg ttgtattaaa aaggtagggc 3900
gcttgttga tcactccaat atagttagaa atgcaatcac caactattga aacatcagca 3960
ccgcctaggg taatcttccc ttgacttgct ttcaagtcgc cacgaagtag actagctaag 4020
10 gtactcttgc cagaaccact ccgcctaaa atagcaattt tttctccttc tttaatatcc 4080
aaatctaaat gatgcaaac ccattctct tgtggcttat actggaaact taaattcttg 4140
acggaaaaat catatggctt attaggcaat t 4171

```

```

15 <210> 33
    <211> 649
    <212> PRT
    <213> Streptococos
20
    <400> 33

```

```

25 Tyr Asp Asn Ile Phe Gln Ser Leu His His Leu Leu Ala Cys Arg Gly
    1      5      10      15
Lys Ser Gly Asn Thr Leu Ile Asp Gln Leu Val Ala Asp Gly Leu Leu
    20      25      30
His Ala Asp Asn His Tyr His Phe Phe Asn Gly Lys Ser Leu Ala Thr
    35      40      45
30 Phe Asn Thr Asn Gln Leu Ile Arg Glu Val Val Tyr Val Glu Ile Ser
    50      55      60
Leu Asp Thr Met Ser Ser Gly Glu His Asp Leu Val Lys Val Asn Ile
    65      70      75      80
35 Ile Arg Pro Thr Thr Glu His Thr Ile Pro Thr Met Met Thr Ala Ser
    85      90      95
Pro Tyr His Gln Gly Ile Asn Asp Pro Ala Ala Asp Gln Lys Thr Tyr
    100      105      110
40 Gln Met Glu Gly Ala Leu Ala Val Lys Gln Pro Lys His Ile Gln Val
    115      120      125
Asp Thr Lys Pro Phe Lys Glu Val Lys His Pro Ser Lys Leu Pro
    130      135      140
Ile Ser Pro Ala Thr Glu Ser Phe Thr His Ile Asp Ser Tyr Ser Leu
    145      150      155      160
45 Asn Asp Tyr Phe Leu Ser Arg Gly Phe Ala Asn Ile Tyr Val Ser Gly
    165      170      175
Val Gly Thr Ala Gly Ser Thr Gly Phe Met Thr Ser Gly Asp Tyr Gln
    180      185      190
50 Gln Ile Gln Ser Phe Lys Ala Val Ile Asp Trp Leu Asn Gly Lys Val
    195      200      205
Thr Ala Phe Thr Ser His Lys Arg Asp Lys Gln Val Lys Ala Asp Trp
    210      215      220
55 Ser Asn Gly Leu Val Ala Thr Thr Gly Lys Ser Tyr Leu Gly Thr Met
    225      230      235      240
Ser Thr Gly Leu Ala Thr Thr Gly Val Glu Gly Leu Lys Val Ile Ile
    245      250      255
Ala Glu Ala Ala Ile Ser Thr Trp Tyr Asp Tyr Tyr Arg Glu Asn Gly
    260      265      270
60 Leu Val Cys Ser Pro Gly Gly Tyr Pro Gly Glu Asp Leu Asp Val Leu
    275      280      285
Thr Glu Leu Thr Tyr Ser Arg Asn Leu Leu Ala Gly Asp Tyr Ile Lys
    290      295      300
65 Asn Asn Asp Cys Tyr Gln Ala Leu Leu Asn Glu Gln Ser Lys Ala Ile

```

ES 2 278 436 T3

	305		310		315		320
	Asp	Arg	Gln	Ser	Gly	Asp	Tyr
					325		330
5	Leu	Thr	His	Val	Asn	Asn	Val
					340		345
	Leu	Gln	Asp	Trp	Asn	Val	Lys
					355		360
10	Ala	Leu	Pro	Gln	Thr	Ile	Lys
					370		375
	His	Val	Tyr	Met	His	Asn	Trp
					385		390
	Asn	Ala	Leu	Leu	Ser	Gln	Glu
					405		410
15	Leu	Glu	Glu	Val	Ile	Trp	Gln
					420		425
	Val	Leu	Asp	Ala	Phe	Gly	Gly
					435		440
20	Asp	Ser	Lys	Lys	Leu	Ile	Asp
					450		455
	Thr	Tyr	Cys	Lys	Asp	Phe	Asn
					465		470
25	Asn	Asn	Lys	Thr	Asn	Gln	Ile
					485		490
	Tyr	Leu	Leu	Asn	Gly	Gln	Cys
					500		505
30	Asp	Lys	Lys	Ala	Ile	Leu	Ser
					515		520
	Lys	Arg	Phe	Lys	Asp	Thr	Pro
					530		535
	Asn	Gly	Lys	Asn	Phe	Ala	Arg
					545		550
35	Lys	Asp	His	Tyr	Arg	Val	Ile
					565		570
	Arg	Thr	Asp	Leu	Leu	Thr	Ile
					580		585
