



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 346 196**

(51) Int. Cl.:

C12N 9/10 (2006.01)

C12N 15/54 (2006.01)

C12P 19/04 (2006.01)

C12P 19/26 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **02804804 .9**

(96) Fecha de presentación : **13.06.2002**

(97) Número de publicación de la solicitud: **1481052**

(97) Fecha de publicación de la solicitud: **01.12.2004**

(54)

Título: **Genes de hialuronano sintasa y expresión de los mismos.**

(30)

Prioridad: **13.06.2001 US 297744 P**
13.06.2001 US 297788 P

(45)

Fecha de publicación de la mención BOPI:
13.10.2010

(45)

Fecha de la publicación del folleto de la patente:
13.10.2010

(73)

Titular/es: **The Board of Regents of the University
of Oklahoma**
660 Parrington Oval
Norman, Oklahoma 73019, US

(72)

Inventor/es: **Deangelis, Paul L.;**
Weigel, Paul, H., y
Kumari, Kshama

(74)

Agente: **Elzaburu Márquez, Alberto**

ES 2 346 196 T3

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

DESCRIPCIÓN

Genes de hialuronano sintasa y expresión de los mismos.

5 Fundamento del invento

1. Campo del invento

El presente invento se refiere a un segmento de ácido nucleico que tiene una región de codificación que codifica una hialuronato sintasa (HAS) enzimáticamente activa, y al uso de este segmento de ácido nucleico en la preparación de células recombinantes que producen hialuronato sintasa y su producto, ácido hialurónico. El hialuronato es también conocido como ácido hialurónico e hialuronano.

2. Breve descripción de la técnica relacionada

La incidencia de infecciones estreptocócicas es un importante problema sanitario y económico mundial, particularmente en los países en desarrollo. Una razón de esto es la capacidad de las bacterias estreptocócicas para crecer sin ser detectadas por las células fagocíticas, es decir, macrófagos y células polimorfonucleares (PMNs; del inglés, polymorphonuclear cells), del organismo. Estas células son responsables de reconocer y engullir microorganismos extraños. Un modo eficaz que tienen las bacterias para eludir la vigilancia es mediante su revestimiento con cápsulas polisacáridas, tal como con una cápsula de ácido hialurónico (HA; del inglés, hyaluronic acid). La estructura del HA es idéntica tanto en procariotas como en eucariotas.

Puesto que el HA es generalmente no inmunogénico, las bacterias encapsuladas no provocan una respuesta inmune y, por lo tanto, no son un objetivo de destrucción. Además, la cápsula ejerce un efecto antifagocítico sobre PMNs *in vitro* y evita la fijación del estreptococo a macrófagos. A causa precisamente de esto, en los estreptococos del Grupo A y el Grupo C, las cápsulas de HA son importantes factores de virulencia en las infecciones naturales y experimentales. Los estreptococos del Grupo A son responsables de numerosas enfermedades humanas, incluyendo faringitis, impétigo, infecciones de tejidos profundos, fiebre reumática y un síndrome de tipo choque tóxico. El *Streptococcus equisimilis* del Grupo C es responsable de osteomielitis, faringitis, abscesos cerebrales y neumonía.

Estructuralmente, el HA es un polisacárido lineal de alto peso molecular compuesto por unidades disacáridas repetitivas que consisten en N-acetilglucosamina (GlcNAc) y ácido glucurónico (GlcUA). En una molécula de HA, el número de disacáridos repetitivos puede exceder de 30.000, un $M_r > 10^7$. El HA es el único glicosaminoglicano sintetizado tanto por células de mamífero como por células bacterianas, particularmente por estreptococos de los Grupos A y C y por *Pasteurella multocida* del Tipo A. Estas cepas generan HA que es secretado al medio, así como cápsulas de HA. El mecanismo por el cual estas bacterias sintetizan HA tiene un gran interés medicinal ya que la producción de la cápsula de HA es un modo muy eficaz e ingenioso que usan los estreptococos para eludir la vigilancia del sistema inmune. Además, moléculas orgánicas o inorgánicas revestidas con HA tienen unas propiedades que les permiten escapar de la detección y destrucción por el sistema inmune del huésped.

El HA es sintetizado por la enzima hialuronato sintasa de células de mamífero y bacterianas, que se ha localizado en la membrana plasmática. Se cree que la síntesis de HA en estos organismos es un proceso de múltiples pasos. El inicio implica la unión de un precursor inicial, UDP-GlcNAc o UDP-GlcUA. Esto va seguido de un alargamiento que implica la adición alterna de los dos azúcares a la cadena oligosacárida en crecimiento. El polímero en crecimiento se extruye a través de la región de la membrana plasmática de la célula y en el espacio extracelular. Aunque el sistema biosintético del HA fue una de las primeras rutas sintéticas de heteropolisacárido de membrana estudiadas, el mecanismo de la síntesis de HA no está aún bien entendido. Esto se puede deber a que los sistemas *in vitro* desarrollados hasta la fecha son inadecuados porque no se ha conseguido la biosíntesis *de novo* del HA.

La dirección del crecimiento del polímero de HA es aún un asunto de desacuerdo entre quienes tienen una experiencia normal en la técnica. La adición de los monosacáridos podría ser al extremo reductor o no reductor de la cadena de HA en crecimiento. Además, quedan cuestiones relativas a (i) si las cadenas nacientes están covalentemente enlazadas a una proteína, a UDP o a un producto intermedio lipídico, (ii) si las cadenas se inician usando un cebador, y (iii) el mecanismo por el cual el polímero maduro se extruye a través de la membrana plasmática del estreptococo. El conocimiento del mecanismo de la biosíntesis de HA puede permitir el desarrollo de estrategias alternativas para controlar las infecciones por estreptococos y *Pasteurella* al interferir en el proceso.

El HA ha sido identificado en casi todos los tejidos de los vertebrados y ha alcanzado un uso muy difundido en diversas aplicaciones clínicas, muy notable y apropiadamente como un suplemento de matriz intraarticular y en cirugía ocular. La bibliografía científica también ha mostrado una transición, desde la percepción original de que el HA es esencialmente un componente estructural pasivo de la matriz de unos pocos tejidos conjuntivos y de la cápsula de ciertas cepas de bacterias, hasta el reconocimiento de que esta macromolécula ubicua está dinámicamente implicada en muchos procesos biológicos: desde modular la migración y la diferenciación celulares durante la embriogénesis para la regulación de la organización y el metabolismo de la matriz extracelular, hasta papeles importantes en los complejos procesos de metástasis, cicatrización de heridas e inflamación. Además, está resultando evidente que el HA es muy metabólicamente activo y que las células ponen mucha atención en los procesos de su síntesis y su catabolismo. Por ejemplo, la semivida del HA en los tejidos varía desde 1 a 3 semanas en el cartílago hasta < 1 día en la epidermis.

También se usa HA en numerosas aplicaciones técnicas (por ejemplo, compuestos lubricantes), cosmética y productos nutracéuticos.

Está ahora claro que una sola proteína utiliza ambos sustratos de azúcar para sintetizar HA; es decir, que las HA sintetas son enzimas únicas que tienen propiedades catalíticas dobles. La abreviatura HAS para la HA sintasa ha obtenido un amplio apoyo para designar esta clase de enzimas. Markovitz *et al.* caracterizaron exitosamente la actividad HAS de *Streptococcus pyogenes* y descubrieron la localización membranal de la enzima y sus necesidades de precursores de azúcar-nucleótido y de Mg^{2+} . Prehm halló que el HA en alargamiento, generado por células B6, era digerido por hialuronidasa añadida al medio y propuso que la HAS reside en la membrana plasmática. Philipson y Schwartz también mostraron que la actividad HAS era cofraccionada con marcadores de la membrana plasmática en células de oligodendroglioma de ratón.

La HAS ensambla HA de alto M_r que es simultáneamente extruido en el espacio extracelular a través de la membrana (o para generar la cápsula celular en el caso de bacterias) conforme transcurre la síntesis del glicosaminoglicano. Este modo de biosíntesis es único entre las macromoléculas ya que los ácidos nucleicos, las proteínas y los lípidos se sintetizan en el núcleo, el retículo endoplásmico/Golgi, el citoplasma o las mitocondrias. La extrusión de la cadena en crecimiento en el espacio extracelular también permite el crecimiento libre del polímero, consiguiéndose por ello el tamaño excepcionalmente largo del HA, mientras que el confinamiento de la síntesis en un compartimento de Golgi o post-Golgi limita la cantidad o longitud global de los polímeros formados. Las elevadas concentraciones de HA en una abertura confinada también pueden crear un ambiente de alta viscosidad que podría ser deletéreo para otras funciones de los orgánulos.

Mediante diversos estudios se ha intentado solubilizar, identificar y purificar la HAS de cepas de estreptococos que generan un revestimiento capsular de HA, así como de células eucarióticas. Aunque las enzimas estreptocócicas y de oligodendroglioma murino fueron exitosamente solubilizadas con detergente y estudiadas, los esfuerzos para purificar una HAS activa para un estudio o una clonación molecular posteriores permanecieron infructuosos durante décadas. Prehm y Mausolf usaron UDP-GlcUA o UDP-GlcNAc oxidados con peryodato para marcar por afinidad una proteína de ~ 52 kDa en membranas estreptocócicas que resultaban copurificadas con HAS. Esto condujo a una comunicación en que se reivindicaba que se había clonado la HAS estreptocócica del Grupo C, lo que desafortunadamente era erróneo. Este estudio fracasó en cuanto a demostrar la expresión de una sintasa activa, y, en realidad, se podía haber clonado un transportador peptídico. Triscott y van de Rijn utilizaron digitonina para solubilizar la HAS de membranas estreptocócicas en una forma activa. Van de Rijn y Drake radiomarcaron selectivamente tres proteínas de membrana estreptocócica, de 42, 33 y 27 kDa, con 5-azido-UDP-GlcUA y sugirieron que la proteína de 33 kDa era HAS. Sin embargo, como se mostró más adelante, la HAS resultó ser la proteína de 42 kDa.

A pesar de estos esfuerzos, el progreso en el conocimiento de la regulación y los mecanismos de la síntesis de HA resultó esencialmente atascado ya que no había sondas moleculares para mRNA de HAS ni proteína HAS. En 1993 ocurrió un avance importante cuando DeAngelis *et al.* presentaron la clonación y caracterización molecular del gen estreptocócico del Grupo A que codifica la proteína HasA. Se supo que este gen era parte de un operón necesario para la síntesis del HA bacteriano, aunque se desconocía la función de esta proteína, que ahora se denomina spHAS (la HAS de *S. pyogenes*). Posteriormente se demostró que spHAS es responsable del alargamiento de HA, y fue la primera glicosaminoglicano sintasa identificada y clonada y luego exitosamente expresada. El operón para la síntesis de HA de *S. pyogenes* codifica otras dos proteínas. HasB es una UDP-glucosa deshidrogenasa, que es necesaria para convertir UDP-glucosa en UDP-GlcUA, uno de los sustratos para la síntesis de HA. HasC es una UDP-glucosa pirofosforilasa, que es necesaria para convertir glucosa-1-fosfato y UTP en UDP-glucosa. La cotransfección de *Enterococcus faecalis* o cepas de *Streptococcus* acapsulares con ambos genes, *hasA* y *hasB*, les confirió la capacidad para sintetizar HA y formar una cápsula. Esto proporcionaba la primera convincente evidencia de que spHAS (*hasA*) era una HA sintasa.

El escurridizo gen de la HA sintasa fue finalmente clonado mediante un procedimiento de mutagénesis por transposones mediante el cual se creaba una cepa acapsular mutante del Grupo A que contenía una interrupción, por transposón, del operón para la síntesis de HA. Secuencias conocidas del transposón permitían que la región de la juntura con DNA estreptocócico fuera identificada y luego clonada a partir de células de tipo silvestre. La spHAS codificada tenía un 5-10% de identidad con una familia de quitina sintetas de levadura y 30% de identidad con la proteína DG42 de *Xenopus laevis* (expresada en el desarrollo durante la gastrulación), cuya función era desconocida en el momento. DeAngelis y Weigel hicieron que se expresara la spHAS recombinante activa en *Escherichia coli* y mostraron que este producto génico purificado individual sintetiza HA de alto M_r cuando se incubaba *in vitro* con UDP-GlcUA y UDP-GlcNAc, mostrando por ello que ambas actividades glicosiltransferasa necesarias para la síntesis de HA son catalizadas por la misma proteína, como se propuso por vez primera en 1959. Utilizando el conocimiento de que (i) spHAS era una sola enzima de acción doble y (ii) las zonas de homología secuencial entre la spHAS, la quitina sintasa y DG42, la casi simultánea identificación de cDNAs de HAS eucariótica en 1996 por cuatro laboratorios consolidó más la hipótesis proteica del inventor de que HAS es una familia multigénica que codifica distintas isozimas. Se descubrieron rápidamente dos genes (*HAS1* y *HAS2*) en mamíferos y más tarde se descubrió un tercer gen, *HAS3*. Se identificó una segunda seHAS estreptocócica, o hialuronato sintasa de *Streptococcus equisimilis*, la cual se describe con detalle en el documento U.S. n.º de serie 09/469.200, presentado el 21 de diciembre de 1999 y publicado como US 6.833.264. La proteína seHAS tiene un elevado nivel de identidad (aproximadamente 70 por ciento) con respecto a la enzima spHAS. Sin embargo, esta identidad es interesante porque el gen de seHAS no presenta hibridación cruzada con el gen de spHAS.

Membranas preparadas a partir de *E. coli* que expresa seHAS recombinante sintetizan HA cuando se proporcionan ambos sustratos. Los resultados confirman que la comunicación previa de Lansing *et al.* en que reivindicaban haber

clonado la HAS del Grupo C era errónea. Desafortunadamente, en diversos estudios se han empleado anticuerpos hacia esta proteína estreptocócica no caracterizada, de 52 kDa, para investigar lo que se creía que era HAS eucariótica.

El gen que codifica la HA sintasa de *Streptococcus uberis* es descrito por Ward *et al.*, Infection and Immunity 69: 392-399 (2001), y en EMBL, número de acceso AJ242946. En la solicitud de patente WO9951265 se describe la HA sintasa de *Pasteurella multocida*.

Itano y Kimata utilizaron la clonación por expresión en una línea celular mutante de carcinoma mamario de ratón, incapaz de sintetizar HA, para clonar el primer supuesto cDNA de HAS de mamífero (mmHAS1). Subclones defectuosos en cuanto a la síntesis de HA fueron divididos en tres clases separadas que eran complementarias para la síntesis de HA en experimentos de fusión de células somáticas, lo que sugiere que se requieren al menos tres proteínas. Dos de estas clases mantenían cierta actividad sintética HA, mientras que una no mostraba ninguna. Se usó esta última línea celular en experimentos de transfección transitoria con cDNA preparado a partir de las células parentales, para identificar una sola proteína que restablecía la actividad sintética HA. Análisis secuenciales revelaron una deducida estructura primaria para una proteína de ~ 65 kDa con una predicha topología de membrana similar a la de spHAS. mmHAS1 tiene una identidad de 30% con respecto a spHAS y una identidad de 55% con respecto a DG42. El mismo mes en que apareció este comunicado, otros tres grupos presentaron papeles que describían cDNAs que codificaban lo que inicialmente se pensó que era la misma enzima de ratón y ser humano. Sin embargo, por una extraordinaria circunstancia, cada uno de los cuatro laboratorios había descubierto una distinta isozima de HAS en ambas especies.

Utilizando un procedimiento de clonación funcional similar al de Itano y Kimata, Shyjan *et al.* identificaron el compuesto homólogo humano de HAS1. Se utilizó un banco de cDNA de ganglios linfáticos mesentéricos para transfectar linfocitos T mucosos murinos que fueron luego explorados en cuanto a su capacidad para adherirse en un ensayo de rosetas. La adhesión de un transfectante fue inhibida por antisueros hacia CD44, una conocida proteína de la superficie celular que se une a HA, y fue directamente anulada mediante un pretratamiento con hialuronidasa. De esta manera, la formación de rosetas por este transfectante requería la síntesis de HA. La clonación y secuenciación del cDNA responsable permitió identificar hsHAS1. Itano y Kimata también presentaron un cDNA de HAS1 humano aislado de un banco de cerebro fetal; sin embargo, los cDNAs de hsHAS1 presentados por los dos grupos difieren en la longitud y codifican una proteína de 578 ó 543 aminoácidos. Sólo se ha demostrado actividad HAS en la forma más larga.

Basándose en la identificación molecular de spHAS como una auténtica HA sintasa y de regiones de identidad próxima entre DG42, spHAS y NodC (un factor de nodulación de β -GlcNAc transferasa en *Rhizobium*), Spicer *et al.* usaron un procedimiento de RT-PCR degenerada para clonar un cDNA de embrión de ratón que codificaba una segunda enzima distinta, que es denominada mmHAS2. La transfección de células COS con cDNA de mmHAS2 dirigía la producción *de novo* de un revestimiento celular de HA detectado mediante un ensayo de exclusión de partículas, lo que proporciona la acusada evidencia de que la proteína HAS2 puede sintetizar HA. Utilizando un procedimiento similar, Watanabe y Yamaguchi exploraron un banco de cDNA de cerebro fetal humano para identificar hsHAS2. Independientemente, Fulop *et al.* usaron una estrategia similar para identificar mmHAS2 en RNA aislado de células de cumulus ovárico que sintetizan activamente HA, un proceso crítico para la expansión normal del cumulus oophorus en el folículo preovular. Se aislaron complejos de célula de cumulus-ovocito de los ratones inmediatamente después del inicio de un ciclo ovular, antes de que comenzara la síntesis de HA, y en momentos posteriores cuando la síntesis de HA recién comenzaba (3 h) o ya era evidente (4 h). La RT-PCR mostró que el mRNA de HAS2 estaba inicialmente ausente pero se expresaba en niveles elevados 3-4 h más tarde, lo que sugiere que la transcripción de HAS2 regula la síntesis de HA en este proceso. Ambos HAS2 tienen una longitud de 552 aminoácidos y una identidad de 98%. mmHAS1 tiene una longitud de 583 aminoácidos y una identidad de 95% con respecto a hsHAS1, que tiene una longitud de 578 aminoácidos.

Muy recientemente, Spicer *et al.* usaron un procedimiento de PCR para identificar un tercer gen de HAS en mamíferos. La proteína mmHAS3 tiene una longitud de 554 aminoácidos y unas identidades de 71, 56 y 28%, respectivamente, en relación con mmHAS1, mmHAS2, DG42 y spHAS. Spicer *et al.* han localizado también los tres genes humanos y de ratón en tres cromosomas diferentes (HAS1 en hsChr 19/mmChr 17; HAS2 en hsChr 8/mmChr 15; HAS3 en hsChr 16/mmChr 8). La localización de los tres genes de HAS en diferentes cromosomas y la aparición de HA por toda la clase de vertebrados sugiere que esta familia génica es antigua y que aparecieron isozimas por duplicación al principio de la evolución de los vertebrados. La elevada identidad (~ 30%) entre los HAS bacterianos y eucarióticos también sugiere que los dos tenían un gen ancestral común. Quizás las bacterias primitivas usurparon el gen de HAS a un antepasado primitivo de los vertebrados antes de que los productos génicos eucarióticos llegaran a ser más grandes y más complejos. Alternativamente, las bacterias podrían haber obtenido un gen más grande de HAS de vertebrado y haber suprimido secuencias reguladoras no esenciales para la actividad enzimática.

El descubrimiento de DG42 de *X. laevis* por Dawid y colaboradores desempeñó un papel significativo en estos desarrollos recientes aun cuando no se sabía que esta proteína era una HA sintasa. No obstante, el que DG42 y spHAS tuvieran una identidad de 30% era crítico para diseñar oligonucleótidos que permitieran la identificación de HAS2 de mamífero. Irónicamente, la evidencia definitiva de que DG42 es una HA sintasa *bona fide* sólo fue comunicada después de los descubrimientos de las isozimas de mamífero, cuando DeAngelis y Achyuthan hicieron que se expresara la proteína recombinante en levadura (un organismo que no puede sintetizar HA) y mostraron que sintetiza HA cuando se proporcionan los dos sustratos a membranas aisladas. Meyer y Kreil también mostraron que lisados de células transfectadas con cDNA de DG42 sintetizan niveles elevados de HA. Por lo tanto, ahora que se conoce su función, DG42 puede ser denominada XIHAS.

Hay características estructurales previstas comunes compartidas por todas las proteínas HAS, incluyendo un gran dominio central y haces de 2-3 dominios transmembranales o asociados a la membrana en ambos extremos de la proteína, el amínico y el carboxílico. El dominio central, que comprende hasta ~ el 88% de las previstas secuencias intracelulares de la proteína HAS, contiene probablemente las regiones catalíticas de la enzima. Este previsto dominio central tiene una longitud de 264 aminoácidos en spHAS (63% de la proteína total) y una longitud de 307-328 restos en los miembros de HAS eucarióticos (54-56% de la proteína total). Aún no se han determinado experimentalmente el número y la orientación exactos de los dominios membranales ni la organización topológica de los bucles extracelulares e intracelulares de ninguna HAS.

