

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 352 704**

21 Número de solicitud: 200802310

51 Int. Cl.:

C12N 15/56

(2006.01)

C12N 9/38

(2006.01)

12

PATENTE DE INVENCION

B1

22 Fecha de presentación: **08.08.2008**

43 Fecha de publicación de la solicitud: **22.02.2011**

Fecha de la concesión: **04.06.2012**

45 Fecha de anuncio de la concesión: **14.06.2012**

45 Fecha de publicación del folleto de la patente:
14.06.2012

73 Titular/es:
**CONSEJO SUPERIOR DE INVESTIGACIONES
CIENTÍFICAS (CSIC)(Titular al 80%)
SERRANO, 117
28006 MADRID, ES
UNIVERSIDAD COMPLUTENSE DE MADRID
(Titular al 20%)**

72 Inventor/es:
**GARCIA LOPEZ, JOSE LUIS;
GARCIA GONZALEZ, PEDRO;
LLULL PEÑALBA, DANIEL;
CAMPUZANO RUIZ, SUSANA;
SERRA FERNANDEZ, BEATRIZ y
PINGARRON CAZARRON, JOSE MANUEL**

74 Agente/Representante:
Pons Ariño, Ángel

54 Título: **SECUENCIA DE NUCLEOTIDO Y PROTEINA β -GALACTOSIDASA DE STREPTOCOCCUS
MITIS, PROCEDIMIENTO DE OBTENCION Y SUS APLICACIONES.**

57 Resumen:

Secuencia de nucleótido y proteína β -galactosidasa de Streptococcus mitis, procedimiento de obtención y sus aplicaciones.

La presente invención describe una proteína inédita β -galactosidasa de Streptococcus mitis que presenta motivos de unión a colina en su estructura y que no se inhibe en presencia de elevadas concentraciones de glucosa. También se describen dos procedimientos de purificación de β -galactosidasas modificadas en un hospedador heterólogo (Escherichia coli).

ES 2 352 704 B1

DESCRIPCIÓN

Secuencia de nucleótido y proteína β -galactosidasa de *Streptococcus mitis*, procedimiento de obtención y sus aplicaciones.

Sector de la técnica

Sector biotecnológico con aplicaciones en la industria agroalimentaria, al análisis de carbohidratos y a la biosíntesis de oligosacáridos, y más concretamente tecnología enzimática aplicada a los productos lácteos.

Estado de la técnica

La β -D-galactosidasa (EC 3.2.1.23) (en lo sucesivo β -galactosidasa), también llamada lactasa, tiene como función primaria la hidrólisis de la lactosa en D-galactosa y D-glucosa. Por ello, la aplicación fundamental de las β -D-galactosidasas se circunscribe por lo general al ámbito de los productos y subproductos lácteos [Pivarnik, L.F.; Senecal, A.G. y Rand, A.G. 1995. "Hydrolytic and transgalactosylic activities of commercial β -galactosidase (lactase) in food processing". Adv. Food Nutrit. Res. 38, 1-102]. Además, esta enzima posee actividad transgalactosilasa y se utiliza para la síntesis de galactosil-oligosacáridos, compuestos con un interés creciente en el área agroalimentaria y farmacéutica.

Como alternativa no enzimática a la hidrólisis con β -galactosidasas se puede utilizar la hidrólisis química, pero produce efectos no deseables (p. ej.: pérdida de nutrientes, defectos organolépticos, etc.) [Gekas, V. y López-Leiva, M. 1985. "Hydrolysis of lactose: A literature review". Process. Biochem. 20, 1-12].

Las β -galactosidasas se emplean en la industria agroalimentaria para reducir el contenido en lactosa de productos lácteos y derivados, lo que tiene un interés muy variado. En primer lugar, el problema más complejo sería hacer accesibles estos alimentos a los afectados de hipolactasia, que constituyen el 70% de la población mundial, para lo que necesitaremos hidrolizar el 100% de la lactosa de la leche [Savaiano, D.A. y Kotz, C. 1988. "Recent advances in the management of lactose intolerance". Contemp. Nutr. 13,10; Wolin, M.J. 1981. "Fermentation in the rumen and human large intestine". Science 213, 1463-1468; Phillips, M.C. y Briggs, G.M. 1975. "Milk and its role in the American diet". J. Dairy Sci. 58, 1751-1763; Birge, S.J.; Keutmann, H.T.; Cuatrecasas, P. y Whedon, G.D. 1963. "Osteoporosis, intestinal lactase deficiency and low dietary calcium intake". N. Engl. J. Med. 276, 445-448].

En las demás utilizaciones, la exigencia de hidrólisis totales puede estar disminuida. Por ejemplo, el uso de β -galactosidasas para revalorizar los sueros de lechería mediante su transformación en jarabes de glucosa-galactosa [Mahoney, R.R. 1985. "Modification of lactose and lactose-containing dairy products with β -galactosidase". En "Developments in Dairy Chemistry: Lactose and minor constituents". Fox, F. (Ed), pp.69-108. Elsevier, Nueva York; Short, J.L. 1975. "Prospects for the utilization of deproteinized whey in New Zealand-A review. N.Z.J. Dairy Sci. Technol. 13,181-194], evitando lo que -por otra parte- sería un grave problema medioambiental, debido a la alta demanda bioquímica de oxígeno de la lactosa de los mencionados sueros, y dado que sólo se procesa la mitad del que se produce vertiéndose el resto a los cauces fluviales [Pivarnik, L.F.; Senecal, A.G. y Rand, A.G. 1995. "Hydrolytic and transgalactosylic activities of commercial β -galactosidase (lactase) in food processing". Adv. Food Nutrit. Res. 38, 1-102] (para hacernos una idea del problema, 1 Kg de queso requiere, como media, 10 L de leche).

Por otro lado, leches con la lactosa hidrolizada son mejores sustratos para la fermentación, de forma que este proceso tiene interés para acelerar la producción de derivados fermentados de la leche (quesos, yogures...). También es interesante considerar que la lactosa tiene una solubilidad reducida cuando se la compara con la de la galactosa y glucosa, con lo que su hidrólisis evita problemas de aparición de cristales de lactosa en leches condensadas, helados, etc, además de añadir un sabor dulce al alimento.

Otra aplicación no hidrolítica sino transglicosílica, sería en la industria alimentaria y farmacéutica: síntesis de galactosil-oligosacáridos por transglicosilación [Zárate, S. y López-Leiva, M.H. 1990. "Oligosaccharide formation during enzymatic lactose hydrolysis: A literature review". J. Food Prot. 53, 262-268], pudiendo emplearse con fines de interés alimentario [Pivarnik, L.F.; Senecal, A.G. y Rand, A.G. 1995. "Hydrolytic and transgalactosylic activities of commercial β -galactosidase (lactase) in food processing". Adv. Food Nutrit. Res. 38, 1-102] y farmacéutico [Nilsson K.G.I. 1988 "Enzymatic synthesis of oligosaccharides" Trends Biotechnol, 6, 256-264].

Una propiedad importante que ha recibido poca atención en la literatura es el nivel de pureza de las preparaciones comerciales de β -galactosidasas, especialmente en lo relacionado a la presencia de otras enzimas, como las proteasas. Estos posibles contaminantes pueden tener un impacto severo en la estabilidad de la enzima, provocando cambios indeseables en los productos lácteos durante su almacenamiento. La elevada afinidad del dominio de unión a colina por este aminoalcohol o sus análogos estructurales tales como el dietilaminoetanol (DEAE) (J.M. Sanz, R. López, J.L. García "Structural requirements of choline derivatives for "conversión" of pneumococcal amidase A new single-step procedure for purification of this autolysin" FEBS Letters 232 (1988) 308-312) permite aumentar la pureza de la β -galactosidasa recombinante de *Streptococcus mitis* mediante cromatografía de afinidad de DEAE-celulosa (A. Vian, A.V. Carrascosa, J.L. García, E. Cortés. "Structure of the β -galactosidase gene from *Thermus* sp T2: Expression in *Escherichia coli* and purification in a single step of an active fusión protein". Applied and Environmental Microbiology 64 (1998) 2187-2191).

Teniendo en cuenta las ventajas que el dominio de unión a colina ofrece para la purificación de enzimas se ha reportado previamente la fusión de diversas β -galactosidasas a dominios de unión a colina:

- β -galactosidasa de *Escherichia coli* fusionada al dominio de unión a colina de N-acetilmuramoyl-L-alanina amidasa (C-LYTA) de *Streptococcus pneumoniae* (J. Madoz, B.A. Kuznetsov, F.J. Medrano, J.L. García, V.M. Fernández "Functionalization of Gold Surfaces for Specific and Reversible Attachment of a Fused β -Galactosidase and Choline-Receptor Protein J. Am. Chem. Soc. 119 (1997) 1043-1051).
- β -galactosidasa de *E. coli* con el dominio de unión a colina de la amidasa autolítica de neumococo, LytA, fusionada a su región N-terminal (J.M. Sánchez-Puelles, J.M. Sanz, J.L. García, E. García "Immobilization and single-step purification of fusion proteins using DEAE-cellulose" Eur. J. Biochem. 203 (1992) 153-159),
- β -galactosidasa termoestable de *Thermus sp.* Strain T2 fusionada al dominio de unión a colina de la principal autolisina de neumococo (A. Vián, A.V. Carrascosa, J.L. García, E. Cortés. "Structure of the β -galactosidase gene from *Thermus sp* T2: Expression in *Escherichia coli* and purification in a single step of an active fusion protein" Applied and Environmental Microbiology 64 (1998) 2187-2191).
- Patente española nº 9701759 (7 de Agosto de 1997): "Un procedimiento para producir β -galactosidasa de *Thermus sp.* (Cepa T2) en células hospedantes, y purificarla en un sólo paso cromatográfico". Inventores: A. Vián; A.V. Carrascosa; J.L. García y E. Cortés.

Descripción

La presente invención se basa en que los inventores han identificado y aislado una secuencia de nucleótidos novedosa correspondiente a la β -galactosidasa de *Streptococcus mitis*. Posteriormente, han creado dos métodos de expresión y purificación de β -galactosidasas modificadas: i) método de expresión y purificación de β -galactosidasa sin péptido señal y ii) método de expresión y purificación de β -galactosidasa con polí-histidinas.

La nueva proteína β -galactosidasa de *S. mitis* es una proteína consistente en 2.411 aminoácidos con un peso molecular aproximado de 268 kDa, presenta 5 motivos repetidos de unión a colina. Estas repeticiones son un rasgo característico de las proteínas de unión a colina de *S. pneumoniae*, lo que permite su purificación de manera sencilla a través de una columna cromatográfica de DEAE-celulosa y su empleo para diferentes aplicaciones.

Más concretamente, se ha aislado la secuencia de nucleótidos (SEQ ID NO1) que corresponde a un gen estructural de *S. mitis* NCTC (226)^T que codifica una β -galactosidasa (EC 3.2.1.23) (localizada entre las posiciones 64190 y 71425). Se facilita la secuencia de aminoácidos (SEQ ID NO2) deducida de la secuencia de nucleótidos. La eliminación de los nucleótidos correspondientes al péptido señal de la β -galactosidasa en la secuencia de nucleótidos SEQ ID NO1 origina la nueva secuencia de nucleótidos SEQ ID NO3, correspondiente a la β -galactosidasa sin péptido señal. La secuencia SEQ ID NO4 es la secuencia de aminoácidos codificada por la secuencia de nucleótidos SEQ ID NO3.

Las secuencias de nucleótidos correspondientes a la β -galactosidasa con el péptido señal (SEQ ID NO1) y sin el péptido señal (SEQ ID NO3) se insertaron en el plásmido pUC19 para la creación de los plásmidos pBSF01 y pBSF02 que expresan, respectivamente, la β -galactosidasa con y sin péptido señal de *S. mitis* en un hospedador heterólogo. El plásmido pUC19 es comercial, tiene la capacidad de producir un alto número de copias y está compuesto por 2.686 pares de bases. Además, tiene una zona que confiere resistencia al antibiótico ampicilina, y otra zona que contiene el gen del regulador de la expresión de Lac Z.

El hospedador heterólogo utilizado es la cepa MC4100, de la especie *E. coli*, que carece de actividad β -galactosidasa por llevar una delección completa del operón de la lactosa, obteniéndose como resultado final la cepa *E. coli* MC4100 (pBSF01 o pBSF02). De estas dos cepas, una de ellas, *E. coli* MC4100 (pBSF02) ha sido depositada en la Colección Española de Cultivos Tipo (CECT) con el nº de depósito 7433.

En el proceso de purificación de la β -galactosidasa sin péptido señal se utilizan columnas de afinidad de DEAE-celulosa, aprovechando la ventaja de disponer de una enzima con un módulo de unión a colina de forma natural. Con objeto de obtener un producto lo más purificado posible, y teniendo en cuenta el elevado peso molecular de la proteína recombinante, se empleó una columna cromatográfica preparativa de exclusión por tamaño.

En otro aspecto de la invención, los inventores han creado una β -galactosidasa en la que se han introducido seis restos de histidina en el extremo N-terminal de la proteína β -galactosidasa sin péptido señal, que denominaremos IMAC/ β -galactosidasa en referencia al proceso cromatográfico de purificación (ver más adelante), cuya secuencia de nucleótidos es la SEQ ID NO5, que se corresponde con la secuencia de aminoácidos SEQ ID NO6. Esta proteína modificada se puede purificar por cromatografía de afinidad a iones metálicos inmovilizados (IMAC, del inglés *immobilized metal ion chromatography*), consiguiendo una mejora en el rendimiento del proceso de purificación respecto al proceso de purificación de la β -galactosidasa sin péptido señal comentados previamente.

Los inventores insertaron esta secuencia SEQ ID NO5 en el plásmido pBSF02, obteniéndose como resultando el plásmido pHGM02. El hospedador heterólogo utilizado en esta ocasión fue nuevamente la cepa *E. coli* MC4100.

Dicho hospedador se transformó con el plásmido pHGM02, obteniéndose como resultado final la cepa *E. coli* MC4100 (pHGM02). Esta cepa ha sido depositada en la CECT con el nº de depósito 7434, y produce la proteína IMAC/ β -galactosidasa con poli-histidinas (SEQ ID NO6).

La producción y purificación de la proteína IMAC/ β -galactosidasa se realizó mediante la obtención de extractos enriquecidos en la proteína siguiendo un procedimiento análogo al descrito previamente para la β -galactosidasa sin péptido señal. La purificación se realizó utilizando soportes cromatográficos IMAC comerciales.

A diferencia de lo que ocurre generalmente con otras β -galactosidasas, la actividad enzimática de las proteínas recombinantes descritas no se inhibe en presencia de elevadas concentraciones de glucosa, lo que hace que estas enzimas encuentren aplicación principalmente en el sector agroalimentario-sanitario, para la hidrólisis de β -galactósidos de mala digestibilidad presentes en diversos alimentos de gran consumo.

La enzima β -galactosidasa sin péptido señal, purificada en columnas cromatográficas de DEAE-celulosa y Sephacryl 300, presenta una actividad específica de 2.209 ó 956 U/mg en condiciones estándar (30°C, 4 min de hidrólisis) cuando se usa ONPG o lactosa, respectivamente, como sustrato y conserva el 100% de su actividad ONPGasa inicial después de mantenerla 55 días a temperatura ambiente (sin estabilizadores).

Por lo tanto, un aspecto de la invención lo constituye una secuencia de nucleótidos (en adelante secuencia de nucleótidos de la invención) de la β -galactosidasa que comprende una de las secuencias pertenecientes al siguiente grupo:

- a) una secuencia de nucleótidos constituida por la secuencia de nucleótidos de la β -galactosidasa (SEQ ID NO1),
- b) una secuencia de nucleótidos análoga a la secuencia de a),
- c) un fragmento de una cualquiera de las secuencias de a) y b).
- d) una secuencia de nucleótidos análoga a la secuencia de a), b) y c).

Tal como se utiliza en la presente invención el término “secuencia de nucleótidos” es una secuencia de ADN, ADNc o ARNm.

Un aspecto más particular de la invención lo constituye la secuencia de nucleótidos de la β -galactosidasa SEQ ID NO1.

Otro aspecto más particular de la invención lo constituye un fragmento de la secuencia de nucleótidos de la β -galactosidasa SEQ ID NO1 en la que el fragmento es la β -galactosidasa sin péptido señal SEQ ID NO3.

Otro aspecto más particular de la invención lo constituye una secuencia de nucleótidos SEQ ID NO5 que comprende la secuencia de nucleótidos SEQ ID NO3 y una secuencia de nucleótidos codificante de seis histidinas en el extremo N-terminal de la SEQ ID NO3.

En el sentido utilizado en esta descripción, el término “análoga” pretende incluir a cualquier secuencia de nucleótidos que pueda ser aislada o construida en base a la secuencia de nucleótidos mostrada en la SEQ. ID NO1, SEQ ID NO3 o SEQ ID NO5 por ejemplo, mediante la introducción de sustituciones de nucleótidos conservativas o no conservativas, incluyendo la inserción de uno o más nucleótidos, la adición de uno o más nucleótidos en cualquiera de los extremos de la molécula o la delección de uno o más nucleótidos en cualquier extremo o en el interior de la secuencia.

En general, una secuencia de nucleótidos análoga es sustancialmente homologa a la secuencia de nucleótidos identificada como la SEQ ID NO1 o SEQ ID NO3 o SEQ ID NO5. En el sentido utilizado en esta descripción, la expresión “sustancialmente homologa” significa que las secuencias de nucleótidos en cuestión tienen un grado de identidad, a nivel de aminoácidos, de, al menos, un 60%, preferentemente de, al menos un 85%, o más preferentemente de, al menos, un 95%.

Una realización particular de la invención lo constituye una construcción genética (en adelante construcción genética de la invención) que comprende la secuencia de nucleótidos de la β -galactosidasa SEQ ID NO1, SEQ ID NO3 o SEQ ID NO5.

Esta construcción genética de la invención, también puede comprender, en caso necesario y para permitir un mejor aislamiento, detección o secreción al exterior de la célula del péptido de interés expresado, una secuencia de nucleótidos que codifica para un péptido susceptible de ser utilizado con fines de aislamiento, detección o secreción de dicho péptido. Por tanto, un objeto particular de la presente invención lo constituye una construcción genética de la invención que comprende cualquier otra secuencia de nucleótidos codificante de un péptido o secuencia peptídica que permita el aislamiento, la detección o la secreción al exterior de la célula del péptido expresado, por ejemplo, a título ilustrativo y sin que limite el alcance de la invención, una secuencia peptídica reconocible por un anticuerpo monoclonal (por

ejemplo, para su identificación, o cualquier otra que sirva para purificar la proteína de fusión resultante por cromatografía de inmunoafinidad: péptidos etiqueta tales como c-myc, HA, E-tag) (Using antibodies: a laboratory manual. Ed. Harlow and David Lañe (1999). Cold Spring Harbor Laboratory Press. New York. Capítulo: Tagging proteins. Pp. 347-377).

La construcción genética de la invención descrita previamente puede aislarse y obtenerse por un experto mediante el empleo de técnicas ampliamente conocidas en el estado de la técnica (Sambrook *et al.* "Molecular cloning, a Laboratory Manual 2nd ed., Cold Sping Harbor Laboratory Press, N.Y., 1989 vol 1-3). Dichas secuencias de nucleótidos pueden estar integradas en un vector de expresión génica que permite la regulación de la expresión de la misma en condiciones adecuadas en el interior de las células.

Por tanto, otro aspecto de la invención lo constituye un vector de expresión (en adelante vector de expresión de la invención) que comprende la secuencia de nucleótidos de la β -galactosidasa SEQ ID NO1, SEQ ID NO3 o SEQ ID NO5.

Un aspecto más particular de la invención lo constituye un vector de expresión de la invención constituido por el plásmido pBSF01, pBSF02 o pHGM02.

En general, un vector de expresión comprende, además de la construcción genética descrita en la invención, un promotor que dirige su transcripción (por ejemplo, pT7, plac, ptrc, ptac, pBAD, ret, etc.), al que está operativamente enlazado, y otras secuencias necesarias o apropiadas que controlan y regulan dicha transcripción y, en su caso, la traducción del producto de interés, por ejemplo, señales de inicio y terminación de transcripción (tlt2, etc.), señal de poliadenilación, origen de replicación, secuencias de unión a ribosomas (RBS), secuencias codificantes de reguladores transcripcionales, (enhancers), silenciadores transcripcionales (silencers), represores, etc. Ejemplos de vectores de expresión apropiados pueden seleccionarse de acuerdo con las condiciones y necesidades de cada caso concreto entre plásmidos de expresión, vectores virales (ADN o ARN), cósmidos, cromosomas artificiales, etc. que pueden contener, además, marcadores utilizables para seleccionar las células transfectadas o transformadas con el gen o genes de interés. La elección del vector dependerá de la célula huésped y del tipo de uso que se quiera realizar. Por tanto, según un modo de realización particular de la presente invención dicho vector es un plásmido o un vector viral. La obtención de dicho vector puede realizarse por métodos convencionales conocidos por los técnicos en la materia al igual que para la transformación de microorganismos y células eucariotas se pueden utilizar diferentes métodos ampliamente conocidas - transformación química, electroporación, microinyección, etc. - descritos en diversos manuales [Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.].

Otro aspecto particular de la invención lo constituye una secuencia de aminoácidos β -galactosidasa (en adelante secuencia de aminoácidos de (a invención) que comprende una de las secuencias de aminoácidos pertenecientes al siguiente grupo:

- a) una secuencia de aminoácidos constituida por la secuencia de aminoácidos de la β -galactosidasa de *S. mitis* (SEQ ID NO2),
- b) una secuencia de aminoácidos análoga a la secuencia a),
- c) un fragmento cualquiera de las secuencias de a) y b),
- d) una secuencia de aminoácidos que comprende una secuencia cualquiera de a), b) y c).

Un aspecto más particular de la invención lo constituye la secuencia de aminoácidos de la invención de la β -galactosidasa SEQ ID NO2.

Otro aspecto más particular de la invención lo constituye un fragmento de la secuencia de aminoácidos de la β -galactosidasa SEQ ID NO2 en la que el fragmento es la β -galactosidasa sin péptido señal SEQ ID NO4.

Otro aspecto más particular de la invención lo constituye una secuencia de aminoácidos SEQ ID NO6 que comprende la secuencia de aminoácidos SEQ ID NO4 y una secuencia de seis histidinas en el extremo N-terminal de la SEQ ID NO4.

Otro aspecto de la invención lo constituye un uso de la secuencia de aminoácidos de la invención (SEQ ID NO2, SEQ ID NO4 o SEQ ID NO6) en la reducción del contenido de lactosa en alimentos o residuos de procesos alimentarios pertenecientes, a título ilustrativo y sin que limite el alcance de la invención, al siguiente grupo: productos lácteos y derivados, sueros de lechería.

Otro aspecto de la invención lo constituye una célula huésped, en adelante célula huésped de la invención, que comprende la construcción genética o el vector de expresión de la invención, preferentemente una célula procariota.

Una realización particular de la invención lo constituye una célula huésped de la invención constituida por una célula *E. coli*, preferentemente células *E. coli* MC4100.

Otra realización particular de la invención lo constituye una célula huésped de la invención constituida por células de la cepa *E. coli* con el número de depósito CECT 7433 o 7434.

Otro aspecto de la invención lo constituye el uso de la célula huésped de la invención en la reducción del contenido de lactosa en alimentos o residuos de procesos alimentarios pertenecientes, a título ilustrativo y sin que limite el alcance de la invención, al siguiente grupo: productos lácteos y derivados, sueros de lechería.

Otro aspecto de la invención lo constituye un procedimiento de obtención (en adelante procedimiento de obtención de la β -galactosidasa de la invención) de la secuencia de aminoácidos de la invención SEQ ID NO2 que comprende las siguientes etapas:

- a) obtención de una construcción genética que comprende la secuencia de nucleótidos SEQ ID NO1, y del vector de expresión que comprende esta construcción genética,
- b) transformación de una célula huésped, preferentemente de *E. coli*, en la que la construcción genética o el vector de expresión de a) bajo control de los promotores adecuados expresa la proteína codificada por la SEQ ID NO1, y purificación del extracto crudo de las células,
- c) inmovilización de la proteína en matrices de DEAE-celulosa mediante el paso del extracto crudo de b) a través de la matriz,
- d) disociación de la proteína de la matriz cromatográfica mediante lavado de la matriz con un medio que comprende colina,
- y, opcionalmente
- e) posterior purificación de la proteína mediante el paso del eluido de d) a través de una matriz cromatográfica de exclusión por tamaño.

Otro aspecto más particular de la invención lo constituye un procedimiento de obtención de β -galactosidasa de la invención en el que la secuencia de nucleótidos de a) es la secuencia de nucleótidos SEQ ID NO3.

Otro aspecto particular de la invención lo constituye un procedimiento de obtención de la secuencia de aminoácidos SEQ ID NO6 caracterizado porque comprende las siguientes etapas:

- a) obtención de una construcción genética que comprende la secuencia de nucleótidos SEQ ID NO5, y del vector de expresión que comprende esta construcción genética,
- b) transformación de una célula huésped, preferentemente de *E. coli*, en la que la construcción genética o el vector de expresión de a) bajo control de los promotores adecuados expresa la proteína codificada por la SEQ ID NO5, y purificación del extracto crudo de las células,
- c) inmovilización de la proteína en matrices de cromatografía de afinidad a iones metálicos inmovilizados mediante el paso del extracto crudo de b) a través de la matriz,
- d) disociación de la proteína de la matriz cromatográfica mediante lavado de la matriz con un medio que comprende imidazol.

Ejemplos de realización

Ejemplo 1

Identificación de la secuencia de nucleótidos de la β -galactosidasa de *S. mitis*

La secuencia del genoma de la cepa tipo de *S. mitis* NCTC (226)^T se conoce sólo en fragmentos no ordenados (llamados *contigs*). El análisis *in silico* de las correspondientes fases de lectura abierta permitió deducir la existencia de, al menos, 14 proteínas de unión a colina. Una de ellas codificaba una proteína de 2411 aminoácidos, con un posible péptido señal, alta similitud con otras β -galactosidasas, y que presentaba 5 motivos repetidos de unión a colina. Estos resultados sugirieron que la proteína podría tratarse de una β -galactosidasa extracelular probablemente diseñada para realizar su función en la envuelta celular, y que poseería características diferentes y especiales con respecto a otras β -galactosidasas intracelulares.

Ejemplo 2

*Expresión de la proteína recombinante β -galactosidasa con y sin péptido señal*5 2.1.- Construcción de plásmidos para expresar la proteína β -galactosidasa de *S. mitis* en el hospedador *E. coli* MC4100 (Figura 1)

Los plásmidos construidos se denominan pBSF01 (con péptido señal) y pBSF02 (sin péptido señal) y para su realización se utilizaron técnicas de ADN recombinante convencionales [Sambrook, J., Fritsch, E.F. y Maniatis, T. 10 1989. Molecular Cloning. CSHL Press, Cold Spring Harbor, Nueva York].

En primer lugar se extrajo el ADN de *S. mitis*, y tras su purificación, se amplificó la secuencia adecuada mediante la técnica de PCR.

15 Para llevar a cabo la amplificación de la parte de ADN de interés, se eligieron unos oligonucleótidos cebadores para que la polimerasa actuara a partir de las posiciones deseadas. Se utilizaron 2 oligonucleótidos diferentes en el sentido 5'-3', uno incluyendo la parte del péptido señal, y el otro sin ella. Además, en los cebadores se incluyeron secuencias de restricción para posteriormente fragmentar el producto amplificado por los sitios adecuados y unirlo posteriormente al vector de expresión.

20 Se eligió introducir una secuencia de corte para *Xba*I en el cebador 5'-3' y *Eco*RV en el cebador 3'-5', puesto que no hay sitios de corte para estas enzimas de restricción en el gen. Además, se introdujo la secuencia que determina el sitio de unión del ribosoma (ribosome binding site, RBS), y codones STOP en las lecturas desfasadas de la secuencia de interés. Finalmente, también se introdujo el codón ATG, que codifica el aminoácido metionina, que indica el inicio del gen.

Después de purificar el producto de PCR obtenido con los oligonucleótidos mencionados con un kit de purificación se realizó la digestión con las enzimas de restricción *Xba*I y *Eco*RV.

30 Para construir los plásmidos recombinantes (pBSF01 y pBSF02) se partió del plásmido pUC19. El plásmido pUC19 se cortó con las enzimas de restricción adecuadas (*Xba*I y *Sma*I), para su posterior ligación al inserto del gen de la β -galactosidasa.

35 Los productos de la digestión del plásmido y del gen amplificado de la β -galactosidasa se purificaron en un gel de agarosa y con un kit de purificación de ADN.

40 Los productos purificados de la digestión del plásmido y del gen amplificado de β -galactosidasa se unieron covalentemente con ADN ligasa. Con la mezcla de ligación antedicha se transformaron células competentes de *E. coli* MC4100. El plásmido resultante, que se denominó pBSF02, hace que las células que lo portan sean resistentes a ampicilina, y es capaz de expresar actividad β -galactosidasa correspondiente al gen estructural de *S. mitis*, SEQ ID NO1.

45 De los varios clones ensayados se eligió el más productivo, que es el clon de *E. coli* MC4100 (pBSF02) depositado en la CECT con el n° de depósito 7433.

El plaqueo de las células competentes de *E. coli* MC4100 (pBSF02) se hizo en agar LB con ampicilina, X-Gal (5-Bromo-4-cloro-3-indolil- β -D-galactopiranosido) e IPTG (isopropil- β -D-1-tiogalactopiranosido), incubándose a 37°C. Se escogió una colonia azul.

50 Tal colonia se reaisló y se hizo una minipreparación de su ADN plasmídico. Tras comprobar que el patrón de restricción con diversas enzimas era el esperado, se hizo una verificación final mediante secuenciación.

55 2.2.- Producción de la proteína recombinante β -galactosidasa sin péptido señal

De las placas con las colonias azules del clon antes mencionado, se eligió una suficientemente aislada, que se inoculó en medio de cultivo LB con ampicilina (100 μ g/ml) y se cultivó toda la noche a 37°C con agitación a 250 rpm. Al día siguiente, un volumen adecuado del cultivo crecido se diluyó en medio LB con ampicilina (100 μ g/ml) e IPTG (1.0x10⁻⁴ M) y se dejó crecer durante toda la noche. Tras medir su densidad óptica a 600 nm, el cultivo se 60 centrifugó a 8000 rpm durante 10 minutos a 4°C. Las células de *E. coli* MC4100 (pBSF02) cosechadas por centrifugación y resuspendidas en un volumen menor del original de tampón fosfato sódico 20 mM (pH 6.9), como método de concentración, se ultrasonicaron para romperlas. Tras una centrifugación adicional, que separa restos celulares, se obtuvo un sobrenadante claro con elevada actividad β -galactosidasa.