40	Asp	Ile	Glu	Phe	Ser	Leu	Gln
					595		600
	Asp	Asn	Leu	Arg	Ile	Ile	Leu
					610		615
	Arg	Asp	Asn	Ala	Ser	Tyr	Ser
					625		630
45	Leu	Thr	Ile	Pro	Thr	Asn	Gln
					645		650

<210> 34

50 <211> 119

<212> PRT

<213> Estreptococos

55 <400> 34

	Met	Lys	Leu	Leu	Thr	Lys	Glu	Arg	Phe	Asp	Asp	Ser	Gln	His	Phe	Trp
	1				5					10					15	
60	Tyr	Gln	Ile	Asn	Leu	Leu	Gln	Glu	Ser	Asn	Phe	Gly	Ala	Val	Phe	Asp
					20				25					30		
	His	Asp	Asn	Lys	Asn	Ile	Pro	Gln	Val	Val	Ala	Thr	Ile	Val	Asp	Asp
					35				40					45		

65

ES 2 278 436 T3

	Leu	Gln	Gly	Ser	Gly	Ser	Ser	Asn	His	Phe	Trp	Tyr	Phe	Gly	Asn	Thr
	50					55						60				
	Thr	Asp	Thr	Ser	Ile	Leu	Met	Ile	Ala	His	Leu	Asn	Arg	Lys	Phe	Tyr
	65				70				75						80	
5	Ile	Gln	Val	Asn	Leu	Lys	Asp	Phe	Asp	Phe	Ala	Leu	Asn	Leu	Ile	Ala
				85					90						95	
	Ile	Asn	Asn	Trp	Lys	Ser	Leu	Leu	Gln	Thr	Gln	Leu	Glu	Ala	Leu	Asn
				100					105					110		
10	Asp	Thr	Leu	Ala	Ile	Phe	Gln									
				115												
	<210> 35															
	<211> 326															
15	<212> PRT															
	<213> Streptococos															
	<400> 35															
20																
	Met	Ser	Ser	Tyr	Trp	Asn	Asn	Tyr	Pro	Glu	Leu	Lys	Lys	Asn	Ile	Asp
	1				5					10					15	
	Glu	Thr	Asn	Gln	Leu	Ile	Gln	Glu	Arg	Ile	Gln	Val	Arg	Asn	Lys	Asp
25			20						25					30		
	Ile	Glu	Ala	Ala	Leu	Ser	Gln	Leu	Thr	Ala	Ala	Gly	Gly	Lys	Gln	Leu
			35				40						45			
	Arg	Pro	Ala	Phe	Phe	Tyr	Leu	Phe	Ser	Gln	Leu	Gly	Asn	Lys	Glu	Asn
	50						55					60				
30	Gln	Asp	Thr	Gln	Gln	Leu	Lys	Lys	Ile	Ala	Ala	Ser	Leu	Glu	Ile	Leu
	65					70				75					80	
	His	Val	Ala	Thr	Leu	Ile	His	Asp	Asp	Val	Ile	Asp	Asp	Ser	Pro	Leu
				85						90					95	
35	Arg	Arg	Gly	Asn	Met	Thr	Ile	Gln	Ser	Lys	Phe	Gly	Lys	Asp	Ile	Ala
				100				105						110		
	Val	Tyr	Thr	Gly	Asp	Leu	Leu	Phe	Thr	Val	Phe	Phe	Asp	Leu	Ile	Leu
			115					120					125			
	Glu	Ser	Met	Thr	Asp	Thr	Pro	Phe	Met	Arg	Ile	Asn	Ala	Lys	Ser	Met
40		130				135						140				
	Arg	Lys	Ile	Leu	Met	Gly	Glu	Leu	Asp	Gln	Met	His	Leu	Arg	Tyr	Asn
	145					150				155					160	
	Gln	Gln	Gln	Gly	Ile	His	His	Tyr	Leu	Arg	Ala	Ile	Ser	Gly	Lys	Thr
				165						170					175	
45	Ala	Glu	Leu	Phe	Lys	Leu	Ala	Ser	Lys	Glu	Gly	Ala	Tyr	Phe	Gly	Gly
				180					185					190		
	Ala	Glu	Lys	Glu	Val	Val	Arg	Leu	Ala	Gly	His	Ile	Gly	Phe	Asn	Ile
			195					200					205			
	Gly	Met	Thr	Phe	Gln	Ile	Leu	Asp	Asp	Ile	Leu	Asp	Tyr	Thr	Ala	Asp
50		210				215						220				
	Lys	Lys	Thr	Phe	Asn	Lys	Pro	Val	Leu	Glu	Asp	Leu	Thr	Gln	Gly	Val
	225					230				235					240	
	Tyr	Ser	Leu	Pro	Leu	Leu	Leu	Ala	Ile	Glu	Glu	Asn	Pro	Asp	Ile	Phe
55				245						250					255	
	Lys	Pro	Ile	Leu	Asp	Lys	Lys	Thr	Asp	Met	Ala	Thr	Glu	Asp	Met	Glu
				260				265						270		
	Lys	Ile	Ala	Tyr	Leu	Val	Val	Ser	His	Arg	Gly	Val	Asp	Lys	Ala	Arg
			275					280					285			
60	His	Leu	Ala	Arg	Lys	Phe	Thr	Glu	Lys	Ala	Ile	Ser	Asp	Ile	Asn	Lys
		290					295					300				
	Leu	Pro	Gln	Asn	Ser	Ala	Lys	Lys	Gln	Leu	Leu	Gln	Leu	Thr	Asn	Tyr
	305					310				315						320
65	Leu	Leu	Lys	Arg	Lys	Ile										
				325												

ES 2 278 436 T3

<210> 36

<211> 247

<212> PRT

5 <213> Estreptococos

<400> 36