La spHAS es un miembro de la familia de HAS que ha sido purificado y parcialmente caracterizado. Estudios iniciales utilizando proteínas de fusión de spHAS/fosfatasa alcalina indican que el extremo N, el extremo C y el gran dominio central de spHAS están, en realidad, dentro de la célula. La spHAS tiene 6 cisteínas, mientras que HAS1, HAS2 y HAS3 tienen 13, 14 y 14 restos de Cys, respectivamente. Dos de los 6 restos de Cys de spHAS están conservados y son idénticos en HAS1 y HAS2. Sólo un resto conservado de Cys se encuentra en la misma posición (Cys-225 en spHAS) en todos los miembros de la familia de HAS. Ésta puede ser una Cys esencial cuya modificación por venenos de sulfhidrilos inhibe parcialmente la actividad de la enzima. En ningún miembro de la familia de Has se ha elucidado aún la posible presencia de enlaces disulfuro ni la identificación de restos críticos de Cys necesarios para cualquiera de las múltiples funciones de HAS indicadas más adelante.

Además del propuesto modo único de síntesis en la membrana plasmática, la familia de la enzima HAS es muy inusual en cuanto al gran número de funciones requeridas para la polimerización global de HA. En la enzima HAS están presentes al menos 6 actividades distintas: sitios de unión para cada uno de los dos diferentes precursores de azúcar-nucleótido (UDP-GlcNAc y UDP-GlcUA), dos actividades glicosiltransferasa diferentes, uno o más sitios de unión que anclan el polímero de HA en crecimiento a la enzima (quizás relacionados con un motivo B-X₇-B), y un mecanismo de transferencia de tipo trinquete que desplaza el polímero en crecimiento uno o dos azúcares a la vez. Esta última actividad es probablemente coincidente con el avance gradual del polímero a través de la membrana. Todas estas funciones, y quizás otras aún desconocidas, están presentes en una proteína relativamente pequeña cuyo tamaño varía de 419 (spHAS) a 588 (XHAS) aminoácidos.

Aunque toda la evidencia disponible respalda la conclusión de que sólo se requiere la proteína spHAS para la biosíntesis de HA en bacterias o *in vitro*, es posible que los miembros eucarióticos más grandes de la familia de HAS sean parte de complejos de múltiples componentes. Puesto que las proteínas HAS eucarióticas son ~ 40% más grandes que spHAS, sus dominios proteínicos adicionales podrían estar implicados en funciones más complicadas, tales como el tráfico y la localización intracelulares, la regulación de la actividad enzimática, e interacciones de mediación con otros componentes celulares.

El inesperado hallazgo de que hay múltiples genes de HAS de vertebrado que codifican diferentes sintasas respalda acusadamente el consenso emergente de que HA es un regulador importante del comportamiento celular y no simplemente un componente estructural de los tejidos. De este modo, en menos de seis meses, el campo se movió de un HAS clonado y conocido (spHAS) al reconocimiento de una familia multigénica que promete futuros avances rápidos, numerosos y excitantes en la comprensión de la síntesis y la biología de HA.

Por ejemplo, más adelante se describen las secuencias de genes de HAS de *Pasteurella multocida*, virus de *Chlorella-Paramecium bursaria* (PBCV-1; del inglés, *Paramecium bursaria Chlorella virus*), *Streptococcus pyogenes*, *Streptococcus uberis*, *Sulfolobus solfataricus*, plásmido pXO1 de *Bacillus anthracis*, y virus de *Ectocarpus siliculosus*. La presencia de hialuronano sintasa en estos sistemas y la purificación y el uso de la hialuronano sintasa de estos diferentes sistemas indican la capacidad para purificar y aislar secuencias de ácido nucleico que codifican hialuronano sintasa enzimáticamente activa en muchas diferentes fuentes procarióticas y virales; en realidad, de fuentes microbianas en general.

La cepa D181 de *Streptococcus equisimilis* del Grupo C sintetiza y secreta ácido hialurónico (HA). Los investigadores han utilizado esta cepa y cepas de *Streptococcus pyogenes* del Grupo A, tales como S43 y A111, para estudiar la biosíntesis de HA y caracterizar la actividad sintetizadora de HA en términos de su necesidad de cationes divalentes, la utilización de precursores (UDP-GlcNAc y UDP-GlcUA) y su pH óptimo.

Tradicionalmente, se ha preparado comercialmente HA por aislamiento a partir de crestas de gallo o de medios extracelulares de cultivos estreptocócicos. Un método que se ha desarrollado para preparar HA es por medio del uso de cultivos de bacterias estreptocócicas que producen HA. En la Patente de EE.UU. n° 4.517.295 se describe dicho procedimiento mediante el cual se fermentan estreptococos productores de HA bajo condiciones anaeróbicas en un medio de crecimiento enriquecido en CO₂. Bajo estas condiciones, se produce HA que puede ser extraído del caldo. Se cree generalmente que el aislamiento de HA a partir de crestas de gallo es laborioso y difícil ya que se empieza con HA en un estado menos puro. La ventaja del aislamiento a partir de crestas de gallo es que el HA producido es de mayor peso molecular. Sin embargo, la preparación de HA por fermentación bacteriana es más sencilla ya que el HA con que se comienza es de mayor pureza. Sin embargo, el peso molecular del HA producido de esta manera es normalmente más pequeño que el de HA procedente de crestas de gallo. Además, el HA preparado por fermentación estreptocócica a menudo provoca respuestas inmunes como lo hace el HA obtenido de crestas de gallo. Por lo tanto, una técnica que permita la producción de HA de alto peso molecular por fermentación bacteriana es una acusada mejoría con respecto a los procedimientos existentes.

Como se mencionó previamente, el HA de alto peso molecular presenta una gran variedad de aplicaciones útiles que van de la cosmética a la cirugía ocular. A causa de su potencial para una elevada viscosidad y de su alta biocompatibilidad, se encuentra una aplicación particular para el HA en la cirugía ocular, como reemplazo del humor vítreo. También se ha usado HA para tratar caballos de carreras con artritis traumática mediante inyecciones intraarticulares de HA, como lubricante en la crema de afeitar, y en una diversidad de productos cosméticos a causa de sus propiedades fisicoquímicas de alta viscosidad y su capacidad para retener la humedad durante largos periodos de tiempo. De hecho, en Agosto de 1997, la "Food and Drug Agency" de EE.UU. aprobó el uso de HA de alto peso molecular en el tratamiento de la artritis grave por medio de la inyección directa de dicho HA de alto peso molecular en las articulaciones afectadas. En general, cuanto mayor es el peso molecular del HA que se emplea, mejor. Esto es porque la viscosidad de la disolución de HA aumenta con el peso molecular medio de las moléculas polímeras de HA individuales en la disolución. Desafortunadamente, mediante los procedimientos de aislamiento actualmente disponibles ha sido difícil obtener HA de peso molecular muy elevado, tal como el que llega hasta 10^7 . Sin embargo, los métodos de producción recombinantes aquí descritos permiten la producción de un HA que tiene un peso molecular de hasta 10^7 y más.

Para abordar éstas u otras dificultades, existe la necesidad de nuevos métodos y construcciones que se puedan usar para producir HA que tenga una o más propiedades mejoradas, tal como mayor pureza o mayor facilidad de preparación. En particular, existe la necesidad de desarrollar una metodología para la producción de cantidades de HA relativamente puro y de peso molecular relativamente elevado mayores que las actualmente disponibles en el comercio. Aún existe otra necesidad: la de poder desarrollar una metodología para la producción de un HA que tenga una distribución de tamaños modificada ($HA_{\Delta\text{tamaño}}$) así como de un HA que tenga una estructura modificada ($HA_{\Delta\text{mod}}$).

La producción industrial de enzimas es de 1.500 millones de dólares por negocio anual, y el setenta por ciento de estos productos se producen a partir de especies de *Bacillus*. La ventaja de usar un sistema de expresión eficaz en cepas de *Bacillus* es que el *Bacillus* es un eficaz secretor de proteínas, y, por lo tanto, la sustitución de *E. coli* o levadura por *Bacillus* en procesos para la producción de proteínas genéticamente diseñadas proporciona una secreción aumentada de la proteína en cuestión. Las cepas de *Bacillus* son además microorganismos "generalmente reconocidos como seguros" (GRAS; del inglés, generally recognized as safe) en comparación con otras cepas bacterianas comúnmente usadas, tales como las de *E. coli*. Durante mucho tiempo se ha usado el *Bacillus* en la industria de alimentos y bebidas y en la producción de antibióticos. Una ventaja del *Bacillus* es que no contiene sustancias pirógenas ni produce toxinas. Hay una amplia experiencia industrial en el uso de *Bacillus* en fermentaciones, tal como en la producción de proteasas y alfa-amilasa para detergentes.

El presente invento aborda una o más limitaciones de la técnica. Se describen métodos para producir, usando la tecnología de DNA recombinante, una HAS enzimáticamente activa en una célula de *Bacillus* en la que se ha introducido un segmento purificado de ácido nucleico que tiene una región de codificación que codifica la HAS enzimáticamente activa, junto con la preparación de células de *Bacillus* recombinantes que producen HAS y su producto, ácido hialurónico.

40 Breve resumen del invento

La presente descripción implica la aplicación de la tecnología de DNA recombinante a la resolución de uno o más problemas de la técnica para preparar ácido hialurónico (HA). Estos problemas se abordan mediante el aislamiento y el uso de un segmento de ácido nucleico que tiene una región de codificación que codifica un gen de hialuronato sintasa (HAS) enzimáticamente activa, un gen responsable de la biosíntesis de la cadena de HA, tal como un gen de HAS de *Streptococcus* del Grupo A o C, *Pasteurella multocida*, *Sulfolobus solfataricus*, y virus de *Ectocarpus siliculosus*. Los genes de HAS aquí descritos fueron clonados a partir de DNA de una apropiada fuente bacteriana y fueron creados en construcciones recombinantes útiles que fueron introducidas en una célula de *Bacillus* para la preparación de HA y para la preparación de grandes cantidades de la propia enzima HAS.

El presente invento proporciona una célula huésped de *Bacillus* recombinante y un método para producir ácido hialurónico usando la célula huésped, como se definen en las Reivindicaciones 1 y 18, respectivamente.

Las expresiones "ácido hialurónico sintasa", "hialuronato sintasa", "hialuronano sintasa" y "HA sintasa" se usan indistintamente para describir una enzima que polimeriza una cadena polisacárida de glicosaminoglicano compuesta de azúcares alternos de ácido glucurónico y N-acetilglucosamina, enlazados en β -1,3 y β -1,4. Por ejemplo, el término "seHAS" describe la enzima HAS procedente de *Streptococcus equisimilis*, en que la expresión del gen que codifica la enzima seHAS se correlaciona con la virulencia de las cepas estreptocócicas del Grupo A y el Grupo C al proporcionar un medio para escapar de la fagocitosis y la vigilancia inmune.

La presente descripción trata del aislamiento y la caracterización de genes de hialuronato o ácido hialurónico sintasa, cDNAs, y productos génicos (HAS), como se pueden usar para la polimerización de ácido glucurónico y N-acetilglucosamina en el glicosaminoglicano ácido hialurónico. En el presente documento se identifica el locus de HAS y se describen las secuencias de ácido nucleico que codifican los genes de HAS enzimáticamente activa procedentes de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, *Pasteurella multocida*, *Sulfolobus solfataricus*, pXO1 de *Bacillus anthracis*, y virus de *Ectocarpus siliculosus*. Los genes de HAS también proporcionan nuevas sondas para evaluar el potencial de especímenes bacterianos para producir ácido hialurónico.

Por medio de la aplicación de técnicas y el conocimiento aquí expuesto, quienes tienen experiencia en la técnica serán capaces de obtener segmentos de ácido nucleico adicionales que codifiquen genes de HAS. Como reconocerán quienes tienen experiencia en la técnica a la luz de la presente descripción, estas ventajas proporcionan una utilidad significativa en cuanto a poder controlar la expresión del gen de HAS y controlar la naturaleza del producto del gen de HAS, la enzima HAS, que se produce.

En consecuencia, la descripción se dirige al aislamiento de un segmento purificado de ácido nucleico que tiene una región de codificación que codifica una HAS enzimáticamente activa, proceda de fuentes procarióticas o proceda de fuentes eucarióticas. Esto es posible porque la enzima, y en realidad el gen, se halla tanto en eucariontes como en algunos procariontes. También se sabe que los eucariontes producen HA y, por lo tanto, tienen genes de HA sintasa que pueden ser empleados.

Los segmentos de ácido nucleico que codifican HA sintasa empleados en el presente invento se definen como aislados exentos de DNA genómico o cromosómico total, por lo que pueden ser fácilmente manipulados mediante técnicas de DNA recombinante. En consecuencia, como aquí se utiliza, la frase “un segmento purificado de ácido nucleico” se refiere a un segmento de DNA aislado exento de DNA genómico o cromosómico no relacionado, y conservado en un estado que le hace útil para la práctica de técnicas recombinantes, tal como DNA en forma de un fragmento de DNA aislado independiente o un vector (por ejemplo, un plásmido, fago o virus) que lleva incorporado dicho fragmento.

El presente invento se refiere a una célula huésped recombinante, en que la célula huésped recombinante es una célula de *Bacillus* que comprende un vector recombinante que comprende un segmento purificado de ácido nucleico que tiene una región de codificación que codifica la hialuronano sintasa enzimáticamente activa de ID. SEC. n° 12. El segmento purificado de ácido nucleico puede comprender una secuencia de nucleótidos de acuerdo con la ID. SEC. n° 11. El vector recombinante puede ser introducido en la célula de *Bacillus* mediante al menos una de las técnicas siguientes: transformación, transfección, transducción y electroporación.

La región de codificación anteriormente descrita está bajo el control de un promotor, tal como un promotor compatible con bacterias Gram-positivas o un promotor compatible con *Bacillus*. La célula huésped recombinante también puede incluir al menos un promotor de RNA polimerasa modificado de modo que, cuando el promotor de RNA polimerasa modificado es reconocido por una RNA polimerasa, la RNA polimerasa es capaz de expresar RNA en una cantidad mayor que en el caso de un promotor de RNA polimerasa endógeno. Dicha modificación puede ser una mutación o la presencia de elementos promotores en tándem, que pueden ser el mismo elemento promotor o diferentes elementos promotores. Además, la célula huésped recombinante puede incluir al menos un elemento estabilizador o desestabilizador de mRNA adicional con respecto a los hallados en una célula de *Bacillus* nativa.

La célula de *Bacillus* puede tener una producción potenciada de al menos uno de UDP-GlcUA y UDP-GlcNAc. La célula huésped recombinante tiene además al menos un segmento purificado de ácido nucleico que tiene una región de codificación que codifica una enzima enzimáticamente activa de la ruta biosintética de UDP-GlcUA, seleccionada del grupo que consiste en UDP-glucosa deshidrogenasa, UDP-glucosa pirofosforilasa y combinaciones de las mismas. Dicho segmento purificado de ácido nucleico puede estar dispuesto en el vector recombinante anteriormente descrito o puede estar dispuesto en un vector recombinante diferente. Cuando está dispuesto en el mismo vector, las dos regiones de codificación pueden estar bajo el control de al menos una copia de al menos un promotor o bajo el control de promotores diferentes. La presencia del al menos un segmento de ácido nucleico que codifica una enzima de la ruta de biosíntesis del precursor de UDP-azúcar proporcionará a la célula huésped recombinante una actividad mayor que en el caso de una célula huésped nativa que expresa una enzima endógena de la ruta de biosíntesis del precursor de UDP-azúcar.

En una alternativa más, la célula huésped recombinante puede incluir al menos un gen mutado de la biosíntesis del precursor de UDP-azúcar, en que la mutación aumenta la semivida de un mRNA transcrito a partir de él, codifica un mRNA que tiene una eficacia de traducción aumentada, o tiene lugar en un sitio de unión a ribosomas presente en el gen para la biosíntesis del precursor de UDP-azúcar de modo que un ribosoma tiene una afinidad ligante aumentada por el sitio de unión al ribosoma.

El presente invento comprende además un método para producir ácido hialurónico, que comprende construir la célula huésped recombinante anteriormente descrita introduciendo el(los) segmento(s) de ácido nucleico purificado(s) anteriormente descrito(s) y cultivando la célula huésped recombinante en un medio para que secrete ácido hialurónico. El huésped de *Bacillus* puede ser cultivado a una temperatura en el intervalo de aproximadamente 25°C a aproximadamente 42°C en medios químicamente definidos, medios complejos o un medio que contiene glucosa y al menos uno de N-acetilglucosamina y glucosamina. El ácido hialurónico secretado es luego recuperado, y el ácido hialurónico recuperado puede ser además extraído del medio y ser luego purificado. Por ejemplo, el ácido hialurónico puede ser separado de las células y los fragmentos celulares mediante al menos una de las técnicas siguientes: filtración, centrifugación y floculación, lo que va seguido de la concentración del ácido hialurónico y luego de la separación del ácido hialurónico concentrado del medio mediante al menos un método seleccionado del grupo que consiste en precipitación, ultrafiltración y diálisis. Esta separación puede incluir además la adición de ácido tricloroacético, el cual facilita la separación del ácido hialurónico de las células y los fragmentos celulares. El agente de precipitación puede ser al menos uno de entre un alcohol, un disolvente o compuesto orgánico y una sal alifática positivamente cargada, y puede ser seleccionado del grupo que consiste en etanol, isopropanol, acetona, bromuro de cetiltrimonio y cloruro de cetilpiridinio.

La presente descripción comprende además ácido hialurónico preparado mediante los métodos anteriormente descritos.

Breve descripción de las diversas vistas de los dibujos

La Figura 1 representa que no tiene lugar hibridación cruzada entre genes de seHAS y spHAS. La sonda del Grupo A utilizada en los carriles 1 y 2 sólo se hibrida con DNA del Grupo A (carril 2), mientras que la sonda del Grupo C utilizada en los carriles 3 y 4 sólo se hibrida con el carril 3.

La Figura 2 representa figuradamente la relación de seHAS con las proteínas HAS bacterianas, víricas y eucarióticas.

La Figura 3 representa figuradamente posibles similitudes y relaciones evolutivas entre algunas de las hialuronano sintasas conocidas.

La Figura 4 representa la distribución de tamaños de HA producida por diversas enzimas HAS estreptocócicas genéticamente modificadas.

La Figura 5 representa figuradamente la sobreexpresión de seHAS y spHAS recombinantes en *E. coli*.

La Figura 6 representa la producción de HA recombinante en *Bacillus subtilis* y *Bacillus licheniformis*.

La Figura 7 representa la purificación de HA sintasa estreptocócica.

La Figura 8 representa un análisis, por filtración en gel, de HA sintetizado por HAS estreptocócica recombinante expresada en membranas de levadura.

La Figura 9 es un análisis, por transferencia Western, de seHAS recombinante utilizando anticuerpos específicos.

La Figura 10 es un análisis cinético de las distribuciones de tamaños de HA producidas por seHAS y spHAS recombinantes.

La Figura 11 representa gráficamente los gráficos de hidropatía para seHAS y previstas regiones asociadas a la membrana.

La Figura 12 es un modelo gráfico para la organización topológica de seHAS en la membrana.

La Figura 13 es una demostración de la síntesis de HA auténtico por la seHAS recombinante.

La Figura 14 representa el reconocimiento de secuencias de ácido nucleico que codifican seHAS, que codifican spHAS o que codifican tanto seHAS como spHAS, usando oligonucleótidos específicos y PCR.

La Figura 15 representa oligonucleótidos utilizados para una hibridación específica por PCR.

La Figura 16A es un gráfico que representa la producción de HA recombinante en *Bacillus subtilis* vivos comparando la producción de HA por *Bacillus subtilis* 168 (vector pSM143 solo) con aquella por *Bacillus subtilis* 168 (pSM143 que contiene seHAS). La Figura 16B es una ampliación de una sección del gráfico de la Figura 16A.

La Figura 17A es un gráfico que representa el control nutricional de la distribución de tamaños de HA recombinante producida por spHAS en *Bacillus subtilis* vivos.

La Figura 17B es un gráfico que representa la producción de HA recombinante en *Bacillus subtilis* 168 vivos en comparación con la de *Bacillus subtilis* que contiene el vector solo.

Las Figuras 18A y 18B son fotomicrografías de *E. coli* recombinante. En la Figura 18A, la tinción con tinta india (ampliación 1000x) revela que células de *E. coli* K5 con pPmHAS producen una cápsula sustancial que aparece como un halo blanco alrededor de las células. En la Figura 18B, el material capsular podía ser separado de las células de *E. coli* K5 (pPmHAS) mediante un breve tratamiento con HA liasa de *Streptomyces*. PmHAS dirige la polimerización del polisacárido HA.

La Figura 19 es un modelo esquemático de la biosíntesis de GAG en bacterias Gram-positivas y Gram-negativas.

La Figura 20 es un gel de agarosa que demuestra la multiplicación del gen de HAS de *Streptococcus uberis* por PCR.

La Figura 21 representa la actividad HA sintasa de *Streptococcus pyogenes*, *Streptococcus equisimilis* y *Streptococcus uberis*.

Descripción detallada del invento

Antes de explicar con detalle al menos una realización del invento, ha de entenderse que el invento no se limita en su aplicación a los detalles de la construcción y la disposición de los componentes expuestos en la descripción siguiente o ilustrados en los dibujos. El invento tiene capacidad para otras realizaciones o para ser llevado a cabo o a la práctica de diversos modos. Además, se ha de entender que la fraseología y la terminología aquí empleadas tienen el fin de descripción y no deben ser consideradas restrictivas.