65

Ejemplo 3

Purificación de la proteína recombinante β -galactosidasa sin péptido señal

5 Obtenido el extracto crudo se siguió su purificación, en una columna cromatográfica de afinidad de DEAE-celulosa [J.M. Sánchez-Puelles, J.M. Sanz, J.L. García, E. García. 1992 "Immobilization and single-step purification of
 10 fusión proteins using DEAE-cellulose" Eur. J. Biochem., 203, 153-159]. Después de cargar el extracto (~ 12.5 mg de proteína por mL de soporte empaquetado) se lavó con tampón fosfato 20 mM pH 6.9. A continuación se eluyó por la columna tampón salino (fosfato 20 mM, pH 6,9 con 1,5 M NaCl) y se recogieron las fracciones de interés con
 un colector, cuando se hizo pasar por la columna tampón con colina (fosfato 20 mM, pH 6,9, con 1,5 M NaCl y 2% colina).

15 Para comprobar la eficacia del proceso de purificación, se realizaron electroforesis en geles de poliacrilamida en presencia de SDS (7.5%) y ensayos de actividad β -galactosidasa y determinación de proteínas por el método de Bradford.

20 Para mejorar la pureza del producto final se utilizó una columna cromatográfica de afinidad por exclusión de tamaño. Tras equilibrar la columna (90x30 cm, Sephacryl-300) con tampón fosfato 20 mM pH 6.9, se cargó el extracto semipurificado y empleando la misma disolución reguladora y con ayuda de una bomba peristáltica (con un flujo de 0.2 mL/min) se recogieron las fracciones de interés. La enzima purificada se guardó a -20°C.

Experimentos de ultracentrifugación analítica demostraron que la enzima recombinante purificada era monomérica (en tampón fosfato 20 mM, pH 6.9, conteniendo 50 mM de NaCl, en ausencia y en presencia de 2% de colina).

Ejemplo 4

Expresión y purificación de la proteína recombinante con poli-histidinas (IMAC/ β -galactosidasa)

30 4.1.- Construcción de un plásmido para expresar la proteína β -galactosidasa de *S. mitis* con poli-histidinas (IMAC/ β -galactosidasa) (Figura 3)

35 El plásmido construido se denomina pHGM02 y para su construcción se emplearon protocolos convencionales, partiendo del plásmido pBSF02. Con el fin de introducir los 6 residuos de histidinas en la región N-terminal se realizó la amplificación de una región de 758 bp del plásmido recombinante pBSF02 con los oligonucleótidos adecuados. En estos oligonucleótidos se incluyeron secuencias de restricción para posteriormente fragmentar el producto amplificado por los sitios adecuados, para su posterior unión al vector de expresión.

40 Se ha elegido introducir una secuencia de corte para *Xba*I en un oligonucleótido y *Spe*I en el otro, puesto que estas enzimas cortan cohesivamente y en posición única el gen de la β -galactosidasa de *S. mitis* pero no el pUC19. Además, se introdujo la secuencia que determina el sitio de unión del ribosoma (ribosome binding site, RBS), y codones STOP en las lecturas desfasadas de la secuencia de interés. Finalmente, también se introdujo el codón ATG, que codifica el aminoácido metionina, que indica el inicio del gen.

45 El fragmento PCR obtenido se purificó y se ligó al vector de clonación pGEMT-Easy (Promega). La selección del plásmido portador del inserto y las posteriores manipulaciones para obtener el plásmido recombinante pHGM02 fueron similares a las explicadas en el ejemplo 2.1 y siguieron el esquema que se explica en Figura 3.

50 4.2.- Producción de la proteína recombinante IMAC/ β -galactosidasa

Para producir la proteína IMAC/ β -galactosidasa se siguió un procedimiento análogo al descrito para la producción de la β -galactosidasa sin péptido señal (ejemplo 2.2). Además, se realizaron estudios bioquímicos comparativos entre ambas proteínas, encontrándose resultados similares a nivel de expresión y características bioquímicas.

55

4.3.- Purificación de la proteína recombinante IMAC/ β -galactosidasa

60 En el proceso de purificación se utilizaron columnas IMAC, en cartuchos comerciales (BioRad) siguiendo el protocolo recomendado. La adsorción de la proteína al soporte se realizó a pH 6.9 y 25°C. Seguidamente se realizó la desorción de la proteína del soporte, usando concentraciones crecientes de imidazol (desde 5 hasta 250 mM), encontrando que a 200 mM eluía el 100% de la enzima de *S. mitis*.

65 Para comprobar la eficacia del proceso de purificación, se realizaron electroforesis en geles de poliacrilamida en presencia de SDS (7.5%) y ensayos de actividad β -galactosidasa y determinación de proteínas por el método de Bradford. La muestra correspondiente a la proteína purificada tras separación electroforética se muestra en la Figura 4.

Ejemplo 5

Caracterización de la actividad enzimática

5 Las actividades enzimáticas de las dos proteínas recombinantes purificadas, la β -galactosidasa sin péptido señal (codificada por el plásmido pBSF02) y la IMAC/ β -galactosidasa (codificada por el plásmido pHGM02), fueron similares en términos de actividad específica.

10 A continuación se presentan los resultados obtenidos a la actividad enzimática de la enzima β -galactosidasa sin péptido señal:

1. *Sustratos*: Hidroliza ONPG con una actividad específica de 2.209 unidades a pH 6,5 y a 30°C. Hidroliza lactosa con una actividad específica de 956 unidades. (Se refieren a las unidades definidas en el ejemplo 1).
- 15 2. *pH de la reacción*: El pH óptimo de la hidrólisis se sitúa entre el 5,5 y 6,5. Se detecta claramente actividad en un rango de pH de 5,5 a 7,5.
- 20 3. *Temperatura óptima*: se alcanzó el máximo de actividad entre 35-40°C a pH 7,0.
4. *Termoestabilidad*: La enzima soluble conserva el 79% de la actividad después de un calentamiento de 5 minutos a 40°C y pH 7,0, y pierde totalmente su actividad enzimática después de calentarla 5 min a 42,5°C y pH 7,0.
- 25 5. *Inhibición por productos de reacción*: En los ensayos con la enzima soluble a 30°C y pH 6,5, se conserva el 100% de la actividad de la enzima en presencia de 200 mM de glucosa. La galactosa, a una concentración de 10 mM, disminuye la actividad en un 64%.
- 30 6. *Conservación*. La enzima soluble se puede conservar a temperatura ambiente (sin estabilizadores) durante al menos 55 días, reteniendo el 100% de su actividad inicial.

Descripción de las figuras

35 Figura 1.- *Esquema del método seguido para construir los plásmidos que contienen el gen de la β -galactosidasa con y sin péptido señal del microorganismo *S. mitis**. Se representan sólo los aspectos más relevantes de los plásmidos. Para detalles, ver el texto (ejemplo 2). Sólo se indican los sitios de restricción y las estructuras más importantes para poder comprender mejor el esquema. La flecha gruesa representa el gen estructural que codifica la β -galactosidasa de *S. mitis*. Rectángulo negro: Péptido señal. PCR: Reacción en cadena de la polimerasa SMC: Sitio múltiple de clonación.

40 Figura 2.- *Secuencias de oligonucleótidos utilizados en los experimentos de PCR*. A) Oligonucleótido utilizado para la construcción del plásmido pBSF01 (1 en Figura 1). B) Oligonucleótido utilizado para la construcción del plásmido pBSF02 (2 en Figura 1). C) Oligonucleótido utilizado para la construcción de los plásmidos pBSF01 y pBSF02 (3 en Figura 1). D) Oligonucleótido utilizado para la construcción del plásmido pHGM02 (4 en Figura 3). E) Oligonucleótido utilizado para la construcción del plásmido pHGM02 (5 en Figura 3).

50 Figura 3.- *Esquema del método seguido para construir un plásmido que contine el gen de la IMAC/ β -galactosidasa del microorganismo *S. mitis**. Se representan sólo los genes más relevantes de los plásmidos. Para detalles, ver el texto (ejemplo 4). Sólo se indican los sitios de restricción y las estructuras más importantes para poder comprender mejor el esquema. La flecha gruesa representa el gen estructural que codifica la β -galactosidasa de *S. mitis*. PCR: Reacción en cadena de la polimerasa.

55 Figura 4.- *Análisis de proteínas por electroforesis en geles de poliacrilamida en condiciones desnaturalizantes al 7,5% (SDS-PAGE)*. Calle x, proteínas marcadoras de peso molecular, indicadas en el margen izquierdo, en kDa. Calle y, β -galactosidasa sin péptido señal de *S. mitis* purificada por DEAE-celulosa y Sephacryl S-300. Calle z, IMAC/ β -galactosidasa de *S. mitis* purificada por IMAC.

REIVINDICACIONES

1. Secuencia de nucleótidos **caracterizada** porque comprende una de las secuencias pertenecientes al siguiente grupo:
 - a) una secuencia de nucleótidos constituida por la secuencia de nucleótidos de la β -galactosidasa SEQ ID NO1,
 - b) una secuencia de nucleótidos análoga a la secuencia de a),
 - c) un fragmento de una cualquiera de las secuencias de a) y b).
 - d) una secuencia de nucleótidos análoga a la secuencia de a), b) y c).
2. Secuencia de nucleótidos según la reivindicación 1 **caracterizada** porque es la secuencia de la β -galactosidasa SEQ ID NO1.
3. Secuencia de nucleótidos según reivindicación 1 **caracterizada** porque la secuencia de nucleótidos de c) es un fragmento de la β -galactosidasa sin péptido señal (SEQ ID NO3).
4. Secuencia de nucleótidos según reivindicación 1 **caracterizada** porque es la secuencia SEQ ID NO5, que comprende la secuencia de nucleótidos SEQ ID NO3 y una secuencia de nucleótidos codificante de seis histidinas en el extremo N-terminal de la SEQ ID NO3.
5. Secuencia de nucleótidos según reivindicaciones 1 a 4 **caracterizada** porque es una secuencia de ADN, ADNc o ARNm.
6. Construcción genética **caracterizada** porque comprende una secuencia de nucleótidos según reivindicaciones 1 a 5.
7. Vector de expresión **caracterizado** porque comprende una secuencia de nucleótidos según reivindicaciones 1 a 6.
8. Vector de expresión según la reivindicación 7 **caracterizado** porque el vector de expresión es el plásmido pBSF01, pBSF02 o pHGM02.
9. Secuencia de aminoácidos **caracterizada** porque comprende una de las secuencias de aminoácidos pertenecientes al siguiente grupo:
 - a) una secuencia de aminoácidos constituida por la secuencia de aminoácidos de la β -galactosidasa de *S. mitis* (SEQ ID NO2),
 - b) una secuencia de aminoácidos análoga a la secuencia a),
 - c) un fragmento cualquiera de las secuencias de a) y b),
 - d) una secuencia de aminoácidos que comprende una secuencia cualquiera de a), b) y c).
10. Secuencia de aminoácidos según reivindicación 9 **caracterizada** porque es la secuencia de la β -galactosidasa SEQ ID NO2.
11. Secuencia de aminoácidos según reivindicación 9 **caracterizada** porque la secuencia de aminoácidos de c) es un fragmento de la β -galactosidasa sin péptido señal (SEQ ID NO4).
12. Secuencia de aminoácidos según reivindicación 9 **caracterizada** porque es la secuencia de la β -galactosidasa modificada SEQ ID NO6, que comprende la secuencia de aminoácidos SEQ ID NO4 y una secuencia de seis histidinas en el extremo N-terminal de la SEQ ID NO4.
13. Uso de la secuencia de aminoácidos según reivindicaciones 9 a 12 en la reducción del contenido de lactosa en alimentos o residuos de procesos alimentarios pertenecientes, a título ilustrativo y sin que limite el alcance de la invención, al siguiente grupo: productos lácteos y derivados, sueros de leche.
14. Célula huésped, preferentemente procariota, **caracterizada** porque comprende una construcción genética según reivindicaciones 1 a 6.

15. Célula huésped, preferentemente procariota, **caracterizada** porque comprende el vector de expresión según reivindicaciones 7 a 8.

16. Célula huésped según reivindicaciones 14 y 15 **caracterizada** porque es una célula de *Escherichia coli*, preferentemente células *E. coli* MC4100.

17. Célula huésped según reivindicación 16 **caracterizada** porque es una célula de la cepa *E. coli* con el número de depósito CECT 7433 o *E. coli* con el número de depósito CECT 7434.

18. Uso de la célula huésped según reivindicaciones 14 a 17 en la reducción del contenido de lactosa en alimentos o residuos de procesos alimentarios pertenecientes, a título ilustrativo y sin que limite el alcance de la invención, al siguiente grupo: productos lácteos y derivados, sueros de lechería.

19. Procedimiento de obtención de la secuencia de aminoácidos SEQ ID NO2 **caracterizado** porque comprende las siguientes etapas:

- a) obtención de una construcción genética que comprende la secuencia de nucleótidos SEQ ID NO1, y del vector de expresión que comprende esta construcción genética,
- b) transformación de una célula huésped, preferentemente de *E. coli*, en la que la construcción genética o el vector de expresión de a) bajo control de los promotores adecuados expresa la proteína codificada por la SEQ ID NO1, y purificación del extracto crudo de las células,
- c) inmovilización de la proteína en matrices de DEAE-celulosa mediante el paso del extracto crudo de b) a través de la matriz,
- d) disociación de la proteína de la matriz cromatográfica mediante lavado de la matriz con un medio que comprende colina,
- y, opcionalmente
- e) posterior purificación de la proteína mediante el paso del eluido de d) a través de una matriz cromatográfica de exclusión por tamaño.

20. Procedimiento de obtención según reivindicación 19 **caracterizado** porque la secuencia de nucleótidos de a) es la secuencia de nucleótidos SEQ ID NO3.

21. Procedimiento de obtención de la secuencia de aminoácidos SEQ ID NO6 **caracterizado** porque comprende las siguientes etapas:

- a) obtención de una construcción genética que comprende la secuencia de nucleótidos SEQ ID NO5, y del vector de expresión que comprende esta construcción genética,
- b) transformación de una célula huésped, preferentemente de *E. coli*, en la que la construcción genética o el vector de expresión de a) bajo control de los promotores adecuados expresa la proteína codificada por la SEQ ID NO5, y purificación del extracto crudo de las células,
- c) inmovilización de la proteína en matrices de cromatografía de afinidad a iones metálicos inmovilizados mediante el paso del extracto crudo de b) a través de la matriz,
- d) disociación de la proteína de la matriz cromatográfica mediante lavado de la matriz con un medio que comprende imidazol.

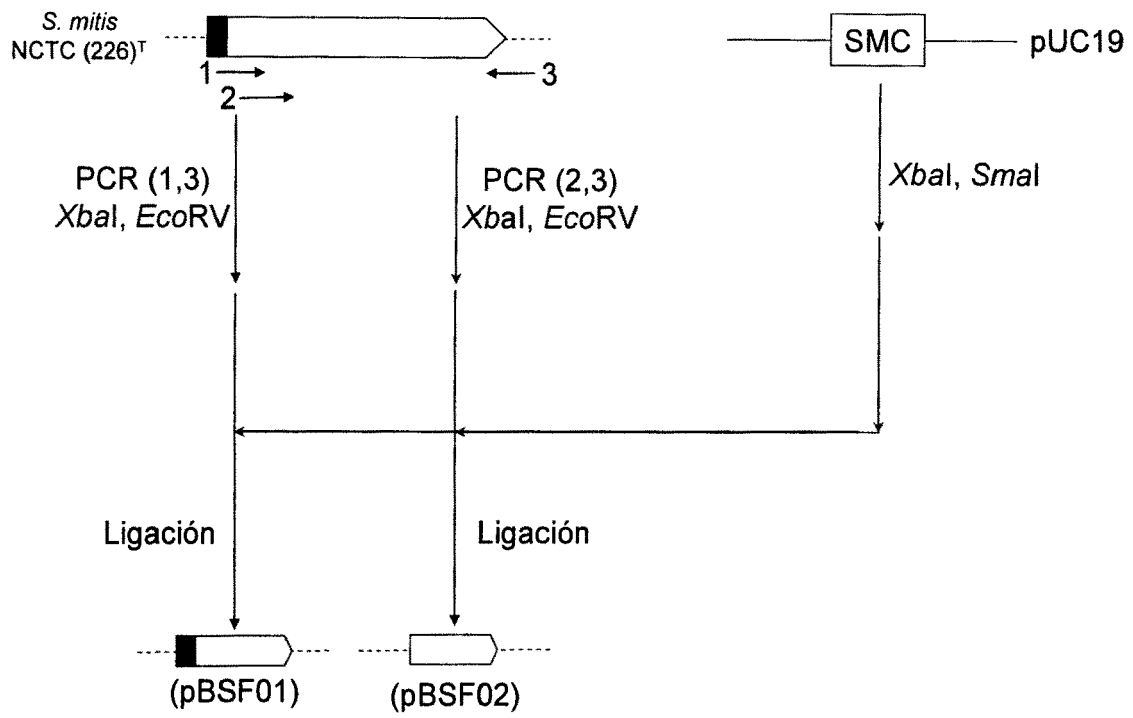


FIG. 1



FIG. 2

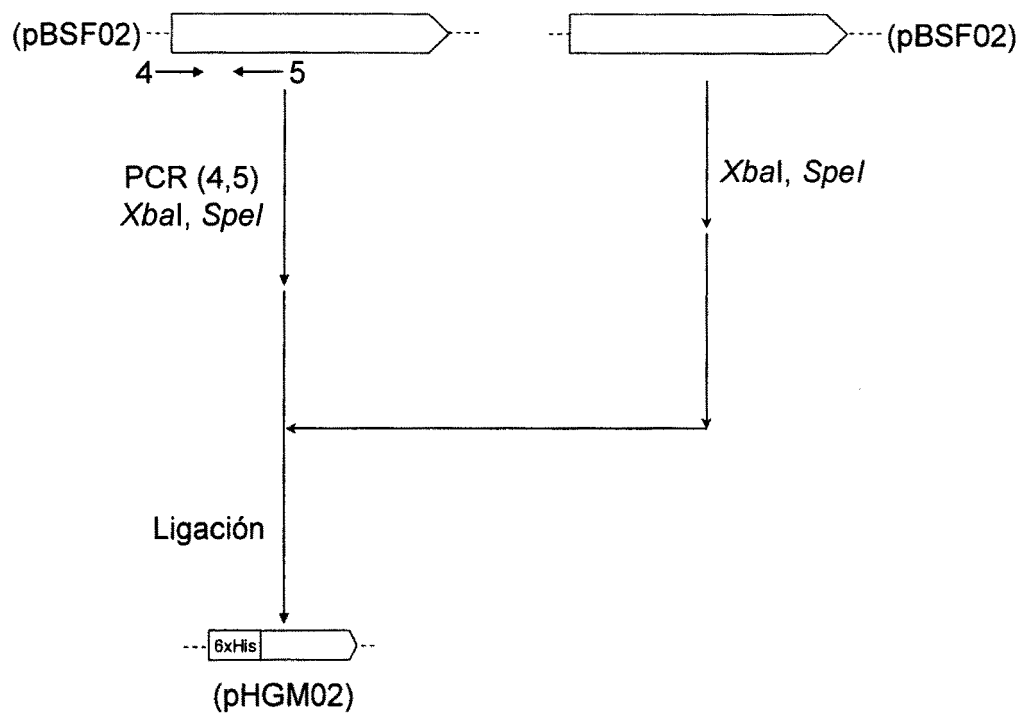


FIG. 3

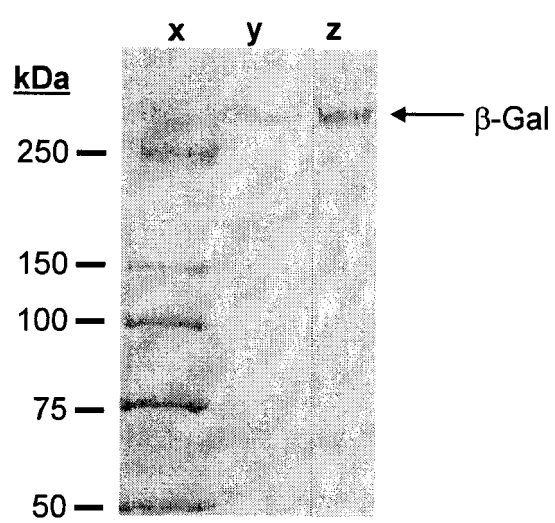


FIG. 4

LISTA DE SECUENCIAS

5 <110> CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
UNIVERSIDAD COMPLUTENSE DE MADRID

<120> SECUENCIAS DE NUCLEÓTIDOS Y PROTEÍNAS DE LA BETA-GALACTOSIDASA DE *STREPTOCOC-*
CUS MITIS, PROCEDIMIENTO DE OBTENCIÓN Y SUS APLICACIONES

10 <130> beta-gal

<160> 6

15 <170> PatentIn version 3.3

<210> 1

<211> 7266

20 <212> DNA

<213> *Streptococcus mitis*

<220>

25 <221> CDS

<222> (31)..(7263)

<223> beta-galactosidasa

30 <220>

<221> misc_feature

<222> (6898)..(6957)

35 <223> Motivo de unión a colina 1

<220>

<221> misc_feature

40 <222> (6958)..(7017)

<223> Motivo de unión a colina 2

<220>

45 <221> misc_feature

<222> (7018)..(7083)

<223> Motivo de unión a colina 3

50 <220>

<221> misc_feature

<222> (7084)..(7149)

55 <223> Motivo de unión a colina 4

<220>

<221> misc_feature

60 <222> (7150)..(7215)

<223> Motivo de unión a colina 5

65

ES 2 352 704 B1

<400> 1

```

aatgaaagcg gttaccaaaa ggaggattac atg caa aaa ggt aat tgg aac aaa
54
Met Gln Lys Gly Asn Trp Asn Lys
5
aaa aga gtg tat agt att cga aaa ttt tca gtc ggt gct tgc tcg gtg
102
Lys Arg Val Tyr Ser Ile Arg Lys Phe Ser Val Gly Ala Cys Ser Val
10
10 15 20
ctt ata ggg aca tgt gct cta tta tta ggg gca ggc gtt agt cta tct
150
Leu Ile Gly Thr Cys Ala Leu Leu Leu Gly Ala Gly Val Ser Leu Ser
15
25 30 35 40
tct tcg gta tat gcc aac gaa aat gct gaa gaa att gtg tta aca aat
198
Ser Ser Val Tyr Ala Asn Glu Asn Ala Glu Glu Ile Val Leu Thr Asn
20
45 50 55
agg act gag cac cag tct gaa aac ctg aaa act act gga ctc tca agt
246
Arg Thr Glu His Gln Ser Glu Asn Leu Lys Thr Thr Gly Leu Ser Ser
25
60 65 70
gaa cat aaa gct gaa ggt gtt acg act gaa aaa gca agt gag caa cct
294
Glu His Lys Ala Glu Gly Val Thr Thr Glu Lys Ala Ser Glu Gln Pro
30
35
75 80 85
gtt aaa gct gaa aat gta tct cta aag gaa aat gcg gat ggt agt gat
342
Val Lys Ala Glu Asn Val Ser Leu Lys Glu Asn Ala Asp Gly Ser Asp
40
90 95 100
gca gta gta aaa cct cac gca gat caa cct tta gtt tca gtt gcc aat
390
Ala Val Val Lys Pro His Ala Asp Gln Pro Leu Val Ser Val Ala Asn
45
105 110 115 120
aat caa tca gtt cct tta gaa gaa ggt aag gct gaa gtc gaa aac aaa
438
Asn Gln Ser Val Pro Leu Glu Glu Gly Lys Ala Glu Val Glu Asn Lys
50
125 130 135
aca gaa gaa agg aac aag cct gag gat aga aat aag ctt gaa gaa aca
486
Thr Glu Glu Arg Asn Lys Pro Glu Asp Arg Asn Lys Leu Glu Glu Thr
55
140 145 150
aat aag act gag gta acg gat aaa tct gaa gat aaa aat aaa tct gaa
534
Asn Lys Thr Glu Val Thr Asp Lys Ser Glu Asp Lys Asn Lys Ser Glu
60
155 160 165
gaa aca aat aaa gca gaa tca caa gat cgt tct gtt tca aaa gat aag
582
Glu Thr Asn Lys Ala Glu Ser Gln Asp Arg Ser Val Ser Lys Asp Lys
65

```

ES 2 352 704 B1

	170	175	180
5	caa gaa ttc aaa tct gct acg aat gaa gta gtt gat aag tta att gaa 630 Gln Glu Phe Lys Ser Ala Thr Asn Glu Val Val Asp Lys Leu Ile Glu		
	185	190	195 200
10	gat aga aac att tcc ttc aac cag aat tgg cac ttt aaa cta aat gcc 678 Asp Arg Asn Ile Ser Phe Asn Gln Asn Trp His Phe Lys Leu Asn Ala		
15		205	210 215
20	aat gca aag gaa gct gta aaa cca gac gca gat att tct tca tgg aaa 726 Asn Ala Lys Glu Ala Val Lys Pro Asp Ala Asp Ile Ser Ser Trp Lys		
	220	225	230
25	aaa atg gat ctt cct cat gac tgg agt att cat ttt gat ttt gac cat 774 Lys Met Asp Leu Pro His Asp Trp Ser Ile His Phe Asp Phe Asp His		
30		235 240	245
35	gat tca cca gct caa aac gaa ggc ggt caa ttg aat ggt gga gat ggc 822 Asp Ser Pro Ala Gln Asn Glu Gly Gly Gln Leu Asn Gly Gly Asp Gly		
	250	255	260
40	tgg tat cgg aaa acc ttt aag tta gat gaa aag gac ttg aat aaa gat 870 Trp Tyr Arg Lys Thr Phe Lys Leu Asp Glu Lys Asp Leu Asn Lys Asp		
	265	270	275 280
45	gtt cgt gta acc ttt gat ggt gtt tac atg gat tca caa gta ttc gtc 918 Val Arg Val Thr Phe Asp Gly Val Tyr Met Asp Ser Gln Val Phe Val		
50		285 290	295
55	aat ggc caa cta gtt ggt cac tat cca aat ggt tac aat caa ttt tca 966 Asn Gly Gln Leu Val Gly His Tyr Pro Asn Gly Tyr Asn Gln Phe Ser		
	300	305	310
60	tat gac att agt aac tat ctc cat aaa gat ggg cga gaa aat gtc att 1014 Tyr Asp Ile Ser Asn Tyr Leu His Lys Asp Gly Arg Glu Asn Val Ile		
65	315	320	325

ES 2 352 704 B1

tct gtt cat gca gtt aat aag caa cca agt agt aga tgg tat tca gga
 1062
 Ser Val His Ala Val Asn Lys Gln Pro Ser Ser Arg Trp Tyr Ser Gly
 5 330 335 340

agt ggt atc tat cga gac gta tca ttg cag gta aca gac aag att cat
 1110
 Ser Gly Ile Tyr Arg Asp Val Ser Leu Gln Val Thr Asp Lys Ile His
 10 345 350 355 360

gtt gaa aag aat ggg aca act att cta act cca aaa cta gaa agt caa
 1158
 Val Glu Lys Asn Gly Thr Thr Ile Leu Thr Pro Lys Leu Glu Ser Gln
 15 365 370 375

caa gga gga aaa gtt gaa acc caa gta agc agt aaa att gtt aat aca
 1206
 Gln Gly Gly Lys Val Glu Thr Gln Val Ser Ser Lys Ile Val Asn Thr
 20 380 385 390

gat aat aaa gac cat gaa att gta gca gag tat caa att ttt gaa cga
 1254
 Asp Asn Lys Asp His Glu Ile Val Ala Glu Tyr Gln Ile Phe Glu Arg
 25 395 400 405

ggt ggt aag gca gtt aca gaa ctt gtt cga act gag agt aaa act cta
 1302
 Gly Gly Lys Ala Val Thr Glu Leu Val Arg Thr Glu Ser Lys Thr Leu
 30 410 415 420

aaa gct aag gaa acc att cat ttg gat tca gca tta gag gtt gaa aaa
 1350
 Lys Ala Lys Glu Thr Ile His Leu Asp Ser Ala Leu Glu Val Glu Lys
 35 425 430 435 440

cca aaa ttg tgg aca gtg tca tct gac aaa cct gct tta tat gag atg
 1398
 Pro Lys Leu Trp Thr Val Ser Ser Asp Lys Pro Ala Leu Tyr Glu Met
 40 445 450 455

gta act cgt gtt tac aaa gat ggt cag ctt gta gat gca aag aaa gat
 1446
 Val Thr Arg Val Tyr Lys Asp Gly Gln Leu Val Asp Ala Lys Lys Asp
 45 460 465 470

tta ttt ggt tat cgc tac tat aac tgg act cca gac caa ggg ttc tct
 1494
 50 55 60 65

ES 2 352 704 B1

Leu Phe Gly Tyr Arg Tyr Tyr Asn Trp Thr Pro Asp Gln Gly Phe Ser
 475 480 485
 5
 tta aat ggt gaa cac atc aag ttc cac ggt gtt tca ttg cac cat gat
 1542
 Leu Asn Gly Glu His Ile Lys Phe His Gly Val Ser Leu His His Asp
 10 490 495 500
 cat gga gcg cta gga gca gaa gaa aac tat aag gca gaa tat cgt cgt
 1590
 His Gly Ala Leu Gly Ala Glu Glu Asn Tyr Lys Ala Glu Tyr Arg Arg
 15 505 510 515 520
 ctg aaa caa atg aag gaa atg gga gtg aat tcg att cgt acg act cac
 1638
 Leu Lys Gln Met Lys Glu Met Gly Val Asn Ser Ile Arg Thr Thr His
 20 525 530 535
 aac cca gca agt cct cag acc ttg cag att gca gct gag ctc ggt ttg
 1686
 Asn Pro Ala Ser Pro Gln Thr Leu Gln Ile Ala Ala Glu Leu Gly Leu
 25 540 545 550
 ctt gtt caa gaa gaa gct ttt gat act tgg tat gga ggt aag aaa cca
 1734
 Leu Val Gln Glu Glu Ala Phe Asp Thr Trp Tyr Gly Gly Lys Lys Pro
 30 555 560 565
 tac gac tat ggt cgt ttc ttt gaa aaa gat gct acc cac cct gaa gca
 1782
 Tyr Asp Tyr Gly Arg Phe Phe Glu Lys Asp Ala Thr His Pro Glu Ala
 35 570 575 580
 aga aaa ggt gaa aaa tgg tcc gac tac gat ctt cga aca atg gta gaa
 1830
 Arg Lys Gly Glu Lys Trp Ser Asp Tyr Asp Leu Arg Thr Met Val Glu
 40 585 590 595 600
 aga gat aaa aac aat cct gct gtt gtt atg tgg tct att ggt aac gaa
 1878
 Arg Asp Lys Asn Asn Pro Ala Val Val Met Trp Ser Ile Gly Asn Glu
 45 605 610 615
 atc ggt gaa gca gat ggg aaa gcc cat tca ctg gta act gtc aaa cgt
 1926
 Ile Gly Glu Ala Asp Gly Lys Ala His Ser Leu Val Thr Val Lys Arg
 50 620 625 630