```

10      Leu Pro Asn Lys Pro Tyr Asp Phe Ser Val Lys Asn Leu Ser Phe Gln
        1          5          10          15
      Tyr Lys Pro Gln Glu Lys Trp Val Leu His His Leu Asp Leu Asp Ile
        20          25          30
15      Lys Glu Gly Glu Lys Ile Ala Ile Leu Gly Arg Ser Gly Ser Gly Lys
        35          40          45
      Ser Thr Leu Ala Ser Leu Leu Arg Gly Asp Leu Lys Ala Ser Gln Gly
        50          55          60
      Lys Ile Thr Leu Gly Gly Ala Asp Val Ser Ile Val Gly Asp Cys Ile
20      65          70          75          80
      Ser Asn Tyr Ile Gly Val Ile Gln Gln Ala Pro Tyr Leu Phe Asn Thr
        85          90          95
      Thr Leu Leu Asn Asn Ile Arg Ile Gly Asn Gln Asp Ala Ser Glu Glu
        100          105          110
25      Asp Val Trp Lys Val Leu Glu Arg Val Gly Leu Lys Glu Met Val Thr
        115          120          125
      Asp Leu Ser Asp Gly Leu Tyr Thr Met Val Asp Glu Ala Gly Leu Arg
        130          135          140
30      Phe Ser Gly Gly Glu Arg His Arg Ile Ala Leu Ala Arg Ile Leu Leu
        145          150          155          160
      Lys Asp Val Pro Ile Val Ile Leu Asp Glu Pro Thr Val Gly Leu Asp
        165          170          175
      Pro Ile Thr Glu Gln Ala Leu Leu Arg Val Phe Met Lys Glu Leu Glu
35      180          185          190
      Gly Lys Thr Leu Val Trp Ile Thr His His Leu Lys Gly Ile Glu His
        195          200          205
      Ala Asp Arg Ile Leu Phe Ile Glu Asn Gly Gln Leu Glu Leu Glu Gly
        210          215          220
40      Ser Pro Gln Glu Leu Ser Gln Ser Ser Gln Arg Tyr Arg Gln Leu Lys
        225          230          235          240
      Ala Ala Asp Asp Gly Asp Leu
        245

```

<210> 37

<211> 3480

<212> ADN

50 <213> Estreptococos

<400> 37