Como aquí se utiliza, las expresiones “segmento de ácido nucleico” y “segmento de DNA” se utilizan indistintamente y se refieren a una molécula de DNA que ha sido aislada exenta de DNA genómico total de una especie concreta. Por lo tanto, como aquí se utiliza, un segmento de DNA o ácido nucleico “purificado” se refiere a un segmento de DNA que contiene una secuencia de codificación de hialuronato sintasa (“HAS”) que es aislada alejada de, o purificada exenta de, DNA genómico no relacionado de la célula fuente. Dentro de la expresión “segmento de DNA” se incluyen segmentos de DNA y fragmentos más pequeños de dichos segmentos, y también vectores recombinantes, incluyendo, por ejemplo, plásmidos, cósmidos, fagos, virus y similares.

Similarmente, un segmento de DNA que comprende un gen de HAS aislado o purificado se refiere a un segmento de DNA que incluye secuencias de codificación de HAS aisladas sustancialmente alejadas de otros genes o secuencias codificadoras de proteína presentes en la naturaleza. A este respecto, por sencillez, el término “gen” se utiliza para referirse a una unidad que codifica una proteína, polipéptido o péptido funcional. Como entenderán los expertos en la técnica, este término funcional incluye secuencias genómicas, secuencias de cDNA y combinaciones de las mismas. “Aislado sustancialmente alejado de otras secuencias de codificación” significa que el gen de interés, de HAS en este caso, forma la parte significativa de la región de codificación del segmento de DNA y que el segmento de DNA no contiene grandes porciones de DNA codificador presente en la naturaleza, tal como fragmentos cromosómicos grandes u otras regiones codificadoras de DNA o genes funcionales. Por supuesto, esto se refiere al segmento de DNA en la forma originalmente aislada y no excluye genes ni regiones de codificación añadidas más tarde a él o dejadas intencionalmente en el segmento por la mano del hombre.

Debido a ciertas ventajas asociadas con el uso de fuentes procarióticas, uno se dará probablemente cuenta de la mayoría de las ventajas tras el aislamiento del gen de HAS de procariontes. En particular, se puede elegir usar una HAS de Clase I o Clase II, tal como una HAS de Clase I de *S. equisimilis* o *S. pyogenes*, o una HAS de Clase II de *P. multocida*.

Los estreptococos se subdividen taxonómicamente en Grupos de Lancefield basándose en los diferentes hidratos de carbono antigénicos de la pared celular. Hay 18 grupos distintos, pero los patógenos más comunes son A, B, C y D. Históricamente, los patógenos más comunes también reciben a menudo nombres de especie específicos, pero el método de ensayo unificado de Lancefield es reconocido por ser un método claro de tipado y, por lo tanto, un esquema de clasificación útil. La especie de *Streptococcus* que se utiliza como fuente del gen de HAS para su empleo en el presente invento es el estreptococo *S. uberis* del Grupo C.

Una de dichas ventajas de aislar el gen de HAS de procariontes es que, típicamente, las enzimas eucarióticas pueden requerir significativas modificaciones postraduccionales que sólo se pueden llevar a cabo en un huésped eucariótico. Esto tenderá a limitar la aplicabilidad de cualquier gen eucariótico de HA sintasa que se obtenga. Además, quienes tienen una experiencia normal en la técnica, se darán probablemente cuenta de las ventajas adicionales en términos de tiempo y de sencillez de manipulación genética cuando se procura emplear un gen de enzima procariótico. Estas ventajas adicionales incluyen (a) la sencillez del aislamiento de un gen procariótico a causa del tamaño relativamente pequeño del genoma y, por lo tanto, de la cantidad reducida de exploración en el correspondiente banco genómico, y (b) la sencillez de manipulación porque el tamaño global de la región de codificación de un gen procariótico es significativamente más pequeño a causa de la ausencia de intrones. Además, si el producto del gen de HAS (es decir, la enzima) requiere modificaciones postraduccionales, éstas se llevarían mejor a cabo en un similar ambiente (huésped) celular procariótico del que se obtuvo el gen.

Preferiblemente, las secuencias de DNA para uso de acuerdo con el presente invento incluirán además regiones de control genético que permitan la expresión de la secuencia en un huésped recombinante seleccionado. Por supuesto, la naturaleza de la región de control empleada variará generalmente dependiendo del uso particular concebido (por ejemplo, la clonación del huésped).

En realizaciones particulares, el invento trata del uso de segmentos de DNA aislados y vectores recombinantes que incorporan secuencias de DNA y que codifican un gen de HAS, que incluye en su secuencia de aminoácidos una secuencia de aminoácidos de acuerdo con la ID. SEC. n° 12. Además, en otras realizaciones particulares, el invento trata del uso de segmentos de DNA aislados y vectores recombinantes que incorporan secuencias de DNA que codifican un gen que incluye en su secuencia de aminoácidos la secuencia de aminoácidos de un gen o DNA de HAS, y, en particular de un gen o cDNA de HAS, que corresponde a HAS de *Streptococcus uberis*. Por ejemplo, cuando el segmento de DNA o el vector codifica una proteína HAS de longitud completa, o se pretende su uso en la expresión de la proteína HAS, las secuencias preferidas son aquéllas que son esencialmente como la expuesta en la ID. SEC. n° 12.

Mediante los métodos aquí descritos se pueden aislar segmentos de ácido nucleico que tengan actividad HA sintasa. La expresión “una secuencia esencialmente como la expuesta en la ID. SEC. n° X” significa que la secuencia

corresponde sustancialmente a una porción de la ID. SEC. n° X y tiene relativamente pocos aminoácidos que no son idénticos a, o un equivalente biológicamente funcional de, los aminoácidos de la ID. SEC. n° X. La expresión “equivalente biológicamente funcional” es bien entendida en la técnica y se define aquí además con detalle como un gen que tiene una secuencia esencialmente como la expuesta en la ID. SEC. n° X y que está asociado con la capacidad de los procariontes para producir HA o un revestimiento de ácido hialurónico.

Por ejemplo, las secuencias de codificación de seHAS y spHAS tienen una identidad de aproximadamente 70% y son ricas en las bases adenina (A) y timina (T). El contenido de bases de seHAS es A-26,71%, C-19,13%, G-20,81% y T-33,33% (A/T = 60%), mientras que el de spHAS es A-31,34%, C-16,42%, G-16,34% y T-35,8% (A/T = 67%). Quienes tienen una experiencia normal en la técnica se sorprenderán de que la secuencia de codificación de seHAS no se hibride con el gen de spHAS y viceversa a pesar de tener una identidad del 70%. Esta inesperada incapacidad para hibridarse cruzadamente podría ser debida a cortas interrupciones de bases desapareadas por todos los marcos de lectura abiertos. La incapacidad de spHAS y seHAS para hibridarse cruzadamente se muestra en la Figura 1. El tramo más largo de nucleótidos idénticos comunes a ambas secuencias de codificación de seHAS y spHAS tiene sólo 20 nucleótidos. Además, las secuencias muy ricas en A-T formarán complejos de hibridación menos estables que las secuencias ricas en G-C. Otra posible explicación podría ser que hubiera varios tramos de As o Ts en ambas secuencias que pudieran hibridarse de un modo mal alineado e inestable. Esto pondría las secuencias génicas de seHAS y spHAS fuera de marco una con respecto a la otra, disminuyendo por ello la probabilidad de una hibridación productiva.

A causa de este fenómeno único de dos genes que no son capaces de hibridarse cruzadamente entre sí y codifican proteínas que tienen una identidad de 70%, es beneficioso pensar en el segmento de ácido nucleico en términos de su función; es decir, un segmento de ácido nucleico que codifica una hialuronato sintasa enzimáticamente activa. Quien tiene una experiencia normal en la técnica apreciará que un segmento de ácido nucleico que codifica una hialuronato sintasa enzimáticamente activa puede contener sustituciones conservadas o semiconservadas con respecto a las secuencias expuestas en las ID. SEC. números 11 y 12 y, sin embargo, estar aún dentro del alcance del invento.

En particular, la técnica está repleta de ejemplos sobre la capacidad de los profesionales para hacer cambios estructurales en un segmento de ácido nucleico (es decir, que codifican sustituciones de aminoácido conservadas o semiconservadas) y conservar aún su actividad enzimática o funcional. Véanse, por ejemplo: (1) Risler *et al.*, “Amino Acid Substitutions in Structurally Related Proteins. A Pattern Recognition Approach”, J. Mol. Biol. 204: 1019-1029 (1988) [“... de acuerdo con la observada capacidad de intercambio de cadenas laterales de aminoácidos, sólo se pudieron esbozar cuatro grupos: (i) Ile y Val; (ii) Leu y Met, (iii) Lys, Arg y Gln, y (iv) Tyr y Phe”]; (2) Niefind *et al.*, “Amino Acid Similarity Coefficients for Protein Modeling and Sequence Alignment Derived from Main-Chain Folding Anoles”, J. Mol. Biol. 219: 481-497 (1991) [“los parámetros de similitud permiten que se diseñen sustituciones de aminoácidos”]; y (3) Overington *et al.*, “Environment-Specific Amino Acid Substitution Tables: Tertiary Templates and Prediction of Protein Folds”, Protein Science 1: 216-226 (1992) [“El análisis del patrón de sustituciones observadas en función del ambiente local muestra que hay distintos patrones... Se pueden hacer cambios compatibles.”].

Estas referencias y otras innumerables indican que quien tiene una experiencia normal en la técnica, dada una secuencia de ácido nucleico, podría hacer sustituciones y cambios en la secuencia de ácido nucleico sin cambiar su funcionalidad. Además, un segmento de ácido nucleico sustituido puede ser muy idéntico y conservar su actividad enzimática con respecto a su precursor no adulterado y, no obstante, fracasar aún a la hora de hibridarse con él.

En el presente documento se describen segmentos de ácido nucleico que codifican hialuronato sintasas enzimáticamente activas, tales como seHAS, spHAS, suHAS y pmHAS. Aunque seHAS y spHAS tienen una identidad de 70% y ambas codifican una hialuronato sintasa enzimáticamente activa, no presentan hibridación cruzada. De esta manera, quien tiene una experiencia normal en la técnica apreciará que se pueden realizar sustituciones en el segmento de ácido nucleico de seHAS expuesto en la ID. SEC. n° 1 sin cambiar su funcionalidad. En la Tabla I se presentan sustituciones normalizadas y aceptadas de aminoácidos funcionalmente equivalentes.

TABLA I

Grupos de aminoácidos	Sustituciones conservativas y semiconservativas
Grupos R apolares	Alanina, valina, leucina, isoleucina, prolina, metionina, fenilalanina, triptófano
Grupos R polares pero no cargados	Glicocola, serina, treonina, cisteína, asparagina, glutamina
Grupos R negativamente cargados	Ácido aspártico, ácido glutámico
Grupos R positivamente cargados	Lisina, arginina, histidina

En el presente invento se emplea un segmento purificado de ácido nucleico que codifica una proteína de acuerdo con la ID. SEC. nº 12, definido además como un vector recombinante. Como aquí se utiliza, la expresión “vector recombinante” se refiere a un vector que ha sido modificado para que contenga un segmento de ácido nucleico que codifica una proteína HAS, o un fragmento del mismo. El vector recombinante puede ser además definido como un vector de expresión que comprende un promotor operativamente unido a dicho segmento de ácido nucleico que codifica HAS.

El presente invento es una célula huésped, hecha recombinante con un vector recombinante que comprende un gen de HAS. Como aquí se usa, con la expresión célula “genéticamente modificada” o “recombinante” se quiere hacer referencia a una célula en que se introducido un gen recombinante, tal como un gen que codifica HAS. Por lo tanto las células genéticamente modificadas son distinguibles de las células presentes en la naturaleza, que no contienen un gen recombinantemente introducido. De este modo, las células genéticamente modificadas son células que tienen un gen o unos genes introducidos por la mano del hombre. Los genes recombinantemente introducidos estarán en forma de un gen de cDNA o de una copia de un gen genómico, o incluirán genes dispuestos adyacentemente a un promotor no naturalmente asociado con el gen introducido concreto.

Cuando se desea utilizar un huésped distinto de *Streptococcus*, como se puede utilizar para producir HA sintasa recombinante, puede resultar ventajoso emplear un sistema procariótico tal como *E. coli*, cepas de *Bacillus*, o *Lactococcus* sp., o incluso sistemas eucarióticos tales como levaduras y células de ovario de hámster chino, células renales de mono verde africano, células VERO y similares. Por supuesto, cuando se acomete esto, será generalmente deseable poner el gen de HA sintasa bajo el control de secuencias que sean funcionales en el huésped alternativo seleccionado. Las secuencias de control de DNA apropiadas, así como su construcción y su uso, son generalmente bien conocidas en la técnica, como se discute más adelante con mayor detalle. Por ejemplo, de acuerdo con el invento, la célula huésped es una célula de *Bacillus*, tal como una célula de *Bacillus subtilis* o *Bacillus licheniformis*, y el vector introducido en ella contiene un promotor compatible con *Bacillus* al que está operativamente unido el gen de HAS.

En una realización más preferida, la célula huésped es una célula de *Bacillus*, tal como de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus metaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis* y *Bacillus thuringiensis*.

En realizaciones preferidas, los segmentos de DNA que codifican Ha sintasa incluyen además secuencias de DNA, conocidas funcionalmente en la técnica como orígenes de replicación o “replicones”, que permiten la replicación de secuencias contiguas por el huésped particular. Dichos orígenes permiten la preparación de plásmidos o segmentos quiméricos replicantes y extracromosómicamente localizados, a los que están ligadas secuencias de DNA de HA sintasa. En casos más preferidos, el origen empleado es uno capaz de replicación en huéspedes bacterianos adecuados para aplicaciones de biotecnología. Sin embargo, para una mayor versatilidad de los segmentos de DNA clonados, puede ser deseable emplear alternativa o incluso adicionalmente orígenes reconocidos por otros sistemas huésped cuyo uso se contempla (tal como en un vector lanzadera).

El aislamiento y el uso de otros orígenes de replicación, tales como los orígenes del virus SV40, polyomavirus o el virus del papiloma bovino, que pueden emplearse para la clonación o expresión en diversos organismos superiores, son bien conocidos por quienes tienen una experiencia normal en la técnica. En ciertos aspectos, la descripción presente se refiere a un vector de transformación recombinante que incluye la secuencia génica que codifica HA sintasa, junto con un origen de replicación apropiado y bajo el control de regiones de control seleccionadas.

De este modo, quienes tienen experiencia en la técnica apreciarán, a la luz de la presente descripción, que se pueden utilizar otros medios para obtener el gen o cDNA de HAS. Por ejemplo, se pueden obtener fragmentos de DNA, producidos por la reacción en cadena de la polimerasa o RT-PCR, que contengan complementos completos de genes o cDNAs de diversas fuentes, incluyendo otras cepas de *Streptococcus*, o de fuentes eucarióticas, tales como bancos de cDNA. Para la generación de fragmentos de DNA de acuerdo con la presente descripción, se puede emplear casi cualquier procedimiento de clonación molecular. De este modo, la única limitación, generalmente sobre el método particular empleado para el aislamiento de DNA, es que los ácidos nucleicos aislados deben codificar una HA sintasa equivalente biológicamente funcional.

Una vez que se ha aislado el DNA, es ligado junto con un vector seleccionado. Para obtener las ventajas de acuerdo con el invento, se puede emplear casi cualquier vector de clonación. Los vectores útiles típicos incluyen plásmidos y fagos para uso en organismos procarióticos, e incluso vectores víricos para uso en organismos eucarióticos. Los ejemplos incluyen pKK223-3, pSA3, lambda recombinante, SV40, polyomavirus, adenovirus, virus del papiloma bovino y retrovirus. Sin embargo, se cree que las ventajas particulares se obtendrán finalmente cuando se empleen vectores capaces de replicación tanto en cepas de *Lactococcus* o *Bacillus* como en *E. coli*.

Vectores tales como estos, ejemplificados por el vector pSA3 de Dao y Ferretti o el vector pAT19 de Trieu-Cuot *et al.*, permiten que uno lleve a cabo una selección de colonias clonales en un huésped fácilmente manipulable tal como *E. coli*, seguida de la subsiguiente transferencia de nuevo a una cepa de *Lactococcus* o *Bacillus* de calidad alimentaria para la producción de HA. Estos son organismos benígnos y bien estudiados que se utilizan en la producción de ciertos alimentos y productos de biotecnología. Estos son ventajosos ya que uno puede aumentar la capacidad de la cepa de *Lactococcus* o *Bacillus* para sintetizar HA a través de una dosificación génica (es decir, proporcionando copias extras

del gen de HA sintasa por multiplicación) y/o la inclusión de genes adicionales para aumentar la disponibilidad de precursores de HA. La capacidad inherente de una bacteria para sintetizar HA puede ser también aumentada a través de la formación de copias extras, o la multiplicación, del plásmido que porta el gen de HA sintasa. Esta multiplicación puede representar un aumento de hasta 10 órdenes de magnitud en el número de copias plasmídicas y, por lo tanto, en el número de copias del gen de HA sintasa.

Otro procedimiento que aumentaría más el número de copias del gen de HA sintasa es la inserción de múltiples copias del gen en el plásmido. Otra técnica incluiría integrar el gen de HAS en DNA cromosómico. Esta multiplicación extra sería especialmente factible ya que el tamaño del gen bacteriano de HA sintasa es pequeño. En algunos escenarios, el vector ligado a DNA cromosómico es empleado para transfectar el huésped que es seleccionado con fines de exploración clonal, tal como *E. coli*, a través del uso de un vector que es capaz de expresar en el huésped elegido el DNA insertado.

En otro aspecto, el gen de HA sintasa se introduce en el cromosoma de la célula huésped por medio de recombinación homóloga o heteróloga. El gen *has* puede ser más estable en esta configuración, especialmente sin selección por fármacos. Para introducir el gen *has* en *Bacillus*, tal como pTLH o pKSV7, o en levadura, tal como YIP211, o en células animales, tales como pcDNA/FRT, se pueden emplear diversos vectores. El DNA es introducido primero en la célula huésped por transformación, transducción o electroporación. El segmento de DNA con el gen *has* se integra luego establemente en el cromosoma del huésped. Por ejemplo, se usó el gen *spHAS* para reparar un cromosoma mutante de *Streptococcus* mediante transducción e integración; esta operación dio lugar a la producción de HA (DeAngelis *et al.*, 1993).

Cuando se emplea una fuente eucariótica, tal como fibroblastos dérmicos o sinoviales o células de cresta de gallo, uno desearía proceder inicialmente preparando un banco de cDNA. Esto se lleva primero a cabo por aislamiento de mRNA de las células anteriores, seguido de la preparación de cDNA de doble cadena utilizando una enzima con actividad transcriptasa inversa y de ligación con el vector seleccionado. Numerosas posibilidades están disponibles y son conocidas en la técnica para la preparación del cDNA de doble cadena, y se cree que todas las citadas técnicas son aplicables. Una técnica preferida implica transcripción inversa. Una vez que se ha obtenido una población de cDNA de doble cadena, se prepara un banco de cDNA mediante técnicas aceptadas en el huésped seleccionado, tales como por ligación en el vector apropiado y multiplicación en el huésped apropiado. Debido al elevado número de clones que se obtienen y a la relativa facilidad para explorar grandes números de clones mediante las técnicas aquí expuestas, uno puede desear emplear vectores de expresión de fagos, tales como λ gt11, λ gt12, λ Gem11 y/o λ ZAP, para la exploración de la clonación y expresión de los clones de cDNA.

En ciertas otras realizaciones, el invento trata del uso de segmentos de DNA aislados y vectores recombinantes que incluyen en sus secuencias una secuencia de ácido nucleico esencialmente como la expuesta en la ID. SEC. n° 11. Por ejemplo, la expresión “esencialmente como la expuesta en la ID. SEC. n° 11” se usa en el mismo sentido que el anteriormente descrito y significa que la secuencia de ácido nucleico corresponde sustancialmente a una porción de la ID. SEC. n° 11 y tiene relativamente pocos codones que no son idénticos, o funcionalmente equivalentes, a los codones de la ID. SEC. n° 11. La expresión “codón funcionalmente equivalente” se utiliza aquí para referirse a codones que codifican el mismo aminoácido, tales como los seis codones para arginina o serina, y también se refiere a codones que codifican aminoácidos biológicamente equivalentes como los expuestos en la Tabla I.

Se entenderá también que las secuencias de aminoácidos y ácido nucleico pueden incluir restos adicionales, tales como aminoácidos N- o C-terminales adicionales o secuencias de ácido nucleico 5' o 3' adicionales, y, no obstante, aún ser esencialmente como la expuesta en una de las secuencias aquí descritas con tal de que la secuencia satisfaga los criterios anteriormente expuestos, incluyendo el mantenimiento de la actividad biológica de la proteína en lo que se refiere a expresión proteica y actividad enzimática. La adición de secuencias terminales se aplica particularmente a secuencias de ácido nucleico que, por ejemplo, pueden incluir diversas secuencias no codificadoras que flanqueen cualquiera de las porciones 5' y 3' de la región codificadora o pueden incluir diversas secuencias internas de las que se sabe que aparecen en los genes. En particular, parece que la secuencia de aminoácidos del producto del gen *has* es en eucariontes un 40% más grande que la hallada en procariontes.