5 ¹⁹⁷⁴ttg gtc aaa gtg gtg aaa tct gtt gat aag aca cgt tat gta acc atg
 Leu Val Lys Val Val Lys Ser Val Asp Lys Thr Arg Tyr Val Thr Met
 635 640 645

10 ²⁰²²gga gct gat aaa ttc cgt ttc gga gat gga act ggg gac cat gaa aag
 Gly Ala Asp Lys Phe Arg Phe Gly Asp Gly Thr Gly Asp His Glu Lys
 650 655 660

15 ²⁰⁷⁰att gca gat gaa ctg gat gca gtt gga ttg aat tac tct gaa gaa aat
 Ile Ala Asp Glu Leu Asp Ala Val Gly Leu Asn Tyr Ser Glu Glu Asn
 20 665 670 675 680

25 ²¹¹⁸tat caa act ctt cat gcg aaa cat ccg aaa tgg ctc atc tat ggt tca
 Tyr Gln Thr Leu His Ala Lys His Pro Lys Trp Leu Ile Tyr Gly Ser
 685 690 695

30 ²¹⁶⁶gaa aca tct tca gca act cgt aca aga gga agt tat ttc cat cca gag
 Glu Thr Ser Ser Ala Thr Arg Thr Arg Gly Ser Tyr Phe His Pro Glu
 700 705 710

35 ²²¹⁴agt gag tgg gta gga agt aac caa tat tgg cgt aat tac gaa caa tcc
 Ser Glu Trp Val Gly Ser Asn Gln Tyr Trp Arg Asn Tyr Glu Gln Ser
 40 715 720 725

45 ²²⁶²gat tat ggt aat gat cgt gtt ggt tgg ggt aaa act gct aca gca tca
 Asp Tyr Gly Asn Asp Arg Val Gly Trp Gly Lys Thr Ala Thr Ala Ser
 730 735 740

50 ²³¹⁰tgg acc ttc gat aga gat cat cca gga tac gct gga caa ttc att tgg
 Trp Thr Phe Asp Arg Asp His Pro Gly Tyr Ala Gly Gln Phe Ile Trp
 55 745 750 755 760

60 ²³⁵⁸acc ggt acg gat tat att ggt gaa cca aca cca tgg cac aat caa aat
 Thr Gly Thr Asp Tyr Ile Gly Glu Pro Thr Pro Trp His Asn Gln Asn
 765 770 775

65 agt aca cca gta aaa agt tcc tac ttt ggt att gtt gat acc gca ggt

ES 2 352 704 B1

2406
Ser Thr Pro Val Lys Ser Ser Tyr Phe Gly Ile Val Asp Thr Ala Gly
780 785 790

5

att cca aaa aat gac tat tat ctc tac caa agt cag tgg gtt tct gcc
2454
Ile Pro Lys Asn Asp Tyr Tyr Leu Tyr Gln Ser Gln Trp Val Ser Ala
795 800 805

10

aag aaa aaa cca atg gtt cat tta ctt cct cac tgg aac tgg gaa aag
2502
Lys Lys Lys Pro Met Val His Leu Leu Pro His Trp Asn Trp Glu Lys
810 815 820

15

act gaa ttg gca gat aat gtt gca gac gct gaa aat cga att cca gtt
2550
Thr Glu Leu Ala Asp Asn Val Ala Asp Ala Glu Asn Arg Ile Pro Val
825 830 835 840

20

aga gct tat tcc aat gct gcc agt gtt gag ttg ttc tta aat aat gag
2598
Arg Ala Tyr Ser Asn Ala Ala Ser Val Glu Leu Phe Leu Asn Asn Glu
845 850 855

25

tca cta ggt ctt aaa aaa ttc aac aaa aaa caa aca agt gac ggt cga
2646
Ser Leu Gly Leu Lys Lys Phe Asn Lys Lys Gln Thr Ser Asp Gly Arg
860 865 870

30

aca tat caa gaa ggt gaa aat cct caa gaa cta tac ttg gag tgg aaa
2694
Thr Tyr Gln Glu Gly Glu Asn Pro Gln Glu Leu Tyr Leu Glu Trp Lys
875 880 885

35

gta gct tat aaa ccg gga act ttg gaa gct gta gct cga gat gag tct
2742
Val Ala Tyr Lys Pro Gly Thr Leu Glu Ala Val Ala Arg Asp Glu Ser
890 895 900

40

ggc aag gaa att gct cgt gat aag att gta act gct ggt gaa cct gca
2790
Gly Lys Glu Ile Ala Arg Asp Lys Ile Val Thr Ala Gly Glu Pro Ala
905 910 915 920

45

ggg gtc cgt ttg gtt aag gaa gaa aat gca atc gct gca gat gga aaa
2838
Gly Val Arg Leu Val Lys Glu Glu Asn Ala Ile Ala Ala Asp Gly Lys

50

55

60

65

ES 2 352 704 B1

	925	930	935
5	gat ttg act tat atc tat tac gaa att gta gat agc aag ggt aat gtt 2886 Asp Leu Thr Tyr Ile Tyr Tyr Glu Ile Val Asp Ser Lys Gly Asn Val		
	940	945	950
10	gtg cca act gct aac aat cta gtt cac ttc caa ctt cat gga caa ggt 2934 Val Pro Thr Ala Asn Asn Leu Val His Phe Gln Leu His Gly Gln Gly		
15	955	960	965
20	caa ctg gtt ggt gtg gat aat ggt gag caa gct agt cgt gaa cgt tat 2982 Gln Leu Val Gly Val Asp Asn Gly Glu Gln Ala Ser Arg Glu Arg Tyr		
	970	975	980
25	aaa gaa caa gca gat ggt tct tgg att cgt aga gcc ttt aat ggt aaa 3030 Lys Glu Gln Ala Asp Gly Ser Trp Ile Arg Arg Ala Phe Asn Gly Lys		
30	985	990	995 1000
35	ggg gta gcg att gtt aaa tcg act gat caa gca ggt aaa ttt acc 3075 Gly Val Ala Ile Val Lys Ser Thr Asp Gln Ala Gly Lys Phe Thr		
	1005	1010	1015
40	ctt aca gca tat tca ggc ttg ttg aaa tca gat caa gtt aca gtt 3120 Leu Thr Ala Tyr Ser Gly Leu Leu Lys Ser Asp Gln Val Thr Val		
	1020	1025	1030
45	ttt aca ggt aag tca gac cag tca gaa aaa act gtt tta gga aca 3165 Phe Thr Gly Lys Ser Asp Gln Ser Glu Lys Thr Val Leu Gly Thr		
50	1035	1040	1045
55	gaa ata cca aaa gtt cgg aca gtt tta gga caa aaa ctt aaa atg 3210 Glu Ile Pro Lys Val Arg Thr Val Leu Gly Gln Lys Leu Lys Met		
	1050	1055	1060
60	cct cag aca gtg cca ttt gta tat agt gat ggt agt cgt gaa aaa 3255 Pro Gln Thr Val Pro Phe Val Tyr Ser Asp Gly Ser Arg Glu Lys		
65	1065	1070	1075

5 cgt cct gta acc tgg tct caa gca gat gtc agt caa gct gga att
 3300
 Arg Pro Val Thr Trp Ser Gln Ala Asp Val Ser Gln Ala Gly Ile
 1080 1085 1090

10 gta act gta aaa ggt atg tct gat gga cgt gaa gtt gag gct cgt
 3345
 Val Thr Val Lys Gly Met Ser Asp Gly Arg Glu Val Glu Ala Arg
 1095 1100 1105

15 gtt gaa gta ctt aca gtt gcg aat gaa ctt cca gct gtt aaa cgc
 3390
 Val Glu Val Leu Thr Val Ala Asn Glu Leu Pro Ala Val Lys Arg
 20 1110 1115 1120

25 atc aca cct gga aca gac ttg agt tct gtg gat aaa ctt gtt tca
 3435
 Ile Thr Pro Gly Thr Asp Leu Ser Ser Val Asp Lys Leu Val Ser
 1125 1130 1135

30 ctt gtt tca aca gat ggt cgt att cac aat tat gaa gtt gaa aat
 3480
 Leu Val Ser Thr Asp Gly Arg Ile His Asn Tyr Glu Val Glu Asn
 1140 1145 1150

35 tgg gaa att tca cca gaa gat aaa gag aag tta tct act ccg ggg
 3525
 Trp Glu Ile Ser Pro Glu Asp Lys Glu Lys Leu Ser Thr Pro Gly
 40 1155 1160 1165

45 gca cat att cag atg act agc caa tta gga gat aga acc atc cac
 3570
 Ala His Ile Gln Met Thr Ser Gln Leu Gly Asp Arg Thr Ile His
 1170 1175 1180

50 gct act ttg ata gtt gat gat ggt aag gaa ttg tct caa gtt aca
 3615
 Ala Thr Leu Ile Val Asp Asp Gly Lys Glu Leu Ser Gln Val Thr
 55 1185 1190 1195

60 cca aat ata gca gta ggt ggt gaa agt gtt acg gaa ctt tca aaa
 3660
 Pro Asn Ile Ala Val Gly Gly Glu Ser Val Thr Glu Leu Ser Lys
 1200 1205 1210

65 gag cat cct gta cat tat cac aaa ctt tct tat gga gca aaa att
 3705
 Glu His Pro Val His Tyr His Lys Leu Ser Tyr Gly Ala Lys Ile

ES 2 352 704 B1

		1215		1220		1225
5	cca gaa gtc agt gct 3750 Pro Glu Val Ser Ala	agt gct gaa aat gcc Ser Ala Glu Asn Ala	gat gtg aac gtg att Asp Val Asn Val Ile			
		1230		1235		1240
10	cag gca aat gaa tcc 3795 Gln Ala Asn Glu Ser	aat gga atg cgt gct Asn Gly Met Arg Ala	agt atc tac att cag Ser Ile Tyr Ile Gln			
15		1245		1250		1255
20	cct aaa gac ggg gga 3840 Pro Lys Asp Gly Gly	cag tta caa acc tat Gln Leu Gln Thr Tyr	gct gtt cag ttc ctt Ala Val Gln Phe Leu			
		1260		1265		1270
25	gtt gat tct cca caa 3885 Val Asp Ser Pro Gln	att gaa agc ttg agt Ile Glu Ser Leu Ser	ctc aga gtg gat caa Leu Arg Val Asp Gln			
30		1275		1280		1285
35	gta tct agc ctc aaa 3930 Val Ser Ser Leu Lys	gaa gat caa act gtt Glu Asp Gln Thr Val	aaa gtg aca gtt tta Lys Val Thr Val Leu			
		1290		1295		1300
40	gcc cgt tat cag gat 3975 Ala Arg Tyr Gln Asp	gga act gaa gct att Gly Thr Glu Ala Ile	tta cca act gat aag Leu Pro Thr Asp Lys			
		1305		1310		1315
45	gtg act ttc tca agc 4020 Val Thr Phe Ser Ser	caa ggc gag gga gga Gln Gly Glu Gly Gly	gtc tct gtt aat aag Val Ser Val Asn Lys			
50		1320		1325		1330
55	gga atg cta gag ttg 4065 Gly Met Leu Glu Leu	cat aaa ccg ggt agt His Lys Pro Gly Ser	ctc act tta caa gct Leu Thr Leu Gln Ala			
		1335		1340		1345
60	caa tat gaa ggt gca 4110 Gln Tyr Glu Gly Ala	aaa gga gaa ctt cct Lys Gly Glu Leu Pro	cta acg atc gca gca Leu Thr Ile Ala Ala			
65		1350		1355		1360

ES 2 352 704 B1

5	aat act gaa gat aaa 4155 Asn Thr Glu Asp Lys	gtt gtt cag gcc att Val Val Gln Ala Ile	cgc cca gtt agc ctt Arg Pro Val Ser Leu
	1365	1370	1375
10	gtg aca gat ttg cac 4200 Val Thr Asp Leu His	caa aca cca aat ctt Gln Thr Pro Asn Leu	cca tct aca gta aca Pro Ser Thr Val Thr
	1380	1385	1390
15	gtt gag tac gat aaa 4245 Val Glu Tyr Asp Lys	gga ttt cca aaa gtt Gly Phe Pro Lys Val	cat aag gta gtt tgg His Lys Val Val Trp
20	1395	1400	1405
25	gat tcg att cca tct 4290 Asp Ser Ile Pro Ser	gaa aaa tta aat caa Glu Lys Leu Asn Gln	tat cat act ttt gaa Tyr His Thr Phe Glu
	1410	1415	1420
30	gta ctt ggt aag gtt 4335 Val Leu Gly Lys Val	gaa gga ttg aat caa Glu Gly Leu Asn Gln	gaa gct cgt gcg aca Glu Ala Arg Ala Thr
	1425	1430	1435
35	gta tca gta gaa gga 4380 Val Ser Val Glu Gly	att gta aaa gtt gaa Ile Val Lys Val Glu	gat gtg agc atc aca Asp Val Ser Ile Thr
40	1440	1445	1450
45	aca cca att tct gag 4425 Thr Pro Ile Ser Glu	gca cca acc ttg cca Ala Pro Thr Leu Pro	gaa agt gtc cga act Glu Ser Val Arg Thr
	1455	1460	1465
50	tat tat tca aat ggc 4470 Tyr Tyr Ser Asn Gly	caa gta tct tca gct Gln Val Ser Ser Ala	aag gtg gct tgg gac Lys Val Ala Trp Asp
55	1470	1475	1480
60	act att gat gca agc 4515 Thr Ile Asp Ala Ser	cag tat gca aaa gaa Gln Tyr Ala Lys Glu	gga atc ttt act gta Gly Ile Phe Thr Val
	1485	1490	1495
65	aca ggc cat gta gaa 4560	gga aca caa ctg aca	acg aaa tta cat gtt

ES 2 352 704 B1

	Thr Gly His Val Glu	Gly Thr Gln Leu Thr	Thr Lys Leu His Val
	1500	1505	1510
5	cgt gta tcg agt caa	aca gaa att ggg aag	aat att tca gat caa
	4605		
	Arg Val Ser Ser Gln	Thr Glu Ile Gly Lys	Asn Ile Ser Asp Gln
10	1515	1520	1525
	tgg aca ggt tca gaa	tta cca ctt gcc ttt	gca tca gat tct cat
	4650		
15	Trp Thr Gly Ser Glu	Leu Pro Leu Ala Phe	Ala Ser Asp Ser His
	1530	1535	1540
20	tct tat gac cca gta	tct aat gta aat gat	aaa gtg att tct tac
	4695		
	Ser Tyr Asp Pro Val	Ser Asn Val Asn Asp	Lys Val Ile Ser Tyr
25	1545	1550	1555
	ggg gat cag cct gca	aat cgt tgg tca aac	tgg aaa cgt tct gaa
	4740		
30	Gly Asp Gln Pro Ala	Asn Arg Trp Ser Asn	Trp Lys Arg Ser Glu
	1560	1565	1570
35	gaa gct tca gta ggg	gta ctc ttt gga gat	tca ggt att ttg aca
	4785		
	Glu Ala Ser Val Gly	Val Leu Phe Gly Asp	Ser Gly Ile Leu Thr
	1575	1580	1585
40	aaa cgt tcg gtc gat	aac ttg aat gta gca	ttt tat gaa gat agc
	4830		
	Lys Arg Ser Val Asp	Asn Leu Asn Val Ala	Phe Tyr Glu Asp Ser
45	1590	1595	1600
	ggg gta ggt gct cca	aaa tct tac gtc att	gag tat tat gtt ggt
	4875		
50	Gly Val Gly Ala Pro	Lys Ser Tyr Val Ile	Glu Tyr Tyr Val Gly
	1605	1610	1615
55	gaa aaa gct cca act	gca ccg aag aat cct	ggg tat gta gag ggt
	4920		
	Glu Lys Ala Pro Thr	Ala Pro Lys Asn Pro	Gly Tyr Val Glu Gly
	1620	1625	1630
60	gaa aac cac gta ttt	aat gac agt aaa aat	tgg aaa cca gtc aca
	4965		
	Glu Asn His Val Phe	Asn Asp Ser Lys Asn	Trp Lys Pro Val Thr
65	1635	1640	1645

ES 2 352 704 B1

	aat ctt aag gct cca gat cag tta aaa gct gga gaa atg aat cat	
	5010	
	Asn Leu Lys Ala Pro Asp Gln Leu Lys Ala Gly Glu Met Asn His	
5	1650	1655 1660
	ttc agt ttt gat aag gtg gat acc tac gcg gta cgt att cgc atg	
	5055	
	Phe Ser Phe Asp Lys Val Asp Thr Tyr Ala Val Arg Ile Arg Met	
10	1665	1670 1675
	aca aga gct gat gat aaa ctt gga aca tcc att aca gaa tta caa	
	5100	
	Thr Arg Ala Asp Asp Lys Leu Gly Thr Ser Ile Thr Glu Leu Gln	
15	1680	1685 1690
	gtt ttc tct aaa aag gta gca cca gct aaa gag gca act act aac	
	5145	
	Val Phe Ser Lys Lys Val Ala Pro Ala Lys Glu Ala Thr Thr Asn	
20	1695	1700 1705
	att aag gtt gct gga aaa gat ctt cct cac ttt aat cca gac tta	
	5190	
	Ile Lys Val Ala Gly Lys Asp Leu Pro His Phe Asn Pro Asp Leu	
25	1710	1715 1720
	aca gat tac tat tta gac tat agt gat gga aaa att cca gaa gta	
	5235	
	Thr Asp Tyr Tyr Leu Asp Tyr Ser Asp Gly Lys Ile Pro Glu Val	
30	1725	1730 1735
	aca gct agt atc aca aat aat gga ctt gct aca gtg gtt cct agc	
	5280	
	Thr Ala Ser Ile Thr Asn Asn Gly Leu Ala Thr Val Val Pro Ser	
35	1740	1745 1750
	gtg aaa gag ggc gat cct gtc cgt gtt att gtg aaa gca gaa aat	
	5325	
	Val Lys Glu Gly Asp Pro Val Arg Val Ile Val Lys Ala Glu Asn	
40	1755	1760 1765
	ggc gat att ctt gga gaa tac aac ctt cac ttt aca aaa gat aaa	
	5370	
	Gly Asp Ile Leu Gly Glu Tyr Asn Leu His Phe Thr Lys Asp Lys	
45	1770	1775 1780
	gat ttg ttg gct cgt aaa cca att gct tct gca aaa cag tct caa	
50		
55		
60		
65		

ES 2 352 704 B1

5415
Asp Leu Leu Ala Arg Lys Pro Ile Ala Ser Ala Lys Gln Ser Gln
1785 1790 1795

5

ttg ttg cag tta ggt caa cct ttg gaa ttg cca aat aag gtt cca
5460
Leu Leu Gln Leu Gly Gln Pro Leu Glu Leu Pro Asn Lys Val Pro
1800 1805 1810

10

gtt tat ttc gca ggt aag agc ggt tat gac gta aaa gat ttg gca
5505
Val Tyr Phe Ala Gly Lys Ser Gly Tyr Asp Val Lys Asp Leu Ala
1815 1820 1825

15

gtt gaa tgg gaa aaa gtc cct gct tct acc tta aac aaa gct ggg
5550
Val Glu Trp Glu Lys Val Pro Ala Ser Thr Leu Asn Lys Ala Gly
1830 1835 1840

20

gaa tat aca att aaa ggc cat gta gtt ggt agt gag gtc cca gta
5595
Glu Tyr Thr Ile Lys Gly His Val Val Gly Ser Glu Val Pro Val
1845 1850 1855

25

gaa tta tct gtt cga gta acg gac aag gtc ggt aga tca ctt tct
5640
Glu Leu Ser Val Arg Val Thr Asp Lys Val Gly Arg Ser Leu Ser
1860 1865 1870

30

gat aat cca gat tat ctt gct act cat agt cag gct ttt gct tca
5685
Asp Asn Pro Asp Tyr Leu Ala Thr His Ser Gln Ala Phe Ala Ser
1875 1880 1885

35

gct act aat gat ttg gat gat aat tcg aac gac cgt gta gac tat
5730
Ala Thr Asn Asp Leu Asp Asp Asn Ser Asn Asp Arg Val Asp Tyr
1890 1895 1900

40

tta aat gat aat gac cga aat caa gac aac cgt tgg aca aac tgg
5775
Leu Asn Asp Asn Asp Arg Asn Gln Asp Asn Arg Trp Thr Asn Trp
1905 1910 1915

45

tcc gct act cca tct gct aat cca gaa gta tca gtt ggt gtt atc
5820
Ser Ala Thr Pro Ser Ala Asn Pro Glu Val Ser Val Gly Val Ile

50

55

60

65

ES 2 352 704 B1

	1920	1925	1930
5	ttt aaa aag aat ggt 5865 Phe Lys Lys Asn Gly	aaa att gtg gat aga Lys Ile Val Asp Arg	aag gtt tca caa gct Lys Val Ser Gln Ala
	1935	1940	1945
10	aaa cta aac ttc ttt 5910 Lys Leu Asn Phe Phe	gca gat agt gga aca Ala Asp Ser Gly Thr	gat gcc cca gct aaa Asp Ala Pro Ala Lys
15	1950	1955	1960
20	ctc gta ttg gaa cga 5955 Leu Val Leu Glu Arg	tac atc ggt cct gac Tyr Ile Gly Pro Asp	ttt gat gtt ccg gtt Phe Asp Val Pro Val
	1965	1970	1975
25	tac tat tca aac tac 6000 Tyr Tyr Ser Asn Tyr	agc tac gaa tca aat Ser Tyr Glu Ser Asn	cat cca ttt aat aac His Pro Phe Asn Asn
30	1980	1985	1990
35	cct gat aat tgg gaa 6045 Pro Asp Asn Trp Glu	ctt gtt cct tat tac Leu Val Pro Tyr Tyr	gca gac aaa gaa att Ala Asp Lys Glu Ile
	1995	2000	2005
40	caa gct gga aat gag 6090 Gln Ala Gly Asn Glu	atc aat gtg acc ttt Ile Asn Val Thr Phe	aaa gct gtg act acc Lys Ala Val Thr Thr
	2010	2015	2020
45	aaa gct atg aga tgg 6135 Lys Ala Met Arg Trp	cgt atg gac aga aaa Arg Met Asp Arg Lys	gca gat aaa aat ggg Ala Asp Lys Asn Gly
50	2025	2030	2035
55	gtt gca cta act gaa 6180 Val Ala Leu Thr Glu	tta agt ttc cga gca Leu Ser Phe Arg Ala	cca agt gag ctt gca Pro Ser Glu Leu Ala
	2040	2045	2050
60	aaa gag agc aca gag 6225 Lys Glu Ser Thr Glu	gga aga att tta gta Gly Arg Ile Leu Val	gat ggt aaa gag ctt Asp Gly Lys Glu Leu
65	2055	2060	2065

ES 2 352 704 B1

	cct gac ttc tct gaa	aat cgt ttg gac tat	gaa ttg att tat aaa
	6270		
	Pro Asp Phe Ser Glu	Asn Arg Leu Asp Tyr	Glu Leu Ile Tyr Lys
5	2070	2075	2080
	ggg gaa cgt cca aaa	atc act gta gaa ggg	aaa gat caa gtc aca
	6315		
10	Gly Glu Arg Pro Lys	Ile Thr Val Glu Gly	Lys Asp Gln Val Thr
	2085	2090	2095
	tct aca att gtg gat	agt gga aat gaa aat	ctt cct gta tta gtc
	6360		
	Ser Thr Ile Val Asp	Ser Gly Asn Glu Asn	Leu Pro Val Leu Val
15	2100	2105	2110
	cgt ctg gta tca gag	agt ggt aag aat gtc	aaa gaa tac cgg att
	6405		
25	Arg Leu Val Ser Glu	Ser Gly Lys Asn Val	Lys Glu Tyr Arg Ile
	2115	2120	2125
	aaa ttg act aag gaa	aaa gtg ata gac gca	aaa ctt gta gct cct
	6450		
30	Lys Leu Thr Lys Glu	Lys Val Ile Asp Ala	Lys Leu Val Ala Pro
	2130	2135	2140
	gct caa tat gac ttg	cct agt ctt gaa gtg	gta gac aaa gaa ttg
	6495		
	Ala Gln Tyr Asp Leu	Pro Ser Leu Glu Val	Val Asp Lys Glu Leu
35	2145	2150	2155
	aac ttc caa act ctt	gaa caa aaa gac gat	act ctg ttt gag gga
	6540		
45	Asn Phe Gln Thr Leu	Glu Gln Lys Asp Asp	Thr Leu Phe Glu Gly
	2160	2165	2170
	gaa gtt cgc ctg ttg	caa gaa ggt aaa gtt	ggg aag gaa cgt atc
	6585		
50	Glu Val Arg Leu Leu	Gln Glu Gly Lys Val	Gly Lys Glu Arg Ile
	2175	2180	2185
	tat aca gaa gta acg	act gat ggc aag aat	aaa gaa aaa tca cgt
	6630		
	Tyr Thr Glu Val Thr	Thr Asp Gly Lys Asn	Lys Glu Lys Ser Arg
55	2190	2195	2200
	gag att ctg gaa gaa	cct gta aat aga att	ctg ttg gta ggt act
	6675		
60	Glu Ile Leu Glu Glu	Pro Val Asn Arg Ile	Leu Leu Val Gly Thr
65			

ES 2 352 704 B1

		2205		2210		2215
5	aag aag aaa aca agt 6720 Lys Lys Lys Thr Ser	aac att gag acg ccg Asn Ile Glu Thr Pro	gcg ggt gga cat caa Ala Gly Gly His Gln			
		2220		2225		2230
10	gac aat ggt agc agt 6765 Asp Asn Gly Ser Ser	caa cct acg gaa caa Gln Pro Thr Glu Gln	aaa cca gga gca aat Lys Pro Gly Ala Asn			
15		2235		2240		2245
20	ccg aat cca aat att 6810 Pro Asn Pro Asn Ile	gtc att cca aat cca Val Ile Pro Asn Pro	tca aat aat ggt gtg Ser Asn Asn Gly Val			
		2250		2255		2260
25	aca agt aac agt tca 6855 Thr Ser Asn Ser Ser	gat ttt aaa aat cat Asp Phe Lys Asn His	cta tca gtt gct ccg Leu Ser Val Ala Pro			
30		2265		2270		2275
35	aac att gtg ata caa 6900 Asn Ile Val Ile Gln	gct cct act cag gaa Ala Pro Thr Gln Glu	agc aaa aaa cct gga Ser Lys Lys Pro Gly			
		2280		2285		2290
40	tgg aac aag gaa gat 6945 Trp Asn Lys Glu Asp	ggg aaa tgg tat tat Gly Lys Trp Tyr Tyr	cgt caa caa aat ggg Arg Gln Gln Asn Gly			
		2295		2300		2305
45	gac ttg gct aag ggt 6990 Asp Leu Ala Lys Gly	tgg gtc aaa gat gga Trp Val Lys Asp Gly	gat act tgg tat tac Asp Thr Trp Tyr Tyr			
50		2310		2315		2320
55	ttt gat caa aca gga 7035 Phe Asp Gln Thr Gly	aaa atg gaa act ggc Lys Met Glu Thr Gly	tgg att aaa gat caa Trp Ile Lys Asp Gln			
		2325		2330		2335
60	tca ggt gca tgg tat 7080 Ser Gly Ala Trp Tyr	tac ttg aat caa tca Tyr Leu Asn Gln Ser	ggg gca atg tca tct Gly Ala Met Ser Ser			
65		2340		2345		2350

ES 2 352 704 B1

```

    aat gag tgg att ttg  gat caa gat ggt aag  tgg tac tat gtt gat
    7125
    Asn Glu Trp Ile Leu  Asp Gln Asp Gly Lys  Trp Tyr Tyr Val Asp
5      2355                      2360                      2365

    gca tca ggt tcg atg  aag aca agt caa tgg  ttc caa gtt gga gac
    7170
    Ala Ser Gly Ser Met  Lys Thr Ser Gln Trp  Phe Gln Val Gly Asp
10     2370                      2375                      2380

    aaa tgg tat tat gtt  gat gca aat gga tcc  ctt gct gtc aat aca
    7215
    Lys Trp Tyr Tyr Val  Asp Ala Asn Gly Ser  Leu Ala Val Asn Thr
15     2385                      2390                      2395

    gta act ccg gat ggt  tat caa gta aat gaa  aac ggt gag tgg atc
    7260
    Val Thr Pro Asp Gly  Tyr Gln Val Asn Glu  Asn Gly Glu Trp Ile
25     2400                      2405                      2410

    gga taa
    7266
    Gly

30

35 <210> 2
    <211> 2411
    <212> PRT
    <213> Streptococcus mitis
40

    <400> 2

45     Met Gln Lys Gly Asn Trp Asn Lys Lys Arg Val Tyr Ser Ile Arg Lys
        1          5          10          15

    Phe Ser Val Gly Ala Cys Ser Val Leu Ile Gly Thr Cys Ala Leu Leu
        20          25          30

50     Leu Gly Ala Gly Val Ser Leu Ser Ser Ser Val Tyr Ala Asn Glu Asn
        35          40          45

55     Ala Glu Glu Ile Val Leu Thr Asn Arg Thr Glu His Gln Ser Glu Asn
        50          55          60

60     Leu Lys Thr Thr Gly Leu Ser Ser Glu His Lys Ala Glu Gly Val Thr
        65          70          75          80

65     Thr Glu Lys Ala Ser Glu Gln Pro Val Lys Ala Glu Asn Val Ser Leu
        85          90          95