```

55      aattctatatt ggagggttttt cttgaataaaa tgggttagtta aggcaagttc cttagttggt      60
        ttaggtggta tggttttatc tgcgggttcc cgagtttttag cggataactta tgcgcgtcca      120
        attgataatg gtagaattac aacagggtttc aatggttatc ctggacattg tgggggtggat      180
        tatgctgttc cgactggaac gattattagg gcagtggcag atggtagtgt gaaatttgca      240
60      ggagctggag ccaacttttc ttggatgaca gacttagcag gaaatttgtt catgattcaa      300
        catgcggatg gaatgcatag tggttacgct catatgtcac gtgtggtggc taggactggg      360
        gaaaaagtca aacaaggaga tatcatcggt tacgtaggag caactggtat ggcgacggga      420

```

65

ES 2 278 436 T3

	cctcaccttc	atcttgaatt	tttaccagct	aaccttaatt	ttcaaaatgg	tttccatgga	480
	cgtatcaatc	caacgtcact	aattgctaac	gttgcgacct	ttagtggaag	aacgcaagca	540
	tcagctccaa	gcattaaagg	attacaatca	gctcctgtac	agaatcaatc	tagtaaatca	600
5	aaagtgtatc	gagtagatga	attacaaaag	gttaaatggg	tttggttagt	caaaaataac	660
	accctaaccg	cgactgggtt	tgatttgaac	gataatggta	taccagcatc	agaaattgat	720
	gaggttgatg	ctaattggtaa	tttgacagct	gaccagggtc	ttcaaaaagg	tggttacttt	780
	atctttaatc	ctaaaactct	taagactgta	gaaaaaccca	tccaagggaac	agctgggtta	840
	acttgggcta	agacacgctt	tgctaattgt	agttcagttt	ggcttcgcgt	tgacaacagt	900
10	caagaactgc	tttacaataa	gtttgaggta	ttgattcatt	gttttaaatg	acagttttgt	960
	tactaactaa	gtacaatttc	tttaaacctg	ctgaaaataa	ttttatagtc	cagtaaagtg	1020
	tgatattata	gtctcggact	aaataaaaag	aaataggaat	tgaagcaatg	aaaatgaata	1080
	aaaagggtact	attgacatcg	acaatggcag	cttcgctatt	atcagtcgca	agtgttcaag	1140
15	cacaagaaac	agatacgacg	tggaacagca	gtactgtttc	agaggtaaa	gctgatttgg	1200
	taaagcaaga	caataaatca	tcatatactg	tgaaatatgg	tgatacacta	agcgtttatt	1260
	cagaagcaat	gtcaattgat	atgaatgtct	tagcaaaaat	taataacatt	gcagatatca	1320
	atcttattta	tcttgagaca	acactgacag	taacttacga	tcagaagagt	catactgcca	1380
	cttcaatgaa	aatagaataa	ccagcaacaa	atgctgctgg	tcaaaccaaca	gctactgtgg	1440
20	atttgaaaac	caatcaaggt	tctgttcgag	accaaaaagt	ttctctcaat	acaatttcgg	1500
	aagggtatgac	accagaagca	gcaacaacga	ttgtttcgcc	aatgaagaca	tattcttctg	1560
	cgccagcttt	gaaatcaaaa	gaagtattag	cacaagagca	agctgttagt	caagcagcag	1620
	ctaatagaac	ggatcaacaa	gtcctgttga	agtcgattac	ttcagaagtt	ccagcagcta	1680
	aagaggaagt	taaaccaact	cagacgtcag	tcagtcagtc	aacaacagta	tcaccagctt	1740
25	ctgttgccgc	tgaaacacca	gtcccagtag	ctaaagtagc	accggttaaga	actgtagcag	1800
	cccctagagt	ggcaagtgtt	aaagttagtc	ctcctaaagt	agaaactggg	gcacaccag	1860
	agcatgtatc	agctccagca	gttctgttga	ctacgacttc	aacagctaca	gacagtaagt	1920
	tacaagcgac	tgaaagttaa	agcgttcctg	tagcacaata	agctccaaca	gcaacaccgg	1980
30	tagcacaact	agcttcaaca	acaaatgcag	tagctgcaca	tcctgaaaat	gcagggtccc	2040
	aacctcatgt	tgacagctta	aaagaaaaag	tagcgtcaac	ttatggagtt	aatgaattca	2100
	gtacataacc	tcaggttgat	ccaggtgatc	atggtaaaag	tttagcagtc	gactttattg	2160
	taggttaaaa	ccaagcactt	ggtaaatgaag	ttgcacagta	ctctacacaa	aatatggcag	2220
	caaataacat	ttcatatggt	atctggcaac	aaaagtttta	ctcaaatata	aatagtattt	2280
35	atggacctgc	taatacttgg	aatgcaatgc	cagatcgtgg	tggtggttact	gccaacattt	2340
	atgaccatgt	tcacgtatca	tttaacaaat	aatataaaaa	aggaagctat	ttgggtctct	2400
	ttttatatgc	cttgaataga	ctttcaaggt	tcttatctaa	tttttatata	attgaggaga	2460
	ttaagctata	agtcctgaac	tactttcacg	ttaaccgtga	ctaaatcaaa	acgttaaaac	2520
40	taaaatctaa	gtctgtaaa	attattgaaa	acgttttaaa	aacagatata	ataaggtttg	2580
	tagatatcta	aaattaaaaa	agataaggaa	gtgagaatat	gccacatcta	agtaaaagag	2640
	cttttaaaaa	gcaaataaaa	aatggcatta	ttgtgtcatg	tcaagctttg	cctggggagc	2700
	ctctttatag	tgaaagtggg	ggtgttatgc	ctcttttagc	tttggcagct	caagaagcag	2760
	gagcgggttg	tataagagcc	aatagtgtcc	gcgacattaa	ggaaattcaa	gaagtactta	2820
45	atttacctat	catcggcatt	attaaacgtg	aatatcctcc	acaagaacca	tttatcactg	2880
	ctacgatgac	agagggtgat	caattagcta	gttttagatat	tgagtaata	gccttagatt	2940
	gtacacttag	agagcgtcat	gatgggttga	gtgtagctga	gtttattcaa	aagataaaag	3000
	ggaaatatcc	tgaacagttg	ctaattggctg	atataagtac	ttttgaagaa	ggtaaaaatg	3060
	cttttgaagc	aggagttagt	tttgtgggtg	caactctatc	tggtatcaca	gattacagcc	3120
50	gccaaagaaga	aggaccggat	atagaactcc	ttaataagct	ttgtcaagcc	ggatatagatg	3180
	tgattgcgga	aggtaaaaat	catactccta	agcaagctaa	tgaaattaat	catatagggtg	3240
	ttgcaggaa	tgtagtgtgt	ggtgctatca	ctagaccaaa	agaaatagcg	gagcgtttca	3300
	tctcaggact	tagttaaaa	tgttactcaa	aatcaaaaat	caaaaataaa	aagggaataa	3360
	gttatgagta	tcaaaaaaag	tgtgattggt	ttttgcctcg	gagctgcagc	attatcaatg	3420
55	tttgcttggtg	tagacagtag	tcaatctggt	atggctgcgc	agaaggataa	agtcgaaatt	3480

<210> 38

<211> 306

60 <212> PRT

<213> Streptococcus

65

ES 2 278 436 T3

<400> 38