Teniendo en cuenta la degeneración del código genético así como las sustituciones conservadas y semiconservadas, las secuencias que tienen entre aproximadamente 40% y aproximadamente 80%, o más preferiblemente entre aproximadamente 80% y aproximadamente 90%, o incluso más preferiblemente entre aproximadamente 90% y aproximadamente 99%, de nucleótidos que son idénticos a los nucleótidos de la ID. SEC. n° 11 serán secuencias que son “esencialmente como la expuesta en la ID. SEC. n° 11”. Las secuencias que son esencialmente las mismas que la expuesta en la ID. SEC. n° 11 pueden ser también funcionalmente definidas como secuencias que son capaces de hibridarse con un segmento de ácido nucleico que contiene el complemento de la ID. SEC. n° 11, bajo condiciones de hibridación estándares o menos rigurosas. Las condiciones de hibridación estándares adecuadas serán conocidas por los expertos en la técnica y son claramente expuestas aquí.

Como aquí se utiliza, la expresión “condiciones de hibridación estándares” se usa para describir aquellas condiciones bajo las cuales segmentos de ácido nucleico sustancialmente complementarios formarán apareamientos de bases de Watson-Crick estándares. Se conocen diversos factores que determinan la especificidad de la unión o la hibridación, tales como el pH, la temperatura, la concentración salina, la presencia de agentes, tales como formamida y dimetilsulfóxido, la longitud de los segmentos que se hibridan, y similares. Cuando se contempla que se van a utilizar

segmentos de ácido nucleico más cortos para la hibridación, por ejemplo, fragmentos de entre aproximadamente 14 y aproximadamente 100 nucleótidos, las condiciones preferidas para hibridación en cuanto a sal y temperatura incluirán tampón de alto fosfato 1,2-1,8 x (HPB; del inglés, high phosphate buffer) a 40-50°C.

5 Naturalmente, el presente invento también abarca el uso de segmentos de DNA que son complementarios, o esencialmente complementarios, de la secuencia expuesta en la ID. SEC. n° 11. Las secuencias de ácido nucleico que son “complementarias” son aquellas que son capaces de presentar apareamiento de bases de acuerdo con las reglas de complementariedad estándares de Watson-Crick. Como aquí se usa, la expresión “secuencias complementarias” significa secuencias de ácido nucleico que son sustancialmente complementarias, como se puede estimar mediante la
10 misma comparación de nucleótidos anteriormente expuesta, o como se definen por ser capaces de hibridarse con el segmento de ácido nucleico de la ID. SEC. n° 11.

Los segmentos de ácido nucleico empleados en el presente invento, independientemente de la longitud de la propia secuencia de codificación, pueden ser combinados con otras secuencias de DNA, tales como promotores, señales
15 de poliadenilación, sitios adicionales para enzimas de restricción, múltiples sitios de clonación, etiquetas epitópicas, regiones de polihistidina, otros segmentos de codificación, y similares, por lo que su longitud global puede variar considerablemente. Por lo tanto, se contempla que se pueda emplear un fragmento de ácido nucleico de casi cualquier longitud, estando preferiblemente limitada la longitud total por la facilidad de preparación y de uso en el protocolo de DNA recombinante previsto.

20 Naturalmente, también se entenderá que este invento no se limita a las particulares secuencias de ácido nucleico y aminoácidos de la ID. SEC. n° 11 y la ID. SEC. n° 12, respectivamente. Por lo tanto, los vectores recombinantes y los segmentos de DNA aislados pueden incluir variadamente las propias regiones de codificación de HAS, regiones de codificación que llevan alteraciones o modificaciones seleccionadas en la región de codificación básica como se
25 define en la Reivindicación 1, o pueden codificar polipéptidos más grandes que, sin embargo, incluyen regiones de codificación de HAS o pueden codificar proteínas o péptidos equivalentes biológicamente funcionales que tienen secuencias de aminoácidos variantes como se define en la Reivindicación 1.

30 Durante mucho tiempo se sospechó que la cápsula de *P. multocida* del tipo A de Carter contenía ácido hialurónico (HA). La caracterización de la HA sintasa de *P. multocida* condujo a interesantes diferencias enzimológicas entre ella y las proteínas seHAS y spHAS.

Las células de *P. multocida* producen una cápsula de HA extracelular fácilmente visible y, puesto que las dos HASs estreptocócicas son proteínas de membrana, se ensayaron preparaciones de membrana del patógeno del cólera
35 aviar. En las primeras pruebas, las fracciones crudas de membrana obtenidas sólo por ultrasonificación poseían niveles bajísimos de incorporación de UDP-[¹⁴C]GlcUA, dependiente de UDP-GlcNAc, a HA {~0,2 picomoles de transferencia de GlcUA [(μg de proteínas)⁻¹h⁻¹]} cuando se ensayaba bajo unas condiciones similares a aquellas para medir la actividad HAS estreptocócica. La enzima de *E. coli* con el plásmido *hasA* recombinante también resultó refractaria al aislamiento al principio. Estos resultados contrastaban con las cantidades fácilmente detectables obtenidas de
40 *Streptococcus* mediante métodos similares.

Un protocolo alternativo para la preparación en que se usaba un tratamiento con lisozima enfriada con hielo en presencia de inhibidores de proteasas y junto con ultrasonificación permitió la recuperación sustancial de actividad HAS de ambas especies de bacterias Gram-negativas. Con el nuevo método, se obtuvieron rutinariamente actividades espe-
45 cíficas para HAS de 5-10 picomoles de GlcUA transferido [(μg de proteínas)⁻¹h⁻¹] en el caso de membranas crudas de *P. multocida* de tipo silvestre. En ausencia de UDP-GlcNAc, casi no se incorporó radiactividad (< 1% del ensayo idéntico con ambos precursores de azúcar) de UDP-[¹⁴C]GlcUA a un material de mayor peso molecular. Las membranas preparadas a partir del mutante acapsular, TnA, no poseían actividad HAS detectable cuando se complementaba con ambos precursores de azúcar-nucleótido (datos no mostrados). Un análisis por filtración en gel utilizando una columna
50 de Sephacryl S-200 indica que la masa molecular de la mayoría del producto marcado con ¹⁴C y sintetizado *in vitro* es ≥ 8 x 10⁴ Da ya que el material es eluido en los volúmenes muertos; dicho valor corresponde a una molécula de HA compuesta de al menos 400 monómeros. Este producto es sensible a una digestión con hialuronidasa de *Streptomyces* pero resistente al tratamiento con proteasas.

55 Se variaron los parámetros del ensayo de HAS para maximizar la incorporación de UDP-azúcares al polisacárido por las membranas de *P. multocida*. La spHAS estreptocócica requiere Mg²⁺ y, por lo tanto, se incluyó este ion metálico en los ensayos iniciales de membranas de *P. multocida*. La HAS de *P. multocida* (pmHAS) era relativamente activa en un pH de 6,5 a 8,6 en tampones de tipo Tris, con un valor óptimo en un pH de 7. La actividad HAS era lineal con respecto al tiempo de incubación en un pH neutro durante al menos 1 hora. La pmHAS era aparentemente menos
60 activa en mayores fuerzas iónicas ya que la adición de NaCl 100 mM a la mezcla de reacción que contenía Tris 50 mM, pH de 7, y MgCl₂ 20 mM reducía la incorporación de azúcar en ~ 50%.

Se evaluó la especificidad de la pmHAS para iones metálicos en un pH de 7. Bajo condiciones exentas de metales en presencia de EDTA, no fue detectable incorporación alguna del precursor radiomarcado al polisacárido (< 0,5% de la señal máxima). Para los metales ensayados (Mg, Mn, Co, Cu y Ni), Mn²⁺ proporcionó las mayores velocidades de
65 incorporación a las menores concentraciones iónicas. Mg²⁺ proporcionó aproximadamente el 50% de la estimulación del Mn²⁺ pero en concentraciones 10 veces mayores. Co²⁺ y Ni²⁺ en una concentración 10 mM mantenían menores niveles de actividad (20% y 9%, respectivamente, de los ensayos con Mn²⁺ 1 mM), pero las membranas a las que se su-

ministraba Cu^{2+} 10 mM eran inactivas. En realidad, el mezclamiento de Cu^{2+} 10 mM y Mg^{2+} 20 mM con la preparación de membrana daba lugar a casi ninguna incorporación de marcador al polisacárido ($< 0,8\%$ del valor de Mg solo).

Se llevó a cabo la caracterización inicial de la pmHAS en presencia de Mg^{2+} . Se evaluó la afinidad ligante de la enzima por sus precursores de azúcar-nucleótido midiendo el valor aparente de K_M . Se determinó la incorporación de $[^{14}\text{C}]\text{GlcUA}$ o $[^3\text{H}]\text{GlcNAc}$ al polisacárido con varias concentraciones de UDP-GlcNAc y UDP-GlcUA, respectivamente. En tampones que contenían Mg^{2+} , se determinaron unos valores aparentes de K_M de $\sim 20 \mu\text{M}$ para UDP-GlcUA y $\sim 75 \mu\text{M}$ para UDP-GlcNAc utilizando representaciones de Hanes-Woolf ($[S]/v$ frente a $[S]$) de los datos de titulación. Los valores de $V_{\max}$ eran iguales para ambos azúcares porque las pendientes, correspondientes a $1/V_{\max}$, de las representaciones de Hanes-Woolf eran equivalentes. En comparación con los resultados de los ensayos con Mg^{2+} , el valor de K_M para UDP-GlcNAc aumentaba en aproximadamente 25-50% hasta $\sim 105 \mu\text{M}$ y la $V_{\max}$ aumentaba en un factor de 2-3 en presencia de Mn^{2+} .

Las enzimas HA sintasa de *P. multocida*, *S. equisimilis* y *P. pyogenes* utilizan UDP-azúcares pero poseen valores cinéticos óptimos algo diferentes con respecto a la dependencia del pH y los iones metálicos y a los valores de K_M . Las enzimas son muy activas en un pH de 7; sin embargo, la pmHAS presenta supuestamente más actividad en un pH ligeramente ácido y es relativamente inactiva en un pH superior a 7,4. La pmHAS utiliza Mn^{2+} más eficazmente que Mg^{2+} bajo las condiciones de ensayo *in vitro*, pero se desconoce la identidad del cofactor metálico fisiológico en la célula bacteriana. En comparación, en estudios previos con la enzima estreptocócica, Mg^{2+} era mucho mejor que Mn^{2+} pero el efecto no obstante más pequeño de Mn^{2+} era máximo en concentraciones ~ 10 veces menores que la concentración óptima de Mg^{2+} . La pmHAS se une aparentemente a los UDP-azúcares más fuertemente que la spHAS. Los valores medidos de K_M para la pmHAS en membranas crudas son para cada sustrato aproximadamente 2-3 veces menores que los obtenidos con la HAS hallada en membranas estreptocócicas: 50 ó 39 μM para UDP-GlcUA y 500 ó 150 μM para UDP-GlcNAc, respectivamente.

Mediante análisis cinéticos, la $V_{\max}$ de la pmHAS era 2-3 veces mayor en presencia de Mn^{2+} que de Mg^{2+} , pero el valor de K_M para UDP-GlcNAc aumentaba ligeramente en los ensayos con el primer ion. Esta observación de una aparente afinidad reducida sugiere que la velocidad de polimerización aumentada no se debía a una mejor unión del complejo de ion Mn^{2+} /azúcar-nucleótido a el(los) sitio(s) activo(s) de la enzima. Por lo tanto, es posible que Mn^{2+} potencie alguna otra etapa de reacción, altere otro sitio/estructura de la enzima, o modifique el ambiente de la membrana fosfolípida. En las ID. SEC. números 9 y 10 se muestran, respectivamente, la secuencia génica y la secuencia proteica de pmHAS.

El virus PBCV-1 de *Chlorella* codifica una glicosiltransferasa funcional que puede sintetizar hialuronano. Este hallazgo es contrario a la observación general de que los virus: (a) utilizan glicosiltransferasas de la célula huésped para crear nuevas estructuras de hidratos de carbono, o (b) acumulan productos de glicoconjugación de la célula huésped durante la maduración de los viriones. Además, se ha considerado generalmente que el HA está restringido a animales y a algunos de sus patógenos bacterianos virulentos. Aunque se han caracterizado muchos hidratos de carbono vegetales, no se ha detectado previamente HA ni un compuesto relacionado análogo en células de plantas ni protistas.

Las enzimas HAS de vertebrados, bacterias y virus tienen varias regiones con similitud secuencial. Mientras se secuenciaba el genoma de DNA de doble cadena del virus PBCV-1 (virus de *Chlorella-Paramecium bursaria*), se descubrió un marco de lectura abierto (ORF; del inglés, *open reading frame*), A98R (número de acceso: 442580), que codifica una proteína de 567 restos con una identidad de aminoácidos de 28 a 33% con respecto a las diversas HASs. Esta proteína es denominada cvHAS (HA sintasa del virus de *Chlorella*). En las ID. SEC. números 7 y 8 se muestran, respectivamente, la secuencia génica que codifica PBCV-1 y la secuencia proteica que aquella codifica.

PBCV-1 es el prototipo de una familia (Phycodnaviridae) de grandes virus poliédricos (diámetro de 175-190 nm), formadores de placas de lisis, que se replican en ciertas algas verdes unicelulares eucarióticas de tipo *Chlorella*. Los viriones de PBCV-1 contienen al menos 50 proteínas diferentes y un componente lipídico situado dentro de la cápsida glicoproteínica externa. El genoma de PBCV-1 es una molécula lineal no permutada de dsDNA de 330 kb con extremos ahorquillados covalentemente cerrados.

Basándose en su deducida secuencia de aminoácidos, el producto del gen A98R debería ser una proteína integral de membrana. Para ensayar esta hipótesis, se produjo A98R recombinante en *Escherichia coli* y se examinó la fracción de membrana en cuanto a actividad HAS. La fracción de membrana derivada de las células que contenían el gen A98R en un plásmido, pCVHAS (actividad específica media de 2,5 picomoles de transferencia de GlcUA/ μg de proteína/minuto), pero no las muestras de las células testigo ($< 0,001$ picomoles de transferencia de GlcUA/ μg de proteína/minuto), incorporaba UDP-GlcUA y UDP-GlcNAc al polisacárido. No se detectó actividad en la fracción soluble de las células transformadas con pCVHAS. Se requerían simultáneamente UDP-GlcUA y UDP-GlcNAc para la polimerización. La actividad era óptima en tampón Hepes en un pH de 7,2 en presencia de MnCl_2 10 mM, mientras que no se detectaba actividad alguna si se excluía el ión metálico. Mg^{2+} y Co^{2+} eran $\sim 20\%$ tan eficaces como Mn^{2+} en concentraciones similares. La pmHAS tiene una necesidad similar de metales, pero otras HASs prefieren Mg^{2+} .

La enzima A98R recombinante sintetizaba un polisacárido con un peso molecular medio de $3\text{-}6 \times 10^6$ Da, que es más pequeño que el del HA sintetizado por spHAS o DG42 xHAS recombinante *in vitro* ($\sim 10^7$ Da y $\sim 5\text{-}8 \times 10^6$ Da, respectivamente). El polisacárido era completamente degradado por HA liasa de *Streptomyces hyaluroniticus*, una enzima que despolimeriza HA pero no glicosaminoglicanos estructuralmente relacionados tales como heparina y condroitina.

Se examinaron células de *Chlorella* infectadas con PBCV-1 en cuanto a la expresión del gen *A98R*. Apareció un transcrito de *A98R* de ~ 1700 nucleótidos ~ 15 minutos después de la infección, transcrito que desapareció 60 minutos después de la infección, lo que indica que *A98R* es un gen precoz. En consecuencia, fracciones de membrana de células de *Chlorella* no infectadas e infectadas con PBCV-1 fueron examinadas en cuanto a actividad HAS a los 50 y 90 minutos después de la infección. Las células infectadas, pero no las no infectadas, tenían actividad. Como la enzima *A98R* recombinante bacterianamente derivada, la incorporación del radiomarcador UDP-[¹⁴C]GlcUA al polisacárido dependía tanto de Mn²⁺ como de UDP-GlcNAc. Este producto radiomarcado era también degradado por la HA liasa. Los viriones de PBCV-1 alterados no tenían actividad HAS.

Se analizaron células de *Chlorella* infectadas con PBCV-1 en cuanto al polisacárido HA usando una proteína ligante de HA muy específica y marcada con ¹²⁵I. A los 50 y 90 minutos después de la infección, los extractos de las células contenían cantidades sustanciales de HA, pero no los extractos de las algas no infectadas ni los viriones de PBCV-1 alterados. La proteína ligante de HA marcada también interactuaba con las células infectadas intactas a los 50 y 90 minutos después de la infección, pero no con las células sanas. Por lo tanto, una considerable porción del polisacárido HA recién sintetizado estaba inmovilizada en la superficie celular externa de las algas infectadas. El HA extracelular no desempeña ningún papel obvio en la interacción entre el virus y su alga huésped porque no se alteraron ni el tamaño de las placas de lisis ni el número de placas de lisis al incluir hialuronidasa testicular (465 unidades/ml) o polisacárido HA libre (100 µg/ml) en el agar superior del ensayo de placas de lisis para PBCV-1.

El genoma de PBCV-1 también tiene genes adicionales que codifican una UDP-glucosa deshidrogenasa (UDP-Glc DH) y una glutamina:fructosa-6-fosfato aminotransferasa (GFAT). UDP-Glc DH convierte UDP-Glc en UDP-GlcUA, un precursor necesario para la biosíntesis de HA. GFAT convierte fructosa-6-fosfato en glucosamina-6-fosfato, un producto intermedio en la ruta metabólica de UDP-GlcNAc. Estos dos genes de PBCV-1, como el HAS *A98R*, se expresan pronto en la infección y codifican proteínas enzimáticamente activas. La presencia de múltiples enzimas en la ruta de la biosíntesis de HA indica que la producción de HA debe desempeñar una función importante en el ciclo vital de los virus de *Chlorella*.

Las HA sintetasas de *Streptococcus*, vertebrados y PBCV-1 poseen muchos motivos con un patrón de al menos 2 a 4 restos idénticos que se presentan en el mismo orden relativo. Estos motivos conservados reflejan probablemente dominios cruciales para la biosíntesis de HA, como se muestra en la Figura 2. La Figura 2 es un alineamiento de las secuencias proteínicas de suHAS y seHAS del Grupo C de *S. equisimilis* y *S. uberis*, respectivamente; spHAS del Grupo A de *S. pyogenes*; las isozimas mHAS1, mHAS2 y mHAS3 de ratón; las isozimas humanas hHAS1, hHAS2 y hHAS3; las isozimas xHAS1 y xHAS2 de rana; la HAS original del virus PBCV-1, cvHAS, así como las más recientes HASs virales vNC, vMA, vAL y vCA; la rnHAS2 de rata; la ggHAS2 de gallina; la btHAS2 bovina; y las isozimas ocHAS2 y ocHAS3 de conejo. La alineación de la Figura 2 fue realizada utilizando el programa Mutalin, versión 5.4.1, para alineaciones múltiples [propiedad intelectual de I.N.R.A., Francia, 1989, 1991, 1994, 1996; "Multiple sequence alignment with hierarchical clustering", Corpet, *Nucl. Acids Res.* 16: 10.881 (1988)]. Se proporciona una línea de consenso en la línea inferior de la alineación. Las letras mayúsculas de la línea de consenso representan restos idénticos en todas las proteínas HAS enumeradas, mientras que las letras minúsculas representan restos de consenso que no son idénticos en todos los casos. La línea de consenso también contiene los símbolos siguientes: ! es cualquiera de Y y V; \$ es cualquiera de L y M; % es cualquiera de F e Y; # es cualquiera de N, D, Q, E, B y Z. La tabla para comparación de símbolos era blosum62, el peso del hueco era 12 y el peso de la longitud del hueco era 2.

Conforme se secuencian más glicosiltransferasas, se están descubriendo también regiones de similitud entre HASs y otras enzimas que sintetizan polisacáridos β-enlazados a partir de precursores de UDP-azúcar. Los ejemplos incluyen celulosa sintasa bacteriana, quitina sintetasas fúngicas y bacterianas, y las diversas HASs. La significación de estos similares motivos estructurales resultará más evidente conforme se acumulen las estructuras tridimensionales de las glicosiltransferasas.

La Figura 3 representa las posibles relaciones evolutivas entre las hialuronano sintetasas conocidas. El árbol filogenético de la Figura 3 fue generado mediante el algoritmo de Higgins-Sharp usando el programa DNAsis para alineaciones múltiples. Los porcentajes de correspondencia calculados se indican en cada rama del dendrograma.

Los segmentos de DNA usados en el presente invento abarcan proteínas y péptidos HAS equivalentes biológicamente funcionales. Dichas secuencias pueden surgir como consecuencia de la redundancia codónica y la equivalencia funcional de las que se sabe que se producen naturalmente en secuencias de ácido nucleico y las proteínas así codificadas. Alternativamente, se pueden crear proteínas o péptidos funcionalmente equivalentes por medio de la aplicación de la tecnología de DNA recombinante, con la que se pueden crear cambios en la estructura proteica basándose en consideraciones de las propiedades de los aminoácidos que se intercambian. Los cambios diseñados por el hombre pueden ser introducidos mediante la aplicación de técnicas de mutagénesis dirigida al sitio para, por ejemplo, introducir mejoras en la actividad enzimática o la antigenicidad de la proteína HAS o para ensayar mutantes de HAS con objeto de examinar la actividad HA sintasa a nivel molecular.