```

ES 2 352 704 B1

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

ES 2 352 704 B1

Leu Gln Val Thr Asp Lys Ile His Val Glu Lys Asn Gly Thr Thr Ile
 355 360 365
 5 Leu Thr Pro Lys Leu Glu Ser Gln Gln Gly Gly Lys Val Glu Thr Gln
 370 375 380
 10 Val Ser Ser Lys Ile Val Asn Thr Asp Asn Lys Asp His Glu Ile Val
 385 390 395 400
 15 Ala Glu Tyr Gln Ile Phe Glu Arg Gly Gly Lys Ala Val Thr Glu Leu
 405 410 415
 20 Val Arg Thr Glu Ser Lys Thr Leu Lys Ala Lys Glu Thr Ile His Leu
 420 425 430
 25 Asp Ser Ala Leu Glu Val Glu Lys Pro Lys Leu Trp Thr Val Ser Ser
 435 440 445
 30 Asp Lys Pro Ala Leu Tyr Glu Met Val Thr Arg Val Tyr Lys Asp Gly
 450 455 460
 35 Gln Leu Val Asp Ala Lys Lys Asp Leu Phe Gly Tyr Arg Tyr Tyr Asn
 465 470 475 480
 40 Trp Thr Pro Asp Gln Gly Phe Ser Leu Asn Gly Glu His Ile Lys Phe
 485 490 495
 45 His Gly Val Ser Leu His His Asp His Gly Ala Leu Gly Ala Glu Glu
 500 505 510
 50 Asn Tyr Lys Ala Glu Tyr Arg Arg Leu Lys Gln Met Lys Glu Met Gly
 515 520 525
 55 Val Asn Ser Ile Arg Thr Thr His Asn Pro Ala Ser Pro Gln Thr Leu
 530 535 540
 60 Gln Ile Ala Ala Glu Leu Gly Leu Leu Val Gln Glu Glu Ala Phe Asp
 545 550 555 560
 65 Thr Trp Tyr Gly Gly Lys Lys Pro Tyr Asp Tyr Gly Arg Phe Phe Glu
 565 570 575
 70 Lys Asp Ala Thr His Pro Glu Ala Arg Lys Gly Glu Lys Trp Ser Asp
 580 585 590
 75 Tyr Asp Leu Arg Thr Met Val Glu Arg Asp Lys Asn Asn Pro Ala Val
 595 600 605
 80 Val Met Trp Ser Ile Gly Asn Glu Ile Gly Glu Ala Asp Gly Lys Ala
 610 615 620

ES 2 352 704 B1

His Ser Leu Val Thr Val Lys Arg Leu Val Lys Val Val Lys Ser Val
 625 630 635 640
 5 Asp Lys Thr Arg Tyr Val Thr Met Gly Ala Asp Lys Phe Arg Phe Gly
 645 650 655
 10 Asp Gly Thr Gly Asp His Glu Lys Ile Ala Asp Glu Leu Asp Ala Val
 660 665 670
 15 Gly Leu Asn Tyr Ser Glu Glu Asn Tyr Gln Thr Leu His Ala Lys His
 675 680 685
 20 Pro Lys Trp Leu Ile Tyr Gly Ser Glu Thr Ser Ser Ala Thr Arg Thr
 690 695 700
 25 Arg Gly Ser Tyr Phe His Pro Glu Ser Glu Trp Val Gly Ser Asn Gln
 705 710 715 720
 30 Tyr Trp Arg Asn Tyr Glu Gln Ser Asp Tyr Gly Asn Asp Arg Val Gly
 725 730 735
 35 Trp Gly Lys Thr Ala Thr Ala Ser Trp Thr Phe Asp Arg Asp His Pro
 740 745 750
 40 Gly Tyr Ala Gly Gln Phe Ile Trp Thr Gly Thr Asp Tyr Ile Gly Glu
 755 760 765
 45 Pro Thr Pro Trp His Asn Gln Asn Ser Thr Pro Val Lys Ser Ser Tyr
 770 775 780
 50 Phe Gly Ile Val Asp Thr Ala Gly Ile Pro Lys Asn Asp Tyr Tyr Leu
 785 790 795 800
 55 Tyr Gln Ser Gln Trp Val Ser Ala Lys Lys Lys Pro Met Val His Leu
 805 810 815
 60 Leu Pro His Trp Asn Trp Glu Lys Thr Glu Leu Ala Asp Asn Val Ala
 820 825 830
 65 Asp Ala Glu Asn Arg Ile Pro Val Arg Ala Tyr Ser Asn Ala Ala Ser
 835 840 845
 Val Glu Leu Phe Leu Asn Asn Glu Ser Leu Gly Leu Lys Lys Phe Asn
 850 855 860
 Lys Lys Gln Thr Ser Asp Gly Arg Thr Tyr Gln Glu Gly Glu Asn Pro
 865 870 875 880

ES 2 352 704 B1

Gln Glu Leu Tyr Leu₈₈₅ Glu Trp Lys Val Ala₈₉₀ Tyr Lys Pro Gly Thr₈₉₅ Leu
 5 Glu Ala Val Ala Arg Asp Glu Ser Gly₉₀₅ Lys Glu Ile Ala Arg₉₁₀ Asp Lys
 10 Ile Val Thr₉₁₅ Ala Gly Glu Pro Ala₉₂₀ Gly Val Arg Leu Val₉₂₅ Lys Glu Glu
 15 Asn Ala₉₃₀ Ile Ala Ala Asp Gly₉₃₅ Lys Asp Leu Thr Tyr₉₄₀ Ile Tyr Tyr Glu
 Ile Val Asp Ser Lys Gly₉₅₀ Asn Val Val Pro Thr₉₅₅ Ala Asn Asn Leu Val₉₆₀
 20 His Phe Gln Leu His₉₆₅ Gly Gln Gly Gln Leu Val Gly Val Asp Asn₉₇₅ Gly
 25 Glu Gln Ala Ser Arg Glu Arg Tyr Lys₉₈₅ Glu Gln Ala Asp Gly₉₉₀ Ser Trp
 30 Ile Arg Arg₉₉₅ Ala Phe Asn Gly Lys₁₀₀₀ Gly Val Ala Ile Val₁₀₀₅ Lys Ser Thr
 35 Asp Gln₁₀₁₀ Ala Gly Lys Phe Thr₁₀₁₅ Leu Thr Ala Tyr Ser₁₀₂₀ Gly Leu Leu
 40 Lys Ser₁₀₂₅ Asp Gln Val Thr Val₁₀₃₀ Phe Thr Gly Lys Ser₁₀₃₅ Asp Gln Ser
 45 Glu Lys₁₀₄₀ Thr Val Leu Gly Thr₁₀₄₅ Glu Ile Pro Lys Val₁₀₅₀ Arg Thr Val
 50 Leu Gly₁₀₅₅ Gln Lys Leu Lys Met₁₀₆₀ Pro Gln Thr Val Pro₁₀₆₅ Phe Val Tyr
 55 Ser Asp₁₀₇₀ Gly Ser Arg Glu Lys₁₀₇₅ Arg Pro Val Thr Trp₁₀₈₀ Ser Gln Ala
 Asp Val₁₀₈₅ Ser Gln Ala Gly Ile₁₀₉₀ Val Thr Val Lys Gly₁₀₉₅ Met Ser Asp
 60 Gly Arg₁₁₀₀ Glu Val Glu Ala Arg₁₁₀₅ Val Glu Val Leu Thr₁₁₁₀ Val Ala Asn
 65 Glu Leu₁₁₁₅ Pro Ala Val Lys Arg₁₁₂₀ Ile Thr Pro Gly Thr₁₁₂₅ Asp Leu Ser
 Ser Val₁₁₃₀ Asp Lys Leu Val Ser₁₁₃₅ Leu Val Ser Thr Asp₁₁₄₀ Gly Arg Ile

ES 2 352 704 B1

	His	Asn	Tyr	Glu	Val	Glu	Asn	Trp	Glu	Ile	Ser	Pro	Glu	Asp	Lys
		1145					1150					1155			
5	Glu	Lys	Leu	Ser	Thr	Pro	Gly	Ala	His	Ile	Gln	Met	Thr	Ser	Gln
		1160					1165					1170			
10	Leu	Gly	Asp	Arg	Thr	Ile	His	Ala	Thr	Leu	Ile	Val	Asp	Asp	Gly
		1175					1180					1185			
15	Lys	Glu	Leu	Ser	Gln	Val	Thr	Pro	Asn	Ile	Ala	Val	Gly	Gly	Glu
		1190					1195					1200			
20	Ser	Val	Thr	Glu	Leu	Ser	Lys	Glu	His	Pro	Val	His	Tyr	His	Lys
		1205					1210					1215			
25	Leu	Ser	Tyr	Gly	Ala	Lys	Ile	Pro	Glu	Val	Ser	Ala	Ser	Ala	Glu
		1220					1225					1230			
30	Asn	Ala	Asp	Val	Asn	Val	Ile	Gln	Ala	Asn	Glu	Ser	Asn	Gly	Met
		1235					1240					1245			
35	Arg	Ala	Ser	Ile	Tyr	Ile	Gln	Pro	Lys	Asp	Gly	Gly	Gln	Leu	Gln
		1250					1255					1260			
40	Thr	Tyr	Ala	Val	Gln	Phe	Leu	Val	Asp	Ser	Pro	Gln	Ile	Glu	Ser
		1265					1270					1275			
45	Leu	Ser	Leu	Arg	Val	Asp	Gln	Val	Ser	Ser	Leu	Lys	Glu	Asp	Gln
		1280					1285					1290			
50	Thr	Val	Lys	Val	Thr	Val	Leu	Ala	Arg	Tyr	Gln	Asp	Gly	Thr	Glu
		1295					1300					1305			
55	Ala	Ile	Leu	Pro	Thr	Asp	Lys	Val	Thr	Phe	Ser	Ser	Gln	Gly	Glu
		1310					1315					1320			
60	Gly	Gly	Val	Ser	Val	Asn	Lys	Gly	Met	Leu	Glu	Leu	His	Lys	Pro
		1325					1330					1335			
65	Gly	Ser	Leu	Thr	Leu	Gln	Ala	Gln	Tyr	Glu	Gly	Ala	Lys	Gly	Glu
		1340					1345					1350			
	Leu	Pro	Leu	Thr	Ile	Ala	Ala	Asn	Thr	Glu	Asp	Lys	Val	Val	Gln
		1355					1360					1365			
	Ala	Ile	Arg	Pro	Val	Ser	Leu	Val	Thr	Asp	Leu	His	Gln	Thr	Pro
		1370					1375					1380			

ES 2 352 704 B1

	Asn	Leu	Pro	Ser	Thr	Val	Thr	Val	Glu	Tyr	Asp	Lys	Gly	Phe	Pro
		1385					1390					1395			
5	Lys	Val	His	Lys	Val	Val	Trp	Asp	Ser	Ile	Pro	Ser	Glu	Lys	Leu
		1400					1405					1410			
10	Asn	Gln	Tyr	His	Thr	Phe	Glu	Val	Leu	Gly	Lys	Val	Glu	Gly	Leu
		1415					1420					1425			
15	Asn	Gln	Glu	Ala	Arg	Ala	Thr	Val	Ser	Val	Glu	Gly	Ile	Val	Lys
		1430					1435					1440			
	Val	Glu	Asp	Val	Ser	Ile	Thr	Thr	Pro	Ile	Ser	Glu	Ala	Pro	Thr
		1445					1450					1455			
20	Leu	Pro	Glu	Ser	Val	Arg	Thr	Tyr	Tyr	Ser	Asn	Gly	Gln	Val	Ser
		1460					1465					1470			
25	Ser	Ala	Lys	Val	Ala	Trp	Asp	Thr	Ile	Asp	Ala	Ser	Gln	Tyr	Ala
		1475					1480					1485			
30	Lys	Glu	Gly	Ile	Phe	Thr	Val	Thr	Gly	His	Val	Glu	Gly	Thr	Gln
		1490					1495					1500			
35	Leu	Thr	Thr	Lys	Leu	His	Val	Arg	Val	Ser	Ser	Gln	Thr	Glu	Ile
		1505					1510					1515			
40	Gly	Lys	Asn	Ile	Ser	Asp	Gln	Trp	Thr	Gly	Ser	Glu	Leu	Pro	Leu
		1520					1525					1530			
45	Ala	Phe	Ala	Ser	Asp	Ser	His	Ser	Tyr	Asp	Pro	Val	Ser	Asn	Val
		1535					1540					1545			
50	Asn	Asp	Lys	Val	Ile	Ser	Tyr	Gly	Asp	Gln	Pro	Ala	Asn	Arg	Trp
		1550					1555					1560			
55	Ser	Asn	Trp	Lys	Arg	Ser	Glu	Glu	Ala	Ser	Val	Gly	Val	Leu	Phe
		1565					1570					1575			
60	Gly	Asp	Ser	Gly	Ile	Leu	Thr	Lys	Arg	Ser	Val	Asp	Asn	Leu	Asn
		1580					1585					1590			
65	Val	Ala	Phe	Tyr	Glu	Asp	Ser	Gly	Val	Gly	Ala	Pro	Lys	Ser	Tyr
		1595					1600					1605			
70	Val	Ile	Glu	Tyr	Tyr	Val	Gly	Glu	Lys	Ala	Pro	Thr	Ala	Pro	Lys
		1610					1615					1620			
75	Asn	Pro	Gly	Tyr	Val	Glu	Gly	Glu	Asn	His	Val	Phe	Asn	Asp	Ser
		1625					1630					1635			

ES 2 352 704 B1

	Lys	Asn	Trp	Lys	Pro	Val	Thr	Asn	Leu	Lys	Ala	Pro	Asp	Gln	Leu
		1640					1645					1650			
5	Lys	Ala	Gly	Glu	Met	Asn	His	Phe	Ser	Phe	Asp	Lys	Val	Asp	Thr
		1655					1660					1665			
10	Tyr	Ala	Val	Arg	Ile	Arg	Met	Thr	Arg	Ala	Asp	Asp	Lys	Leu	Gly
		1670					1675					1680			
15	Thr	Ser	Ile	Thr	Glu	Leu	Gln	Val	Phe	Ser	Lys	Lys	Val	Ala	Pro
		1685					1690					1695			
20	Ala	Lys	Glu	Ala	Thr	Thr	Asn	Ile	Lys	Val	Ala	Gly	Lys	Asp	Leu
		1700					1705					1710			
25	Pro	His	Phe	Asn	Pro	Asp	Leu	Thr	Asp	Tyr	Tyr	Leu	Asp	Tyr	Ser
		1715					1720					1725			
30	Asp	Gly	Lys	Ile	Pro	Glu	Val	Thr	Ala	Ser	Ile	Thr	Asn	Asn	Gly
		1730					1735					1740			
35	Leu	Ala	Thr	Val	Val	Pro	Ser	Val	Lys	Glu	Gly	Asp	Pro	Val	Arg
		1745					1750					1755			
40	Val	Ile	Val	Lys	Ala	Glu	Asn	Gly	Asp	Ile	Leu	Gly	Glu	Tyr	Asn
		1760					1765					1770			
45	Leu	His	Phe	Thr	Lys	Asp	Lys	Asp	Leu	Leu	Ala	Arg	Lys	Pro	Ile
		1775					1780					1785			
50	Ala	Ser	Ala	Lys	Gln	Ser	Gln	Leu	Leu	Gln	Leu	Gly	Gln	Pro	Leu
		1790					1795					1800			
55	Glu	Leu	Pro	Asn	Lys	Val	Pro	Val	Tyr	Phe	Ala	Gly	Lys	Ser	Gly
		1805					1810					1815			
60	Tyr	Asp	Val	Lys	Asp	Leu	Ala	Val	Glu	Trp	Glu	Lys	Val	Pro	Ala
		1820					1825					1830			
65	Ser	Thr	Leu	Asn	Lys	Ala	Gly	Glu	Tyr	Thr	Ile	Lys	Gly	His	Val
		1835					1840					1845			
	Val	Gly	Ser	Glu	Val	Pro	Val	Glu	Leu	Ser	Val	Arg	Val	Thr	Asp
		1850					1855					1860			
	Lys	Val	Gly	Arg	Ser	Leu	Ser	Asp	Asn	Pro	Asp	Tyr	Leu	Ala	Thr
		1865					1870					1875			

ES 2 352 704 B1

	His	Ser	Gln	Ala	Phe	Ala	Ser	Ala	Thr	Asn	Asp	Leu	Asp	Asp	Asn
		1880					1885					1890			
5	Ser	Asn	Asp	Arg	Val	Asp	Tyr	Leu	Asn	Asp	Asn	Asp	Arg	Asn	Gln
		1895					1900					1905			
10	Asp	Asn	Arg	Trp	Thr	Asn	Trp	Ser	Ala	Thr	Pro	Ser	Ala	Asn	Pro
		1910					1915					1920			
15	Glu	Val	Ser	Val	Gly	Val	Ile	Phe	Lys	Lys	Asn	Gly	Lys	Ile	Val
		1925					1930					1935			
20	Asp	Arg	Lys	Val	Ser	Gln	Ala	Lys	Leu	Asn	Phe	Phe	Ala	Asp	Ser
		1940					1945					1950			
25	Gly	Thr	Asp	Ala	Pro	Ala	Lys	Leu	Val	Leu	Glu	Arg	Tyr	Ile	Gly
		1955					1960					1965			
30	Pro	Asp	Phe	Asp	Val	Pro	Val	Tyr	Tyr	Ser	Asn	Tyr	Ser	Tyr	Glu
		1970					1975					1980			
35	Ser	Asn	His	Pro	Phe	Asn	Asn	Pro	Asp	Asn	Trp	Glu	Leu	Val	Pro
		1985					1990					1995			
40	Tyr	Tyr	Ala	Asp	Lys	Glu	Ile	Gln	Ala	Gly	Asn	Glu	Ile	Asn	Val
		2000					2005					2010			
45	Thr	Phe	Lys	Ala	Val	Thr	Thr	Lys	Ala	Met	Arg	Trp	Arg	Met	Asp
		2015					2020					2025			
50	Arg	Lys	Ala	Asp	Lys	Asn	Gly	Val	Ala	Leu	Thr	Glu	Leu	Ser	Phe
		2030					2035					2040			
55	Arg	Ala	Pro	Ser	Glu	Leu	Ala	Lys	Glu	Ser	Thr	Glu	Gly	Arg	Ile
		2045					2050					2055			
60	Leu	Val	Asp	Gly	Lys	Glu	Leu	Pro	Asp	Phe	Ser	Glu	Asn	Arg	Leu
		2060					2065					2070			
65	Asp	Tyr	Glu	Leu	Ile	Tyr	Lys	Gly	Glu	Arg	Pro	Lys	Ile	Thr	Val
		2075					2080					2085			
70	Glu	Gly	Lys	Asp	Gln	Val	Thr	Ser	Thr	Ile	Val	Asp	Ser	Gly	Asn
		2090					2095					2100			
75	Glu	Asn	Leu	Pro	Val	Leu	Val	Arg	Leu	Val	Ser	Glu	Ser	Gly	Lys
		2105					2110					2115			
80	Asn	Val	Lys	Glu	Tyr	Arg	Ile	Lys	Leu	Thr	Lys	Glu	Lys	Val	Ile
		2120					2125					2130			

ES 2 352 704 B1

	Asp	Ala	Lys	Leu	Val	Ala	Pro	Ala	Gln	Tyr	Asp	Leu	Pro	Ser	Leu
	2135						2140					2145			
5	Glu	Val	Val	Asp	Lys	Glu	Leu	Asn	Phe	Gln	Thr	Leu	Glu	Gln	Lys
	2150						2155					2160			
10	Asp	Asp	Thr	Leu	Phe	Glu	Gly	Glu	Val	Arg	Leu	Leu	Gln	Glu	Gly
	2165						2170					2175			
15	Lys	Val	Gly	Lys	Glu	Arg	Ile	Tyr	Thr	Glu	Val	Thr	Thr	Asp	Gly
	2180						2185					2190			
20	Lys	Asn	Lys	Glu	Lys	Ser	Arg	Glu	Ile	Leu	Glu	Glu	Pro	Val	Asn
	2195						2200					2205			
25	Arg	Ile	Leu	Leu	Val	Gly	Thr	Lys	Lys	Lys	Thr	Ser	Asn	Ile	Glu
	2210						2215					2220			
30	Thr	Pro	Ala	Gly	Gly	His	Gln	Asp	Asn	Gly	Ser	Ser	Gln	Pro	Thr
	2225						2230					2235			
35	Glu	Gln	Lys	Pro	Gly	Ala	Asn	Pro	Asn	Pro	Asn	Ile	Val	Ile	Pro
	2240						2245					2250			
40	Asn	Pro	Ser	Asn	Asn	Gly	Val	Thr	Ser	Asn	Ser	Ser	Asp	Phe	Lys
	2255						2260					2265			
45	Asn	His	Leu	Ser	Val	Ala	Pro	Asn	Ile	Val	Ile	Gln	Ala	Pro	Thr
	2270						2275					2280			
50	Gln	Glu	Ser	Lys	Lys	Pro	Gly	Trp	Asn	Lys	Glu	Asp	Gly	Lys	Trp
	2285						2290					2295			
55	Tyr	Tyr	Arg	Gln	Gln	Asn	Gly	Asp	Leu	Ala	Lys	Gly	Trp	Val	Lys
	2300						2305					2310			
60	Asp	Gly	Asp	Thr	Trp	Tyr	Tyr	Phe	Asp	Gln	Thr	Gly	Lys	Met	Glu
	2315						2320					2325			
65	Thr	Gly	Trp	Ile	Lys	Asp	Gln	Ser	Gly	Ala	Trp	Tyr	Tyr	Leu	Asn
	2330						2335					2340			
70	Gln	Ser	Gly	Ala	Met	Ser	Ser	Asn	Glu	Trp	Ile	Leu	Asp	Gln	Asp
	2345						2350					2355			
75	Gly	Lys	Trp	Tyr	Tyr	Val	Asp	Ala	Ser	Gly	Ser	Met	Lys	Thr	Ser
	2360						2365					2370			

ES 2 352 704 B1

Gln Trp Phe Gln Val Gly Asp Lys Trp Tyr Tyr Val Asp Ala Asn
2375 2380 2385

5 Gly Ser Leu Ala Val Asn Thr Val Thr Pro Asp Gly Tyr Gln Val
2390 2400

10 Asn Glu Asn Gly Glu Trp Ile Gly
2405 2410

<210> 3
<211> 7101
15 <212> DNA
<213> Artificial

<220>
20 <223> beta-galactosidasa sin péptido señal

<220>
<221> CDS
25 <222> (1)..(7101)
<223> beta-galactosidasa sin péptido señal

<400> 3

30 atg gaa aat gct gaa gaa att gtg tta aca aat agg act gag cac cag
48
Met Glu Asn Ala Glu Glu Ile Val Leu Thr Asn Arg Thr Glu His Gln

35 1 5 10 15

tct gaa aac ctg aaa act act gga ctc tca agt gaa cat aaa gct gaa
96
40 Ser Glu Asn Leu Lys Thr Thr Gly Leu Ser Ser Glu His Lys Ala Glu

20 25 30

45 ggt gtt acg act gaa aaa gca agt gag caa cct gtt aaa gct gaa aat
144
Gly Val Thr Thr Glu Lys Ala Ser Glu Gln Pro Val Lys Ala Glu Asn

35 40 45

50 gta tct cta aag gaa aat gcg gat ggt agt gat gca gta gta aaa cct
192
Val Ser Leu Lys Glu Asn Ala Asp Gly Ser Asp Ala Val Val Lys Pro

55 50 55 60

cac gca gat caa cct tta gtt tca gtt gcc aat aat caa tca gtt cct
240
60 His Ala Asp Gln Pro Leu Val Ser Val Ala Asn Asn Gln Ser Val Pro

65 65 70 75 80

65 tta gaa gaa ggt aag gct gaa gtc gaa aac aaa aca gaa gaa agg aac
288
Leu Glu Glu Gly Lys Ala Glu Val Glu Asn Lys Thr Glu Glu Arg Asn

ES 2 352 704 B1

	85	90	95
5	aag cct gag gat aga aat aag ctt gaa gaa aca aat aag act gag gta 336 Lys Pro Glu Asp Arg Asn Lys Leu Glu Glu Thr Asn Lys Thr Glu Val		
	100	105	110
10	acg gat aaa tct gaa gat aaa aat aaa tct gaa gaa aca aat aaa gca 384 Thr Asp Lys Ser Glu Asp Lys Asn Lys Ser Glu Glu Thr Asn Lys Ala		
15	115	120	125
20	gaa tca caa gat cgt tct gtt tca aaa gat aag caa gaa ttc aaa tct 432 Glu Ser Gln Asp Arg Ser Val Ser Lys Asp Lys Gln Glu Phe Lys Ser		
	130	135	140
25	gct acg aat gaa gta gtt gat aag tta att gaa gat aga aac att tcc 480 Ala Thr Asn Glu Val Val Asp Lys Leu Ile Glu Asp Arg Asn Ile Ser		
30	145	150	155
	160		
35	ttc aac cag aat tgg cac ttt aaa cta aat gcc aat gca aag gaa gct 528 Phe Asn Gln Asn Trp His Phe Lys Leu Asn Ala Asn Ala Lys Glu Ala		
	165	170	175
40	gta aaa cca gac gca gat att tct tca tgg aaa aaa atg gat ctt cct 576 Val Lys Pro Asp Ala Asp Ile Ser Ser Trp Lys Lys Met Asp Leu Pro		
	180	185	190
45	cat gac tgg agt att cat ttt gat ttt gac cat gat tca cca gct caa 624 His Asp Trp Ser Ile His Phe Asp Phe Asp His Asp Ser Pro Ala Gln		
50	195	200	205
55	aac gaa ggc ggt caa ttg aat ggt gga gat ggc tgg tat cgg aaa acc 672 Asn Glu Gly Gly Gln Leu Asn Gly Gly Asp Gly Trp Tyr Arg Lys Thr		
	210	215	220
60	ttt aag tta gat gaa aag gac ttg aat aaa gat gtt cgt gta acc ttt 720 Phe Lys Leu Asp Glu Lys Asp Leu Asn Lys Asp Val Arg Val Thr Phe		
	225	230	235
65			240

gat ggt gtt tac atg gat tca caa gta ttc gtc aat ggc caa cta gtt
768
Asp Gly Val Tyr Met Asp Ser Gln Val Phe Val Asn Gly Gln Leu Val
5 245 250 255

ggt cac tat cca aat ggt tac aat caa ttt tca tat gac att agt aac
816
Gly His Tyr Pro Asn Gly Tyr Asn Gln Phe Ser Tyr Asp Ile Ser Asn
10 260 265 270

tat ctc cat aaa gat ggg cga gaa aat gtc att tct gtt cat gca gtt
864
Tyr Leu His Lys Asp Gly Arg Glu Asn Val Ile Ser Val His Ala Val
15 275 280 285

aat aag caa cca agt agt aga tgg tat tca gga agt ggt atc tat cga
912
Asn Lys Gln Pro Ser Ser Arg Trp Tyr Ser Gly Ser Gly Ile Tyr Arg
20 290 295 300

gac gta tca ttg cag gta aca gac aag att cat gtt gaa aag aat ggg
960
Asp Val Ser Leu Gln Val Thr Asp Lys Ile His Val Glu Lys Asn Gly
30 305 310 315 320

aca act att cta act cca aaa cta gaa agt caa caa gga gga aaa gtt
1008
Thr Thr Ile Leu Thr Pro Lys Leu Glu Ser Gln Gln Gly Gly Lys Val
35 325 330 335

gaa acc caa gta agc agt aaa att gtt aat aca gat aat aaa gac cat
1056
Glu Thr Gln Val Ser Ser Lys Ile Val Asn Thr Asp Asn Lys Asp His
40 340 345 350

gaa att gta gca gag tat caa att ttt gaa cga ggt ggt aag gca gtt
1104
Glu Ile Val Ala Glu Tyr Gln Ile Phe Glu Arg Gly Gly Lys Ala Val
45 355 360 365

aca gaa ctt gtt cga act gag agt aaa act cta aaa gct aag gaa acc
1152
Thr Glu Leu Val Arg Thr Glu Ser Lys Thr Leu Lys Ala Lys Glu Thr
50 370 375 380

att cat ttg gat tca gca tta gag gtt gaa aaa cca aaa ttg tgg aca
1200
Ile His Leu Asp Ser Ala Leu Glu Val Glu Lys Pro Lys Leu Trp Thr
55 60 65

ES 2 352 704 B1

	385	390	395	400
5	gtg tca tct gac aaa cct gct tta tat gag atg gta act cgt gtt tac 1248 Val Ser Ser Asp Lys Pro Ala Leu Tyr Glu Met Val Thr Arg Val Tyr			
	405	410	415	
10	aaa gat ggt cag ctt gta gat gca aag aaa gat tta ttt ggt tat cgc 1296 Lys Asp Gly Gln Leu Val Asp Ala Lys Lys Asp Leu Phe Gly Tyr Arg			
15	420	425	430	
20	tac tat aac tgg act cca gac caa ggg ttc tct tta aat ggt gaa cac 1344 Tyr Tyr Asn Trp Thr Pro Asp Gln Gly Phe Ser Leu Asn Gly Glu His			
	435	440	445	
25	atc aag ttc cac ggt gtt tca ttg cac cat gat cat gga gcg cta gga 1392 Ile Lys Phe His Gly Val Ser Leu His His Asp His Gly Ala Leu Gly			
30	450	455	460	
35	gca gaa gaa aac tat aag gca gaa tat cgt cgt ctg aaa caa atg aag 1440 Ala Glu Glu Asn Tyr Lys Ala Glu Tyr Arg Arg Leu Lys Gln Met Lys			
	465	470	475	480
40	gaa atg gga gtg aat tcg att cgt acg act cac aac cca gca agt cct 1488 Glu Met Gly Val Asn Ser Ile Arg Thr Thr His Asn Pro Ala Ser Pro			
	485	490	495	
45	cag acc ttg cag att gca gct gag ctc ggt ttg ctt gtt caa gaa gaa 1536 Gln Thr Leu Gln Ile Ala Ala Glu Leu Gly Leu Leu Val Gln Glu Glu			
50	500	505	510	
55	gct ttt gat act tgg tat gga ggt aag aaa cca tac gac tat ggt cgt 1584 Ala Phe Asp Thr Trp Tyr Gly Gly Lys Lys Pro Tyr Asp Tyr Gly Arg			
	515	520	525	
60	ttc ttt gaa aaa gat gct acc cac cct gaa gca aga aaa ggt gaa aaa 1632 Phe Phe Glu Lys Asp Ala Thr His Pro Glu Ala Arg Lys Gly Glu Lys			
65	530	535	540	

5 **1680**
 tgg tcc gac tac gat ctt cga aca atg gta gaa aga gat aaa aac aat
 Trp Ser Asp Tyr Asp Leu Arg Thr Met Val Glu Arg Asp Lys Asn Asn
 545 550 555 560