```

5      Asn Ser Ile Trp Arg Phe Phe Leu Asn Lys Trp Leu Val Lys Ala Ser
      1          5          10          15
      Ser Leu Val Val Leu Gly Gly Met Val Leu Ser Ala Gly Ser Arg Val
      20          25          30
      Leu Ala Asp Thr Tyr Val Arg Pro Ile Asp Asn Gly Arg Ile Thr Thr
10      35          40          45
      Gly Phe Asn Gly Tyr Pro Gly His Cys Gly Val Asp Tyr Ala Val Pro
      50          55          60
      Thr Gly Thr Ile Ile Arg Ala Val Ala Asp Gly Thr Val Lys Phe Ala
15      65          70          75          80
      Gly Ala Gly Ala Asn Phe Ser Trp Met Thr Asp Leu Ala Gly Asn Cys
      85          90          95
      Val Met Ile Gln His Ala Asp Gly Met His Ser Gly Tyr Ala His Met
      100          105          110
20      Ser Arg Val Val Ala Arg Thr Gly Glu Lys Val Lys Gln Gly Asp Ile
      115          120          125
      Ile Gly Tyr Val Gly Ala Thr Gly Met Ala Thr Gly Pro His Leu His
      130          135          140
      Phe Glu Phe Leu Pro Ala Asn Pro Asn Phe Gln Asn Gly Phe His Gly
25      145          150          155          160
      Arg Ile Asn Pro Thr Ser Leu Ile Ala Asn Val Ala Thr Phe Ser Gly
      165          170          175
      Lys Thr Gln Ala Ser Ala Pro Ser Ile Lys Pro Leu Gln Ser Ala Pro
      180          185          190
30      Val Gln Asn Gln Ser Ser Lys Leu Lys Val Tyr Arg Val Asp Glu Leu
      195          200          205
      Gln Lys Val Asn Gly Val Trp Leu Val Lys Asn Asn Thr Leu Thr Pro
      210          215          220
35      Thr Gly Phe Asp Trp Asn Asp Asn Gly Ile Pro Ala Ser Glu Ile Asp
      225          230          235          240
      Glu Val Asp Ala Asn Gly Asn Leu Thr Ala Asp Gln Val Leu Gln Lys
      245          250          255
      Gly Gly Tyr Phe Ile Phe Asn Pro Lys Thr Leu Lys Thr Val Glu Lys
40      260          265          270
      Pro Ile Gln Gly Thr Ala Gly Leu Thr Trp Ala Lys Thr Arg Phe Ala
      275          280          285
      Asn Gly Ser Ser Val Trp Leu Arg Val Asp Asn Ser Gln Glu Leu Leu
45      290          295          300
      Tyr Lys
      305

```

<210> 39

50 <211> 434

<212> PRT

<213> Streptococcus

55 <400> 39