Además, cambios específicos en la secuencia de codificación de HAS pueden dar lugar a la producción de un HA que tenga una distribución de tamaños o una configuración estructural modificadas. Quien tiene una experiencia normal en la técnica apreciará que la secuencia de codificación de HAS puede ser manipulada de modo que se produzca un hialuronato sintasa alterada que, a su vez, sea capaz de producir ácidos hialurónicos que tengan tamaños de polímero y/o capacidades funcionales diferentes. Por ejemplo, la secuencia de codificación de HAS puede ser alterada de tal

modo que la hialuronato sintasa tenga una especificidad alterada por el sustrato de azúcar para que la hialuronato sintasa cree un nuevo polímero de tipo ácido hialurónico que lleve incorporada una estructura diferente, tal como un azúcar o derivado de azúcar previamente no incorporado. Este azúcar recién incorporado podría dar lugar a un ácido hialurónico modificado que tuviera diferentes propiedades funcionales, un ácido hialurónico que tuviera un tamaño de polímero/peso molecular más pequeño o más grande, o ambas cosas. Como apreciará quien tiene una experiencia normal en la técnica, dadas las secuencias de codificación de HAS, se pueden hacer cambios y/o sustituciones en la secuencia de codificación de HAS para que se puedan alcanzar estas deseadas modificaciones de propiedades y/o tamaños. En la Tabla II se enumeran la especificidad de azúcar-nucleótido y la necesidad de ion magnesio de la seHAS recombinante.

TABLA II

Especificidad de azúcar-nucleótido y necesidad de ion magnesio de la seHAS recombinante

Segundo azúcar-nucleótido presente (μM)	Síntesis de HA*	
	UDP-[¹⁴ C]GlcUA dpm (%)	UDP-[³ H]GlcNAc dpm (%)
Ninguno	90 (2,1%)	8 (1,2%)
UDP-GlcNAc (300)	4134 (100%)	----
UDP-GlcUA (120)	----	635 (100%)
UDP-Glc (160)	81 (1,9%)	10 (1,5%)
UDP-GalNAc (280)	74 (1,7%)	19 (2,9%)
UDP-GalA (150)	58 (1,4%)	19 (2,9%)
UDP-GlcNAc + EDTA	31 (0,7%)	----
UDP-GlcUA + EDTA	----	22 (3,4%)

* Se incubaron membranas (324 ng de proteína) a 37 °C durante 1 hora con UDP-[¹⁴C]GlcUA 120 μM (2,8x10⁴ dpm) o UDP-[³H]GlcNAc 300 μM (2x10⁴ dpm). Se usó el azúcar-nucleótido radiomarcado en presencia del segundo azúcar-nucleótido no marcado indicado. Se determinó la actividad HA sintasa del modo descrito en la solicitud.

Como aquí se usa, la expresión “estructura modificada” significa un polímero de ácido hialurónico que contiene un azúcar o derivado no hallado normalmente en el polisacárido de HA presente en la naturaleza. La expresión “distribución de tamaños modificada” se refiere a la síntesis de moléculas de ácido hialurónico con una distribución de tamaños no hallada normalmente en la enzima nativa; el tamaño genéticamente modificado podría ser mucho más pequeño o mucho más grande que el normal.

Los diversos productos de ácido hialurónico de diferente tamaño tienen aplicación en varias áreas, tal como en la distribución de fármacos. Las aplicaciones en la angiogénesis y la cicatrización de heridas son potencialmente grandes si se pueden preparar polímeros de ácido hialurónico de aproximadamente 4 a aproximadamente 20 monosacáridos en grandes cantidades. Otra aplicación particular para oligosacáridos de ácido hialurónico pequeños es en la estabilización de proteínas humanas recombinantes utilizadas con fines médicos. Un problema importante con dichas proteínas es su aclaramiento de la sangre y una corta semivida biológica. Una presente solución a este problema es acoplar un revestimiento de moléculas pequeñas que evite que la proteína sea aclarada demasiado rápidamente de la circulación. El ácido hialurónico de muy pequeño peso molecular es bien adecuado para este papel y sería no inmunogénico y biocompatible. Se puede utilizar ácido hialurónico de mayor peso molecular fijado a un fármaco o una proteína para dirigirlo al sistema celular reticuloendotelial, que tiene receptores endocíticos para el ácido hialurónico.

Dada esta descripción, quien tiene una experiencia normal en la técnica apreciará que hay varias formas en que se podría regular la distribución de tamaños del polímero de ácido hialurónico generado por la hialuronato sintasa, para obtener diferentes tamaños. En primer lugar, se puede alterar el control cinético del tamaño del producto disminuyendo la temperatura, disminuyendo el tiempo de la acción enzimática y disminuyendo la concentración de uno o ambos

sustratos de azúcar-nucleótido. La disminución de cualquiera de, o todas, estas variables proporcionará cantidades menores y tamaños más pequeños del producto de ácido hialurónico. Las desventajas de estos planteamientos son que también disminuirá la producción de producto y que puede ser difícil conseguir una reproducibilidad de día a día o de lote a lote.

En segundo lugar, la distribución de tamaños del polímero de HA puede ser regulada alterando la capacidad intrínseca de la enzima para sintetizar un producto de ácido hialurónico grande. Se pueden producir genéticamente cambios en la proteína mediante la tecnología de DNA recombinante, incluyendo la sustitución, delección y adición de aminoácidos específicos (o incluso la introducción de grupos prostéticos a través de un procesamiento metabólico). Dichos cambios que dan lugar a una enzima intrínsecamente más lenta permiten entonces un control más reproducible del tamaño del ácido hialurónico por medios cinéticos. La distribución de tamaños final del ácido hialurónico viene determinada por ciertas características de la enzima que se basan en aminoácidos particulares de la secuencia. Entre el 20% de los restos absolutamente conservados entre las enzimas estreptocócicas y las hialuronato sintasas eucarióticas, hay un conjunto de aminoácidos en posiciones únicas que controlan, o influyen en gran medida en, el tamaño del polímero de ácido hialurónico que puede hacer la enzima. Cambios específicos en cualquiera de estos restos pueden producir una HAS modificada que produzca un producto de HA que tenga una distribución de tamaños modificada. Los cambios genéticamente producidos en seHAS, spHAS, suHAS, pmHAS o cvHAS que disminuyan el tamaño intrínseco del ácido hialurónico que la enzima puede hacer antes de que el ácido hialurónico sea liberado proporcionarán potentes medios para producir un producto de ácido hialurónico con un tamaño más pequeño o potencialmente más grande que el de la enzima nativa.

Finalmente, se puede degradar el ácido hialurónico de mayor peso molecular con hialuronidasas específicas para obtener ácido hialurónico de menor peso molecular. Sin embargo, con esta práctica es muy difícil conseguir reproducibilidad y se debe repurificar meticulosamente el ácido hialurónico para separar la hialuronidasa y los productos de digestión indeseados.

Como se muestra en la Figura 4, se puede modificar genéticamente una hialuronano sintasa para que produzca polímeros de ácido hialurónico de diferente tamaño, en particular más pequeños que los de la enzima normal de tipo silvestre. En la figura se muestra la distribución de tamaños de HA (en millones de dáltones, una medida de peso molecular) para una serie de enzimas spHAS, cada una de las cuales fue genéticamente modificada por mutagénesis dirigida al sitio para que tuviera un único cambio de aminoácido con respecto a la enzima nativa. Cada una tiene un diferente resto de cisteína sustituido por alanina. El haz de cinco curvas con símbolos claros representa las siguientes proteínas spHAS: tipo silvestre, C124A, C261A, C366A y C402A. Los círculos oscuros representan la proteína C225A mal expresada, que es sólo parcialmente activa.

Los triángulos oscuros representan la proteína spHAS C280A, de la que se ha hallado que sintetiza un intervalo mucho más pequeño de polímeros de HA que la enzima normal o las otras variantes mostradas. En la práctica, esta reducción muestra que es factible modificar genéticamente la enzima hialuronato sintasa para que sintetice un intervalo deseado de tamaños del producto HA. Cualquiera de los genes de HAS que codifican hialuronato sintasa aquí descritos puede ser también manipulado por mutagénesis dirigida al sitio para que produzca una enzima que sintetice un intervalo deseado de tamaños del producto HA.

El ácido hialurónico estructuralmente modificado no es conceptualmente diferente que una alteración de la distribución de tamaños del producto de ácido hialurónico al cambiar aminoácidos particulares en la deseada HAS o la spHAS. Se espera que sean particularmente útiles los derivados de UDP-GlcNAc en que el grupo N-acetilo falta (UDP-GlcN) o ha sido sustituido por otro grupo químicamente útil. La potente especificidad de sustrato debe depender de un subconjunto particular de aminoácidos de entre el 20% que están conservados. Cambios específicos en uno o más de estos restos crea una sintasa funcional que interacciona con uno o más de los sustratos menos específicamente que la enzima nativa. Esta enzima alterada podría utilizar luego azúcar-nucleótidos naturales o especiales alternos para incorporar derivados de azúcar destinados a permitir que se empleen químicas diferentes con los fines siguientes: (i) copular covalentemente fármacos, proteínas o toxinas específicas con el ácido hialurónico estructuralmente modificado, para una distribución general o dirigida de fármacos, procedimientos radiológicos, etc., (ii) entrecruzar covalentemente el propio ácido hialurónico con otros soportes para obtener un gel, u otro biomaterial tridimensional con propiedades físicas más acusadas, y (iii) enlazar covalentemente el ácido hialurónico a una superficie para crear una película o monocapa biocompatible.

También se pueden modificar genéticamente bacterias para que produzcan ácido hialurónico. Por ejemplo, hemos creado cepas de *B. subtilis* que contienen un gen de HAS así como el gen para uno de los precursores de azúcar-nucleótido. Escogimos esta bacteria ya que se utiliza frecuentemente en la industria biotecnológica para la producción de productos para uso humano. Estas bacterias fueron proyectadas como prototipos de primera generación para la generación de una bacteria capaz de producir ácido hialurónico en cantidades mayores que las actualmente disponibles al utilizar una cepa natural de tipo silvestre.

Por ejemplo, se construyeron tres cepas de *Bacillus subtilis* para que contuvieran uno o los dos genes de *Streptococcus pyogenes* para hialuronano sintasa (*spHAS*) y UDP-glucosa deshidrogenasa (*hasB*), los resultados de las cuales se muestran en la Tabla III. Basándose en un sensible ensayo radiométrico comercial para detectar y cuantificar HA, se determinó que la cepa con ambos genes (cepa nº 3) produce HA y los secreta al medio. La cepa parental o cepa con sólo el gen de la deshidrogenasa (cepa nº 1) no produce HA. La cepa nº 2, que sólo contiene el gen *spHAS*, produce

ES 2 346 196 T3

HA pero sólo aproximadamente el 10% de lo que produce la cepa nº 3. Una electroforesis en gel de agarosa mostró que el HA secretado al medio por la cepa nº 3 tiene un elevadísimo peso molecular.

Los datos de la Tabla III demuestran que se puede modificar genéticamente *B. subtilis* 168 para que produzca y secrete HA mediante la introducción del gen de spHAS y el gen *hasB* por técnicas de DNA recombinante. Aunque esta cepa modificada produce HA aun sin la inclusión del segundo gen, el nivel de HA producido con él es muy elevado. *B. subtilis* 168 contiene dos genes (*tauD* y *gtaB*) que aumentan los niveles de ambos azúcar-nucleótidos necesarios para la síntesis de HA. En la Tabla IV se demuestra que *B. subtilis* 168 también produce HA, aún en ausencia del gen *hasB*, cuando es genéticamente modificado para que contenga y exprese (en el plásmido pSM143, ATCC) seHAS así como variantes específicas de seHAS genéticamente modificadas para que produzcan HA con un tamaño diferente al correspondiente al tipo silvestre. En particular, las variantes seHAS(C226A) y seHAS(C281A) sustentaban la síntesis de HA en células de *B. subtilis* 168 vivas. El nivel de síntesis de HA en estos últimos casos fue inferior al observado con células que expresan spHAS y el gen *hasB*, debido al menor nivel endógeno de los dos precursores necesarios para la síntesis de HA.

Experimentos *in vitro* usando membranas aisladas de células de *B. subtilis* 168 transformadas con plásmidos que contienen *hasB* y la variante seHAS(C226A) o seHAS(C281A) demostraron que las distribuciones de tamaños de HA producidas por estas enzimas HAS modificadas eran más grandes y más pequeñas, respectivamente, que la producida por la seHAS de tipo silvestre. El tamaño aproximado del HA producido en *B. subtilis* a partir de seHAS de tipo silvestre es 1,5 MDa.

La producción de HA recombinante a partir del gen de spHAS ha sido también demostrada en *Bacillus licheniformis* y *Enterococcus faecalis*. La Figura 6 es una imagen digital de un gel de agarosa teñido para HA. Se puede ver producción de HA en cepas de *B. subtilis*, *B. licheniformis* y *E. faecalis* que tienen incorporado un plásmido que codifica spHAS (pPD41Δ5). Como testigo negativo, se introdujo en cada cepa un plásmido que contenía un gen *hasA* no funcional (pPD41ΔEcoRV), y ninguna de estas cepas fue capaz de producir HA.

TABLA III

Producción de HA en B. subtilis 168 que contiene los genes de spHAS y hasB

Nº de cepa	Células	Medio (*)	Cepa con genes	Densidad celular (A ₆₀₀)
(µg de HA por ml de cultivo)				
1	0	0	hasB	4,8
2	4	35	spHAS	3,9
3	≥ 10	≥ 250	spHAS + hasB	3,2

(*) La mayor parte del HA está en el medio, pero algo estaba celularmente asociado. Se determinó el HA usando el kit HA Test 50 de Pharmacia.

TABLA IV

Hialuronano (HA) producido en B. subtilis 168

Cepas recombinantes	HA (g/l) por método ELISA
spHAS de tipo silvestre en pPD41Δ5	5,16
seHAS de tipo silvestre en pSM143	0,37
seHAS(C226A) en pSM143	0,25
seHAS(C281A) en pSM143	0,32
pSM143 en <i>B. subtilis</i> 168 (vector solo)	0,05

Se desarrollaron cultivos de 100 ml durante la noche. Se analizó el medio mediante un kit para ensayo de HA de Corgenix (ácido hialurónico "Chugai").

En estos experimentos se utilizaron los promotores estreptocócicos normalmente hallados con estos genes para conducir la expresión proteica. Se espera que la construcción de cepas con el marco de lectura de spHAS o seHAS bajo el control de un promotor compatible con *Bacillus* o bacterias Gram-positivas produzca resultados aún más superiores. El vector utilizado es un vector lanzadera de bacteria Gram-positiva/*E. coli* que tiene un número medio de copias en *B. subtilis* y un gen para resistencia a eritromicina (que permite resistencia a 8 µg/ml en *B. subtilis* o 175 µg/ml en *E. coli*). La cepa huésped de *B. subtilis* utilizada es 1A1 de BGSC, que tiene necesidad de triptófano pero, por lo demás, es de tipo silvestre y puede esporular. El crecimiento celular y la producción de HA fueron en Medio Mínimo de Spizizen más triptófano, glucosa, oligoelementos y eritromicina (8 µg/ml). El crecimiento fue a 32°C con agitación enérgica hasta que se agotó el medio (~ 36 horas).

En las Tablas III y IV se demuestra que estas células genéticamente modificadas, que normalmente no producirían ácido hialurónico, se vuelven competentes para esto cuando son transformadas con el gen de spHAS o seHAS. Cualquiera de los genes de HAS aquí descritos también podría ser introducido en una bacteria no productora de ácido hialurónico para crear una cepa bacteriana genéticamente modificada capaz de producir ácido hialurónico.

Volviendo a la expresión de uno de los genes de HAS aquí descritos, sea de DNA genómico o un cDNA, uno puede proceder a preparar un sistema de expresión para la preparación recombinante de la proteína HAS. La modificación genética de un(unos) segmento(s) de DNA para expresión en un sistema procariótico o eucariótico puede ser llevada a cabo mediante técnicas generalmente conocidas por quienes tienen experiencia en la expresión recombinante.

Se puede expresar exitosamente HAS en sistemas de expresión eucarióticos; sin embargo, los inventores afirman que se pueden usar sistemas de expresión bacterianos para la preparación de HAS para todos los fines. Se cree que la expresión bacteriana tendrá finalmente ventajas con respecto a la expresión eucariótica en términos de facilidad de uso, coste de producción, y cantidad de material obtenido mediante ella.

En la Tabla V y la Figura 7 se muestra la purificación de hialuronano sintasa estreptocócica (seHAS). Se analizaron las fracciones de las diversas etapas del esquema de purificación mediante SDS-PAGE en un gel al 12,5%, el cual fue luego teñido con Azul Brillante de Coomassie R250. Carriles: marcadores de pesos moleculares; 1, membranas de *E. coli* completas que contienen la seHAS-H6 recombinante; 2, fracción insoluble después de la solubilización de las membranas con detergente; 3, fracción solubilizada con detergente; 4, flujo a través de la resina de Ni-NTA para cromatografía; 5-9, cinco lavados sucesivos de la columna (cada uno de dos volúmenes de columna); 10, la HA sintasa pura eluida que es una sola banda. La purificación de spHAS fue idéntica a la mostrada para seHAS (Tlapak-Simmons, 1999).

TABLA V

Etapas	Proteína total (µg)	Actividad específica (milimoles/µg/h)	Actividad total (nanomoles de UDP-GlcUA)	Rendimiento (%)	Purificación (nº de veces)
Membranas	3690	1,0	3649	100	1,0
Extracto	2128	2,2	4725	129	2,2
Columna de afinidad	39	13	500	14	13,1

Se propone que la transformación de células huésped con segmentos de DNA que codifican HAS proporcionará un medio conveniente para obtener una proteína HAS. También se propone que cDNA, secuencias genómicas y combinaciones de las mismas son adecuadas para la expresión eucariótica ya que, por supuesto, la célula huésped procesará los transcritos genómicos para producir mRNA funcional para su traducción a proteína.

Otro aspecto del presente invento es un método para preparar una composición proteica, que comprende cultivar una célula huésped recombinante que comprende un vector que codifica una proteína que incluye una secuencia de aminoácidos de acuerdo con la ID. SEC. nº 12. La célula huésped será cultivada bajo unas condiciones que permitan la expresión de ácido nucleico y la producción de proteína seguidas de la recuperación de la proteína así producida. La producción de HAS y finalmente HA, incluyendo la célula huésped, y las condiciones que permiten la expresión de ácido nucleico, la producción de proteína y la recuperación, serán conocidas por los expertos en la técnica a la luz de la presente descripción y los métodos aquí descritos.

Los huéspedes para la expresión de ácido hialurónico son comúnmente procariontes, tales como *S. equisimilis*, y otros miembros adecuados de las especies de *Streptococcus*. Sin embargo, se sabe también que el HA puede ser sintetizado por células huésped heterólogas que expresan HA sintasa recombinante, tales como miembros de las especies de los géneros *Bacillus*, *Enterococcus* e incluso *Escherichia*.

Los huéspedes para uso en los métodos para expresión de HA de acuerdo con el presente invento son especies de *Bacillus* porque dichas células son secretoras industriales eficaces, y diversas especies han sido designadas organismos GRAS. Los ejemplos de células de *Bacillus* que se pueden utilizar en los métodos del presente invento incluyen, pero no se limitan a, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus metaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis* y *Bacillus thuringiensis*. Un huésped muy preferido es *Bacillus subtilis*.

Se cree similarmente que se puede utilizar casi cualquier sistema de expresión eucariótico para la expresión de HAS; por ejemplo, se podrían emplear sistemas basados en baculovirus, basados en glutamina sintasa, basados en dihidrofolato reductasa, basados en SV-40, basados en adenovirus, basados en citomegalovirus, basados en levadura, y similares. Para la expresión de esta manera, se situarían las secuencias de codificación adyacentemente a, y bajo el control de, el promotor. Se entiende en la técnica que, al poner una secuencia de codificación bajo el control de dicho promotor, se sitúa el extremo 5' del sitio de inicio de la transcripción del marco de lectura transcripcional de la proteína entre aproximadamente 1 y aproximadamente 50 nucleótidos "cadena abajo" de (es decir, en 3' de) el promotor elegido. Además, sistemas de vectores de expresión de la levadura *Saccharomyces cerevisiae*, tal como pYES2, producirán también HAS bajo el control del promotor de GAL, como se muestra en la Figura 8. En la Figura 8 se muestra que se produjo sPHAS o la enzima x1HAS en una levadura recombinante utilizando el plásmido pYES2. Cuando se suministra con UDP-GlcUA y UDP-GlcNAc, cualquier enzima produce HA de alto peso molecular, como se observa en estos perfiles de cromatografía por filtración en gel (el pico de HA va de aproximadamente 13 ml a aproximadamente 25 ml).