10 **1728**
 cct gct gtt gtt atg tgg tct att ggt aac gaa atc ggt gaa gca gat
 Pro Ala Val Val Met Trp Ser Ile Gly Asn Glu Ile Gly Glu Ala Asp
 565 570 575

15 **1776**
 ggg aaa gcc cat tca ctg gta act gtc aaa cgt ttg gtc aaa gtg gtg
 Gly Lys Ala His Ser Leu Val Thr Val Lys Arg Leu Val Lys Val Val
 580 585 590

20 **1824**
 aaa tct gtt gat aag aca cgt tat gta acc atg gga gct gat aaa ttc
 Lys Ser Val Asp Lys Thr Arg Tyr Val Thr Met Gly Ala Asp Lys Phe
 595 600 605

25 **1872**
 cgt ttc gga gat gga act ggg gac cat gaa aag att gca gat gaa ctg
 Arg Phe Gly Asp Gly Thr Gly Asp His Glu Lys Ile Ala Asp Glu Leu
 610 615 620

30 **1920**
 gat gca gtt gga ttg aat tac tct gaa gaa aat tat caa act ctt cat
 Asp Ala Val Gly Leu Asn Tyr Ser Glu Glu Asn Tyr Gln Thr Leu His
 625 630 635 640

35 **1968**
 gcg aaa cat ccg aaa tgg ctc atc tat ggt tca gaa aca tct tca gca
 Ala Lys His Pro Lys Trp Leu Ile Tyr Gly Ser Glu Thr Ser Ser Ala
 645 650 655

40 **2016**
 act cgt aca aga gga agt tat ttc cat cca gag agt gag tgg gta gga
 Thr Arg Thr Arg Gly Ser Tyr Phe His Pro Glu Ser Glu Trp Val Gly
 660 665 670

45 **2064**
 agt aac caa tat tgg cgt aat tac gaa caa tcc gat tat ggt aat gat
 Ser Asn Gln Tyr Trp Arg Asn Tyr Glu Gln Ser Asp Tyr Gly Asn Asp
 675 680 685

50 **2112**
 cgt gtt ggt tgg ggt aaa act gct aca gca tca tgg acc ttc gat aga
 685 690 695

ES 2 352 704 B1

	Arg Val Gly Trp Gly Lys Thr Ala Thr Ala Ser Trp Thr Phe Asp Arg	
	690	695 700
5		
	gat cat cca gga tac gct gga caa ttc att tgg acc ggt acg gat tat	
	2160	
	Asp His Pro Gly Tyr Ala Gly Gln Phe Ile Trp Thr Gly Thr Asp Tyr	
10	705	710 715 720
	att ggt gaa cca aca cca tgg cac aat caa aat agt aca cca gta aaa	
	2208	
15	Ile Gly Glu Pro Thr Pro Trp His Asn Gln Asn Ser Thr Pro Val Lys	
	725	730 735
20	agt tcc tac ttt ggt att gtt gat acc gca ggt att cca aaa aat gac	
	2256	
	Ser Ser Tyr Phe Gly Ile Val Asp Thr Ala Gly Ile Pro Lys Asn Asp	
25	740	745 750
	tat tat ctc tac caa agt cag tgg gtt tct gcc aag aaa aaa cca atg	
	2304	
30	Tyr Tyr Leu Tyr Gln Ser Gln Trp Val Ser Ala Lys Lys Lys Pro Met	
	755	760 765
35	gtt cat tta ctt cct cac tgg aac tgg gaa aag act gaa ttg gca gat	
	2352	
	Val His Leu Leu Pro His Trp Asn Trp Glu Lys Thr Glu Leu Ala Asp	
	770	775 780
40	aat gtt gca gac gct gaa aat cga att cca gtt aga gct tat tcc aat	
	2400	
	Asn Val Ala Asp Ala Glu Asn Arg Ile Pro Val Arg Ala Tyr Ser Asn	
45	785	790 795 800
	gct gcc agt gtt gag ttg ttc tta aat aat gag tca cta ggt ctt aaa	
	2448	
50	Ala Ala Ser Val Glu Leu Phe Leu Asn Asn Glu Ser Leu Gly Leu Lys	
	805	810 815
55	aaa ttc aac aaa aaa caa aca agt gac ggt cga aca tat caa gaa ggt	
	2496	
	Lys Phe Asn Lys Lys Gln Thr Ser Asp Gly Arg Thr Tyr Gln Glu Gly	
	820	825 830
60		
	gaa aat cct caa gaa cta tac ttg gag tgg aaa gta gct tat aaa ccg	
	2544	
	Glu Asn Pro Gln Glu Leu Tyr Leu Glu Trp Lys Val Ala Tyr Lys Pro	
65	835	840 845

5 gga act ttg gaa gct gta gct cga gat gag tct ggc aag gaa att gct
 2592
 Gly Thr Leu Glu Ala Val Ala Arg Asp Glu Ser Gly Lys Glu Ile Ala
 850 855 860

10 cgt gat aag att gta act gct ggt gaa cct gca ggt gtc cgt ttg gtt
 2640
 Arg Asp Lys Ile Val Thr Ala Gly Glu Pro Ala Gly Val Arg Leu Val
 865 870 875 880

15 aag gaa gaa aat gca atc gct gca gat gga aaa gat ttg act tat atc
 2688
 Lys Glu Glu Asn Ala Ile Ala Ala Asp Gly Lys Asp Leu Thr Tyr Ile
 885 890 895

20 tat tac gaa att gta gat agc aag ggt aat gtt gtg cca act gct aac
 2736
 Tyr Tyr Glu Ile Val Asp Ser Lys Gly Asn Val Val Pro Thr Ala Asn
 900 905 910

25 aat cta gtt cac ttc caa ctt cat gga caa ggt caa ctg gtt ggt gtg
 2784
 Asn Leu Val His Phe Gln Leu His Gly Gln Gly Gln Leu Val Gly Val
 915 920 925

30 gat aat ggt gag caa gct agt cgt gaa cgt tat aaa gaa caa gca gat
 2832
 Asp Asn Gly Glu Gln Ala Ser Arg Glu Arg Tyr Lys Glu Gln Ala Asp
 930 935 940

35 ggt tct tgg att cgt aga gcc ttt aat ggt aaa ggg gta gcg att gtt
 2880
 Gly Ser Trp Ile Arg Arg Ala Phe Asn Gly Lys Gly Val Ala Ile Val
 945 950 955 960

40 aaa tcg act gat caa gca ggt aaa ttt acc ctt aca gca tat tca ggc
 2928
 Lys Ser Thr Asp Gln Ala Gly Lys Phe Thr Leu Thr Ala Tyr Ser Gly
 965 970 975

45 ttg ttg aaa tca gat caa gtt aca gtt ttt aca ggt aag tca gac cag
 2976
 Leu Leu Lys Ser Asp Gln Val Thr Val Phe Thr Gly Lys Ser Asp Gln
 980 985 990

50 tca gaa aaa act gtt tta gga aca gaa ata cca aaa gtt cgg aca gtt
 3024

3024
 Ser Glu Lys Thr Val Leu Gly Thr Glu Ile Pro Lys Val Arg Thr Val
 995 1000 1005

5

tta gga caa aaa ctt aaa atg cct cag aca gtg cca ttt gta tat
 3069
 Leu Gly Gln Lys Leu Lys Met Pro Gln Thr Val Pro Phe Val Tyr
 1010 1015 1020

10

agt gat ggt agt cgt gaa aaa cgt cct gta acc tgg tct caa gca
 3114
 Ser Asp Gly Ser Arg Glu Lys Arg Pro Val Thr Trp Ser Gln Ala
 1025 1030 1035

15

gat gtc agt caa gct gga att gta act gta aaa ggt atg tct gat
 3159
 Asp Val Ser Gln Ala Gly Ile Val Thr Val Lys Gly Met Ser Asp
 1040 1045 1050

20

gga cgt gaa gtt gag gct cgt gtt gaa gta ctt aca gtt gcg aat
 3204
 Gly Arg Glu Val Glu Ala Arg Val Glu Val Leu Thr Val Ala Asn
 1055 1060 1065

25

gaa ctt cca gct gtt aaa cgc atc aca cct gga aca gac ttg agt
 3249
 Glu Leu Pro Ala Val Lys Arg Ile Thr Pro Gly Thr Asp Leu Ser
 1070 1075 1080

30

tct gtg gat aaa ctt gtt tca ctt gtt tca aca gat ggt cgt att
 3294
 Ser Val Asp Lys Leu Val Ser Leu Val Ser Thr Asp Gly Arg Ile
 1085 1090 1095

35

cac aat tat gaa gtt gaa aat tgg gaa att tca cca gaa gat aaa
 3339
 His Asn Tyr Glu Val Glu Asn Trp Glu Ile Ser Pro Glu Asp Lys
 1100 1105 1110

40

gag aag tta tct act ccg ggg gca cat att cag atg act agc caa
 3384
 Glu Lys Leu Ser Thr Pro Gly Ala His Ile Gln Met Thr Ser Gln
 1115 1120 1125

45

tta gga gat aga acc atc cac gct act ttg ata gtt gat gat ggt
 3429
 Leu Gly Asp Arg Thr Ile His Ala Thr Leu Ile Val Asp Asp Gly
 1130 1135 1140

50

55

60

65

ES 2 352 704 B1

	1130	1135	1140
5	aag gaa ttg tct caa gtt aca cca aat ata gca gta ggt ggt gaa 3474 Lys Glu Leu Ser Gln Val Thr Pro Asn Ile Ala Val Gly Gly Glu		
	1145	1150	1155
10	agt gtt acg gaa ctt tca aaa gag cat cct gta cat tat cac aaa 3519 Ser Val Thr Glu Leu Ser Lys Glu His Pro Val His Tyr His Lys		
15	1160	1165	1170
20	ctt tct tat gga gca aaa att cca gaa gtc agt gct agt gct gaa 3564 Leu Ser Tyr Gly Ala Lys Ile Pro Glu Val Ser Ala Ser Ala Glu		
	1175	1180	1185
25	aat gcc gat gtg aac gtg att cag gca aat gaa tcc aat gga atg 3609 Asn Ala Asp Val Asn Val Ile Gln Ala Asn Glu Ser Asn Gly Met		
30	1190	1195	1200
35	cgt gct agt atc tac att cag cct aaa gac ggg gga cag tta caa 3654 Arg Ala Ser Ile Tyr Ile Gln Pro Lys Asp Gly Gly Gln Leu Gln		
	1205	1210	1215
40	acc tat gct gtt cag ttc ctt gtt gat tct cca caa att gaa agc 3699 Thr Tyr Ala Val Gln Phe Leu Val Asp Ser Pro Gln Ile Glu Ser		
	1220	1225	1230
45	ttg agt ctc aga gtg gat caa gta tct agc ctc aaa gaa gat caa 3744 Leu Ser Leu Arg Val Asp Gln Val Ser Ser Leu Lys Glu Asp Gln		
50	1235	1240	1245
55	act gtt aaa gtg aca gtt tta gcc cgt tat cag gat gga act gaa 3789 Thr Val Lys Val Thr Val Leu Ala Arg Tyr Gln Asp Gly Thr Glu		
	1250	1255	1260
60	gct att tta cca act gat aag gtg act ttc tca agc caa ggc gag 3834 Ala Ile Leu Pro Thr Asp Lys Val Thr Phe Ser Ser Gln Gly Glu		
65	1265	1270	1275

ES 2 352 704 B1

gga gga gtc tct gtt aat aag gga atg cta gag ttg cat aaa ccg
 3879
 Gly Gly Val Ser Val Asn Lys Gly Met Leu Glu Leu His Lys Pro
 5 1280 1285 1290

ggt agt ctc act tta caa gct caa tat gaa ggt gca aaa gga gaa
 3924
 Gly Ser Leu Thr Leu Gln Ala Gln Tyr Glu Gly Ala Lys Gly Glu
 10 1295 1300 1305

ctt cct cta acg atc gca gca aat act gaa gat aaa gtt gtt cag
 3969
 Leu Pro Leu Thr Ile Ala Ala Asn Thr Glu Asp Lys Val Val Gln
 15 1310 1315 1320

gcc att cgc cca gtt agc ctt gtg aca gat ttg cac caa aca cca
 4014
 Ala Ile Arg Pro Val Ser Leu Val Thr Asp Leu His Gln Thr Pro
 25 1325 1330 1335

aat ctt cca tct aca gta aca gtt gag tac gat aaa gga ttt cca
 4059
 Asn Leu Pro Ser Thr Val Thr Val Glu Tyr Asp Lys Gly Phe Pro
 30 1340 1345 1350

aaa gtt cat aag gta gtt tgg gat tcg att cca tct gaa aaa tta
 4104
 Lys Val His Lys Val Val Trp Asp Ser Ile Pro Ser Glu Lys Leu
 35 1355 1360 1365

aat caa tat cat act ttt gaa gta ctt ggt aag gtt gaa gga ttg
 4149
 Asn Gln Tyr His Thr Phe Glu Val Leu Gly Lys Val Glu Gly Leu
 40 1370 1375 1380

aat caa gaa gct cgt gcg aca gta tca gta gaa gga att gta aaa
 4194
 Asn Gln Glu Ala Arg Ala Thr Val Ser Val Glu Gly Ile Val Lys
 50 1385 1390 1395

gtt gaa gat gtg agc atc aca aca cca att tct gag gca cca acc
 4239
 Val Glu Asp Val Ser Ile Thr Thr Pro Ile Ser Glu Ala Pro Thr
 55 1400 1405 1410

ttg cca gaa agt gtc cga act tat tat tca aat ggc caa gta tct
 4284
 Leu Pro Glu Ser Val Arg Thr Tyr Tyr Ser Asn Gly Gln Val Ser
 60 1420 1425 1430

65

ES 2 352 704 B1

	1415	1420	1425
5	tca gct aag gtg gct tgg gac 4329 Ser Ala Lys Val Ala Trp Asp	act att gat gca agc Thr Ile Asp Ala Ser	cag tat gca Gln Tyr Ala
	1430	1435	1440
10	aaa gaa gga atc ttt act gta 4374 Lys Glu Gly Ile Phe Thr Val	aca ggc cat gta gaa Thr Gly His Val Glu	gga aca caa Gly Thr Gln
15	1445	1450	1455
20	ctg aca acg aaa tta cat gtt 4419 Leu Thr Thr Lys Leu His Val	cgt gta tcg agt caa Arg Val Ser Ser Gln	aca gaa att Thr Glu Ile
	1460	1465	1470
25	ggg aag aat att tca gat caa 4464 Gly Lys Asn Ile Ser Asp Gln	tgg aca ggt tca gaa Trp Thr Gly Ser Glu	tta cca ctt Leu Pro Leu
30	1475	1480	1485
35	gcc ttt gca tca gat tct cat 4509 Ala Phe Ala Ser Asp Ser His	tct tat gac cca gta Ser Tyr Asp Pro Val	tct aat gta Ser Asn Val
	1490	1495	1500
40	aat gat aaa gtg att tct tac 4554 Asn Asp Lys Val Ile Ser Tyr	ggg gat cag cct gca Gly Asp Gln Pro Ala	aat cgt tgg Asn Arg Trp
	1505	1510	1515
45	tca aac tgg aaa cgt tct gaa 4599 Ser Asn Trp Lys Arg Ser Glu	gaa gct tca gta ggg Glu Ala Ser Val Gly	gta ctc ttt Val Leu Phe
50	1520	1525	1530
55	gga gat tca ggt att ttg aca 4644 Gly Asp Ser Gly Ile Leu Thr	aaa cgt tcg gtc gat Lys Arg Ser Val Asp	aac ttg aat Asn Leu Asn
	1535	1540	1545
60	gta gca ttt tat gaa gat agc 4689 Val Ala Phe Tyr Glu Asp Ser	ggg gta ggt gct cca Gly Val Gly Ala Pro	aaa tct tac Lys Ser Tyr
	1550	1555	1560
65			

gtc att gag tat tat gtt ggt gaa aaa gct cca act gca ccg aag
 4734
 Val Ile Glu Tyr Tyr Val Gly Glu Lys Ala Pro Thr Ala Pro Lys
 5 1565 1570 1575

aat cct ggt tat gta gag ggt gaa aac cac gta ttt aat gac agt
 4779
 Asn Pro Gly Tyr Val Glu Gly Glu Asn His Val Phe Asn Asp Ser
 10 1580 1585 1590

aaa aat tgg aaa cca gtc aca aat ctt aag gct cca gat cag tta
 4824
 Lys Asn Trp Lys Pro Val Thr Asn Leu Lys Ala Pro Asp Gln Leu
 15 1595 1600 1605

aaa gct gga gaa atg aat cat ttc agt ttt gat aag gtg gat acc
 4869
 Lys Ala Gly Glu Met Asn His Phe Ser Phe Asp Lys Val Asp Thr
 20 1610 1615 1620

tac gcg gta cgt att cgc atg aca aga gct gat gat aaa ctt gga
 4914
 Tyr Ala Val Arg Ile Arg Met Thr Arg Ala Asp Asp Lys Leu Gly
 25 1625 1630 1635

aca tcc att aca gaa tta caa gtt ttc tct aaa aag gta gca cca
 4959
 Thr Ser Ile Thr Glu Leu Gln Val Phe Ser Lys Lys Val Ala Pro
 30 1640 1645 1650

gct aaa gag gca act act aac att aag gtt gct gga aaa gat ctt
 5004
 Ala Lys Glu Ala Thr Thr Asn Ile Lys Val Ala Gly Lys Asp Leu
 35 1655 1660 1665

cct cac ttt aat cca gac tta aca gat tac tat tta gac tat agt
 5049
 Pro His Phe Asn Pro Asp Leu Thr Asp Tyr Tyr Leu Asp Tyr Ser
 40 1670 1675 1680

gat gga aaa att cca gaa gta aca gct agt atc aca aat aat gga
 5094
 Asp Gly Lys Ile Pro Glu Val Thr Ala Ser Ile Thr Asn Asn Gly
 45 1685 1690 1695

ctt gct aca gtg gtt cct agc gtg aaa gag ggc gat cct gtc cgt
 5139

65

ES 2 352 704 B1

	Leu	Ala	Thr	Val	Val	Pro	Ser	Val	Lys	Glu	Gly	Asp	Pro	Val	Arg
	1700						1705					1710			
5	gtt	att	gtg	aaa	gca	gaa	aat	ggc	gat	att	ctt	gga	gaa	tac	aac
	5184														
	Val	Ile	Val	Lys	Ala	Glu	Asn	Gly	Asp	Ile	Leu	Gly	Glu	Tyr	Asn
10	1715						1720					1725			
	ctt	cac	ttt	aca	aaa	gat	aaa	gat	ttg	ttg	gct	cgt	aaa	cca	att
	5229														
15	Leu	His	Phe	Thr	Lys	Asp	Lys	Asp	Leu	Leu	Ala	Arg	Lys	Pro	Ile
	1730						1735					1740			
	gct	tct	gca	aaa	cag	tct	caa	ttg	ttg	cag	tta	ggc	caa	cct	ttg
	5274														
20	Ala	Ser	Ala	Lys	Gln	Ser	Gln	Leu	Leu	Gln	Leu	Gly	Gln	Pro	Leu
	1745						1750					1755			
25															
	gaa	ttg	cca	aat	aag	gtt	cca	gtt	tat	ttc	gca	ggc	aag	agc	ggc
	5319														
30	Glu	Leu	Pro	Asn	Lys	Val	Pro	Val	Tyr	Phe	Ala	Gly	Lys	Ser	Gly
	1760						1765					1770			
	tat	gac	gta	aaa	gat	ttg	gca	gtt	gaa	tg	gaa	aaa	gtc	cct	gct
	5364														
35	Tyr	Asp	Val	Lys	Asp	Leu	Ala	Val	Glu	Trp	Glu	Lys	Val	Pro	Ala
	1775						1780					1785			
40	tct	acc	tta	aac	aaa	gct	ggg	gaa	tat	aca	att	aaa	ggc	cat	gta
	5409														
	Ser	Thr	Leu	Asn	Lys	Ala	Gly	Glu	Tyr	Thr	Ile	Lys	Gly	His	Val
45	1790						1795					1800			
	gtt	ggc	agt	gag	gtc	cca	gta	gaa	tta	tct	gtt	cga	gta	acg	gac
	5454														
50	Val	Gly	Ser	Glu	Val	Pro	Val	Glu	Leu	Ser	Val	Arg	Val	Thr	Asp
	1805						1810					1815			
	aag	gtc	ggc	aga	tca	ctt	tct	gat	aat	cca	gat	tat	ctt	gct	act
	5499														
55	Lys	Val	Gly	Arg	Ser	Leu	Ser	Asp	Asn	Pro	Asp	Tyr	Leu	Ala	Thr
	1820						1825					1830			
60															
	cat	agt	cag	gct	ttt	gct	tca	gct	act	aat	gat	ttg	gat	gat	aat
	5544														
65	His	Ser	Gln	Ala	Phe	Ala	Ser	Ala	Thr	Asn	Asp	Leu	Asp	Asp	Asn
	1835						1840					1845			

tcg aac gac cgt gta gac tat tta aat gat aat gac cga aat caa
 5589
 Ser Asn Asp Arg Val Asp Tyr Leu Asn Asp Asn Asp Arg Asn Gln
 5 1850 1855 1860

gac aac cgt tgg aca aac tgg tcc gct act cca tct gct aat cca
 5634
 Asp Asn Arg Trp Thr Asn Trp Ser Ala Thr Pro Ser Ala Asn Pro
 10 1865 1870 1875

gaa gta tca gtt ggt gtt atc ttt aaa aag aat ggt aaa att gtg
 5679
 Glu Val Ser Val Gly Val Ile Phe Lys Lys Asn Gly Lys Ile Val
 15 1880 1885 1890

gat aga aag gtt tca caa gct aaa cta aac ttc ttt gca gat agt
 5724
 Asp Arg Lys Val Ser Gln Ala Lys Leu Asn Phe Phe Ala Asp Ser
 20 1895 1900 1905

gga aca gat gcc cca gct aaa ctc gta ttg gaa cga tac atc ggt
 5769
 Gly Thr Asp Ala Pro Ala Lys Leu Val Leu Glu Arg Tyr Ile Gly
 25 1910 1915 1920

cct gac ttt gat gtt ccg gtt tac tat tca aac tac agc tac gaa
 5814
 Pro Asp Phe Asp Val Pro Val Tyr Tyr Ser Asn Tyr Ser Tyr Glu
 30 1925 1930 1935

tca aat cat cca ttt aat aac cct gat aat tgg gaa ctt gtt cct
 5859
 Ser Asn His Pro Phe Asn Asn Pro Asp Asn Trp Glu Leu Val Pro
 35 1940 1945 1950

tat tac gca gac aaa gaa att caa gct gga aat gag atc aat gtg
 5904
 Tyr Tyr Ala Asp Lys Glu Ile Gln Ala Gly Asn Glu Ile Asn Val
 40 1955 1960 1965

acc ttt aaa gct gtg act acc aaa gct atg aga tgg cgt atg gac
 5949
 Thr Phe Lys Ala Val Thr Thr Lys Ala Met Arg Trp Arg Met Asp
 45 1970 1975 1980

aga aaa gca gat aaa aat ggg gtt gca cta act gaa tta agt ttc
 50 1985 1990 1995

5994
Arg Lys Ala Asp Lys Asn Gly Val Ala Leu Thr Glu Leu Ser Phe
1985 1990 1995

5

cga gca cca agt gag ctt gca aaa gag agc aca gag gga aga att
6039
Arg Ala Pro Ser Glu Leu Ala Lys Glu Ser Thr Glu Gly Arg Ile
2000 2005 2010

10

tta gta gat ggt aaa gag ctt cct gac ttc tct gaa aat cgt ttg
6084
Leu Val Asp Gly Lys Glu Leu Pro Asp Phe Ser Glu Asn Arg Leu
2015 2020 2025

15

gac tat gaa ttg att tat aaa ggg gaa cgt cca aaa atc act gta
6129
Asp Tyr Glu Leu Ile Tyr Lys Gly Glu Arg Pro Lys Ile Thr Val
2030 2035 2040

25

gaa ggg aaa gat caa gtc aca tct aca att gtg gat agt gga aat
6174
Glu Gly Lys Asp Gln Val Thr Ser Thr Ile Val Asp Ser Gly Asn
2045 2050 2055

30

gaa aat ctt cct gta tta gtc cgt ctg gta tca gag agt ggt aag
6219
Glu Asn Leu Pro Val Leu Val Arg Leu Val Ser Glu Ser Gly Lys
2060 2065 2070

35

aat gtc aaa gaa tac cgg att aaa ttg act aag gaa aaa gtg ata
6264
Asn Val Lys Glu Tyr Arg Ile Lys Leu Thr Lys Glu Lys Val Ile
2075 2080 2085

40

gac gca aaa ctt gta gct cct gct caa tat gac ttg cct agt ctt
6309
Asp Ala Lys Leu Val Ala Pro Ala Gln Tyr Asp Leu Pro Ser Leu
2090 2095 2100

45

gaa gtg gta gac aaa gaa ttg aac ttc caa act ctt gaa caa aaa
6354
Glu Val Val Asp Lys Glu Leu Asn Phe Gln Thr Leu Glu Gln Lys
2105 2110 2115

50

gac gat act ctg ttt gag gga gaa gtt cgc ctg ttg caa gaa ggt
6399
Asp Asp Thr Leu Phe Glu Gly Glu Val Arg Leu Leu Gln Glu Gly
2105 2110 2115

55

60

65

ES 2 352 704 B1

	2120	2125	2130
5	aaa gtt ggt aag gaa cgt atc 6444 Lys Val Gly Lys Glu Arg Ile	tat aca gaa gta acg Tyr Thr Glu Val Thr	act gat ggc Thr Asp Gly
	2135	2140	2145
10	aag aat aaa gaa aaa tca cgt 6489 Lys Asn Lys Glu Lys Ser Arg	gag att ctg gaa gaa Glu Ile Leu Glu Glu	cct gta aat Pro Val Asn
15	2150	2155	2160
20	aga att ctg ttg gta ggt act 6534 Arg Ile Leu Leu Val Gly Thr	aag aag aaa aca agt Lys Lys Lys Thr Ser	aac att gag Asn Ile Glu
	2165	2170	2175
25	acg ccg gcg ggt gga cat caa 6579 Thr Pro Ala Gly Gly His Gln	gac aat ggt agc agt Asp Asn Gly Ser Ser	caa cct acg Gln Pro Thr
30	2180	2185	2190
35	gaa caa aaa cca gga gca aat 6624 Glu Gln Lys Pro Gly Ala Asn	ccg aat cca aat att Pro Asn Pro Asn Ile	gtc att cca Val Ile Pro
	2195	2200	2205
40	aat cca tca aat aat ggt gtg 6669 Asn Pro Ser Asn Asn Gly Val	aca agt aac agt tca Thr Ser Asn Ser Ser	gat ttt aaa Asp Phe Lys
	2210	2215	2220
45	aat cat cta tca gtt gct ccg 6714 Asn His Leu Ser Val Ala Pro	aac att gtg ata caa Asn Ile Val Ile Gln	gct cct act Ala Pro Thr
50	2225	2230	2235
55	cag gaa agc aaa aaa cct gga 6759 Gln Glu Ser Lys Lys Pro Gly	tgg aac aag gaa gat Trp Asn Lys Glu Asp	ggg aaa tgg Gly Lys Trp
	2240	2245	2250
60	tat tat cgt caa caa aat ggg 6804 Tyr Tyr Arg Gln Gln Asn Gly	gac ttg gct aag ggt Asp Leu Ala Lys Gly	tgg gtc aaa Trp Val Lys
65	2255	2260	2265

ES 2 352 704 B1

```

gat gga gat act tgg tat tac ttt gat caa aca gga aaa atg gaa
6849
Asp Gly Asp Thr Trp Tyr Tyr Phe Asp Gln Thr Gly Lys Met Glu
5      2270      2275      2280

act ggc tgg att aaa gat caa tca ggt gca tgg tat tac ttg aat
6894
Thr Gly Trp Ile Lys Asp Gln Ser Gly Ala Trp Tyr Tyr Leu Asn
10     2285      2290      2295

caa tca ggt gca atg tca tct aat gag tgg att ttg gat caa gat
6939
Gln Ser Gly Ala Met Ser Ser Asn Glu Trp Ile Leu Asp Gln Asp
15     2300      2305      2310

ggt aag tgg tac tat gtt gat gca tca ggt tcg atg aag aca agt
6984
Gly Lys Trp Tyr Tyr Val Asp Ala Ser Gly Ser Met Lys Thr Ser
25     2315      2320      2325

caa tgg ttc caa gtt gga gac aaa tgg tat tat gtt gat gca aat
7029
Gln Trp Phe Gln Val Gly Asp Lys Trp Tyr Tyr Val Asp Ala Asn
30     2330      2335      2340

gga tcc ctt gct gtc aat aca gta act ccg gat ggt tat caa gta
7074
Gly Ser Leu Ala Val Asn Thr Val Thr Pro Asp Gly Tyr Gln Val
35     2345      2350      2355

aat gaa aac ggt gag tgg atc gga taa
7101
Asn Glu Asn Gly Glu Trp Ile Gly
45     2360      2365

<210> 4
<211> 2366
<212> PRT
<213> Artificial
55

<220>
<223> Synthetic Construct

60 <400> 4

Met Glu Asn Ala Glu Glu Ile Val Leu Thr Asn Arg Thr Glu His Gln
1      5      10      15

Ser Glu Asn Leu Lys Thr Thr Gly Leu Ser Ser Glu His Lys Ala Glu
20     25     30