```

      Met Lys Met Asn Lys Lys Val Leu Leu Thr Ser Thr Met Ala Ala Ser
      1          5          10          15
60      Leu Leu Ser Val Ala Ser Val Gln Ala Glu Thr Asp Thr Thr Trp
      20          25          30
      Thr Ala Arg Thr Val Ser Glu Val Lys Ala Asp Leu Val Lys Gln Asp
      35          40          45
65      Asn Lys Ser Ser Tyr Thr Val Lys Tyr Gly Asp Thr Leu Ser Val Ile

```

ES 2 278 436 T3

	50		55		60											
	Ser	Glu	Ala	Met	Ser	Ile	Asp	Met	Asn	Val	Leu	Ala	Lys	Ile	Asn	Asn
	65				70					75					80	
5	Ile	Ala	Asp	Ile	Asn	Leu	Ile	Tyr	Pro	Glu	Thr	Thr	Leu	Thr	Val	Thr
				85						90				95		
	Tyr	Asp	Gln	Lys	Ser	His	Thr	Ala	Thr	Ser	Met	Lys	Ile	Glu	Thr	Pro
				100					105					110		
10	Ala	Thr	Asn	Ala	Ala	Gly	Gln	Thr	Thr	Ala	Thr	Val	Asp	Leu	Lys	Thr
			115				120						125			
	Asn	Gln	Val	Ser	Val	Ala	Asp	Gln	Lys	Val	Ser	Leu	Asn	Thr	Ile	Ser
		130				135					140					
	Glu	Gly	Met	Thr	Pro	Glu	Ala	Ala	Thr	Thr	Ile	Val	Ser	Pro	Met	Lys
	145				150					155					160	
15	Thr	Tyr	Ser	Ser	Ala	Pro	Ala	Leu	Lys	Ser	Lys	Glu	Val	Leu	Ala	Gln
				165						170				175		
	Glu	Gln	Ala	Val	Ser	Gln	Ala	Ala	Ala	Asn	Glu	Gln	Val	Ser	Thr	Ala
			180					185					190			
20	Pro	Val	Lys	Ser	Ile	Thr	Ser	Glu	Val	Pro	Ala	Ala	Lys	Glu	Glu	Val
		195					200					205				
	Lys	Pro	Thr	Gln	Thr	Ser	Val	Ser	Gln	Ser	Thr	Thr	Val	Ser	Pro	Ala
		210				215					220					
25	Ser	Val	Ala	Ala	Glu	Thr	Pro	Ala	Pro	Val	Ala	Lys	Val	Ala	Pro	Val
	225				230					235					240	
	Arg	Thr	Val	Ala	Ala	Pro	Arg	Val	Ala	Ser	Val	Lys	Val	Val	Thr	Pro
				245					250					255		
	Lys	Val	Glu	Thr	Gly	Ala	Ser	Pro	Glu	His	Val	Ser	Ala	Pro	Ala	Val
			260					265					270			
30	Pro	Val	Thr	Thr	Thr	Ser	Thr	Ala	Thr	Asp	Ser	Lys	Leu	Gln	Ala	Thr
		275					280					285				
	Glu	Val	Lys	Ser	Val	Pro	Val	Ala	Gln	Lys	Ala	Pro	Thr	Ala	Thr	Pro
		290				295					300					
35	Val	Ala	Gln	Pro	Ala	Ser	Thr	Thr	Asn	Ala	Val	Ala	Ala	His	Pro	Glu
	305				310					315					320	
	Asn	Ala	Gly	Leu	Gln	Pro	His	Val	Ala	Ala	Tyr	Lys	Glu	Lys	Val	Ala
				325					330					335		
40	Ser	Thr	Tyr	Gly	Val	Asn	Glu	Phe	Ser	Thr	Tyr	Arg	Ala	Gly	Asp	Pro
			340					345					350			
	Gly	Asp	His	Gly	Lys	Gly	Leu	Ala	Val	Asp	Phe	Ile	Val	Gly	Lys	Asn
		355				360						365				
	Gln	Ala	Leu	Gly	Asn	Glu	Val	Ala	Gln	Tyr	Ser	Thr	Gln	Asn	Met	Ala
		370			375				380							
45	Ala	Asn	Asn	Ile	Ser	Tyr	Val	Ile	Trp	Gln	Gln	Lys	Phe	Tyr	Ser	Asn
	385				390				395						400	
	Thr	Asn	Ser	Ile	Tyr	Gly	Pro	Ala	Asn	Thr	Trp	Asn	Ala	Met	Pro	Asp
				405				410					415			
50	Arg	Gly	Gly	Val	Thr	Ala	Asn	His	Tyr	Asp	His	Val	His	Val	Ser	Phe
			420					425				430				
	Asn	Lys														

55 <210> 40
 <211> 232
 <212> PRT
 <213> Streptococos

60

65

ES 2 278 436 T3

<400> 40

```

Met Pro His Leu Ser Lys Glu Ala Phe Lys Lys Gln Ile Lys Asn Gly
 1           5           10           15
5 Ile Ile Val Ser Cys Gln Ala Leu Pro Gly Glu Pro Leu Tyr Thr Glu
 20           25           30
Ser Gly Gly Val Met Pro Leu Leu Ala Leu Ala Ala Gln Glu Ala Gly
 35           40           45
10 Ala Val Gly Ile Arg Ala Asn Ser Val Arg Asp Ile Lys Glu Ile Gln
 50           55           60
Glu Val Thr Asn Leu Pro Ile Ile Gly Ile Ile Lys Arg Glu Tyr Pro
65           70           75           80
Pro Gln Glu Pro Phe Ile Thr Ala Thr Met Thr Glu Val Asp Gln Leu
 85           90           95
15 Ala Ser Leu Asp Ile Ala Val Ile Ala Leu Asp Cys Thr Leu Arg Glu
 100          105          110
Arg His Asp Gly Leu Ser Val Ala Glu Phe Ile Gln Lys Ile Lys Gly
 115          120          125
20 Lys Tyr Pro Glu Gln Leu Leu Met Ala Asp Ile Ser Thr Phe Glu Glu
 130          135          140
Gly Lys Asn Ala Phe Glu Ala Gly Val Asp Phe Val Gly Thr Thr Leu
145          150          155          160
Ser Gly Tyr Thr Asp Tyr Xaa Arg Gln Glu Gly Pro Asp Ile Glu
 165          170          175
25 Leu Leu Asn Lys Leu Cys Gln Ala Gly Ile Asp Val Ile Ala Glu Gly
 180          185          190
Lys Ile His Thr Pro Lys Gln Ala Asn Glu Ile Asn His Ile Gly Val
 195          200          205
30 Ala Gly Ile Val Val Gly Gly Ala Ile Thr Arg Pro Lys Glu Ile Ala
 210          215          220
Glu Arg Phe Ile Ser Gly Leu Ser
225          230

```

<210> 41

<211> 39

<212> PRT

<213> Streptococcus

<400> 41