En la Figura 8 se muestra un análisis, por filtración en gel, de ácido hialurónico sintetizado por HASs recombinantes expresadas en membranas de levadura. Un fragmento de DNA que codifica (a) el marco de lectura abierto de 419 restos de aminoácido que corresponden a sPHAS (con el codón de Val original cambiado por el de Met) o (b) la proteína x1HAS fue subclonado mediante métodos estándares en el vector de expresión pYES2 de levadura (de Invitrogen) para producir pYES/HA. Se prepararon membranas de células con esta construcción mediante agitación con glóbulos de vidrio. Las muestras procedentes de construcciones pYES/HA contenían una sustancial actividad HA sintasa, y se detectó la proteína HAS de "42 kDa" mediante un análisis Western usando anticuerpos específicos; las membranas de células con el vector solo no poseían actividad ni la banda inmunorreactiva (no mostrado). Se incubaron primero las membranas (315 µg de proteína) con UDP-[¹⁴C]GlcUA (37 kBq de ¹⁴C) exento de vehículo y UDP-GlcNAc no marcado 900 µM en Tris 50 mM, pH de 7, MgCl₂ 20 mM, DTT 1 mM y NaCl 0,05 M (volumen de reacción de 450 µl) a 30 grados Celsius durante 1,5 minutos. Después de este periodo de pulsos de radiomarcador, se añadió UDP-GlcUA no radiomarcado hasta concentraciones finales de 900 µM. Se tomaron muestras (100 µl) después del pulso a 1,5 minutos (círculo oscuro), y 15 (cuadrado negro) y 45 (triángulo negro) minutos después de la "caza". Se terminaron las reacciones mediante la adición de SDS a 2% y calentamiento a 95 grados Celsius durante 1 minuto. Las muestras fueron clarificadas por centrifugación (10.000 x g, 5 minutos) antes de la inyección de la mitad de la muestra a una columna de Sephacryl S-500HR para filtración en gel (Pharmacia; 1 x 50 cm), equilibrada en NaCl 0,2 M, Tris 5 mM, pH de 8.

Se sometió la columna a elución a 0,5 ml/min y se cuantificó la radiactividad de las fracciones (1 ml) mediante la cuenta de centelleo en estado líquido después de añadir el cóctel BioSafeII (4,5 ml, Research Products Intl.). El volumen muerto y el volumen totalmente incluido estaban en los volúmenes de elución de 14 ml y 35,5 ml, respectivamente. El pico del azul de dextrano (tamaño medio de 2x10⁶ Da) era eluido a 25-27 ml. La HAS recombinante expresada en las células eucarióticas de levadura produce ácido hialurónico de alto peso molecular *in vitro*.

Cuando se contempla la expresión eucariótica, también se deseará típicamente incorporar, a la unidad de transcripción que incluye el DNA o gen de HAS, un sitio de poliadenilación apropiado (por ejemplo, 5'-AATAAA-3') si no hubiera ya uno contenido en el segmento clonado original. Típicamente, el sitio de adición de poliA está situado de aproximadamente 30 a 2000 nucleótidos "cadena abajo" del sitio de terminación de la proteína, en una posición previa a la terminación de la transcripción.

Se contempla que, en relación con la expresión de HAS, se pueda usar casi cualquiera de las células huésped eucarióticas comúnmente empleadas. Los ejemplos de líneas celulares para expresar cDNA de HAS incluyen líneas celulares típicamente empleadas para expresión eucariótica, tales como las líneas celulares 293, AtT-20, HepG2, VERO, HeLa, CHO, WI 38, BHK, COS-7, RIN y MDCK. Esto incluirá generalmente las operaciones de proporcionar un huésped recombinante que porte el segmento de DNA recombinante que codifica la enzima HAS y sea capaz de expresar la enzima; cultivar el huésped recombinante en un medio bajo unas condiciones que permitan la transcripción del cDNA o gen de HAS clonado y sean apropiadas para la producción del ácido hialurónico; y separar del huésped recombinante y purificar la enzima HAS o el ácido hialurónico secretado.

Generalmente, las condiciones apropiadas para la expresión del cDNA o gen de HAS clonado dependerán del promotor, el vector, y el sistema huésped que se emplee. Por ejemplo, cuando se emplea el promotor *lac*, se deseará inducir la transcripción por medio de la inclusión de un material que estimule la transcripción de *lac*, tal como isopropiltoiogalactósido. Por ejemplo, los genes de seHAS y spHAS clonados se expresan como proteínas que contienen HIS₆ en *E. coli* con el fin de purificación de la HAS, como se muestra en la Figura 5. Cuando se emplean otros promotores, se pueden necesitar diferentes materiales para inducir o, de otro modo, suprarregular la transcripción.

En la Figura 5 se representa la sobreexpresión de seHAS y spHAS recombinantes en *E. coli*. Se fraccionaron proteínas de membrana (5 µg por carril) mediante SDS-PAGE utilizando un gel al 10% (peso/volumen) bajo condiciones reductoras. El gel fue teñido con azul de Coomassie R-250, fotografiado, escaneado y cuantificado utilizando un densitómetro personal de Molecular Dynamics (modelo PDSI P60). La posición de la HA sintasa se señala con la flecha. El carril A es spHAS nativa (Grupo A); el carril C es seHAS nativa; el carril E es seHAS recombinante; el carril P es spHAS recombinante; y el carril V es el vector solo. Los patrones usados eran de bajo M_r de Bio-Rad y se muestran en kDa.

Además de obtener la expresión de la sintasa, se deseará preferiblemente proporcionar un ambiente que sea conducente a la síntesis de HA al incluir genes apropiados que codifican enzimas necesarias para la biosíntesis de precursores de azúcar-nucleótido, o usar medios de crecimiento que contengan sustratos para las enzimas que suministran precursores, tales como N-acetilglucosamina o glucosamina (GlcNac o GlcNH₂) y glucosa (Glc).

Se puede desear además incorporar el gen a un huésped que sea defectuoso en la enzima hialuronidasa, para que el producto sintetizado por la enzima no se degrade en el medio. Además, se escogería un huésped para optimizar la producción de HA. Por ejemplo, un huésped adecuado sería uno que produjera grandes cantidades de los precursores de azúcar-nucleótido para sostener la enzima HAS y permitir que produzca grandes cantidades de HA. Dicho huésped puede hallarse naturalmente o puede ser preparado mediante una diversidad de técnicas que incluyen la mutagénesis y la tecnología de DNA recombinante. También se podrían aislar los genes para las enzimas que sintetizan azúcar-nucleótido, particularmente la UDP-Glc deshidrogenasa necesaria para producir UDP-GlcUA, e incorporarlos a un vector junto con el cDNA o gen de HAS. Una realización preferida del presente invento es un huésped que contiene estos cDNAs o genes recombinantes auxiliares, lo que permite una producción aumentada de los azúcar-nucleótidos y, por lo tanto, de HA.

No se cree que el medio empleado para el cultivo de la célula huésped sea particularmente crucial. Para detalles útiles, se puede querer referirse a la descripción de la Patente de EE.UU. n° 4.517.295, 4.801.539, 4.784.990 ó 4.780.414. Cuando se emplea un huésped procariótico, tal como *S. equisimilis*, se puede desear emplear una fermentación de las bacterias bajo condiciones anaerobias en medios de crecimiento de caldo enriquecido en CO₂. Esto permite una producción mayor de HA que bajo condiciones aerobias. Otra consideración es que las células estreptocócicas cultivadas anaerobiamente no producen exotoxinas pirógenas. Para otros huéspedes procarióticos se pueden adaptar las condiciones de crecimiento apropiadas a la luz de la presente descripción, como conocerán quienes tienen experiencia en la técnica.

Una vez que se ha construido el huésped apropiado y se ha cultivado bajo unas condiciones apropiadas para la producción de HA, se deseará separar el HA así producido. Típicamente, el HA será secretado o, de lo contrario, desprendido al medio circundante por el organismo recombinante, lo que permite el rápido aislamiento del HA del medio por técnicas conocidas. Por ejemplo, el HA puede ser separado de las células y los fragmentos celulares mediante al menos una de las técnicas siguientes: filtración, centrifugación y floculación, y la adición de ácido tricloroacético puede facilitar más la separación del ácido hialurónico de las células y los fragmentos celulares. El HA es luego separado del medio mediante al menos una de las técnicas siguientes: precipitación, ultrafiltración y diálisis. Los agentes de precipitación incluyen alcoholes tales como etanol e isopropanol, disolventes o compuestos orgánicos, tal como acetona, y sales orgánicas de amonio cuaternario (alifáticas positivamente cargadas) tales como cloruro de cetilpiridinio (CPC) y bromuro de cetiltriamonio (CTB; del inglés, cetyl triammonium bromide).

En la Patente de EE.UU. n° 4.517.295 se describe una técnica preferida para el aislamiento de HA, en la que se añade el ácido carboxílico orgánico, el ácido tricloroacético, a la suspensión bacteriana al final de la fermentación. El ácido tricloroacético hace que las células bacterianas se agrupen y mueran y facilita la separación del HA, el producto deseado, de estas células y los fragmentos celulares asociados. El sobrenadante clarificado es concentrado y es dializado para separar los contaminantes de bajo peso molecular, incluyendo el ácido orgánico. En el procedimiento anteriormente mencionado se utiliza la filtración a través de cartuchos filtrantes, tales como los que contienen filtros con un tamaño de poro de 0,22 µm. Se continúa la diafiltración hasta que la conductividad de la disolución disminuye hasta aproximadamente 0,5 megahmios.

El HA concentrado es precipitado al añadir un exceso de etanol u otro disolvente orgánico de calidad "reactivo", y el HA precipitado es luego secado mediante lavado con etanol y secado en vacío y es liofilizado para eliminar el alcohol. El HA puede ser luego redisoluto en un tampón de borato, pH de 8, y ser precipitado con CPC o ciertas sales de amonio orgánicas distintas, tal como CETAB, una disolución mixta de bromuro de trimetilamonio, a 4 grados Celsius. El HA precipitado es recuperado por filtración grosera, resuspendido en NaCl 1 M, diafiltrado y concentrado del modo adicionalmente descrito en la anterior patente referida. El HA resultante es esterilizado por filtración y está listo para ser convertido en una sal, polvo seco o disolución estéril apropiada, dependiendo del deseado uso final.

A. Métodos de ingeniería genética típicos que se pueden emplear

Si se usan células sin las formidables barreras de la membrana celular como células huésped, la transfección se lleva a cabo mediante el método de precipitación con fosfato cálcico, bien conocido por los expertos en la técnica. Sin embargo, también se pueden utilizar otros métodos para introducir DNA en células, tales como mediante inyección nuclear, lípidos catiónicos, electroporación, fusión de protoplastos y mediante el sistema de distribución de biopartículas BIOLISTIC™ desarrollado por DuPont (1989). La ventaja de utilizar el sistema de DuPont es una elevada eficacia de transformación. Si se usan células procarióticas o células que contienen construcciones de pared celular sustanciales, el método preferido para la transfección es el tratamiento con calcio utilizando cloruro cálcico para inducir competencia, o la electroporación.

En la construcción de vectores adecuados que contienen las deseadas secuencias de codificación y control se emplean técnicas de ligación estándares. Los plásmidos o fragmentos de DNA aislados son escindidos, adaptados y vueltos a ligar en la forma deseada para construir los plásmidos requeridos. La escisión se lleva a cabo por tratamiento con una(s) enzima(s) de restricción en un tampón adecuado. En general, se usa aproximadamente 1 µg de fragmentos de DNA o plásmido con aproximadamente 1 unidad de enzima en aproximadamente 20 µl de disolución tampón. Los tampones y las cantidades de sustrato apropiados para las particulares enzimas de restricción vienen especificados por el fabricante. Son operativos unos tiempos de incubación de aproximadamente 1 hora a 37°C.

Después de las incubaciones, se separa la proteína mediante extracción con fenol y cloroformo y se recupera el ácido nucleico de la fracción acuosa mediante precipitación con etanol. Si se requieren extremos romos, la preparación es tratada durante 15 minutos a 15°C con 10 unidades de polimerasa I (Klenow), extraída con fenol-cloroformo y precipitada con etanol. Para la ligación, se tratan cantidades aproximadamente equimolares de los componentes deseados, con los extremos adecuadamente adaptados para obtener el apareamiento correcto, con aproximadamente 10 unidades de DNA ligasa de T4 por 0,5 µg de DNA. Cuando se utilizan vectores escindidos como componentes, puede ser útil evitar la religación del vector escindido mediante un pretratamiento con fosfatasa alcalina bacteriana.

Para el análisis para confirmar las secuencias funcionales en los plásmidos construidos, la primera etapa fue multiplicar el DNA plasmídico por clonación en células de *E. coli* SURE específicamente competentes (Stratagene), realizando la transformación a 30-32°C. En segundo lugar, se utiliza el plásmido recombinante para transformar la cepa Bi8337-41 de *E. coli* K5, que puede producir el precursor UDP-GlcUA, y se seleccionan los transformantes exitosos por resistencia a antibióticos según sea apropiado. Los plásmidos del banco de transformantes son luego explorados en cuanto a colonias bacterianas que presentan producción de HA. Estas colonias son escogidas y multiplicadas, y los plásmidos son purificados y son analizados mediante un mapeo de restricción. Los plásmidos que muestran indicaciones de un gen de HAS funcional son luego adicionalmente caracterizados mediante cualquier número de técnicas para análisis de secuencias, que son conocidas por quienes tienen experiencia normal en este campo técnico.

B. Fuente y vectores y cultivos de células huésped

En general, se utilizaron procariontes para la clonación inicial de las secuencias de DNA y la construcción de los vectores útiles en el invento. Se cree que una fuente adecuada puede ser células Gram-positivas, particularmente las derivadas de las cepas estreptocócicas del Grupo C. Las bacterias con una sola membrana pero con una gruesa pared celular, tales como los estafilococos y estreptococos, son Gram-positivas. Las bacterias Gram-negativas, tales como *E. coli*, contienen dos membranas separadas en lugar de una, que rodean la célula. Los organismos Gram-negativos tienden a tener paredes celulares más delgadas. La membrana única de los organismos Gram-positivos es análoga a la membrana plasmática interna de las bacterias Gram-negativas.

En general, los vectores plasmídicos que contienen orígenes de replicación y secuencias de control que proceden de especies compatibles con la célula huésped se usan en relación con estos huéspedes. El vector lleva normalmente un origen de replicación así como secuencias marcadoras que permiten proporcionar la selección fenotípica en las células transformadas. Por ejemplo, se transforma típicamente *E. coli* utilizando pBR322, un plásmido derivado de una especie de *E. coli*. pBR322 contiene genes para resistencia a ampicilina y tetraciclina y, de este modo, proporciona un medio sencillo para identificar células transformadas. Un plásmido pBR o un plásmido pUC, u otro fago o plásmido microbiano, también debe contener, o debe ser modificado para que contenga, promotores que pueden ser utilizados por el organismo microbiano para la expresión de sus propias proteínas. Tales promotores pueden ser promotores heterólogos, es decir, promotores de otro organismo, con tal de que el promotor sea compatible con la célula huésped, es decir, sea reconocido por la RNA polimerasa de la célula huésped para que la RNA polimerasa transcriba el gen al que está fijado el promotor compatible con el huésped.

Los promotores muy comúnmente utilizados en la construcción de DNA recombinante incluyen el promotor *lacZ*, el promotor *tac*, el promotor del bacteriófago T7 y el sistema promotor de triptófano (*trp*). Aunque estos son los más comúnmente utilizados, se han descubierto y utilizado otros promotores microbianos y se han publicado detalles relativos a sus secuencias de nucleótidos, lo que permite que un trabajador experto les ligue funcionalmente con vectores plasmídicos.

Además, el promotor puede ser un promotor de RNA polimerasa modificado que tenga una actividad promotora aumentada. La modificación del promotor puede ser una mutación o la adición de dos o más elementos promotores en tándem. Cuando dos o más elementos promotores están dispuestos en tándem, los dos o más elementos promotores en tándem pueden ser el mismo elemento promotor o dos o más elementos promotores diferentes. Se entenderá que la expresión “elementos promotores en tándem”, como aquí se usa, significa dos o más secuencias promotoras cada una de las cuales está operativamente enlazada con una secuencia de codificación de modo que las secuencias promotoras dirigen la producción de un polipéptido codificado por la secuencia de codificación al mediar en la transcripción de la secuencia de codificación en mRNA. En las Patentes de EE.UU. números 5.955.310 y 6.255.076, concedidas a Widner *et al.* el 21 de Septiembre de 1999 y el 3 de Julio de 2001, respectivamente, se describen con detalle promotores en tándem así como construcciones y métodos para uso de los mismos para expresión en células de *Bacillus*.

Además, cuando un vector recombinante para uso en el presente invento contiene más de un segmento de ácido nucleico en que cada uno tiene una región de codificación que codifica una proteína, tal como, por ejemplo, un segmento de ácido nucleico que tiene una región de codificación que codifica hialuronano sintasa enzimáticamente activa y un segmento de ácido nucleico que tiene una región de codificación que codifica UDP-glucosa deshidrogenasa enzimáticamente activa, cada uno de los segmentos de ácido nucleico está operativamente enlazado a un promotor. Los dos o más segmentos de ácido nucleico pueden estar enlazados al mismo promotor, y este promotor único puede conducir la expresión de ambos genes, o los dos o más segmentos de ácido nucleico pueden estar enlazados a promotores diferentes.

Además de procariontes, se pueden usar también microbios eucarióticos, tales como levaduras en cultivo. *Saccharomyces cerevisiae*, o levadura común del panadero, es el más comúnmente utilizado entre los microorganismos eucarióticos, aunque se dispone comúnmente de otras diversas cepas. Para expresión en *Saccharomyces*, se utiliza comúnmente, por ejemplo, el plásmido YRp7. Este plásmido ya contiene el gen *trp1* que proporciona un marcador de selección para una cepa mutante de levadura que carece de capacidad para crecer sin triptófano, por ejemplo, ATCC n° 44076 o PEP4-1. La presencia de la lesión de *trp1* como una característica del genoma de la célula huésped de levadura proporciona entonces un ambiente eficaz para detectar la transformación por crecimiento en ausencia de triptófano. Las secuencias promotoras adecuadas en vectores de levadura incluyen los promotores de los genes para utilización de galactosa, la 3-fosfoglicerato cinasa u otras enzimas glicolíticas, tales como enolasa, gliceraldehído-3-fosfato deshidrogenasa, hexocinasa, piruvato descarboxilasa, fosfofructocinasa, glucosa-6-fosfato isomerasa, 3-fosfoglicerato mutasa, piruvato cinasa, triosafosfato isomerasa, fosfoglucosa isomerasa y glucocinasa.

En la construcción de plásmidos de expresión adecuados, las secuencias de terminación asociadas con estos genes están también ligadas al vector de expresión en 3' de la secuencia que se desea que se exprese, para proporcionar la poliadenilación del mRNA y la terminación. Otros promotores, que tienen la ventaja adicional de una transcripción controlada por las condiciones de crecimiento, son la región promotora para alcohol deshidrogenasa 2, citocromo C, fosfatasa ácida, enzimas degradantes asociadas con el metabolismo del nitrógeno, y la susodicha gliceraldehído-3-fosfato deshidrogenasa, y enzimas responsables de la utilización de maltosa y galactosa. Es adecuado cualquier vector plasmídico que contenga un promotor compatible con levadura, un origen de replicación y secuencias de terminación.

Además de microorganismos, también se pueden usar cultivos de células derivadas de organismos multicelulares como huéspedes. En principio, cualquiera de dichos cultivos celulares es operativo, sea un cultivo de vertebrado o sea de invertebrado. Sin embargo, el interés ha sido máximo en las células de vertebrado, y la propagación de células de vertebrado en cultivo ha llegado a ser un procedimiento rutinario en los últimos años. Los ejemplos de dichas líneas celulares huésped útiles son las células VERO y HeLa, las líneas celulares de ovario de hámster chino (CHO; del inglés, *chinese hamster ovary*) y las líneas celulares WI38, BHK, COS y MDCK.

Para uso en células de mamífero, las funciones de control de los vectores de expresión son a menudo proporcionadas por un material vírico. Por ejemplo, los promotores comúnmente utilizados proceden de poliomavirus, adenovirus 2, virus del papiloma bovino (BPV; del inglés, *bovine papilloma virus*) y, muy frecuentemente, el virus 40 de simios (SV40; del inglés, *simian virus 40*). Los promotores precoz y tardío del virus SV40 son particularmente útiles porque ambos se obtienen fácilmente del virus como un fragmento que también contiene el origen vírico de replicación de SV40. También se pueden utilizar fragmentos de SV40 más pequeños o más grandes con tal de que esté incluida la secuencia de aproximadamente 250 pares de bases (bp; del inglés, *base pairs*) que se extiende desde el sitio *Hind III* hacia el sitio *Bgl I* situado en el origen vírico de replicación.

Además, también es posible, y a menudo deseable, utilizar secuencias promotoras o de control normalmente asociadas con la secuencia génica deseada, con tal de que dichas secuencias de control sean compatibles con los sistemas de la célula huésped. Se puede proporcionar un origen de replicación mediante la construcción del vector para que incluya un origen exógeno, tal como el que puede proceder de SV40 u otra fuente vírica (por ejemplo, poliomavirus, adenovirus, BPV), o puede ser proporcionado por el mecanismo de replicación cromosómica de la célula huésped. Si el vector se integra en el cromosoma de la célula huésped, el último mecanismo es a menudo suficiente.

C. Aislamiento de un gen de *Ha* sintasa *bona fide* de una cepa muy encapsulada de *Streptococcus equisimilis* del Grupo C

La proteína codificada, denominada seHAS, tiene 417 aminoácidos (peso molecular calculado de 47.778 y pI de 9,1) y es el miembro más pequeño de la familia de HAS identificado hasta ahora (Figura 2). La seHAS también migra anómalamente rápida en SDS-PAGE ($M_r \sim 42$ kDa) (Figuras 5 y 9).