```

ES 2 352 704 B1

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

ES 2 352 704 B1

	Asn	Lys	Gln	Pro	Ser	Ser	Arg	Trp	Tyr	Ser	Gly	Ser	Gly	Ile	Tyr	Arg
		290					295					300				
5	Asp	Val	Ser	Leu	Gln	Val	Thr	Asp	Lys	Ile	His	Val	Glu	Lys	Asn	Gly
	305					310					315					320
10	Thr	Thr	Ile	Leu	Thr	Pro	Lys	Leu	Glu	Ser	Gln	Gln	Gly	Gly	Lys	Val
					325					330					335	
15	Glu	Thr	Gln	Val	Ser	Ser	Lys	Ile	Val	Asn	Thr	Asp	Asn	Lys	Asp	His
				340					345					350		
20	Glu	Ile	Val	Ala	Glu	Tyr	Gln	Ile	Phe	Glu	Arg	Gly	Gly	Lys	Ala	Val
			355					360					365			
25	Thr	Glu	Leu	Val	Arg	Thr	Glu	Ser	Lys	Thr	Leu	Lys	Ala	Lys	Glu	Thr
		370					375					380				
30	Ile	His	Leu	Asp	Ser	Ala	Leu	Glu	Val	Glu	Lys	Pro	Lys	Leu	Trp	Thr
	385					390					395					400
35	Val	Ser	Ser	Asp	Lys	Pro	Ala	Leu	Tyr	Glu	Met	Val	Thr	Arg	Val	Tyr
					405					410					415	
40	Lys	Asp	Gly	Gln	Leu	Val	Asp	Ala	Lys	Lys	Asp	Leu	Phe	Gly	Tyr	Arg
				420					425					430		
45	Tyr	Tyr	Asn	Trp	Thr	Pro	Asp	Gln	Gly	Phe	Ser	Leu	Asn	Gly	Glu	His
			435					440					445			
50	Ile	Lys	Phe	His	Gly	Val	Ser	Leu	His	His	Asp	His	Gly	Ala	Leu	Gly
		450					455					460				
55	Ala	Glu	Glu	Asn	Tyr	Lys	Ala	Glu	Tyr	Arg	Arg	Leu	Lys	Gln	Met	Lys
	465					470					475					480
60	Glu	Met	Gly	Val	Asn	Ser	Ile	Arg	Thr	Thr	His	Asn	Pro	Ala	Ser	Pro
					485					490					495	
65	Gln	Thr	Leu	Gln	Ile	Ala	Ala	Glu	Leu	Gly	Leu	Leu	Val	Gln	Glu	Glu
				500					505					510		
70	Ala	Phe	Asp	Thr	Trp	Tyr	Gly	Gly	Lys	Lys	Pro	Tyr	Asp	Tyr	Gly	Arg
			515					520					525			
75	Phe	Phe	Glu	Lys	Asp	Ala	Thr	His	Pro	Glu	Ala	Arg	Lys	Gly	Glu	Lys
		530					535					540				
80	Trp	Ser	Asp	Tyr	Asp	Leu	Arg	Thr	Met	Val	Glu	Arg	Asp	Lys	Asn	Asn
	545					550					555					560

ES 2 352 704 B1

Pro Ala Val Val Met Trp Ser Ile Gly Asn Glu Ile Gly Glu Ala Asp
 565 570 575
 5 Gly Lys Ala His Ser Leu Val Thr Val Lys Arg Leu Val Lys Val Val
 580 585 590
 10 Lys Ser Val Asp Lys Thr Arg Tyr Val Thr Met Gly Ala Asp Lys Phe
 595 600 605
 15 Arg Phe Gly Asp Gly Thr Gly Asp His Glu Lys Ile Ala Asp Glu Leu
 610 615 620
 20 Asp Ala Val Gly Leu Asn Tyr Ser Glu Glu Asn Tyr Gln Thr Leu His
 625 630 635 640
 25 Ala Lys His Pro Lys Trp Leu Ile Tyr Gly Ser Glu Thr Ser Ser Ala
 645 650 655
 30 Thr Arg Thr Arg Gly Ser Tyr Phe His Pro Glu Ser Glu Trp Val Gly
 660 665 670
 35 Ser Asn Gln Tyr Trp Arg Asn Tyr Glu Gln Ser Asp Tyr Gly Asn Asp
 675 680 685
 40 Arg Val Gly Trp Gly Lys Thr Ala Thr Ala Ser Trp Thr Phe Asp Arg
 690 695 700
 45 Asp His Pro Gly Tyr Ala Gly Gln Phe Ile Trp Thr Gly Thr Asp Tyr
 705 710 715 720
 50 Ile Gly Glu Pro Thr Pro Trp His Asn Gln Asn Ser Thr Pro Val Lys
 725 730 735
 55 Ser Ser Tyr Phe Gly Ile Val Asp Thr Ala Gly Ile Pro Lys Asn Asp
 740 745 750
 60 Tyr Tyr Leu Tyr Gln Ser Gln Trp Val Ser Ala Lys Lys Lys Pro Met
 755 760 765
 65 Val His Leu Leu Pro His Trp Asn Trp Glu Lys Thr Glu Leu Ala Asp
 770 775 780
 70 Asn Val Ala Asp Ala Glu Asn Arg Ile Pro Val Arg Ala Tyr Ser Asn
 785 790 795 800
 75 Ala Ala Ser Val Glu Leu Phe Leu Asn Asn Glu Ser Leu Gly Leu Lys
 805 810 815

ES 2 352 704 B1

Lys Phe Asn Lys Lys Gln Thr Ser Asp Gly Arg Thr Tyr Gln Glu Gly
 820 825 830
 5 Glu Asn Pro Gln Glu Leu Tyr Leu Glu Trp Lys Val Ala Tyr Lys Pro
 835 840 845
 10 Gly Thr Leu Glu Ala Val Ala Arg Asp Glu Ser Gly Lys Glu Ile Ala
 850 855 860
 15 Arg Asp Lys Ile Val Thr Ala Gly Glu Pro Ala Gly Val Arg Leu Val
 865 870 875 880
 20 Lys Glu Glu Asn Ala Ile Ala Ala Asp Gly Lys Asp Leu Thr Tyr Ile
 885 890 895
 25 Tyr Tyr Glu Ile Val Asp Ser Lys Gly Asn Val Val Pro Thr Ala Asn
 900 905 910
 30 Asn Leu Val His Phe Gln Leu His Gly Gln Gly Gln Leu Val Gly Val
 915 920 925
 35 Asp Asn Gly Glu Gln Ala Ser Arg Glu Arg Tyr Lys Glu Gln Ala Asp
 930 935 940
 40 Gly Ser Trp Ile Arg Arg Ala Phe Asn Gly Lys Gly Val Ala Ile Val
 945 950 955 960
 45 Lys Ser Thr Asp Gln Ala Gly Lys Phe Thr Leu Thr Ala Tyr Ser Gly
 965 970 975
 50 Leu Leu Lys Ser Asp Gln Val Thr Val Phe Thr Gly Lys Ser Asp Gln
 980 985 990
 55 Ser Glu Lys Thr Val Leu Gly Thr Glu Ile Pro Lys Val Arg Thr Val
 995 1000 1005
 60 Leu Gly Gln Lys Leu Lys Met Pro Gln Thr Val Pro Phe Val Tyr
 1010 1015 1020
 65 Ser Asp Gly Ser Arg Glu Lys Arg Pro Val Thr Trp Ser Gln Ala
 1025 1030 1035
 70 Asp Val Ser Gln Ala Gly Ile Val Thr Val Lys Gly Met Ser Asp
 1040 1045 1050
 75 Gly Arg Glu Val Glu Ala Arg Val Glu Val Leu Thr Val Ala Asn
 1055 1060 1065
 80 Glu Leu Pro Ala Val Lys Arg Ile Thr Pro Gly Thr Asp Leu Ser
 1070 1075 1080

ES 2 352 704 B1

Ser Val Asp Lys Leu Val Ser Leu Val Ser Thr Asp Gly Arg Ile
 1085 1090 1095
 5 His Asn Tyr Glu Val Glu Asn Trp Glu Ile Ser Pro Glu Asp Lys
 1100 1105 1110
 10 Glu Lys Leu Ser Thr Pro Gly Ala His Ile Gln Met Thr Ser Gln
 1115 1120 1125
 15 Leu Gly Asp Arg Thr Ile His Ala Thr Leu Ile Val Asp Asp Gly
 1130 1135 1140
 20 Lys Glu Leu Ser Gln Val Thr Pro Asn Ile Ala Val Gly Gly Glu
 1145 1150 1155
 25 Ser Val Thr Glu Leu Ser Lys Glu His Pro Val His Tyr His Lys
 1160 1165 1170
 30 Leu Ser Tyr Gly Ala Lys Ile Pro Glu Val Ser Ala Ser Ala Glu
 1175 1180 1185
 35 Asn Ala Asp Val Asn Val Ile Gln Ala Asn Glu Ser Asn Gly Met
 1190 1195 1200
 40 Arg Ala Ser Ile Tyr Ile Gln Pro Lys Asp Gly Gly Gln Leu Gln
 1205 1210 1215
 45 Thr Tyr Ala Val Gln Phe Leu Val Asp Ser Pro Gln Ile Glu Ser
 1220 1225 1230
 50 Leu Ser Leu Arg Val Asp Gln Val Ser Ser Leu Lys Glu Asp Gln
 1235 1240 1245
 55 Thr Val Lys Val Thr Val Leu Ala Arg Tyr Gln Asp Gly Thr Glu
 1250 1255 1260
 60 Ala Ile Leu Pro Thr Asp Lys Val Thr Phe Ser Ser Gln Gly Glu
 1265 1270 1275
 65 Gly Gly Val Ser Val Asn Lys Gly Met Leu Glu Leu His Lys Pro
 1280 1285 1290
 Gly Ser Leu Thr Leu Gln Ala Gln Tyr Glu Gly Ala Lys Gly Glu
 1295 1300 1305
 60 Leu Pro Leu Thr Ile Ala Ala Asn Thr Glu Asp Lys Val Val Gln
 1310 1315 1320

ES 2 352 704 B1

	Ala	Ile	Arg	Pro	Val	Ser	Leu	Val	Thr	Asp	Leu	His	Gln	Thr	Pro
		1325					1330					1335			
5	Asn	Leu	Pro	Ser	Thr	Val	Thr	Val	Glu	Tyr	Asp	Lys	Gly	Phe	Pro
		1340					1345					1350			
10	Lys	Val	His	Lys	Val	Val	Trp	Asp	Ser	Ile	Pro	Ser	Glu	Lys	Leu
		1355					1360					1365			
15	Asn	Gln	Tyr	His	Thr	Phe	Glu	Val	Leu	Gly	Lys	Val	Glu	Gly	Leu
		1370					1375					1380			
20	Asn	Gln	Glu	Ala	Arg	Ala	Thr	Val	Ser	Val	Glu	Gly	Ile	Val	Lys
		1385					1390					1395			
25	Val	Glu	Asp	Val	Ser	Ile	Thr	Thr	Pro	Ile	Ser	Glu	Ala	Pro	Thr
		1400					1405					1410			
30	Leu	Pro	Glu	Ser	Val	Arg	Thr	Tyr	Tyr	Ser	Asn	Gly	Gln	Val	Ser
		1415					1420					1425			
35	Ser	Ala	Lys	Val	Ala	Trp	Asp	Thr	Ile	Asp	Ala	Ser	Gln	Tyr	Ala
		1430					1435					1440			
40	Lys	Glu	Gly	Ile	Phe	Thr	Val	Thr	Gly	His	Val	Glu	Gly	Thr	Gln
		1445					1450					1455			
45	Leu	Thr	Thr	Lys	Leu	His	Val	Arg	Val	Ser	Ser	Gln	Thr	Glu	Ile
		1460					1465					1470			
50	Gly	Lys	Asn	Ile	Ser	Asp	Gln	Trp	Thr	Gly	Ser	Glu	Leu	Pro	Leu
		1475					1480					1485			
55	Ala	Phe	Ala	Ser	Asp	Ser	His	Ser	Tyr	Asp	Pro	Val	Ser	Asn	Val
		1490					1495					1500			
60	Asn	Asp	Lys	Val	Ile	Ser	Tyr	Gly	Asp	Gln	Pro	Ala	Asn	Arg	Trp
		1505					1510					1515			
65	Ser	Asn	Trp	Lys	Arg	Ser	Glu	Glu	Ala	Ser	Val	Gly	Val	Leu	Phe
		1520					1525					1530			
70	Gly	Asp	Ser	Gly	Ile	Leu	Thr	Lys	Arg	Ser	Val	Asp	Asn	Leu	Asn
		1535					1540					1545			
75	Val	Ala	Phe	Tyr	Glu	Asp	Ser	Gly	Val	Gly	Ala	Pro	Lys	Ser	Tyr
		1550					1555					1560			
80	Val	Ile	Glu	Tyr	Tyr	Val	Gly	Glu	Lys	Ala	Pro	Thr	Ala	Pro	Lys
		1565					1570					1575			

ES 2 352 704 B1

	Asn	Pro	Gly	Tyr	Val	Glu	Gly	Glu	Asn	His	Val	Phe	Asn	Asp	Ser
		1580					1585					1590			
5	Lys	Asn	Trp	Lys	Pro	Val	Thr	Asn	Leu	Lys	Ala	Pro	Asp	Gln	Leu
		1595					1600					1605			
10	Lys	Ala	Gly	Glu	Met	Asn	His	Phe	Ser	Phe	Asp	Lys	Val	Asp	Thr
		1610					1615					1620			
15	Tyr	Ala	Val	Arg	Ile	Arg	Met	Thr	Arg	Ala	Asp	Asp	Lys	Leu	Gly
		1625					1630					1635			
20	Thr	Ser	Ile	Thr	Glu	Leu	Gln	Val	Phe	Ser	Lys	Lys	Val	Ala	Pro
		1640					1645					1650			
25	Ala	Lys	Glu	Ala	Thr	Thr	Asn	Ile	Lys	Val	Ala	Gly	Lys	Asp	Leu
		1655					1660					1665			
30	Pro	His	Phe	Asn	Pro	Asp	Leu	Thr	Asp	Tyr	Tyr	Leu	Asp	Tyr	Ser
		1670					1675					1680			
35	Asp	Gly	Lys	Ile	Pro	Glu	Val	Thr	Ala	Ser	Ile	Thr	Asn	Asn	Gly
		1685					1690					1695			
40	Leu	Ala	Thr	Val	Val	Pro	Ser	Val	Lys	Glu	Gly	Asp	Pro	Val	Arg
		1700					1705					1710			
45	Val	Ile	Val	Lys	Ala	Glu	Asn	Gly	Asp	Ile	Leu	Gly	Glu	Tyr	Asn
		1715					1720					1725			
50	Leu	His	Phe	Thr	Lys	Asp	Lys	Asp	Leu	Leu	Ala	Arg	Lys	Pro	Ile
		1730					1735					1740			
55	Ala	Ser	Ala	Lys	Gln	Ser	Gln	Leu	Leu	Gln	Leu	Gly	Gln	Pro	Leu
		1745					1750					1755			
60	Glu	Leu	Pro	Asn	Lys	Val	Pro	Val	Tyr	Phe	Ala	Gly	Lys	Ser	Gly
		1760					1765					1770			
65	Tyr	Asp	Val	Lys	Asp	Leu	Ala	Val	Glu	Trp	Glu	Lys	Val	Pro	Ala
		1775					1780					1785			
	Ser	Thr	Leu	Asn	Lys	Ala	Gly	Glu	Tyr	Thr	Ile	Lys	Gly	His	Val
		1790					1795					1800			
	Val	Gly	Ser	Glu	Val	Pro	Val	Glu	Leu	Ser	Val	Arg	Val	Thr	Asp
		1805					1810					1815			

	Lys	Val	Gly	Arg	Ser	Leu	Ser	Asp	Asn	Pro	Asp	Tyr	Leu	Ala	Thr
		1820					1825					1830			
5	His	Ser	Gln	Ala	Phe	Ala	Ser	Ala	Thr	Asn	Asp	Leu	Asp	Asp	Asn
		1835					1840					1845			
10	Ser	Asn	Asp	Arg	Val	Asp	Tyr	Leu	Asn	Asp	Asn	Asp	Arg	Asn	Gln
		1850					1855					1860			
15	Asp	Asn	Arg	Trp	Thr	Asn	Trp	Ser	Ala	Thr	Pro	Ser	Ala	Asn	Pro
		1865					1870					1875			
20	Glu	Val	Ser	Val	Gly	Val	Ile	Phe	Lys	Lys	Asn	Gly	Lys	Ile	Val
		1880					1885					1890			
25	Asp	Arg	Lys	Val	Ser	Gln	Ala	Lys	Leu	Asn	Phe	Phe	Ala	Asp	Ser
		1895					1900					1905			
30	Gly	Thr	Asp	Ala	Pro	Ala	Lys	Leu	Val	Leu	Glu	Arg	Tyr	Ile	Gly
		1910					1915					1920			
35	Pro	Asp	Phe	Asp	Val	Pro	Val	Tyr	Tyr	Ser	Asn	Tyr	Ser	Tyr	Glu
		1925					1930					1935			
40	Ser	Asn	His	Pro	Phe	Asn	Asn	Pro	Asp	Asn	Trp	Glu	Leu	Val	Pro
		1940					1945					1950			
45	Tyr	Tyr	Ala	Asp	Lys	Glu	Ile	Gln	Ala	Gly	Asn	Glu	Ile	Asn	Val
		1955					1960					1965			
50	Thr	Phe	Lys	Ala	Val	Thr	Thr	Lys	Ala	Met	Arg	Trp	Arg	Met	Asp
		1970					1975					1980			
55	Arg	Lys	Ala	Asp	Lys	Asn	Gly	Val	Ala	Leu	Thr	Glu	Leu	Ser	Phe
		1985					1990					1995			
60	Arg	Ala	Pro	Ser	Glu	Leu	Ala	Lys	Glu	Ser	Thr	Glu	Gly	Arg	Ile
		2000					2005					2010			
65	Leu	Val	Asp	Gly	Lys	Glu	Leu	Pro	Asp	Phe	Ser	Glu	Asn	Arg	Leu
		2015					2020					2025			
70	Asp	Tyr	Glu	Leu	Ile	Tyr	Lys	Gly	Glu	Arg	Pro	Lys	Ile	Thr	Val
		2030					2035					2040			
75	Glu	Gly	Lys	Asp	Gln	Val	Thr	Ser	Thr	Ile	Val	Asp	Ser	Gly	Asn
		2045					2050					2055			
80	Glu	Asn	Leu	Pro	Val	Leu	Val	Arg	Leu	Val	Ser	Glu	Ser	Gly	Lys
		2060					2065					2070			

	Asn	Val	Lys	Glu	Tyr	Arg	Ile	Lys	Leu	Thr	Lys	Glu	Lys	Val	Ile
		2075					2080					2085			
5	Asp	Ala	Lys	Leu	Val	Ala	Pro	Ala	Gln	Tyr	Asp	Leu	Pro	Ser	Leu
		2090					2095					2100			
10	Glu	Val	Val	Asp	Lys	Glu	Leu	Asn	Phe	Gln	Thr	Leu	Glu	Gln	Lys
		2105					2110					2115			
15	Asp	Asp	Thr	Leu	Phe	Glu	Gly	Glu	Val	Arg	Leu	Leu	Gln	Glu	Gly
		2120					2125					2130			
20	Lys	Val	Gly	Lys	Glu	Arg	Ile	Tyr	Thr	Glu	Val	Thr	Thr	Asp	Gly
		2135					2140					2145			
25	Lys	Asn	Lys	Glu	Lys	Ser	Arg	Glu	Ile	Leu	Glu	Glu	Pro	Val	Asn
		2150					2155					2160			
30	Arg	Ile	Leu	Leu	Val	Gly	Thr	Lys	Lys	Lys	Thr	Ser	Asn	Ile	Glu
		2165					2170					2175			
35	Thr	Pro	Ala	Gly	Gly	His	Gln	Asp	Asn	Gly	Ser	Ser	Gln	Pro	Thr
		2180					2185					2190			
40	Glu	Gln	Lys	Pro	Gly	Ala	Asn	Pro	Asn	Pro	Asn	Ile	Val	Ile	Pro
		2195					2200					2205			
45	Asn	Pro	Ser	Asn	Asn	Gly	Val	Thr	Ser	Asn	Ser	Ser	Asp	Phe	Lys
		2210					2215					2220			
50	Asn	His	Leu	Ser	Val	Ala	Pro	Asn	Ile	Val	Ile	Gln	Ala	Pro	Thr
		2225					2230					2235			
55	Gln	Glu	Ser	Lys	Lys	Pro	Gly	Trp	Asn	Lys	Glu	Asp	Gly	Lys	Trp
		2240					2245					2250			
60	Tyr	Tyr	Arg	Gln	Gln	Asn	Gly	Asp	Leu	Ala	Lys	Gly	Trp	Val	Lys
		2255					2260					2265			
65	Asp	Gly	Asp	Thr	Trp	Tyr	Tyr	Phe	Asp	Gln	Thr	Gly	Lys	Met	Glu
		2270					2275					2280			
70	Thr	Gly	Trp	Ile	Lys	Asp	Gln	Ser	Gly	Ala	Trp	Tyr	Tyr	Leu	Asn
		2285					2290					2295			
75	Gln	Ser	Gly	Ala	Met	Ser	Ser	Asn	Glu	Trp	Ile	Leu	Asp	Gln	Asp
		2300					2305					2310			

Gly Lys Trp Tyr Tyr Val Asp Ala Ser Gly Ser Met Lys Thr Ser
 2315 2320 2325
 5 Gln Trp Phe Gln Val Gly Asp Lys Trp Tyr Tyr Val Asp Ala Asn
 2330 2335 2340
 10 Gly Ser Leu Ala Val Asn Thr Val Thr Pro Asp Gly Tyr Gln Val
 2345 2350 2355
 Asn Glu Asn Gly Glu Trp Ile Gly
 2360 2365
 15 <210> 5
 <211> 7119
 <212> DNA
 20 <213> Artificial
 <220>
 <223> beta-galactosidasa sin péptido señal con poli-histidina
 25 <220>
 <221> CDS
 <222> (1)..(7119)
 30 <223> beta-galactosidasa sin péptido señal con poli-histidina
 <220>
 <221> misc_feature
 35 <222> (4)..(21)
 <223> Etiqueta poli-his
 <400> 5
 40 atg cac cac cac cac cac cac gaa aat gct gaa gaa att gtg tta aca
 48 Met His His His His His His Glu Asn Ala Glu Glu Ile Val Leu Thr
 45 1 5 10 15
 aat agg act gag cac cag tct gaa aac ctg aaa act act gga ctc tca
 96 Asn Arg Thr Glu His Gln Ser Glu Asn Leu Lys Thr Thr Gly Leu Ser
 50 20 25 30
 agt gaa cat aaa gct gaa ggt gtt acg act gaa aaa gca agt gag caa
 144 Ser Glu His Lys Ala Glu Gly Val Thr Thr Glu Lys Ala Ser Glu Gln
 55 35 40 45
 cct gtt aaa gct gaa aat gta tct cta aag gaa aat gcg gat ggt agt
 192 Pro Val Lys Ala Glu Asn Val Ser Leu Lys Glu Asn Ala Asp Gly Ser
 60 50 55 60
 65 gat gca gta gta aaa cct cac gca gat caa cct tta gtt tca gtt gcc
 240

ES 2 352 704 B1

Asp Ala Val Val Lys Pro His Ala Asp Gln Pro Leu Val Ser Val Ala
 65 70 75 80
 5
 aat aat caa tca gtt cct tta gaa gaa ggt aag gct gaa gtc gaa aac
 288
 Asn Asn Gln Ser Val Pro Leu Glu Glu Gly Lys Ala Glu Val Glu Asn
 10 85 90 95
 aaa aca gaa gaa agg aac aag cct gag gat aga aat aag ctt gaa gaa
 336
 15 Lys Thr Glu Glu Arg Asn Lys Pro Glu Asp Arg Asn Lys Leu Glu Glu
 100 105 110
 aca aat aag act gag gta acg gat aaa tct gaa gat aaa aat aaa tct
 384
 20 Thr Asn Lys Thr Glu Val Thr Asp Lys Ser Glu Asp Lys Asn Lys Ser
 115 120 125
 25
 gaa gaa aca aat aaa gca gaa tca caa gat cgt tct gtt tca aaa gat
 432
 30 Glu Glu Thr Asn Lys Ala Glu Ser Gln Asp Arg Ser Val Ser Lys Asp
 130 135 140
 aag caa gaa ttc aaa tct gct acg aat gaa gta gtt gat aag tta att
 480
 35 Lys Gln Glu Phe Lys Ser Ala Thr Asn Glu Val Val Asp Lys Leu Ile
 145 150 155 160
 40
 gaa gat aga aac att tcc ttc aac cag aat tgg cac ttt aaa cta aat
 528
 45 Glu Asp Arg Asn Ile Ser Phe Asn Gln Asn Trp His Phe Lys Leu Asn
 165 170 175
 gcc aat gca aag gaa gct gta aaa cca gac gca gat att tct tca tgg
 576
 50 Ala Asn Ala Lys Glu Ala Val Lys Pro Asp Ala Asp Ile Ser Ser Trp
 180 185 190
 aaa aaa atg gat ctt cct cat gac tgg agt att cat ttt gat ttt gac
 624
 55 Lys Lys Met Asp Leu Pro His Asp Trp Ser Ile His Phe Asp Phe Asp
 195 200 205
 60
 cat gat tca cca gct caa aac gaa ggc ggt caa ttg aat ggt gga gat
 672
 65 His Asp Ser Pro Ala Gln Asn Glu Gly Gly Gln Leu Asn Gly Gly Asp
 210 215 220

ES 2 352 704 B1

	ggc	tgg	tat	cgg	aaa	acc	ttt	aag	tta	gat	gaa	aag	gac	ttg	aat	aaa
	720 Gly	Trp	Tyr	Arg	Lys	Thr	Phe	Lys	Leu	Asp	Glu	Lys	Asp	Leu	Asn	Lys
5	225					230					235					240
	gat	gtt	cgt	gta	acc	ttt	gat	ggg	gtt	tac	atg	gat	tca	caa	gta	ttc
10	768 Asp	Val	Arg	Val	Thr	Phe	Asp	Gly	Val	Tyr	Met	Asp	Ser	Gln	Val	Phe
					245					250					255	
15	gtc	aat	ggc	caa	cta	gtt	ggg	cac	tat	cca	aat	ggg	tac	aat	caa	ttt
	816 Val	Asn	Gly	Gln	Leu	Val	Gly	His	Tyr	Pro	Asn	Gly	Tyr	Asn	Gln	Phe
				260					265					270		
20																
	tca	tat	gac	att	agt	aac	tat	ctc	cat	aaa	gat	ggg	cga	gaa	aat	gtc
25	864 Ser	Tyr	Asp	Ile	Ser	Asn	Tyr	Leu	His	Lys	Asp	Gly	Arg	Glu	Asn	Val
			275					280					285			
30	att	tct	gtt	cat	gca	gtt	aat	aag	caa	cca	agt	agt	aga	tgg	tat	tca
	912 Ile	Ser	Val	His	Ala	Val	Asn	Lys	Gln	Pro	Ser	Ser	Arg	Trp	Tyr	Ser
		290						295				300				
35																
	gga	agt	ggg	atc	tat	cga	gac	gta	tca	ttg	cag	gta	aca	gac	aag	att
40	960 Gly	Ser	Gly	Ile	Tyr	Arg	Asp	Val	Ser	Leu	Gln	Val	Thr	Asp	Lys	Ile
	305					310					315					320
45																
	cat	gtt	gaa	aag	aat	ggg	aca	act	att	cta	act	cca	aaa	cta	gaa	agt
	1008 His	Val	Glu	Lys	Asn	Gly	Thr	Thr	Ile	Leu	Thr	Pro	Lys	Leu	Glu	Ser
					325					330					335	
50																
	caa	caa	gga	gga	aaa	gtt	gaa	acc	caa	gta	agc	agt	aaa	att	gtt	aat
	1056 Gln	Gln	Gly	Gly	Lys	Val	Glu	Thr	Gln	Val	Ser	Ser	Lys	Ile	Val	Asn
				340					345					350		
55																
	aca	gat	aat	aaa	gac	cat	gaa	att	gta	gca	gag	tat	caa	att	ttt	gaa
60	1104 Thr	Asp	Asn	Lys	Asp	His	Glu	Ile	Val	Ala	Glu	Tyr	Gln	Ile	Phe	Glu
			355					360					365			
	cga	ggg	ggg	aag	gca	gtt	aca	gaa	ctt	gtt	cga	act	gag	agt	aaa	act

ES 2 352 704 B1

1152
Arg Gly Gly Lys Ala Val Thr Glu Leu Val Arg Thr Glu Ser Lys Thr
370 375 380

5

cta aaa gct aag gaa acc att cat ttg gat tca gca tta gag gtt gaa
1200
Leu Lys Ala Lys Glu Thr Ile His Leu Asp Ser Ala Leu Glu Val Glu
385 390 395 400

10

aaa cca aaa ttg tgg aca gtg tca tct gac aaa cct gct tta tat gag
1248
Lys Pro Lys Leu Trp Thr Val Ser Ser Asp Lys Pro Ala Leu Tyr Glu
405 410 415

15

atg gta act cgt gtt tac aaa gat ggt cag ctt gta gat gca aag aaa
1296
Met Val Thr Arg Val Tyr Lys Asp Gly Gln Leu Val Asp Ala Lys Lys
420 425 430

20

gat tta ttt ggt tat cgc tac tat aac tgg act cca gac caa ggg ttc
1344
Asp Leu Phe Gly Tyr Arg Tyr Tyr Asn Trp Thr Pro Asp Gln Gly Phe
435 440 445

25

tct tta aat ggt gaa cac atc aag ttc cac ggt gtt tca ttg cac cat
1392
Ser Leu Asn Gly Glu His Ile Lys Phe His Gly Val Ser Leu His His
450 455 460

30

gat cat gga gcg cta gga gca gaa gaa aac tat aag gca gaa tat cgt
1440
Asp His Gly Ala Leu Gly Ala Glu Glu Asn Tyr Lys Ala Glu Tyr Arg
465 470 475 480

35

cgt ctg aaa caa atg aag gaa atg gga gtg aat tcg att cgt acg act
1488
Arg Leu Lys Gln Met Lys Glu Met Gly Val Asn Ser Ile Arg Thr Thr
485 490 495

40

cac aac cca gca agt cct cag acc ttg cag att gca gct gag ctc ggt
1536
His Asn Pro Ala Ser Pro Gln Thr Leu Gln Ile Ala Ala Glu Leu Gly
500 505 510

45

ttg ctt gtt caa gaa gaa gct ttt gat act tgg tat gga ggt aag aaa
1584
Leu Leu Val Gln Glu Glu Ala Phe Asp Thr Trp Tyr Gly Gly Lys Lys