```

Met Ser Ile Lys Lys Ser Val Ile Gly Phe Cys Leu Gly Ala Ala Ala
 1           5           10           15
45 Leu Ser Met Phe Ala Cys Val Asp Ser Ser Gln Ser Val Met Ala Ala
 20           25           30
Glu Lys Asp Lys Val Glu Ile
 35

```

<210> 42

<211> 1305

<212> ADN

<213> Streptococcus

<400> 42

```

atgaaaatga ataaaaaggt actattgaca tcgacaatgg cagcttcgct attatcagtc 60
gcaagtgttc aagcacaaga aacagatacg acgtggacag cacgtactgt ttcagaggta 120
aaggctgatt tggtaaagca agacaataaa tcatcatata ctgtgaaata tggatgataca 180
ctaagcgtaa tttcagaagc aatgtcaatt gatatgaatg tcttagcaaa aattaataac 240
attgcagata tcaatcttat ttatcctgag acaaacactga cagtaactta cgatcagaag 300
agtcatactg ccacttcaat gaaaatagaa acaccagcaa caaatgctgc tgggtcaaaca 360

```

ES 2 278 436 T3

```

5  acagctactg tggatttgaa aaccaatcaa gtttctgttg cagaccaaaa agtttctctc 420
   aatacaattt cggaaggatg gacaccagaa gcagcaacaa cgattgtttc gccaatgaag 480
   acataattctt ctgcgccagc tttgaaatca aaagaagtat tagcacaaga gcaagctggt 540
   agtcaagcag cagctaataga acaggtatca acagctcctg tgaagtcgat tacttcagaa 600
   gttccagcag ctaaagagga agttaaacca actcagacgt cagtcagtcg gtcaacaaca 660
   gtatcaccag cttctgtttg cgctgaaaca ccagctccag tagctaaagt agcaccggta 720
   agaactgtag cagcccctag agtggcaagt gttaaagtag tcactcctaa agtagaaact 780
10 ggtgcatcac cagagcatgt atcagctcca gcagttcctg tgactacgac ttcaacagct 840
   acagacagta agttacaagc gactgaagtt aagagcggtc cggtagcaca aaaagctcca 900
   acagcaacac cggtagcaca accagcttca acaacaaatg cagtagctgc acatcctgaa 960
   aatgcagggc tccaacctca tgttcagctt tataaagaaa aagtagcgtc aacttatgga 1020
   gttaattgaat tcagtagata ccgtgcaggt gatccaggtg atcatggtaa aggttttagca 1080
15 gtcgacttta ttgtaggtaa aaaccaagca cttggtaatg aagttgcaca gtactctaca 1140
   caaaatatgg cagcaataaa catttcatat gttatctggc aacaaaagtt ttactcaaat 1200
   acaaatagta tttatggacc tgctaatact tggaaatgcaa tgccagatcg tgggtggcgtt 1260
   actgccaaac attatgacca tgttcacgta tcatttaaca aataa 1305

```

```

20 <210> 43
   <211> 1230
   <212> ADN
25 <213> Streptococos

   <400> 43

```

```

30  caagaaacag atacgacgtg gacagcacgt actgtttcag aggtaaaggc tgatttggtg 60
   aagcaagaca ataaatcatc atatactgtg aaatatgggtg atacactaag cgttatttca 120
   gaagcaatgt caattgatat gaatgtctta gcaaaaatta ataacattgc agatatcaat 180
   cttatattatc ctgagacaac actgacagta acttacgatc agaagagtca tactgccact 240
   tcaatgaaaa tagaaacacc agcaacaaat gctgctgggtc aaacaacagc tactgtggat 300
35  ttgaaaacca atcaagtttc tgttcgagac caaaaagttt ctctcaatac aatttcggaa 360
   ggtatgacac cagaagcagc aacaacgatt gtttcgccaa tgaagacata ttcttctgcg 420
   ccagctttga aatcaaaaaga agtatttagca caagagcaag ctggttagtca agcagcagct 480
   aatgaacagg tatcaacagc tcctgtgaag tcgattactt cagaagttcc agcagctaaa 540
   gaggaagtta aaccaactca gacgtcagtc agtcagtcac caacagtatc accagcttct 600
40  gttgccgctg aaacaccagc tccagtagct aaagtagcac cggtaaagaac tgtagcagcc 660
   cctagagtgg caagtgttaa agtagtcact cctaaagtag aaactgggtg atcaccagag 720
   catgtatcag ctccagcagt tcctgtgact acgacttcaa cagctacaga cagtaagtta 780
   caagcgactg aagttaagag cgttccggtg gcacaaaaag ctccaacagc aacaccggta 840
   gcacaaccag cttcaacaac aaatgcagta gctgcacatc ctgaaaatgc agggctccaa 900
45  cctcatgttg cagcttataa agaaaaagta gcgtaactct atggagttaa tgaattcagt 960
   acataccgtg caggtgatcc aggtgatcat ggtaaagggt tagcagtcga ctttattgta 1020
   ggtaaaaacc aagcacttgg taatgaagtt gcacagtact ctacacaaaa tatggcagca 1080
   aataacattt catatgttat ctggcaacaa aagttttact caaatacaaa tagtatttat 1140
50  ggacctgcta atacttgga tgcaatgcca gatcgtgggt gcgttactgc caaccattat 1200
   gacctgttcc acgtatcatt taacaaataa

```