La Figura 9 es una representación gráfica de un análisis, por transferencia Western, de seHAS recombinante utilizando anticuerpos específicos. Se fraccionaron membranas estreptocócicas del Grupo C (C; carril 1) o el Grupo A (A; carril 4) y membranas de *E. coli* (9 μ g/carril) que contenían seHAS (E; carriles 2, 7 y 9) o spHAS (P; carriles 3, 6, 8 y 10) recombinante mediante SDS-PAGE reductora y se electrotransferieron a nitrocelulosa. Se sondaron y desarrollaron tiras de nitrocelulosa del modo descrito en la solicitud utilizando fracciones purificadas de IgG generadas contra las siguientes regiones de spHAS: péptido E¹⁴⁷-T¹⁶¹ del dominio central (carriles 1-4); péptido del extremo C (carriles 5-6); la proteína completa (carriles 7 y 8); y el dominio central recombinante (carriles 9 y 10). IgG no inmune o membranas de células transformadas con el vector solo no proporcionaron tinción alguna, como en el carril 5.

Las secuencias que codifican las proteínas seHAS y spHAS (ID. SEC. números 1 y 13, respectivamente) tienen una identidad de 72%. La deducida secuencia proteica de seHAS fue confirmada por reactividad con un anticuerpo contra un péptido sintético (Figura 9). La seHAS recombinante expresada en *E. coli* fue recuperada de membranas como una proteína principal (Figura 5) y sintetizaba HA de peso molecular muy grande en presencia de UDP-GlcNAc y UDP-GlcUA *in vitro* (Figura 10).

La Figura 10 muestra un análisis cinético de las distribuciones de tamaños de HA producidas por seHAS y spHAS. Se incubaron membranas de *E. coli*, que contenían cantidades iguales de proteína seHAS o spHAS, a 37°C, con UDP-[¹⁴C]GlcUA 1,35 mM (1,3 x 10³ dpm/nanomol) y UDP-GlcNAc 3,0 mM del modo descrito en la solicitud. Estas concentraciones de sustrato son mayores que 15 veces los respectivos valores de Km. Muestras tomadas a los 0,5, 1,0 y 60 minutos fueron tratadas con SDS y sometidas a cromatografía sobre Sephacryl S400 HR. Los perfiles de HA en el intervalo de fraccionamiento de la columna (fracciones 12-24) se normalizan con respecto al porcentaje de HA total en cada fracción. Los valores por encima de las flechas en el gráfico superior son los pesos moleculares (en millones) de HA determinados directamente en un experimento separado usando un instrumento Dawn para dispersión de luz láser de ángulo múltiple (Wyatt Technology Corp.). Como se indica, se muestran las distribuciones de tamaños del HA sintetizado por seHAS (●, ■, ▲) y spHAS (○, □, ▴) a los 0,5 minutos (○, ●), 1,0 minutos (□, ■) y 60 minutos (Δ, ▲). El análisis mostró que seHAS y spHAS son esencialmente idénticas en cuanto a la distribución de tamaños de las cadenas de HA que sintetizan (Figura 10). seHAS es dos veces más rápida que spHAS en su capacidad para producir HA.

C.1. Cepas bacterianas y vectores

Se obtuvo la cepa D181 mucoide del Grupo C (*Streptococcus equisimilis*) de la Rockefeller University Collection). Las cepas Sure y XL1-Blue MRF⁺ del *E. coli* huésped eran de Stratagene, y la cepa Top10 F⁺ era de Invitrogen. A menos que se indique otra cosa, los estreptococos se cultivaron en THY y las cepas de *E. coli* se cultivaron en medio LB. El vector de expresión pKK-223 era de Pharmacia, el vector de clonación pCR 2.1 era de Invitrogen, y el vector λ Zap Express TM *Bam* HI/CIAP predigerido era de Stratagene.

C.2. DNA recombinante y clonación

DNA genómico de elevado peso molecular de *Streptococcus equisimilis*, aislado mediante el método de Caparon y Scott (como saben quienes tienen una experiencia normal en la técnica), fue parcialmente digerido con *Sau3A1* hasta un tamaño medio de 2-12 kb. El DNA digerido fue precipitado con etanol, lavado y ligado al vector de expresión *Bam* HI/CIAP λ Zap Express. El DNA ligado fue empaquetado en un fago con un extracto PackageneTM obtenido de Promega. El título del banco de fago empaquetado fue comprobado usando *E. coli* XL1-Blue MRF⁺ como huésped.

C.3. Multiplicación por PCR degenerada

Se diseñaron oligonucleótidos degenerados basándose en secuencias conservadas entre spHAS (*Streptococcus pyogenes*), DG42 (HAS de *Xenopus laevis*; 19) y nodC (un factor de nodulación de *Rhizobium meliloti*; 20) y se usaron para una multiplicación por PCR con DNA genómico de D181 como molde. Las condiciones de multiplicación fueron 34 ciclos a: 94°C durante 1 minuto, 44°C durante 1 minuto, 72°C durante 1,5 minutos, seguidos de una extensión final a 72°C durante 10 minutos. El oligonucleótido *HADRF1*, 5'-GAY MGA YRT YTX ACX AAT TAY GCT ATH GAY TTR GG-3' (cadena sentido) corresponde a la secuencia D²⁵⁹RCLTNYAIDL (spHAS). El oligonucleótido *HACTR1*, 5'-ACG WGT WCC CCA NTC XGY ATT TTT NAD XGT RCA-3' (cadena antisentido) corresponde a la región C⁴⁰⁴TIKNTWGT (spHAS). La degeneración de bases en ciertas posiciones se representa mediante la nomenclatura

adoptada por la “International Union for Pure and Applied Chemistry” (IUPAC) en sus códigos para bases degeneradas enumerados en la Tabla VI.

TABLA VI

Códigos – Bases degeneradas de la IUPAC

La IUPAC ha establecido una designación estándar de una sola letra para las bases degeneradas. Ésta es:

B = C+G+T

D = A+G+T

H = A+C+T

K = T+G

M = A+C

N = A+C+G+T

R = A+G

S = G+C

W = A+T

V = A+C+G

X = unas bases menores (especificadas en otra parte)

Y = C+T

Estos dos oligonucleótidos proporcionaron un producto de PCR de 459 bp que fue sometido a separación en un gel de agarosa y purificado utilizando el kit BIO-101 GeneClean. Este fragmento fue luego clonado en el vector pCR2.1 utilizando células TOP 10 F⁺ como huésped de acuerdo con las instrucciones del fabricante. Se purificó DNA plasmídico de doble cadena procedente de *E. coli* (Top 10 F⁺) utilizando el kit QIAfilter Plasmid Midi (Qiagen). También se sintetizaron otros dos cebadores sentido degenerados: *HAVAF1*, 5'-GTN GCT GCT GTW RTX CCW WSX TWT AAY GAR GA-3' (que corresponde a la región V⁶⁶ AAVIPSYNE de spHAS), y *HAVDF1*, 5'-GTX RWT GAY GGN WSX WSN RAX GAT GAX GC-3' (basado en V¹⁰⁰ DDGSSNTD de spHAS). Se sintetizaron dos únicos cebadores antisentido basándose en la secuencia del producto de PCR de 459 bp. Estos eran: *D181.2*, 5'-GAA GGA CTT GTT CCA GCG GT-3', y *D181.4*, 5'-TGA ATG TTC CGA CAC AGG GC-3'. Cada uno de los dos cebadores sentido degenerados, cuando se utilizaron con D181.2 o D181.4 para multiplicar DNA genómico de D181, proporcionaron los productos de PCR con el tamaño esperado. Los cuatro productos de PCR fueron clonados y secuenciados utilizando la misma estrategia que antes. Para cada producto de PCR, se compararon las secuencias obtenidas de seis clones diferentes con objeto de derivar una secuencia de consenso. De esta manera, se obtuvo una secuencia de 1042 bp con un ORF continuo con elevada homología con respecto a spHAS.

C.4. Exploración de bancos

Se usaron dos sondas moleculares para explorar el banco: el producto de PCR clonado de 459 bp y el oligonucleótido *D181.5* (5'-GCTTGATAGGTCACCAG-TGTCACG-3'; derivado de la secuencia de 1042 bp). El producto de PCR de 459 bp fue radiomarcado utilizando el kit de marcación con cebadores aleatorios Prime-It 11 (Stratagene) de acuerdo con las instrucciones del fabricante. Se marcaron oligonucleótidos mediante el kit de fosforilación KinA-ce-It (Stratagene) utilizando [γ ³²P]ATP. Los productos radiomarcados fueron separados del material no marcado, en columnas de empuje NucTrap (Stratagene). La oligosonda se hibridaba específicamente con un producto de digestión genómico de D181 en transferencias Southern. Para explorar el banco del fago λ , se utilizó XLBLUE MRF⁺ como huésped (3000 placas de lisis/placa) sobre membranas de nitrocelulosa que contenían el fago adsorbido, realizándose una prehibridación a 60°C y la hibridación con el oligonucleótido 5'-terminalmente marcado, D181.5, en disolución para hibridación QuikHyb (Stratagene) a 80°C de acuerdo con las instrucciones.

Las membranas fueron luego lavadas con tampón SSC 2x y SDS al 0,1% (peso/volumen) a temperatura ambiental durante 15 minutos y con tampón SSC 0,1x y SDS al 0,1% (peso/volumen) a 60°C durante 30 minutos, secadas y luego expuestas a una película Bio-Max MS a -70°C durante la noche. Las placas de lisis positivas fueron vueltas a sembrar y vueltas a explorar dos veces. Los fagos positivos puros fueron guardados en tampón SM con cloroformo. Una PCR sobre estos fagos con cebadores del vector reveló tres diferentes tamaños de inserto.

Una PCR con una combinación de cebadores del vector y cebadores de diferentes regiones de la secuencia de 1042 bp clonada reveló que sólo uno de los tres fagos diferentes tenía el gen de HAS completo. En este fago, el tamaño

del inserto era 6,5 kb. Fallaron los intentos para subclonar el inserto en forma plasmídica por autoescisión del clon seleccionado del banco de fago. Por lo tanto, se aplicó de nuevo una estrategia de PCR sobre el DNA del fago positivo puro para obtener los extremos 5' y 3' del ORF. Los cebadores oligonucleotídicos *D181.3* (5'-GCCCTGTGTCGGAA-CATTCA-3') y T3 (cebador del vector) multiplicaron un producto de 3 kb, y los oligonucleótidos *D181.5* y T7 (cebador del vector) multiplicaron un producto de 2,5 kb. Al secuenciar estos dos productos anteriores se obtuvieron las secuencias de los extremos 5' y 3' del ORF. El análisis de todas las secuencias de los productos de PCR nos permitió reconstruir el ORF del gen de seHAS de 1254 bp.

C.5. Clonación de expresión de la seHAS

Se diseñaron cebadores en la regiones de los codones de inicio y parada de seHAS para que contuvieran un sitio de restricción *EcoR1* en el oligonucleótido sentido (5'-AGGATCCGAATTCATGAGAACATTAACCTC-3') y un sitio *Pst1* en el oligonucleótido antisentido (5'-AGAATTCTGCAGTTATAATAATTTTACGTGT-3'). Esto se hizo multiplicando un producto de PCR de 1,2 kb procedente de DNA genómico de D181 así como procedente del fago puro positivo en cuanto a hibridación. El producto de 1,2 kb fue purificado mediante electroforesis en gel de agarosa, digerido con *Pst1* y *EcoR1*, y clonado direccionalmente en el vector pKK223 digerido con *Pst1* y *EcoR1*. Se transformaron células *E. coli* SURE con el vector ligado, células que fueron luego cultivadas a 30°C. Esta operación era prácticamente importante ya que otras células huésped o temperaturas superiores daban lugar a delecciones del inserto clonado. Se aislaron las colonias y se purificó su pDNA. De las seis colonias (denominadas a, b, c, d, e, f), cinco tenían el inserto de tamaño correcto, mientras que una no tenía inserto.

C.6. Actividad HA sintasa

Se examinó la actividad HA sintasa en membranas preparadas a partir de los 5 clones anteriores. Se recogieron células frescas en fase logarítmica a 3000g y se lavaron a 4°C con PBS, y se aislaron las membranas mediante una modificación de un método con protoplastos, como conocen quienes tienen una experiencia normal en la técnica. También se obtuvieron preparaciones de membrana de *Streptococcus pyogenes* y *Streptococcus equisimilis* mediante una modificación de un diferente procedimiento con protoplastos. Se incubaron las membranas a 37°C en fosfato sódico y potásico 50 mM, pH de 7,0, con MgCl₂ 20 mM, DTE 1 mM, UDP-GlcUA 120 μM y UDP-GlcNAc 300 μM. Se controló la incorporación de azúcar usando UDP-[¹⁴C]GlcUA (11.766 MBq/milimol; ICN) y/o UDP-[³H]GlcNAc (1080,4 MBq/milimol; NEN). Se terminaron las reacciones mediante la adición de SDS a una concentración final de 2% (peso/volumen). El HA producido fue separado de los precursores mediante cromatografía descendente en papel y fue medido determinando la radiactividad incorporada en el origen.

C.7. Análisis por filtración en gel

El HA radiomarcado producido *in vitro* por membranas que contenían seHAS o spHAS recombinante fue analizado por cromatografía en una columna (0,9 x 40 cm) de Sephacryl S500 HR (Pharmacia Biotech Inc.). Se eluyeron muestras (0,4 ml en NaCl 200 mM, Tris-HCl 5 mM, pH de 8,0, más SDS al 0,5%) con NaCl 200 mM, Tris-HCl 5 mM, pH de 8,0, y se evaluaron fracciones de 0,5 ml en cuanto a la radiactividad por ¹⁴C y/o ³H. La autenticidad del polisacárido HA fue evaluada por tratamiento de una muestra idéntica independiente con la hialuronato liasa de *Streptomyces hyalurolyticus* (EC 4.2.2.1), específica de HA, a 37°C durante 3 horas. El producto de digestión fue luego sometido a filtración en gel.

C.8. SDS-PAGE y transferencia Western

Se llevó a cabo una SDS-PAGE de acuerdo con el método de Laemmli. Se llevaron a cabo electrotransferencias a nitrocelulosa en tampón estándar para transferencias con metanol al 20%, utilizando un dispositivo Mini Transblot de Bio-Rad. Las transferencias fueron bloqueadas con BSA al 2% en TBS. Para la detección, se utilizaron un producto de conjugación de proteína A/G y fosfatasa alcalina (Pierce) y azul de p-nitrotetrazolio/sal p-toluidínica del fosfato de 5-bromo-4-cloro-3-indolilo.

C.9. Secuencia y análisis de DNA

Se secuenciaron ambas cadenas de los plásmidos usando cebadores fluorescentes marcados del vector. Las reacciones de secuenciación se llevaron a cabo usando un kit ThermosequenaseTM para cebadores fluorescentes marcados (con 7-desazaG). Se sometieron las muestras a electroforesis en un secuenciador de DNA ALF Express de Pharmacia y se analizaron los datos mediante el software ALF Manager v3.02. Se secuenciaron las regiones internas de los insertos con cebadores internos usando el aparato ABI Prism 377 (versión de software 2.1.1). Las regiones ambiguas fueron secuenciadas manualmente usando 7-desaza - DNA polimerasa SequenaseTM, mezcla magistral de 7-desaza-GTP (USB) y [α -³⁵S]-dATP (Amersham Life Sciences). Las secuencias obtenidas fueron compiladas y analizadas usando DNASIS, v2.1 (Hitachi Software Engineering Co., Ltd.). Se compararon las secuencias de nucleótidos y aminoácidos con otras secuencias de Genbank y otras bases de datos.

C.10. Identificación de seHAS

Se llevó a cabo la identificación de seHAS utilizando un procedimiento de PCR con cebadores oligonucleotídicos basados en diversas regiones de elevada identidad entre spHAS, DG42 (del que no se sabía que era una HAS de

X. laevis regulada en el desarrollo, y denominada xlHAS) y NodC (una β -GlcNAc transferasa de *Rhizobium*). Las proteínas xlHAS y NodC tienen una identidad de $\sim 50\%$ y $\sim 10\%$ en relación con spHAS, respectivamente. Esta estrategia proporcionó un producto de PCR de 459 bp cuya secuencia tenía una identidad de 66,4% con respecto a spHAS, lo que indicaba que se había identificado un compuesto homólogo del Grupo C (seHAS), del gen de la HA sintasa del Grupo A (spHAS). Luego se reconstruyó la región de codificación completa del gen utilizando una similar estrategia basada en PCR. Luego se utilizó un conjunto final de cebadores de PCR para multiplicar el ORF completo a partir de DNA genómico. Cuando este fragmento de PCR de 1,2 kb fue incorporado al vector de expresión pKK223 y usado para transformar células de *E. coli* SURE, se demostró actividad sintética de HA en las membranas aisladas de 5 de las 5 colonias examinadas.

El ORF del gen reconstruido codifica una nueva proteína prevista de 417 aminoácidos que no estaba en la base de datos y que tiene dos aminoácidos menos que spHAS. Las dos proteínas bacterianas tienen una identidad de 72%, y las secuencias de ácido nucleico tienen una identidad de 70%. El previsto peso molecular de la proteína seHAS es 47.778 y el previsto punto isoeléctrico está en un pH de 9,1. Las HASs de mamífero recién identificadas, tales como las isozimas de ratón y ser humano (mHAS1, 2 y 3, y hHAS1, 2 y 3, respectivamente, en la Figura 2), son similares a las proteínas bacterianas. La identidad global entre los dos grupos es ~ 28 -31%, y, además, muchos aminoácidos de seHAS están muy conservados con aquellos de las HASs eucarióticas (por ejemplo, sustituciones K/R o D/E). A98R, la HAS de PBCV-1, tiene una identidad de 28-33 por ciento con respecto a las HASs de mamífero y se prevé que tenga una topología similar en la membrana lipídica. En las especies de mamífero, los miembros de la misma familia son casi completamente idénticos (por ejemplo, muHAS1 y huHAS1 tienen una identidad de 95%; muHAS2 y huHAS2 tienen una identidad de 98%). Sin embargo, y como se muestra en la Figura 3, incluso en la misma especie, los miembros de las diferentes familias de HAS son más divergentes (por ejemplo, muHAS1 y muHAS2 tienen una identidad de 53%; muHAS1 y muHAS3 tienen una identidad de 57%; y muHAS2 y muHAS3 tienen una identidad de 71%).

En la Figura 11 se muestra el gráfico de hidropatía para seHAS y la prevista topología de membrana. El gráfico de hidrofiliia para la HAS estreptocócica del Grupo C fue generado mediante el método de Kyte y Doolittle (J. Mol. Biol. 157, 105, 1982) usando DNASIS. Se prevé que la proteína sea una proteína integral de membrana.

En la Figura 12 se muestra un modelo para la organización topológica de seHAS en la membrana. La propuesta topología para la proteína se ajusta a la regla de cargas internas y pone dentro el gran dominio central. Es probable que este dominio contenga la mayoría de las funciones ligantes de sustrato y catalíticas de la enzima. En el dominio central se muestran Cys²²⁶ de seHAS, que está conservada en los miembros de todas las familias de HAS, así como las otras tres cisteínas. Cys²⁸¹ es un resto crítico cuya alteración puede alterar drásticamente la distribución de tamaños del producto de HA sintetizado por la enzima.

La topología global de membrana prevista para seHAS es idéntica a aquella para spHAS y para las HASs eucarióticas hasta ahora presentadas. La proteína tienen dos supuestos dominios transmembranales en el extremo amínico y 2-3 dominios transmembranales o asociados a la membrana en el extremo carboxílico. Los gráficos de hidropatía para las dos enzimas estreptocócicas son casi idénticos e ilustran la dificultad para predecir la topología de la región muy hidrófoba de ~ 90 restos en K³¹³-R⁴⁰⁶ de seHAS (K³¹³-K⁴⁰⁵ de spHAS).

La seHAS se expresaba eficazmente en células de *E. coli*. Aproximadamente el 10% de la proteína total de membrana era seHAS, según se valoró por tinción de geles de SDS-PAGE (Figura 5). La prominente banda de seHAS a 42 kDa está cuantitativamente ausente en el carril testigo con vector solo. Este nivel inusualmente elevado de expresión para una proteína de membrana también se encuentra para spHAS al utilizar el mismo vector en células de SURE. En las células de *E. coli* SURE, aproximadamente el 8% de la proteína de membrana es spHAS. Por contraste, la cantidad de seHAS en las membranas del Grupo C no es superior al 1% de la proteína total de membrana. Apenas es detectable la spHAS en las membranas del Grupo A. La seHAS recombinante expresada en células de *E. coli* SURE no sintetiza HA *in vivo* ya que estas células carecen de UDP-GlcUA, uno de los sustratos requeridos. Sin embargo, las membranas que contienen la proteína seHAS recombinante sintetizan HA cuando se les suministran los sustratos UDP-GlcNAc y UDP-GlcUA (Figura 13).