50

55

60

65

ES 2 352 704 B1

	515	520	525
5	cca tac gac tat ggt cgt ttc ttt gaa aaa gat gct acc cac cct gaa 1632 Pro Tyr Asp Tyr Gly Arg Phe Phe Glu Lys Asp Ala Thr His Pro Glu		
	530	535	540
10	gca aga aaa ggt gaa aaa tgg tcc gac tac gat ctt cga aca atg gta 1680 Ala Arg Lys Gly Glu Lys Trp Ser Asp Tyr Asp Leu Arg Thr Met Val		
15	545	550	555 560
20	gaa aga gat aaa aac aat cct gct gtt gtt atg tgg tct att ggt aac 1728 Glu Arg Asp Lys Asn Asn Pro Ala Val Val Met Trp Ser Ile Gly Asn	565	570 575
25	gaa atc ggt gaa gca gat ggg aaa gcc cat tca ctg gta act gtc aaa 1776 Glu Ile Gly Glu Ala Asp Gly Lys Ala His Ser Leu Val Thr Val Lys	580	585 590
30			
35	cgt ttg gtc aaa gtg gtg aaa tct gtt gat aag aca cgt tat gta acc 1824 Arg Leu Val Lys Val Val Lys Ser Val Asp Lys Thr Arg Tyr Val Thr	595	600 605
40	atg gga gct gat aaa ttc cgt ttc gga gat gga act ggg gac cat gaa 1872 Met Gly Ala Asp Lys Phe Arg Phe Gly Asp Gly Thr Gly Asp His Glu	610	615 620
45	aag att gca gat gaa ctg gat gca gtt gga ttg aat tac tct gaa gaa 1920 Lys Ile Ala Asp Glu Leu Asp Ala Val Gly Leu Asn Tyr Ser Glu Glu		
50	625	630	635 640
55	aat tat caa act ctt cat gcg aaa cat ccg aaa tgg ctc atc tat ggt 1968 Asn Tyr Gln Thr Leu His Ala Lys His Pro Lys Trp Leu Ile Tyr Gly	645	650 655
60	tca gaa aca tct tca gca act cgt aca aga gga agt tat ttc cat cca 2016 Ser Glu Thr Ser Ser Ala Thr Arg Thr Arg Gly Ser Tyr Phe His Pro	660	665 670
65			

ES 2 352 704 B1

gag agt gag tgg gta gga agt aac caa tat tgg cgt aat tac gaa caa
 2064
 Glu Ser Glu Trp Val Gly Ser Asn Gln Tyr Trp Arg Asn Tyr Glu Gln
 5 675 680 685

 tcc gat tat ggt aat gat cgt gtt ggt tgg ggt aaa act gct aca gca
 2112
 Ser Asp Tyr Gly Asn Asp Arg Val Gly Trp Gly Lys Thr Ala Thr Ala
 10 690 695 700

 tca tgg acc ttc gat aga gat cat cca gga tac gct gga caa ttc att
 2160
 Ser Trp Thr Phe Asp Arg Asp His Pro Gly Tyr Ala Gly Gln Phe Ile
 15 705 710 715 720

 tgg acc ggt acg gat tat att ggt gaa cca aca cca tgg cac aat caa
 2208
 Trp Thr Gly Thr Asp Tyr Ile Gly Glu Pro Thr Pro Trp His Asn Gln
 25 725 730 735

 aat agt aca cca gta aaa agt tcc tac ttt ggt att gtt gat acc gca
 2256
 Asn Ser Thr Pro Val Lys Ser Ser Tyr Phe Gly Ile Val Asp Thr Ala
 30 740 745 750

 ggt att cca aaa aat gac tat tat ctc tac caa agt cag tgg gtt tct
 2304
 Gly Ile Pro Lys Asn Asp Tyr Tyr Leu Tyr Gln Ser Gln Trp Val Ser
 35 755 760 765

 gcc aag aaa aaa cca atg gtt cat tta ctt cct cac tgg aac tgg gaa
 2352
 Ala Lys Lys Lys Pro Met Val His Leu Leu Pro His Trp Asn Trp Glu
 45 770 775 780

 aag act gaa ttg gca gat aat gtt gca gac gct gaa aat cga att cca
 2400
 Lys Thr Glu Leu Ala Asp Asn Val Ala Asp Ala Glu Asn Arg Ile Pro
 50 785 790 795 800

 gtt aga gct tat tcc aat gct gcc agt gtt gag ttg ttc tta aat aat
 2448
 Val Arg Ala Tyr Ser Asn Ala Ala Ser Val Glu Leu Phe Leu Asn Asn
 55 805 810 815

 gag tca cta ggt ctt aaa aaa ttc aac aaa aaa caa aca agt gac ggt
 2496
 Glu Ser Leu Gly Leu Lys Lys Phe Asn Lys Lys Gln Thr Ser Asp Gly
 65

ES 2 352 704 B1

	820	825	830
5	cga aca tat caa gaa ggt gaa aat cct caa gaa cta tac ttg gag tgg 2544 Arg Thr Tyr Gln Glu Gly Glu Asn Pro Gln Glu Leu Tyr Leu Glu Trp		
	835	840	845
10	aaa gta gct tat aaa ccg gga act ttg gaa gct gta gct cga gat gag 2592 Lys Val Ala Tyr Lys Pro Gly Thr Leu Glu Ala Val Ala Arg Asp Glu		
15	850	855	860
20	tct ggc aag gaa att gct cgt gat aag att gta act gct ggt gaa cct 2640 Ser Gly Lys Glu Ile Ala Arg Asp Lys Ile Val Thr Ala Gly Glu Pro		
	865	870	875 880
25	gca ggt gtc cgt ttg gtt aag gaa gaa aat gca atc gct gca gat gga 2688 Ala Gly Val Arg Leu Val Lys Glu Glu Asn Ala Ile Ala Ala Asp Gly		
30	885	890	895
35	aaa gat ttg act tat atc tat tac gaa att gta gat agc aag ggt aat 2736 Lys Asp Leu Thr Tyr Ile Tyr Tyr Glu Ile Val Asp Ser Lys Gly Asn		
	900	905	910
40	gtt gtg cca act gct aac aat cta gtt cac ttc caa ctt cat gga caa 2784 Val Val Pro Thr Ala Asn Asn Leu Val His Phe Gln Leu His Gly Gln		
	915	920	925
45	ggc caa ctg gtt ggt gtg gat aat ggt gag caa gct agt cgt gaa cgt 2832 Gly Gln Leu Val Gly Val Asp Asn Gly Glu Gln Ala Ser Arg Glu Arg		
50	930	935	940
55	tat aaa gaa caa gca gat ggt tct tgg att cgt aga gcc ttt aat ggt 2880 Tyr Lys Glu Gln Ala Asp Gly Ser Trp Ile Arg Arg Ala Phe Asn Gly		
	945	950	955 960
60	aaa ggg gta gcg att gtt aaa tcg act gat caa gca ggt aaa ttt acc 2928 Lys Gly Val Ala Ile Val Lys Ser Thr Asp Gln Ala Gly Lys Phe Thr		
65	965	970	975

ctt aca gca tat tca ggc ttg ttg aaa tca gat caa gtt aca gtt ttt
 2976
 Leu Thr Ala Tyr Ser Gly Leu Leu Lys Ser Asp Gln Val Thr Val Phe
 5 980 985 990

 aca ggt aag tca gac cag tca gaa aaa act gtt tta gga aca gaa ata
 3024
 Thr Gly Lys Ser Asp Gln Ser Glu Lys Thr Val Leu Gly Thr Glu Ile
 10 995 1000 1005

 cca aaa gtt cgg aca gtt tta gga caa aaa ctt aaa atg cct cag
 3069
 Pro Lys Val Arg Thr Val Leu Gly Gln Lys Leu Lys Met Pro Gln
 15 1010 1015 1020

 aca gtg cca ttt gta tat agt gat ggt agt cgt gaa aaa cgt cct
 3114
 Thr Val Pro Phe Val Tyr Ser Asp Gly Ser Arg Glu Lys Arg Pro
 20 1025 1030 1035

 gta acc tgg tct caa gca gat gtc agt caa gct gga att gta act
 3159
 Val Thr Trp Ser Gln Ala Asp Val Ser Gln Ala Gly Ile Val Thr
 25 1040 1045 1050

 gta aaa ggt atg tct gat gga cgt gaa gtt gag gct cgt gtt gaa
 3204
 Val Lys Gly Met Ser Asp Gly Arg Glu Val Glu Ala Arg Val Glu
 30 1055 1060 1065

 gta ctt aca gtt gcg aat gaa ctt cca gct gtt aaa cgc atc aca
 3249
 Val Leu Thr Val Ala Asn Glu Leu Pro Ala Val Lys Arg Ile Thr
 35 1070 1075 1080

 cct gga aca gac ttg agt tct gtg gat aaa ctt gtt tca ctt gtt
 3294
 Pro Gly Thr Asp Leu Ser Ser Val Asp Lys Leu Val Ser Leu Val
 40 1085 1090 1095

 tca aca gat ggt cgt att cac aat tat gaa gtt gaa aat tgg gaa
 3339
 Ser Thr Asp Gly Arg Ile His Asn Tyr Glu Val Glu Asn Trp Glu
 45 1100 1105 1110

 att tca cca gaa gat aaa gag aag tta tct act ccg ggg gca cat
 3384
 50 55 60 65

ES 2 352 704 B1

	Ile Ser	Pro Glu Asp Lys Glu	Lys Leu Ser Thr Pro	Gly Ala His
	1115	1120	1125	
5				
	att cag	atg act agc caa tta	gga gat aga acc atc	cac gct act
	3429			
	Ile Gln	Met Thr Ser Gln Leu	Gly Asp Arg Thr Ile	His Ala Thr
10	1130	1135	1140	
	ttg ata	gtt gat gat ggt aag	gaa ttg tct caa gtt	aca cca aat
	3474			
15	Leu Ile	Val Asp Asp Gly Lys	Glu Leu Ser Gln Val	Thr Pro Asn
	1145	1150	1155	
20	ata gca	gta ggt ggt gaa agt	gtt acg gaa ctt tca	aaa gag cat
	3519			
	Ile Ala	Val Gly Gly Glu Ser	Val Thr Glu Leu Ser	Lys Glu His
	1160	1165	1170	
25				
	cct gta	cat tat cac aaa ctt	tct tat gga gca aaa	att cca gaa
	3564			
	Pro Val	His Tyr His Lys Leu	Ser Tyr Gly Ala Lys	Ile Pro Glu
30	1175	1180	1185	
35	gtc agt	gct agt gct gaa aat	gcc gat gtg aac gtg	att cag gca
	3609			
	Val Ser	Ala Ser Ala Glu Asn	Ala Asp Val Asn Val	Ile Gln Ala
	1190	1195	1200	
40	aat gaa	tcc aat gga atg cgt	gct agt atc tac att	cag cct aaa
	3654			
	Asn Glu	Ser Asn Gly Met Arg	Ala Ser Ile Tyr Ile	Gln Pro Lys
45	1205	1210	1215	
50	gac ggg	gga cag tta caa acc	tat gct gtt cag ttc	ctt gtt gat
	3699			
	Asp Gly	Gly Gln Leu Gln Thr	Tyr Ala Val Gln Phe	Leu Val Asp
	1220	1225	1230	
55	tct cca	caa att gaa agc ttg	agt ctc aga gtg gat	caa gta tct
	3744			
	Ser Pro	Gln Ile Glu Ser Leu	Ser Leu Arg Val Asp	Gln Val Ser
	1235	1240	1245	
60				
	agc ctc	aaa gaa gat caa act	gtt aaa gtg aca gtt	tta gcc cgt
	3789			
	Ser Leu	Lys Glu Asp Gln Thr	Val Lys Val Thr Val	Leu Ala Arg
65	1250	1255	1260	

	tat cag gat gga act gaa gct att tta cca act gat aag gtg act	
	3834 Tyr Gln Asp Gly Thr Glu Ala Ile Leu Pro Thr Asp Lys Val Thr	
5	1265	1270 1275
	ttc tca agc caa ggc gag gga gga gtc tct gtt aat aag gga atg	
	3879 Phe Ser Ser Gln Gly Glu Gly Gly Val Ser Val Asn Lys Gly Met	
10	1280	1285 1290
	cta gag ttg cat aaa ccg ggt agt ctc act tta caa gct caa tat	
	3924 Leu Glu Leu His Lys Pro Gly Ser Leu Thr Leu Gln Ala Gln Tyr	
15	1295	1300 1305
	gaa ggt gca aaa gga gaa ctt cct cta acg atc gca gca aat act	
	3969 Glu Gly Ala Lys Gly Glu Leu Pro Leu Thr Ile Ala Ala Asn Thr	
20	1310	1315 1320
	gaa gat aaa gtt gtt cag gcc att cgc cca gtt agc ctt gtg aca	
	4014 Glu Asp Lys Val Val Gln Ala Ile Arg Pro Val Ser Leu Val Thr	
25	1325	1330 1335
	gat ttg cac caa aca cca aat ctt cca tct aca gta aca gtt gag	
	4059 Asp Leu His Gln Thr Pro Asn Leu Pro Ser Thr Val Thr Val Glu	
30	1340	1345 1350
	tac gat aaa gga ttt cca aaa gtt cat aag gta gtt tgg gat tcg	
	4104 Tyr Asp Lys Gly Phe Pro Lys Val His Lys Val Val Trp Asp Ser	
35	1355	1360 1365
	att cca tct gaa aaa tta aat caa tat cat act ttt gaa gta ctt	
	4149 Ile Pro Ser Glu Lys Leu Asn Gln Tyr His Thr Phe Glu Val Leu	
40	1370	1375 1380
	ggt aag gtt gaa gga ttg aat caa gaa gct cgt gcg aca gta tca	
	4194 Gly Lys Val Glu Gly Leu Asn Gln Glu Ala Arg Ala Thr Val Ser	
45	1385	1390 1395
	gta gaa gga att gta aaa gtt gaa gat gtg agc atc aca aca cca	
50		
55		
60		
65		

	4239	Val Glu Gly Ile Val Lys Val Glu Asp Val Ser Ile Thr Thr Pro
5	1400	1405 1410
	att tct gag gca cca acc ttg cca gaa agt gtc cga act tat tat	
10	4284	Ile Ser Glu Ala Pro Thr Leu Pro Glu Ser Val Arg Thr Tyr Tyr
	1415	1420 1425
	tca aat ggc caa gta tct tca gct aag gtg gct tgg gac act att	
15	4329	Ser Asn Gly Gln Val Ser Ser Ala Lys Val Ala Trp Asp Thr Ile
	1430	1435 1440
20		
	gat gca agc cag tat gca aaa gaa gga atc ttt act gta aca ggc	
25	4374	Asp Ala Ser Gln Tyr Ala Lys Glu Gly Ile Phe Thr Val Thr Gly
	1445	1450 1455
	cat gta gaa gga aca caa ctg aca acg aaa tta cat gtt cgt gta	
30	4419	His Val Glu Gly Thr Gln Leu Thr Thr Lys Leu His Val Arg Val
	1460	1465 1470
35		
	tcg agt caa aca gaa att ggg aag aat att tca gat caa tgg aca	
40	4464	Ser Ser Gln Thr Glu Ile Gly Lys Asn Ile Ser Asp Gln Trp Thr
	1475	1480 1485
	ggg tca gaa tta cca ctt gcc ttt gca tca gat tct cat tct tat	
45	4509	Gly Ser Glu Leu Pro Leu Ala Phe Ala Ser Asp Ser His Ser Tyr
	1490	1495 1500
	gac cca gta tct aat gta aat gat aaa gtg att tct tac ggt gat	
50	4554	Asp Pro Val Ser Asn Val Asn Asp Lys Val Ile Ser Tyr Gly Asp
	1505	1510 1515
55		
	cag cct gca aat cgt tgg tca aac tgg aaa cgt tct gaa gaa gct	
60	4599	Gln Pro Ala Asn Arg Trp Ser Asn Trp Lys Arg Ser Glu Glu Ala
	1520	1525 1530
	tca gta ggg gta ctc ttt gga gat tca ggt att ttg aca aaa cgt	
65	4644	Ser Val Gly Val Leu Phe Gly Asp Ser Gly Ile Leu Thr Lys Arg

ES 2 352 704 B1

	1535	1540	1545
5	tcg gtc 4689 Ser Val	gat aac ttg aat gta Asp Asn Leu Asn Val	gca ttt tat gaa gat agc ggt gta Ala Phe Tyr Glu Asp Ser Gly Val
	1550	1555	1560
10	ggt gct 4734 Gly Ala	cca aaa tct tac gtc Pro Lys Ser Tyr Val	att gag tat tat gtt ggt gaa aaa Ile Glu Tyr Tyr Val Gly Glu Lys
15	1565	1570	1575
20	gct cca 4779 Ala Pro	act gca ccg aag aat Thr Ala Pro Lys Asn	cct ggt tat gta gag ggt gaa aac Pro Gly Tyr Val Glu Gly Glu Asn
	1580	1585	1590
25	cac gta 4824 His Val	ttt aat gac agt aaa Phe Asn Asp Ser Lys	aat tgg aaa cca gtc aca aat ctt Asn Trp Lys Pro Val Thr Asn Leu
30	1595	1600	1605
35	aag gct 4869 Lys Ala	cca gat cag tta aaa Pro Asp Gln Leu Lys	gct gga gaa atg aat cat ttc agt Ala Gly Glu Met Asn His Phe Ser
	1610	1615	1620
40	ttt gat 4914 Phe Asp	aag gtg gat acc tac Lys Val Asp Thr Tyr	gcg gta cgt att cgc atg aca aga Ala Val Arg Ile Arg Met Thr Arg
	1625	1630	1635
45	gct gat 4959 Ala Asp	gat aaa ctt gga aca Asp Lys Leu Gly Thr	tcc att aca gaa tta caa gtt ttc Ser Ile Thr Glu Leu Gln Val Phe
50	1640	1645	1650
55	tct aaa 5004 Ser Lys	aag gta gca cca gct Lys Val Ala Pro Ala	aaa gag gca act act aac att aag Lys Glu Ala Thr Thr Asn Ile Lys
	1655	1660	1665
60	gtt gct 5049 Val Ala	gga aaa gat ctt cct Gly Lys Asp Leu Pro	cac ttt aat cca gac tta aca gat His Phe Asn Pro Asp Leu Thr Asp
65	1670	1675	1680

ES 2 352 704 B1

	tac tat	tta gac tat agt gat	gga aaa att cca gaa	gta aca gct
	5094			
	Tyr Tyr	Leu Asp Tyr Ser Asp	Gly Lys Ile Pro Glu	Val Thr Ala
5		1685	1690	1695
	agt atc	aca aat aat gga ctt	gct aca gtg gtt cct	agc gtg aaa
	5139			
10	Ser Ile	Thr Asn Asn Gly Leu	Ala Thr Val Val Pro	Ser Val Lys
		1700	1705	1710
	gag ggc	gat cct gtc cgt gtt	att gtg aaa gca gaa	aat ggc gat
	5184			
15	Glu Gly	Asp Pro Val Arg Val	Ile Val Lys Ala Glu	Asn Gly Asp
		1715	1720	1725
20				
	att ctt	gga gaa tac aac ctt	cac ttt aca aaa gat	aaa gat ttg
	5229			
25	Ile Leu	Gly Glu Tyr Asn Leu	His Phe Thr Lys Asp	Lys Asp Leu
		1730	1735	1740
	ttg gct	cgt aaa cca att gct	tct gca aaa cag tct	caa ttg ttg
	5274			
30	Leu Ala	Arg Lys Pro Ile Ala	Ser Ala Lys Gln Ser	Gln Leu Leu
		1745	1750	1755
35				
	cag tta	ggg caa cct ttg gaa	ttg cca aat aag gtt	cca gtt tat
	5319			
40	Gln Leu	Gly Gln Pro Leu Glu	Leu Pro Asn Lys Val	Pro Val Tyr
		1760	1765	1770
	ttc gca	ggg aag agc ggt tat	gac gta aaa gat ttg	gca gtt gaa
	5364			
45	Phe Ala	Gly Lys Ser Gly Tyr	Asp Val Lys Asp Leu	Ala Val Glu
		1775	1780	1785
	tgg gaa	aaa gtc cct gct tct	acc tta aac aaa gct	ggg gaa tat
	5409			
50	Trp Glu	Lys Val Pro Ala Ser	Thr Leu Asn Lys Ala	Gly Glu Tyr
		1790	1795	1800
55				
	aca att	aaa ggc cat gta gtt	ggg agt gag gtc cca	gta gaa tta
	5454			
60	Thr Ile	Lys Gly His Val Val	Gly Ser Glu Val Pro	Val Glu Leu
		1805	1810	1815
	tct gtt	cga gta acg gac aag	gtc ggt aga tca ctt	tct gat aat
	5499			
65	Ser Val	Arg Val Thr Asp Lys	Val Gly Arg Ser Leu	Ser Asp Asn

ES 2 352 704 B1

	1820	1825	1830
5	cca gat tat ctt gct act cat 5544 Pro Asp Tyr Leu Ala Thr His	agt cag gct ttt gct Ser Gln Ala Phe Ala	tca gct act Ser Ala Thr
	1835	1840	1845
10	aat gat ttg gat gat aat tcg 5589 Asn Asp Leu Asp Asp Asn Ser	aac gac cgt gta gac Asn Asp Arg Val Asp	tat tta aat Tyr Leu Asn
15	1850	1855	1860
20	gat aat gac cga aat caa gac 5634 Asp Asn Asp Arg Asn Gln Asp	aac cgt tgg aca aac Asn Arg Trp Thr Asn	tgg tcc gct Trp Ser Ala
	1865	1870	1875
25	act cca tct gct aat cca gaa 5679 Thr Pro Ser Ala Asn Pro Glu	gta tca gtt ggt gtt Val Ser Val Gly Val	atc ttt aaa Ile Phe Lys
30	1880	1885	1890
35	aag aat ggt aaa att gtg gat 5724 Lys Asn Gly Lys Ile Val Asp	aga aag gtt tca caa Arg Lys Val Ser Gln	gct aaa cta Ala Lys Leu
	1895	1900	1905
40	aac ttc ttt gca gat agt gga 5769 Asn Phe Phe Ala Asp Ser Gly	aca gat gcc cca gct Thr Asp Ala Pro Ala	aaa ctc gta Lys Leu Val
	1910	1915	1920
45	ttg gaa cga tac atc ggt cct 5814 Leu Glu Arg Tyr Ile Gly Pro	gac ttt gat gtt ccg Asp Phe Asp Val Pro	gtt tac tat Val Tyr Tyr
50	1925	1930	1935
55	tca aac tac agc tac gaa tca 5859 Ser Asn Tyr Ser Tyr Glu Ser	aat cat cca ttt aat Asn His Pro Phe Asn	aac cct gat Asn Pro Asp
	1940	1945	1950
60	aat tgg gaa ctt gtt cct tat 5904 Asn Trp Glu Leu Val Pro Tyr	tac gca gac aaa gaa Tyr Ala Asp Lys Glu	att caa gct Ile Gln Ala
65	1955	1960	1965

ES 2 352 704 B1

	gga aat 5949 Gly Asn	gag atc aat gtg acc Glu Ile Asn Val Thr	ttt aaa gct gtg act Phe Lys Ala Val Thr	acc aaa gct Thr Lys Ala
5	1970	1975	1980	
	atg aga 5994 Met Arg	tgg cgt atg gac aga Trp Arg Met Asp Arg	aaa gca gat aaa aat Lys Ala Asp Lys Asn	ggg gtt gca Gly Val Ala
10	1985	1990	1995	
	cta act 6039 Leu Thr	gaa tta agt ttc cga Glu Leu Ser Phe Arg	gca cca agt gag ctt Ala Pro Ser Glu Leu	gca aaa gag Ala Lys Glu
15	2000	2005	2010	
	agc aca 6084 Ser Thr	gag gga aga att tta Glu Gly Arg Ile Leu	gta gat ggt aaa gag Val Asp Gly Lys Glu	ctt cct gac Leu Pro Asp
25	2015	2020	2025	
	ttc tct 6129 Phe Ser	gaa aat cgt ttg gac Glu Asn Arg Leu Asp	tat gaa ttg att tat Tyr Glu Leu Ile Tyr	aaa ggg gaa Lys Gly Glu
30	2030	2035	2040	
	cgt cca 6174 Arg Pro	aaa atc act gta gaa Lys Ile Thr Val Glu	ggg aaa gat caa gtc Gly Lys Asp Gln Val	aca tct aca Thr Ser Thr
35	2045	2050	2055	
	att gtg 6219 Ile Val	gat agt gga aat gaa Asp Ser Gly Asn Glu	aat ctt cct gta tta Asn Leu Pro Val Leu	gtc cgt ctg Val Arg Leu
45	2060	2065	2070	
	gta tca 6264 Val Ser	gag agt ggt aag aat Glu Ser Gly Lys Asn	gtc aaa gaa tac cgg Val Lys Glu Tyr Arg	att aaa ttg Ile Lys Leu
50	2075	2080	2085	
	act aag 6309 Thr Lys	gaa aaa gtg ata gac Glu Lys Val Ile Asp	gca aaa ctt gta gct Ala Lys Leu Val Ala	cct gct caa Pro Ala Gln
55	2090	2095	2100	
	tat gac 6354	ttg cct agt ctt gaa	gtg gta gac aaa gaa	ttg aac ttc
65				

ES 2 352 704 B1

	Tyr Asp	Leu Pro Ser	Leu Glu	Val Val	Asp Lys	Glu	Leu Asn Phe
	2105		2110			2115	
5	caa act 6399	ctt gaa caa	aaa gac	gat act	ctg ttt	gag	gga gaa gtt
	Gln Thr	Leu Glu Gln	Lys Asp	Asp Thr	Leu Phe	Glu	Gly Glu Val
10	2120		2125			2130	
	cgc ctg 6444	ttg caa gaa	ggg aaa	gtt ggt	aag gaa	cgt	atc tat aca
15	Arg Leu	Leu Gln Glu	Gly Lys	Val Gly	Lys Glu	Arg	Ile Tyr Thr
	2135		2140			2145	
20	gaa gta 6489	acg act gat	ggc aag	aat aaa	gaa aaa	tca	cgt gag att
	Glu Val	Thr Thr Asp	Gly Lys	Asn Lys	Glu Lys	Ser	Arg Glu Ile
25	2150		2155			2160	
	ctg gaa 6534	gaa cct gta	aat aga	att ctg	ttg gta	ggg	act aag aag
30	Leu Glu	Glu Pro Val	Asn Arg	Ile Leu	Leu Val	Gly	Thr Lys Lys
	2165		2170			2175	
35	aaa aca 6579	agt aac att	gag acg	ccg gcg	ggg gga	cat	caa gac aat
	Lys Thr	Ser Asn Ile	Glu Thr	Pro Ala	Gly Gly	His	Gln Asp Asn
	2180		2185			2190	
40	ggg agc 6624	agt caa cct	acg gaa	caa aaa	cca gga	gca	aat ccg aat
	Gly Ser	Ser Gln Pro	Thr Glu	Gln Lys	Pro Gly	Ala	Asn Pro Asn
45	2195		2200			2205	
	cca aat 6669	att gtc att	cca aat	cca tca	aat aat	ggg	gtg aca agt
50	Pro Asn	Ile Val Ile	Pro Asn	Pro Ser	Asn Asn	Gly	Val Thr Ser
	2210		2215			2220	
55	aac agt 6714	tca gat ttt	aaa aat	cat cta	tca gtt	gct	ccg aac att
	Asn Ser	Ser Asp Phe	Lys Asn	His Leu	Ser Val	Ala	Pro Asn Ile
	2225		2230			2235	
60	gtg ata 6759	caa gct cct	act cag	gaa agc	aaa aaa	cct	gga tgg aac
	Val Ile	Gln Ala Pro	Thr Gln	Glu Ser	Lys Lys	Pro	Gly Trp Asn
65	2240		2245			2250	

ES 2 352 704 B1

5 aag gaa gat ggt aaa tgg tat tat cgt caa caa aat ggg gac ttg
 6804
 Lys Glu Asp Gly Lys Trp Tyr Tyr Arg Gln Gln Asn Gly Asp Leu
 2255 2260 2265

 10 gct aag ggt tgg gtc aaa gat gga gat act tgg tat tac ttt gat
 6849
 Ala Lys Gly Trp Val Lys Asp Gly Asp Thr Trp Tyr Tyr Phe Asp
 2270 2275 2280

 15 caa aca gga aaa atg gaa act ggc tgg att aaa gat caa tca ggt
 6894
 Gln Thr Gly Lys Met Glu Thr Gly Trp Ile Lys Asp Gln Ser Gly
 2285 2290 2295

 20 gca tgg tat tac ttg aat caa tca ggt gca atg tca tct aat gag
 6939
 Ala Trp Tyr Tyr Leu Asn Gln Ser Gly Ala Met Ser Ser Asn Glu
 2300 2305 2310

 25 tgg att ttg gat caa gat ggt aag tgg tac tat gtt gat gca tca
 6984
 Trp Ile Leu Asp Gln Asp Gly Lys Trp Tyr Tyr Val Asp Ala Ser
 2315 2320 2325

 30 ggt tcg atg aag aca agt caa tgg ttc caa gtt gga gac aaa tgg
 7029
 Gly Ser Met Lys Thr Ser Gln Trp Phe Gln Val Gly Asp Lys Trp
 2330 2335 2340

 35 tat tat gtt gat gca aat gga tcc ctt gct gtc aat aca gta act
 7074
 Tyr Tyr Val Asp Ala Asn Gly Ser Leu Ala Val Asn Thr Val Thr
 2345 2350 2355

 40 ccg gat ggt tat caa gta aat gaa aac ggt gag tgg atc gga taa
 7119
 Pro Asp Gly Tyr Gln Val Asn Glu Asn Gly Glu Trp Ile Gly
 2360 2365 2370