```

   <210> 44
55 <211> 409
   <212> PRT
   <213> Streptococos

60 <400> 44

```

```

      Gln Glu Thr Asp Thr Thr Trp Thr Ala Arg Thr Val Ser Glu Val Lys
      1          5          10          15
      Ala Asp Leu Val Lys Gln Asp Asn Lys Ser Ser Tyr Thr Val Lys Tyr
      20          25          30
      Gly Asp Thr Leu Ser Val Ile Ser Glu Ala Met Ser Ile Asp Met Asn

```

ES 2 278 436 T3

		35		40		45										
	Val	Leu	Ala	Lys	Ile	Asn	Asn	Ile	Ala	Asp	Ile	Asn	Leu	Ile	Tyr	Pro
		50					55					60				
5	Glu	Thr	Thr	Leu	Thr	Val	Thr	Tyr	Asp	Gln	Lys	Ser	His	Thr	Ala	Thr
	65					70					75				80	
	Ser	Met	Lys	Ile	Glu	Thr	Pro	Ala	Thr	Asn	Ala	Ala	Gly	Gln	Thr	Thr
					85					90				95		
10	Ala	Thr	Val	Asp	Leu	Lys	Thr	Asn	Gln	Val	Ser	Val	Ala	Asp	Gln	Lys
					100				105					110		
	Val	Ser	Leu	Asn	Thr	Ile	Ser	Glu	Gly	Met	Thr	Pro	Glu	Ala	Ala	Thr
					115				120				125			
	Thr	Ile	Val	Ser	Pro	Met	Lys	Thr	Tyr	Ser	Ser	Ala	Pro	Ala	Leu	Lys
15		130					135					140				
	Ser	Lys	Glu	Val	Leu	Ala	Gln	Glu	Gln	Ala	Val	Ser	Gln	Ala	Ala	Ala
	145					150				155					160	
	Asn	Glu	Gln	Val	Ser	Thr	Ala	Pro	Val	Lys	Ser	Ile	Thr	Ser	Glu	Val
					165					170				175		
20	Pro	Ala	Ala	Lys	Glu	Glu	Val	Lys	Pro	Thr	Gln	Thr	Ser	Val	Ser	Gln
					180				185				190			
	Ser	Thr	Thr	Val	Ser	Pro	Ala	Ser	Val	Ala	Ala	Glu	Thr	Pro	Ala	Pro
					195				200				205			
25	Val	Ala	Lys	Val	Ala	Pro	Val	Arg	Thr	Val	Ala	Ala	Pro	Arg	Val	Ala
		210					215					220				
	Ser	Val	Lys	Val	Val	Thr	Pro	Lys	Val	Glu	Thr	Gly	Ala	Ser	Pro	Glu
	225					230				235					240	
	His	Val	Ser	Ala	Pro	Ala	Val	Pro	Val	Thr	Thr	Thr	Ser	Thr	Ala	Thr
30					245					250					255	
	Asp	Ser	Lys	Leu	Gln	Ala	Thr	Glu	Val	Lys	Ser	Val	Pro	Val	Ala	Gln
					260				265				270			
	Lys	Ala	Pro	Thr	Ala	Thr	Pro	Val	Ala	Gln	Pro	Ala	Ser	Thr	Thr	Asn
					275				280				285			
35	Ala	Val	Ala	Ala	His	Pro	Glu	Asn	Ala	Gly	Leu	Gln	Pro	His	Val	Ala
		290					295					300				
	Ala	Tyr	Lys	Glu	Lys	Val	Ala	Ser	Thr	Tyr	Gly	Val	Asn	Glu	Phe	Ser
	305					310				315					320	
40	Thr	Tyr	Arg	Ala	Gly	Asp	Pro	Gly	Asp	His	Gly	Lys	Gly	Leu	Ala	Val
					325					330					335	
	Asp	Phe	Ile	Val	Gly	Lys	Asn	Gln	Ala	Leu	Gly	Asn	Glu	Val	Ala	Gln
					340				345				350			
	Tyr	Ser	Thr	Gln	Asn	Met	Ala	Ala	Asn	Asn	Ile	Ser	Tyr	Val	Ile	Trp
45					355				360				365			
	Gln	Gln	Lys	Phe	Tyr	Ser	Asn	Thr	Asn	Ser	Ile	Tyr	Gly	Pro	Ala	Asn
		370					375					380				
	Thr	Trp	Asn	Ala	Met	Pro	Asp	Arg	Gly	Gly	Val	Thr	Ala	Asn	His	Tyr
	385					390				395					400	
50	Asp	His	Val	His	Val	Ser	Phe	Asn	Lys							
					405											