 45 <210> 6
 <211> 2372
 50 <212> PRT
 <213> Artificial

 55 <220>
 <223> Synthetic Construct

ES 2 352 704 B1

<400> 6

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

ES 2 352 704 B1

Val Asn Gly Gln Leu Val Gly His Tyr Pro Asn Gly Tyr Asn Gln Phe
260 265 270

5 Ser Tyr Asp Ile Ser Asn Tyr Leu His Lys Asp Gly Arg Glu Asn Val
275 280 285

10 Ile Ser Val His Ala Val Asn Lys Gln Pro Ser Ser Arg Trp Tyr Ser
290 295 300

15 Gly Ser Gly Ile Tyr Arg Asp Val Ser Leu Gln Val Thr Asp Lys Ile
305 310 315 320

His Val Glu Lys Asn Gly Thr Thr Ile Leu Thr Pro Lys Leu Glu Ser
325 330 335

20 Gln Gln Gly Gly Lys Val Glu Thr Gln Val Ser Ser Lys Ile Val Asn
340 345 350

25 Thr Asp Asn Lys Asp His Glu Ile Val Ala Glu Tyr Gln Ile Phe Glu
355 360 365

30 Arg Gly Gly Lys Ala Val Thr Glu Leu Val Arg Thr Glu Ser Lys Thr
370 375 380

35 Leu Lys Ala Lys Glu Thr Ile His Leu Asp Ser Ala Leu Glu Val Glu
385 390 395 400

Lys Pro Lys Leu Trp Thr Val Ser Ser Asp Lys Pro Ala Leu Tyr Glu
405 410 415

40 Met Val Thr Arg Val Tyr Lys Asp Gly Gln Leu Val Asp Ala Lys Lys
420 425 430

45 Asp Leu Phe Gly Tyr Arg Tyr Tyr Asn Trp Thr Pro Asp Gln Gly Phe
435 440 445

50 Ser Leu Asn Gly Glu His Ile Lys Phe His Gly Val Ser Leu His His
450 455 460

55 Asp His Gly Ala Leu Gly Ala Glu Glu Asn Tyr Lys Ala Glu Tyr Arg
465 470 475 480

Arg Leu Lys Gln Met Lys Glu Met Gly Val Asn Ser Ile Arg Thr Thr
485 490 495

60 His Asn Pro Ala Ser Pro Gln Thr Leu Gln Ile Ala Ala Glu Leu Gly
500 505 510

65 Leu Leu Val Gln Glu Glu Ala Phe Asp Thr Trp Tyr Gly Gly Lys Lys

ES 2 352 704 B1

	515		520		525												
5	Pro	Tyr	Asp	Tyr	Gly	Arg	Phe	Phe	Glu	Lys	Asp	Ala	Thr	His	Pro	Glu	
	530						535					540					
10	Ala	Arg	Lys	Gly	Glu	Lys	Trp	Ser	Asp	Tyr	Asp	Leu	Arg	Thr	Met	Val	
	545					550					555					560	
15	Glu	Arg	Asp	Lys	Asn	Asn	Pro	Ala	Val	Val	Met	Trp	Ser	Ile	Gly	Asn	
					565						570				575		
20	Glu	Ile	Gly	Glu	Ala	Asp	Gly	Lys	Ala	His	Ser	Leu	Val	Thr	Val	Lys	
				580					585					590			
25	Arg	Leu	Val	Lys	Val	Val	Lys	Ser	Val	Asp	Lys	Thr	Arg	Tyr	Val	Thr	
			595					600					605				
30	Met	Gly	Ala	Asp	Lys	Phe	Arg	Phe	Gly	Asp	Gly	Thr	Gly	Asp	His	Glu	
	610						615					620					
35	Lys	Ile	Ala	Asp	Glu	Leu	Asp	Ala	Val	Gly	Leu	Asn	Tyr	Ser	Glu	Glu	
	625					630					635					640	
40	Asn	Tyr	Gln	Thr	Leu	His	Ala	Lys	His	Pro	Lys	Trp	Leu	Ile	Tyr	Gly	
					645					650					655		
45	Ser	Glu	Thr	Ser	Ser	Ala	Thr	Arg	Thr	Arg	Gly	Ser	Tyr	Phe	His	Pro	
				660					665					670			
50	Glu	Ser	Glu	Trp	Val	Gly	Ser	Asn	Gln	Tyr	Trp	Arg	Asn	Tyr	Glu	Gln	
			675					680					685				
55	Ser	Asp	Tyr	Gly	Asn	Asp	Arg	Val	Gly	Trp	Gly	Lys	Thr	Ala	Thr	Ala	
	690						695					700					
60	Ser	Trp	Thr	Phe	Asp	Arg	Asp	His	Pro	Gly	Tyr	Ala	Gly	Gln	Phe	Ile	
	705					710					715				720		
65	Trp	Thr	Gly	Thr	Asp	Tyr	Ile	Gly	Glu	Pro	Thr	Pro	Trp	His	Asn	Gln	
					725					730					735		
70	Asn	Ser	Thr	Pro	Val	Lys	Ser	Ser	Tyr	Phe	Gly	Ile	Val	Asp	Thr	Ala	
				740					745					750			
75	Gly	Ile	Pro	Lys	Asn	Asp	Tyr	Tyr	Leu	Tyr	Gln	Ser	Gln	Trp	Val	Ser	
			755					760					765				
80	Ala	Lys	Lys	Lys	Pro	Met	Val	His	Leu	Leu	Pro	His	Trp	Asn	Trp	Glu	
	770						775					780					

ES 2 352 704 B1

Lys Thr Glu Leu Ala Asp Asn Val Ala Asp Ala Glu Asn Arg Ile Pro
 785 790 795 800
 5 Val Arg Ala Tyr Ser Asn Ala Ala Ser Val Glu Leu Phe Leu Asn Asn
 805 810 815
 10 Glu Ser Leu Gly Leu Lys Lys Phe Asn Lys Lys Gln Thr Ser Asp Gly
 820 825 830
 15 Arg Thr Tyr Gln Glu Gly Glu Asn Pro Gln Glu Leu Tyr Leu Glu Trp
 835 840 845
 20 Lys Val Ala Tyr Lys Pro Gly Thr Leu Glu Ala Val Ala Arg Asp Glu
 850 855 860
 25 Ser Gly Lys Glu Ile Ala Arg Asp Lys Ile Val Thr Ala Gly Glu Pro
 865 870 875 880
 30 Ala Gly Val Arg Leu Val Lys Glu Glu Asn Ala Ile Ala Ala Asp Gly
 885 890 895
 35 Lys Asp Leu Thr Tyr Ile Tyr Tyr Glu Ile Val Asp Ser Lys Gly Asn
 900 905 910
 40 Val Val Pro Thr Ala Asn Asn Leu Val His Phe Gln Leu His Gly Gln
 915 920 925
 45 Gly Gln Leu Val Gly Val Asp Asn Gly Glu Gln Ala Ser Arg Glu Arg
 930 935 940
 50 Tyr Lys Glu Gln Ala Asp Gly Ser Trp Ile Arg Arg Ala Phe Asn Gly
 945 950 955 960
 55 Lys Gly Val Ala Ile Val Lys Ser Thr Asp Gln Ala Gly Lys Phe Thr
 965 970 975
 60 Leu Thr Ala Tyr Ser Gly Leu Leu Lys Ser Asp Gln Val Thr Val Phe
 980 985 990
 65 Thr Gly Lys Ser Asp Gln Ser Glu Lys Thr Val Leu Gly Thr Glu Ile
 995 1000 1005
 70 Pro Lys Val Arg Thr Val Leu Gly Gln Lys Leu Lys Met Pro Gln
 1010 1015 1020
 75 Thr Val Pro Phe Val Tyr Ser Asp Gly Ser Arg Glu Lys Arg Pro
 1025 1030 1035
 80 Val Thr Trp Ser Gln Ala Asp Val Ser Gln Ala Gly Ile Val Thr

ES 2 352 704 B1

	1040		1045		1050	
5	Val	Lys 1055	Gly Met Ser Asp	Gly 1060	Arg Glu Val Glu	Ala 1065 Arg Val Glu
10	Val	Leu 1070	Thr Val Ala Asn	Glu 1075	Leu Pro Ala Val	Lys 1080 Arg Ile Thr
15	Pro	Gly 1085	Thr Asp Leu Ser	Ser 1090	Val Asp Lys Leu	Val 1095 Ser Leu Val
20	Ser	Thr 1100	Asp Gly Arg Ile	His 1105	Asn Tyr Glu Val	Glu 1110 Asn Trp Glu
25	Ile	Ser 1115	Pro Glu Asp Lys	Glu 1120	Lys Leu Ser Thr	Pro 1125 Gly Ala His
30	Ile	Gln 1130	Met Thr Ser Gln	Leu 1135	Gly Asp Arg Thr	Ile 1140 His Ala Thr
35	Leu	Ile 1145	Val Asp Asp Gly	Lys 1150	Glu Leu Ser Gln	Val 1155 Thr Pro Asn
40	Ile	Ala 1160	Val Gly Gly Glu	Ser 1165	Val Thr Glu Leu	Ser 1170 Lys Glu His
45	Pro	Val 1175	His Tyr His Lys	Leu 1180	Ser Tyr Gly Ala	Lys 1185 Ile Pro Glu
50	Val	Ser 1190	Ala Ser Ala Glu	Asn 1195	Ala Asp Val Asn	Val 1200 Ile Gln Ala
55	Asn	Glu 1205	Ser Asn Gly Met	Arg 1210	Ala Ser Ile Tyr	Ile 1215 Gln Pro Lys
60	Asp	Gly 1220	Gly Gln Leu Gln	Thr 1225	Tyr Ala Val Gln	Phe 1230 Leu Val Asp
65	Ser	Pro 1235	Gln Ile Glu Ser	Leu 1240	Ser Leu Arg Val	Asp 1245 Gln Val Ser
	Ser	Leu 1250	Lys Glu Asp Gln	Thr 1255	Val Lys Val Thr	Val 1260 Leu Ala Arg
	Tyr	Gln 1265	Asp Gly Thr Glu	Ala 1270	Ile Leu Pro Thr	Asp 1275 Lys Val Thr
	Phe	Ser 1280	Ser Gln Gly Glu	Gly 1285	Gly Val Ser Val	Asn 1290 Lys Gly Met

ES 2 352 704 B1

	Leu	Glu	Leu	His	Lys	Pro	Gly	Ser	Leu	Thr	Leu	Gln	Ala	Gln	Tyr
		1295					1300					1305			
5	Glu	Gly	Ala	Lys	Gly	Glu	Leu	Pro	Leu	Thr	Ile	Ala	Ala	Asn	Thr
		1310					1315					1320			
10	Glu	Asp	Lys	Val	Val	Gln	Ala	Ile	Arg	Pro	Val	Ser	Leu	Val	Thr
		1325					1330					1335			
15	Asp	Leu	His	Gln	Thr	Pro	Asn	Leu	Pro	Ser	Thr	Val	Thr	Val	Glu
		1340					1345					1350			
20	Tyr	Asp	Lys	Gly	Phe	Pro	Lys	Val	His	Lys	Val	Val	Trp	Asp	Ser
		1355					1360					1365			
25	Ile	Pro	Ser	Glu	Lys	Leu	Asn	Gln	Tyr	His	Thr	Phe	Glu	Val	Leu
		1370					1375					1380			
30	Gly	Lys	Val	Glu	Gly	Leu	Asn	Gln	Glu	Ala	Arg	Ala	Thr	Val	Ser
		1385					1390					1395			
35	Val	Glu	Gly	Ile	Val	Lys	Val	Glu	Asp	Val	Ser	Ile	Thr	Thr	Pro
		1400					1405					1410			
40	Ile	Ser	Glu	Ala	Pro	Thr	Leu	Pro	Glu	Ser	Val	Arg	Thr	Tyr	Tyr
		1415					1420					1425			
45	Ser	Asn	Gly	Gln	Val	Ser	Ser	Ala	Lys	Val	Ala	Trp	Asp	Thr	Ile
		1430					1435					1440			
50	Asp	Ala	Ser	Gln	Tyr	Ala	Lys	Glu	Gly	Ile	Phe	Thr	Val	Thr	Gly
		1445					1450					1455			
55	His	Val	Glu	Gly	Thr	Gln	Leu	Thr	Thr	Lys	Leu	His	Val	Arg	Val
		1460					1465					1470			
60	Ser	Ser	Gln	Thr	Glu	Ile	Gly	Lys	Asn	Ile	Ser	Asp	Gln	Trp	Thr
		1475					1480					1485			
65	Gly	Ser	Glu	Leu	Pro	Leu	Ala	Phe	Ala	Ser	Asp	Ser	His	Ser	Tyr
		1490					1495					1500			
70	Asp	Pro	Val	Ser	Asn	Val	Asn	Asp	Lys	Val	Ile	Ser	Tyr	Gly	Asp
		1505					1510					1515			
75	Gln	Pro	Ala	Asn	Arg	Trp	Ser	Asn	Trp	Lys	Arg	Ser	Glu	Glu	Ala
		1520					1525					1530			
80	Ser	Val	Gly	Val	Leu	Phe	Gly	Asp	Ser	Gly	Ile	Leu	Thr	Lys	Arg

ES 2 352 704 B1

	1535					1540					1545				
5	Ser	Val 1550	Asp	Asn	Leu	Asn	Val 1555	Ala	Phe	Tyr	Glu	Asp 1560	Ser	Gly	Val
10	Gly	Ala 1565	Pro	Lys	Ser	Tyr	Val 1570	Ile	Glu	Tyr	Tyr	Val 1575	Gly	Glu	Lys
15	Ala	Pro 1580	Thr	Ala	Pro	Lys	Asn 1585	Pro	Gly	Tyr	Val	Glu 1590	Gly	Glu	Asn
20	His	Val 1595	Phe	Asn	Asp	Ser	Lys 1600	Asn	Trp	Lys	Pro	Val 1605	Thr	Asn	Leu
25	Lys	Ala 1610	Pro	Asp	Gln	Leu	Lys 1615	Ala	Gly	Glu	Met	Asn 1620	His	Phe	Ser
30	Phe	Asp 1625	Lys	Val	Asp	Thr	Tyr 1630	Ala	Val	Arg	Ile	Arg 1635	Met	Thr	Arg
35	Ala	Asp 1640	Asp	Lys	Leu	Gly	Thr 1645	Ser	Ile	Thr	Glu	Leu 1650	Gln	Val	Phe
40	Ser	Lys 1655	Lys	Val	Ala	Pro	Ala 1660	Lys	Glu	Ala	Thr	Thr 1665	Asn	Ile	Lys
45	Val	Ala 1670	Gly	Lys	Asp	Leu	Pro 1675	His	Phe	Asn	Pro	Asp 1680	Leu	Thr	Asp
50	Tyr	Tyr 1685	Leu	Asp	Tyr	Ser	Asp 1690	Gly	Lys	Ile	Pro	Glu 1695	Val	Thr	Ala
55	Ser	Ile 1700	Thr	Asn	Asn	Gly	Leu 1705	Ala	Thr	Val	Val	Pro 1710	Ser	Val	Lys
60	Glu	Gly 1715	Asp	Pro	Val	Arg	Val 1720	Ile	Val	Lys	Ala	Glu 1725	Asn	Gly	Asp
65	Ile	Leu 1730	Gly	Glu	Tyr	Asn	Leu 1735	His	Phe	Thr	Lys	Asp 1740	Lys	Asp	Leu
	Leu	Ala 1745	Arg	Lys	Pro	Ile	Ala 1750	Ser	Ala	Lys	Gln	Ser 1755	Gln	Leu	Leu
	Gln	Leu 1760	Gly	Gln	Pro	Leu	Glu 1765	Leu	Pro	Asn	Lys	Val 1770	Pro	Val	Tyr
	Phe	Ala 1775	Gly	Lys	Ser	Gly	Tyr 1780	Asp	Val	Lys	Asp	Leu 1785	Ala	Val	Glu

ES 2 352 704 B1

	Trp	Glu 1790	Lys	Val	Pro	Ala	Ser 1795	Thr	Leu	Asn	Lys	Ala 1800	Gly	Glu	Tyr
5	Thr	Ile 1805	Lys	Gly	His	Val	Val 1810	Gly	Ser	Glu	Val	Pro 1815	Val	Glu	Leu
10	Ser	Val 1820	Arg	Val	Thr	Asp	Lys 1825	Val	Gly	Arg	Ser	Leu 1830	Ser	Asp	Asn
15	Pro	Asp 1835	Tyr	Leu	Ala	Thr	His 1840	Ser	Gln	Ala	Phe	Ala 1845	Ser	Ala	Thr
	Asn	Asp 1850	Leu	Asp	Asp	Asn	Ser 1855	Asn	Asp	Arg	Val	Asp 1860	Tyr	Leu	Asn
20	Asp	Asn 1865	Asp	Arg	Asn	Gln	Asp 1870	Asn	Arg	Trp	Thr	Asn 1875	Trp	Ser	Ala
25	Thr	Pro 1880	Ser	Ala	Asn	Pro	Glu 1885	Val	Ser	Val	Gly	Val 1890	Ile	Phe	Lys
30	Lys	Asn 1895	Gly	Lys	Ile	Val	Asp 1900	Arg	Lys	Val	Ser	Gln 1905	Ala	Lys	Leu
35	Asn	Phe 1910	Phe	Ala	Asp	Ser	Gly 1915	Thr	Asp	Ala	Pro	Ala 1920	Lys	Leu	Val
	Leu	Glu 1925	Arg	Tyr	Ile	Gly	Pro 1930	Asp	Phe	Asp	Val	Pro 1935	Val	Tyr	Tyr
40	Ser	Asn 1940	Tyr	Ser	Tyr	Glu	Ser 1945	Asn	His	Pro	Phe	Asn 1950	Asn	Pro	Asp
45	Asn	Trp 1955	Glu	Leu	Val	Pro	Tyr 1960	Tyr	Ala	Asp	Lys	Glu 1965	Ile	Gln	Ala
50	Gly	Asn 1970	Glu	Ile	Asn	Val	Thr 1975	Phe	Lys	Ala	Val	Thr 1980	Thr	Lys	Ala
55	Met	Arg 1985	Trp	Arg	Met	Asp	Arg 1990	Lys	Ala	Asp	Lys	Asn 1995	Gly	Val	Ala
	Leu	Thr 2000	Glu	Leu	Ser	Phe	Arg 2005	Ala	Pro	Ser	Glu	Leu 2010	Ala	Lys	Glu
60	Ser	Thr 2015	Glu	Gly	Arg	Ile	Leu 2020	Val	Asp	Gly	Lys	Glu 2025	Leu	Pro	Asp
65	Phe	Ser	Glu	Asn	Arg	Leu	Asp	Tyr	Glu	Leu	Ile	Tyr	Lys	Gly	Glu

ES 2 352 704 B1

	2030						2035					2040				
5	Arg	Pro	Lys	Ile	Thr	Val	Glu	Gly	Lys	Asp	Gln	Val	Thr	Ser	Thr	
	2045						2050					2055				
10	Ile	Val	Asp	Ser	Gly	Asn	Glu	Asn	Leu	Pro	Val	Leu	Val	Arg	Leu	
	2060						2065					2070				
15	Val	Ser	Glu	Ser	Gly	Lys	Asn	Val	Lys	Glu	Tyr	Arg	Ile	Lys	Leu	
	2075						2080					2085				
20	Thr	Lys	Glu	Lys	Val	Ile	Asp	Ala	Lys	Leu	Val	Ala	Pro	Ala	Gln	
	2090						2095					2100				
25	Tyr	Asp	Leu	Pro	Ser	Leu	Glu	Val	Val	Asp	Lys	Glu	Leu	Asn	Phe	
	2105						2110					2115				
30	Gln	Thr	Leu	Glu	Gln	Lys	Asp	Asp	Thr	Leu	Phe	Glu	Gly	Glu	Val	
	2120						2125					2130				
35	Arg	Leu	Leu	Gln	Glu	Gly	Lys	Val	Gly	Lys	Glu	Arg	Ile	Tyr	Thr	
	2135						2140					2145				
40	Glu	Val	Thr	Thr	Asp	Gly	Lys	Asn	Lys	Glu	Lys	Ser	Arg	Glu	Ile	
	2150						2155					2160				
45	Leu	Glu	Glu	Pro	Val	Asn	Arg	Ile	Leu	Leu	Val	Gly	Thr	Lys	Lys	
	2165						2170					2175				
50	Lys	Thr	Ser	Asn	Ile	Glu	Thr	Pro	Ala	Gly	Gly	His	Gln	Asp	Asn	
	2180						2185					2190				
55	Gly	Ser	Ser	Gln	Pro	Thr	Glu	Gln	Lys	Pro	Gly	Ala	Asn	Pro	Asn	
	2195						2200					2205				
60	Pro	Asn	Ile	Val	Ile	Pro	Asn	Pro	Ser	Asn	Asn	Gly	Val	Thr	Ser	
	2210						2215					2220				
65	Asn	Ser	Ser	Asp	Phe	Lys	Asn	His	Leu	Ser	Val	Ala	Pro	Asn	Ile	
	2225						2230					2235				
70	Val	Ile	Gln	Ala	Pro	Thr	Gln	Glu	Ser	Lys	Lys	Pro	Gly	Trp	Asn	
	2240						2245					2250				
75	Lys	Glu	Asp	Gly	Lys	Trp	Tyr	Tyr	Arg	Gln	Gln	Asn	Gly	Asp	Leu	
	2255						2260					2265				
80	Ala	Lys	Gly	Trp	Val	Lys	Asp	Gly	Asp	Thr	Trp	Tyr	Tyr	Phe	Asp	
	2270						2275					2280				

ES 2 352 704 B1

	Gln	Thr	Gly	Lys	Met	Glu	Thr	Gly	Trp	Ile	Lys	Asp	Gln	Ser	Gly
		2285					2290					2295			
5	Ala	Trp	Tyr	Tyr	Leu	Asn	Gln	Ser	Gly	Ala	Met	Ser	Ser	Asn	Glu
		2300					2305					2310			
10	Trp	Ile	Leu	Asp	Gln	Asp	Gly	Lys	Trp	Tyr	Tyr	Val	Asp	Ala	Ser
		2315					2320					2325			
15	Gly	Ser	Met	Lys	Thr	Ser	Gln	Trp	Phe	Gln	Val	Gly	Asp	Lys	Trp
		2330					2335					2340			
20	Tyr	Tyr	Val	Asp	Ala	Asn	Gly	Ser	Leu	Ala	Val	Asn	Thr	Val	Thr
		2345					2350					2355			
25	Pro	Asp	Gly	Tyr	Gln	Val	Asn	Glu	Asn	Gly	Glu	Trp	Ile	Gly	
		2360					2365					2370			
30															
35															
40															
45															
50															
55															
60															
65															



OFICINA ESPAÑOLA
DE PATENTES Y MARCAS

ESPAÑA

②① N.º solicitud: 200802310

②② Fecha de presentación de la solicitud: 08.08.2008

③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤① Int. Cl.: **C12N15/56** (01.01.2006)
C12N9/38 (01.01.2006)

DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
X	WO 2004 092209 A2 (INTERCELL AG) 28.10.2004, secuencia 34.	1
X	VICKERMAN, M.M. et al. Genome-wide transcriptional changes in <i>Streptococcus gordonii</i> in response to competence signaling peptide. JOURNAL OF BACTERIOLOGY. 2007. Volumen 189, No. 21, páginas: 7799-7807. DOI: 10.1128/JB.01023-07. Ver resumen, introducción, resultados y discusión.	9
X	PARMJIT S PANESAR et al. Microbial production, immobilization and applications of β -D-galactosidase. JOURNAL OF CHEMICAL TECHNOLOGY AND BIOTECHNOLOGY. 04.2006. Vol. 81, pág. 560-543. DOI:10.1002/jctb.1453. Ver resumen e introducción.	13,18
X	PIVARNIK, L.F. et al. Hidrolitic and transgalactosylic activities of commercial β -Galactosidase in food processing. ADVANCES FOOD AND NUTRITION RESEARCH. 1995. vol. 38, pág. 1-102.	13,18
A	MADOZ, J. et al. Functionalization of Gold Surfaces for Specific and Reversible Attachment of a Fused β -Galactosidase and Choline-Receptor Protein. J. AMERICAN CHEMICAL SOCIETY. 05.02.1997. Volumen 119, pág. 1043-1051. DOI: 10.1021/ja963465r. Resumen e introducción.	1-21

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

☒ para todas las reivindicaciones

☐ para las reivindicaciones nº:

Fecha de realización del informe
16.11.2010

Examinador
I. Abad Gurumeta

Página
1/5

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C12N

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, BIOSIS, EMBASE, EMBL ALL, WIPI, NPL

Fecha de Realización de la Opinión Escrita: 16.11.2010

Declaración**Novedad (Art. 6.1 LP 11/1986)**

Reivindicaciones 2-8; 10-21

SI

Reivindicaciones 1; 9

NO**Actividad inventiva (Art. 8.1 LP11/1986)**

Reivindicaciones 2-8; 10-12; 14-17

SI

Reivindicaciones 1; 9; 13; 18; 19-21

NO

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

Base de la Opinión.-

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

1. Documentos considerados.-

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	WO 2004 092209 A2 (INTERCELL AG).	28.10.2004
D02	VICKERMAN, M.M. et al. Genome-wide transcriptional changes in <i>Streptococcus gordonii</i> in response to competence signaling peptide. JOURNAL OF BACTERIOLOGY. Volumen 189, No. 21, páginas: 7799-7807.	2007
D03	PARMJIT S PANESAR et al. Microbial production, immobilization and applications of β -D-galactosidase. JOURNAL OF CHEMICAL TECHNOLOGY AND BIOTECHNOLOGY. Vol. 81, pág. 560-543.	04.2006
D04	PIVARNIK, L.F. et al. Hidrolitic and transgalactosylic activities of commercial β -Galactosidase in food processing. ADVANCES FOOD AND NUTRITION RESEARCH.	1995
D05	MADOZ, J. et al. Functionalization of Gold Surfaces for Specific and Reversible Attachment of a Fused β -Galactosidase and Choline-Receptor Protein. J. AMERICAN CHEMICAL SOCIETY. Volumen 119, pág. 1043-1051.	05.02.1997

2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración

El objeto de la invención, tal y como se recoge en las reivindicaciones 1-21, es una secuencia de nucleótidos de β -galactosidasa, SEQ ID NO1, sin péptido señal, SEQ ID NO3, o modificada, SEQ ID NO5, (ver reivindicaciones 1-5), su construcción genética (reivindicación 6) y vector de expresión (reivindicación 7-8). La secuencia de aminoácidos de β -galactosidasa de *S.mits*, SEQ ID NO2, sin péptido señal, SEQ ID NO4, o modificada SEQ ID NO6 (reivindicaciones 9-12). Así como las células huésped resultantes de las anteriores (reivindicaciones 14-16). También tiene por objeto el uso de la secuencia de aminoácidos y célula huésped para la reducción del contenido de lactosa en alimentos (reivindicación 13 y 18). Otro objeto de la invención es el procedimiento de obtención de la secuencia de aminoácidos SEQ ID NO2 y SEQ ID NO6 (reivindicaciones 19-21)

El documento D01 publica las secuencias nucleótidas aisladas que codifican un suero de reactivos antígenos o fragmentos de *Streptococcus pneumoniae*. (Ver resumen y secuencia 34).

El documento D02 publica el genoma de *Streptococcus gordinii* y su transcripción peptídica. (Ver resumen, introducción, resultados y discusión)

El documento D03 publica el uso de la β -galactosidasa como enzima en procesos alimenticios, ya que cataliza la hidrólisis de lactosa. El aislamiento, inmovilización y purificación en el problema sobre el que se centra, así como su utilización para hidrólisis de la lactosa en la leche y otras aplicaciones. (Ver resumen e introducción)

El documento D04 divulga el uso de la β -galactosidasa en procesos de hidrólisis de la lactosa en su aplicación en alimentación, en particular, en la obtención de productos y subproductos lácteos. (Ver páginas 1-102)

El documento D05 publica un método de inmovilización de proteínas en electrodos de oro, usado para la purificación de proteínas por unión a colinas. (Ver resumen e introducción)

1. NOVEDAD (ART. 6.1 Ley 11/1986)

El documento D01 divulga la secuencia SEQ ID NO:1 con una identidad superior al 73 % en la secuenciación de las secuencias nucleótidas de *Streptococcus pneumoniae*. (Ver secuencia 34)

La reivindicación 1 de la solicitud reivindica la secuencia de nucleótidos de la β -galactosidasa SEQ ID NO:1, secuencia análoga, un fragmento de cualquiera de las dos o una secuencia análoga de cualquiera de las tres anteriores. Por tanto, la secuencia objeto de la invención recogido en la reivindicación 1 ha sido divulgado en el documento D01, ya que un fragmento de la misma ha sido divulgado. Por consiguiente, la reivindicación 1 no es nueva en el sentido del artículo 6.1 Ley 11/1986

El documento D02 publica la secuencia SEQ ID NO:2 con una identidad superior al 73 % al realizar la transcripción peptídica del genoma de *Streptococcus gordinii* (Ver resumen, introducción, resultados y discusión).

La reivindicación 9 de la solicitud reivindica la secuencia de aminoácidos de la β -galactosidasa SEQ ID NO:2, secuencia análoga, un fragmento de cualquiera de las dos o una secuencia análoga de cualquiera de las tres anteriores. Por tanto, la secuencia objeto de la invención recogido en la reivindicación 9 ha sido divulgado en el documento D02, ya que un fragmento de la misma ha sido divulgado. Por consiguiente, la reivindicación 9 no es nueva en el sentido del artículo 6.1 Ley 11/1986.

Tanto las secuencias de DNA: SEQ ID NO1, SEQ ID NO3 y SEQ ID NO5; como las secuencias proteicas: SEQ ID NO2, SEQ ID NO4 y SEQ ID NO6, no han sido divulgadas en el estado de la técnica. Por lo que el objeto de la invención tal y como se recoge en las reivindicaciones 2-8 y 10-21 es nuevo en el sentido del artículo 6.1 Ley 11/1986.

2. ACTIVIDAD INVENTIVA (ART. 8.1 Ley 11/1986)

El documento D03 publica el uso de la β -galactosidasa como enzima en procesos alimenticios, ya que cataliza la hidrólisis de lactosa. (Ver resumen e introducción)

El documento D04 divulga el uso de la β -galactosidasa en procesos de hidrólisis de la lactosa en su aplicación en alimentación, en particular, en la obtención de productos y subproductos lácteos. (Ver páginas 1-102)

La invención definida en las reivindicaciones 13 y 18 acerca del uso para la reducción del contenido de lactosa en los alimentos o residuos de procesos alimentarios, no difiere de la técnica conocida descrita en los documentos D03 o D04 en ninguna forma esencial. Por lo tanto, la invención según las reivindicaciones 13 y 18 se considera obvia para un experto en la materia. Por consiguiente, la invención según las reivindicaciones 13 y 18 no se consideran que impliquen actividad inventiva en el sentido del artículo 8.1 Ley 11/1986.

Como ya se ha indicado en el apartado de novedad, las reivindicaciones 1 y 9 no son nuevas, por tanto carecen de actividad inventiva en el sentido del artículo 8.1 Ley 11/1986.

Los métodos de acuerdo con las reivindicaciones 19-21 son conocidos para un experto en la materia como métodos de expresión y purificación de proteínas. Así la invención reivindicada en las reivindicaciones 19-21 no es nueva y por lo tanto carece de actividad inventiva en el sentido del artículo 8.1 Ley 11/1986.

Por lo que el objeto de la invención tal y como se recoge en las reivindicaciones 2-8; 10-12 y 14-17 tiene actividad inventiva en el sentido del artículo 8.1 Ley 11/1986.

En conclusión las reivindicaciones 1, 9, 13, 18 y 19-21 carecen de actividad inventiva en el sentido del artículo 8.1 Ley 11/1986.