En la Figura 13 se muestra la síntesis de HA auténtico por seHAS recombinante. Se incubaron membranas de *E. coli* (69 μ g) preparadas a partir de células que contenían seHAS recombinante o vector solo, a 37°C durante 1 hora, con UDP-[³H]GlcNAc 700 μ M (2,78 x 10³ dpm/nanomol; □, ■) y UDP-[¹⁴C]GlcUA 300 μ M (3,83 x 10³ dpm/nanomol; o, ●) en un volumen final de 200 μ l, como aquí se describe. La reacción enzimática fue detenida mediante la adición de EDTA a una concentración final de 25 mM. La mitad de la mezcla de reacción fue tratada con hialuronidasa de *Streptomyces* a 37°C durante 3 horas. Se añadió SDS (2%, peso/volumen) a las muestras tratadas (o, □) y no tratadas (●, ■) con hialuronidasa, las cuales fueron calentadas a 90°C durante 1 minuto. Se diluyeron las muestras hasta 500 μ l con tampón de columna (Tris 5 mM, NaCl 0,2 M, pH de 8,0) y se clarificaron por centrifugación, y se inyectaron 200 μ l en una columna de Sephacryl S-500 HR. Se recogieron fracciones (1 ml) y se determinó la radiactividad. BD es la posición del pico de elución del azul de dextrano (del inglés, blue dextran) ($\sim 2 \times 10^6$ Da; Pharmacia). V₀ indica el volumen excluido y V_i el volumen incluido. La relación de [¹⁴C]GlcUA : [³H]GlcNAc incorporados a la cantidad total de HA fraccionado en la columna es 1,4, que es idéntica a la relación de actividades específicas de los dos sustratos. Por lo tanto, las relaciones molares de los azúcares incorporados al producto es 1:1, como se predijo para el HA auténtico. Las membranas de las células transformadas con vector solo no sintetizaron HA.

Utilizando UDP-GlcUA 120 μ M y UDP-GlcNAc 300 μ M, la síntesis de HA era lineal con proteína de membrana (en $\leq 0,2 \mu$ g) y durante al menos 1 hora. Además, las membranas preparadas a partir de células no transformadas o de células transformadas con vector solo no tienen actividad HAS detectable. La síntesis de HA es insignificante si se forma un quelato de Mg^{2+} con EDTA ($< 5\%$ del testigo) o si se prescinde de cualquiera de los dos sustratos ($\sim 2\%$ del testigo). La seHAS recombinante también mostraba la especificidad esperada hacia los sustratos de azúcar-nucleótido, siendo incapaz de copolimerizar UDP-GalA, UDP-Glc o UDP-GalNAc con cualquiera de los dos sustratos normales (Tabla II).

Basándose en el análisis por filtración en gel, la masa media del HA sintetizado por seHAS en membranas aisladas es $5-10 \times 10^6$ Da. Se estima que el producto de la seHAS recombinante es HA auténtico basándose en la incorporación equimolar de ambos azúcares y en su sensibilidad a la degradación por la hialuronidasa específica de *Streptomyces* (Figura 13). Aunque las condiciones para la síntesis total de HA no eran óptimas (ya que se incorporaba $\sim 90\%$ de un sustrato al producto), la enzima producía una amplia distribución de longitudes de cadena de HA. La fracción del pico corresponde a una masa de HA de $7,5 \times 10^6$ Da, que es un polímero que contiene aproximadamente 36.000 azúcares monómeros. La distribución de los tamaños de HA resueltos en esta columna variaba de 2 a 20×10^6 Da.

La deducida secuencia proteica de seHAS fue confirmada por la capacidad de anticuerpos hacia la proteína spHAS para reaccionar cruzadamente con la proteína del Grupo C (Figura 9). Anticuerpos policlonales hacia la proteína spHAS completa o hacia justo el dominio central de spHAS también reaccionaban con la proteína seHAS. Un anticuerpo antipéptido hacia el extremo C de spHAS no reaccionaba cruzadamente con esta región algo divergente de la proteína seHAS. Sin embargo, un anticuerpo antipéptido dirigido contra la secuencia $E^{147}-T^{161}$ de spHAS reconocía la misma prevista secuencia de seHAS. El anticuerpo antipéptido también reacciona con las proteínas seHAS y spHAS nativas de membranas estreptocócicas y confirma que las enzimas nativa y recombinante de ambas especies tienen un tamaño idéntico. Como la proteína spHAS, seHAS migra anómalamente deprisa en SDS-PAGE. Aunque la masa calculada es 47.778 Da, el M_r por SDS-PAGE es consistentemente ~ 42 kDa.

A causa de la identidad secuencial en sus regiones del dominio central y a la idéntica estructura global prevista para las dos enzimas bacterianas, el anticuerpo específico de péptido contra la región $E^{147}-T^{161}$ puede ser utilizado para normalizar la expresión de la proteína HAS en membranas preparadas a partir de células transformadas con genes para las dos diferentes enzimas. Utilizando este planteamiento, se compararon membranas con cantidades esencialmente idénticas de spHAS o seHAS recombinante, con respecto a la velocidad inicial de síntesis de HA y a la distribución del tamaño del producto, HA.

Como se muestra para spHAS, la síntesis de las cadenas de HA por seHAS es progresiva. Parece que las enzimas permanecen asociadas con una cadena de HA en crecimiento hasta que es liberada como un producto final. Por lo tanto, es posible comparar las velocidades de alargamiento de HA por seHAS y spHAS controlando la distribución de tamaños de las cadenas de HA producidas en los primeros momentos, durante el primer ciclo de la síntesis de las cadenas de HA. Basándose en el análisis, por filtración en gel, de los tamaños del producto de HA en diversos momentos, estimamos que la velocidad media de alargamiento por seHAS es aproximadamente 9000 monosacáridos/minuto a 37°C (Figura 10). En cinco minutos, las enzimas pueden polimerizar una cadena de HA de $5-10 \times 10^6$ Da. Por lo tanto, durante una incubación de 60 minutos, cada molécula de enzima podría iniciar, completar y liberar potencialmente del orden de 5-8 de dichas moléculas grandes de HA. En los primeros momentos (por ejemplo, ≤ 1 minuto), que reflejan el alargamiento de las primeras cadenas de HA, la distribución de tamaños del HA producido por seHAS se desplazaba a especies más grandes en comparación con spHAS. En 60 minutos, las dos distribuciones de tamaños del producto de HA eran indistinguibles.

La seHAS clonada representa la HA sintasa auténtica del Grupo C. Por lo tanto, las proteínas del “Grupo C” previamente presentadas o descritas no son las verdaderas HAS del Grupo C. La proteína seHAS es homóloga a nueve de las actualmente conocidas HA sintasas de bacterias, vertebrados y un virus que ahora comprenden esta familia de HA sintasas en rápido crecimiento. Esta homología se muestra particularmente en la Figura 2. En mamíferos, se han identificado tres genes, denominados *HAS1*, *HAS2* y *HAS3*, y se han mapeado en tres cromosomas diferentes tanto en ser humano como en ratón. En anfibios, la única proteína HAS hasta ahora identificada es la DG42 regulada en el desarrollo, que fue clonada en 1988 y de la que, por análisis de la proteína recombinante en membranas de levadura, se ha mostrado recientemente que codifica la actividad HA sintasa. Probablemente, pronto se identificarán otros genes de HAS de *X. laevis*.

Un modelo de evolución divergente sugiere que se pudo haber usurpado un precursor bacteriano primitivo de HAS al principio, durante el desarrollo de los vertebrados, o que la estrategia patogénica bacteriana de producir una cápsula de HA se desarrolló cuando unas bacterias primitivas capturaron una HAS primordial. La evolución convergente de las enzimas HAS bacterianas y eucarióticas hacia una solución estructural común parece improbable, pero pudo haber ocurrido.

Se pronostica que al menos diez proteínas HAS identificadas son proteínas de membrana con una topología similar. La síntesis de HA tiene lugar en la membrana plasmática, y el HA se desprende al medio o permanece celularmente asociado para formar la cápsula bacteriana o un revestimiento pericelular eucariótico. Se utilizan los sustratos de azúcar-nucleótido del citoplasma para ensamblar cadenas de HA que son extruidas al espacio externo a través de la membrana.

Parece que la topología proteica de la porción carboxílica muy hidrófoba de la proteína HAS es crítica para comprender cómo las enzimas prolongan la cadena de HA en crecimiento conforme se extruye simultáneamente a través de la membrana. Por ejemplo, la insólita actividad enzimática puede requerir interacciones poco frecuentes y complejas de la proteína con la bicapa lipídica. Los resultados preliminares basados en el análisis de proteínas de fusión de spHAS-fosfatasa alcalina indican que los extremos amínico y carboxílico y los grandes dominios centrales son todos intracelulares, como se muestra en las Figuras 11 y 12. La proteína seHAS también contiene un gran dominio central (~ 63% de la proteína total) que parece contener los dos sitios ligantes de sustrato y las dos actividades glicosiltransferasa necesarias para la síntesis de HA. Aunque los actuales programas informáticos no pueden predecir fiablemente el número ni la naturaleza de los dominios asociados a la membrana en el largo tramo hidrófobo C-terminal, la propuesta disposición topológica está de acuerdo con la presente evidencia y se aplica también a las enzimas eucarióticas, que son ~ 40% más grandes debido esencialmente a la extensión del extremo C-terminal de la proteína con 2 previstos dominios transmembranales adicionales. Cuatro de los seis restos de Cys de spHAS están conservados en relación con seHAS. Sólo la Cys225 de spHAS y la Cys224 de seHAS están conservadas en todos los miembros de la familia de HAS. Puesto que los agentes que reaccionan con sulfhidrilos, tales como p-mercurobenzoato y NEM, inhiben en gran medida la actividad HAS, es probable que esta Cys conservada sea necesaria o importante para la actividad de la enzima. Sin embargo, los resultados iniciales de estudios por mutagénesis dirigida al sitio indican que un mutante C225S de spHAS no es inactivo, sino que conserva el 5-10% de la actividad de tipo silvestre.

En la Figura 14 se muestra el reconocimiento de las secuencias de ácido nucleico que codifican sólo seHAS, sólo spHAS, o tanto seHAS como spHAS, usando oligonucleótidos específicos. Se diseñaron tres pares de oligonucleótidos sentido-antisentido basándose en la secuencia de la ID. SEC. n° 1 y la secuencia de codificación para spHAS. En la Figura 15 se indican los segmentos de ácido nucleico basados en seHAS (se1-se2 y sesp1-sep2, ID. SEC. números 3-6, respectivamente). Estos tres pares de oligonucleótidos fueron hibridados bajo reacciones PCR típicas con DNA genómico de estreptococos del Grupo C (seHAS; carriles 2, 4 y 6) o Grupo A (spHAS; carriles 3, 5 y 7). Los carriles 1 y 8 indican las posiciones de patrones de peso molecular en kilobases (kb). Las reacciones PCR se llevaron a cabo utilizando DNA polimerasa *Taq* (de Promega) durante 25 ciclos del modo siguiente: 94 grados Celsius durante 1 minuto para conseguir la desnaturalización del DNA, 48 grados Celsius (42 grados Celsius para los cebadores sesp comunes más pequeños) durante 1 minuto para permitir la hibridación, y 72 grados Celsius durante 1,5 minutos para la síntesis de DNA. Las mezclas de reacción PCR fueron luego sometidas a separación mediante electroforesis en un gel de agarosa al 1%.

Se diseñó la pareja de cebadores se1-se2 para que fuera únicamente específica para la HAS del Grupo C (seHAS). Se diseñó la pareja de cebadores sp1-sp2 para que fuera únicamente específica para la HAS del Grupo A (spHAS). Se diseñó la pareja de cebadores sesp1-sep2 para que se hibridara con las secuencias de ácido nucleico de HAS tanto del Grupo A como del Grupo C. Las tres parejas de cebadores se comportaron del modo esperado, lo que mostraba la apropiada capacidad para hibridarse cruzadamente y soportar la generación de productos de PCR que eran específicos y/o únicos.

En la Figura 15 se muestran los oligonucleótidos utilizados para la PCR o la hibridación específicas. Se indican las regiones correspondientes de la ID. SEC. n° 1 para cada uno de los oligonucleótidos sintéticos de las ID. SEC. números 3, 4, 5 y 6. Cada uno de los cuatro oligonucleótidos se hibridará específicamente con la secuencia de seHAS, y las parejas apropiadas de cebadores sentido/antisentido son adecuadas para uso en la reacción en cadena de la polimerasa, como se muestra en la Figura 14.

Expresión de seHAS en *Bacillus subtilis*

En las Figuras 16A y 16B se demuestra la producción de HA recombinante por seHAS en una cepa de *Bacillus subtilis*, *B. subtilis* 168, según se evidencia mediante cromatografía por filtración en gel. La Figura 16A es un gráfico en que se compara la producción de HA en *Bacillus subtilis* 168 transformado con el vector pSM143 solo, con aquella en *Bacillus subtilis* 168 transformado con pSM143 que contenía seHAS. La producción de HA puede ser visualizada por el pico entre aproximadamente 13,5 minutos y aproximadamente 16 minutos. La Figura 16B es una ampliación de este pico para evitar el gran pico causado por la proteína radiomarcada y el azúcar que no se incorporó a HA, que se puede ver entre aproximadamente 16,5 minutos y aproximadamente 20 minutos en la Figura 16A.

Análisis, por filtración en gel, de la producción de HA recombinante por spHAS en *Bacillus subtilis*

En la Figura 17A se demuestra el control nutricional de la distribución de tamaños del HA recombinante producido por spHAS en *Bacillus subtilis*. Se cultivó *Bacillus subtilis* 168 (pPD41_Δ5), que codifica la enzima spHAS, en caldo Luria Bertani (LB), y esta bacteria produjo HA que resultaba eluido de la columna para filtración en gel en un punto temporal anterior (pico a 13,48 minutos) al de la misma cepa cultivada en medio de Spizzen (Sp) (pico aproximadamente 14,43 minutos). Estos dos cultivos fueron desarrollados en paralelo, pero las bacterias cultivadas en LB producen un HA más grande. Se cuantifica la radiactividad del HA tritiado en desintegraciones por segundo (dps). Este control negativo muestra que *B. subtilis* no produce normalmente HA. El pico de HA producido por *Bacillus subtilis* transformado con spHAS es sensible a la HA liasa específica pero no a la proteasa.

Por lo tanto, se puede alterar el peso molecular del HA producido en una célula huésped recombinante variando el medio en que se cultiva la célula huésped. Por ejemplo, cultivando la célula huésped recombinante en un medio complejo, tal como medio LB (Luria-Bertani), Caldo Terrific, NZCYM, SOB, SOC o 2xYT, se producirá una molécula

de HA de mayor peso molecular en comparación con el HA producido por una célula huésped recombinante cultivada en un medio químicamente definido, tal como medio de Spizizen o medio mínimo M9. También se puede variar el tamaño del HA mediante la fuente de carbono suministrada, tal como glucosa.

- 5 En la Figura 17B se muestra la diferencia resultante en el aspecto de los picos cuando se usa el *Bacillus subtilis* 168 que contiene spHAS y el *Bacillus subtilis* que contiene el vector solo.

Mediante un análisis por filtración en gel, se analizaron muestras de medios obtenidas después de la marcación *in vivo* de *Bacillus subtilis* con ^3H -glucosamina. Utilizando este método, es posible determinar el tamaño relativo y la cantidad del ácido hialurónico (HA) producido por las bacterias. Todas las muestras se clarificaron por centrifugación a 16.000xg durante 5 minutos antes de la filtración en gel. Los componentes radiactivos se detectaron con un Detector de Radioflujo LB508 (EG & G Berthold) y un cóctel Zinsser (1,8 ml/minuto).

El tamaño de los polímeros de HA fue analizado por cromatografía en una columna Phenomenex PolySep-GFC-P 5000 ó 6000 (300 x 7,8 mm), realizándose la elución con nitrato sódico 0,2 M a 0,6 ml/minuto en un sistema Waters 600E. Las columnas fueron normalizadas con dextranos de diversos tamaños (580, 145, 50 y 20 kDa) o HA marcado con MANT (DeAngelis, 2001) con pesos moleculares medios de 600 y 80 kDa por dispersión de luz láser de ángulo múltiple (MALLS; del inglés, multiple angle laser light scattering). Para la MALLS, se cargaron primero los polímeros de HA (100 μg) en dos columnas Toso Biosep TSK-GEL en tándem (6000PWXL seguida de 4000PWXL; cada una de 7,8 mm por 30 cm; Japón) y luego se eluyeron con fosfato sódico 50 mM, NaCl 150 mM, pH de 7, a 0,5 ml/minuto. El eluyente fluía a través de un refractómetro interferométrico Optilab DSP y luego de un fotómetro láser Dawn DSF (632,8 nm; Wyatt Technology, Santa Bárbara, California, EE.UU.) en modo de ángulo múltiple. Se utilizó el paquete informático del fabricante para determinar el peso molecular medio absoluto utilizando un coeficiente dn/dC de 0,153.

Se prepararon los patrones de HA por marcación subestequiométrica (1 MANT/~ 50 monosacáridos) de grupos hidroxilo del polisacárido de HA estreptocócico con anhídrido N-metil-isatoico. Se obtuvo el patrón de 600 kDa por subfraccionamiento del grueso del HA utilizando HPLC preparativa. Para producir el patrón de 80 kDa, se utilizó una ultrasonificación prolongada (intervalos de 2 minutos para un total de 30 minutos, acetona al 1% en agua, sobre hielo) del grueso del HA con un sonicador Heat Systems-Ultrasonic W-380 con una micropunta (ajuste de potencia: 4).

30 *Expresión heteróloga de HAS de P. multocida*

El ORF de PmHAS en el inserto pPm7A fue multiplicado mediante 13 ciclos de PCR con polimerasa Taq (Fisher) y cebadores correspondientes a la secuencia próxima a los deducidos extremos amínico y carboxílico (codones en letras mayúsculas: sentido, 5'-gcgaattcaaaggacagaaaATGAACACATTATCACAAAG-3', y antisentido, 5'-gggaattctgcagttatAGAGTTTACTATTAATAATGAAC-3'; codones de inicio y parada, respectivamente, en letra negrita). El codón 2 (T a C) fue alterado (letra minúscula cursiva) para aumentar la producción de proteína en *E. coli*. Los cebadores también contenían sitios de restricción para *EcoRI* y *PstI* (subrayados) para facilitar la clonación en el plásmido de expresión pKK223-3 (promotor tac; Pharmacia). Se transformaron células de *E. coli* SURE (Stratagene) con la construcción recombinante resultante, pPmHAS, y se usó esta cepa como fuente de preparaciones de membrana para ensayos de HAS *in vitro*. Antes de la recolección, los cultivos en fase logarítmica (caldo LB, 30°C) fueron estimulados con isopropiltiogalactósido (IPTG) 0,5 mM durante 3 horas. También se utilizó el plásmido para transformar *E. coli* K5; la cepa resultante fue examinada en cuanto a la presencia de una cápsula mediante microscopía óptica y centrifugación en gradiente de densidades. Los cultivos de bacterias K5 no fueron rutinariamente estimulados ya que la adición de IPTG no aumentaba significativamente los niveles de HA en LB ni en medio definido. Las bacterias K5 son huéspedes extraños útiles porque contienen proteínas para transporte de polisacáridos y una maquinaria que interacciona con PmHAS durante la síntesis de HA; éstas proteínas facilitan el transporte de HA fuera de la célula.

Membranas derivadas de células de *E. coli* SURE que contenían el plásmido pPmHAS, pero no las muestras de células con el vector pKK223-3 solo, sintetizaban HA *in vitro* cuando se les suministraba tanto UDP-GlcUA como UDP-GlcNAc {25 frente a * 1,5 picomoles de transferencia de GlcUA ([mg de proteína] $^{-1}$ [h] $^{-1}$), respectivamente}. No se observó incorporación de [^{14}C]GlcA si se prescindía de UDP-GlcNAc o si se quelaban los iones metálicos divalentes con EDTA. La actividad HAS derivada de HAS recombinante era similar a la de la enzima obtenida de membranas de *P. multocida* de tipo silvestre ya que Mn^{2+} estimulaba al menos diez veces más actividad que Mg^{2+} .

También se ensayaron cultivos de *E. coli* recombinante en cuanto a la presencia del polisacárido HA mediante un ensayo radiométrico en que se usaba una proteína ligante de HA marcada. *E. coli* K5 con pPmHAS producía 460 $\mu\text{g}/\text{ml}$ de HA por A600. Las células de K5 con el vector pKK223-3 solo no producían HA (* 0,01 $\mu\text{g}/\text{ml}$ de HA por A600). Por comparación, *P. multocida* P-1059 de tipo silvestre cultivada en el mismo medio producía 1100 $\mu\text{g}/\text{ml}$ de HA por A600. *E. coli* K5 con pPmHAS producía niveles tan elevados de HA que las células llegaban a encapsularse (Figura 18A). El radio de la cápsula de la cepa recombinante era ~ 0,2-0,5 μm (suponiendo una anchura de célula bacteriana de 0,5 μm). Esta cápsula podía ser separada por tratamiento con hialuronidasa testicular ovina o HA liasa de *Streptomyces* (Figura 18B). Ni la cepa huésped K5 nativa ni los transformantes que contenían el vector pKK223-3 poseían una cápsula fácilmente observable, según se determinaba por microscopía óptica. Mediante centrifugación en gradiente de densidades, también se consideró que las células de K5 con pPmHAS estaban encapsuladas. Las células recombinantes flotaban en la parte superior del colchón de Percoll al 58%, mientras que las células con el vector testigo o las células recombinantes tratadas con hialuronidasa se iban al fondo a través del colchón de Percoll (no mostrado).

Papel de las glicosiltransferasas en el transporte durante la biosíntesis capsular

Las glicosiltransferasas catalizan la formación de la cadena principal de GAG repetitiva pero, en ciertos casos, estos mismos polipéptidos pueden desempeñar también papeles en el transporte del polímero a través de la membrana celular. Los estreptococos de los Grupos A y C Gram-positivos poseen sólo una membrana lipídica, y el operón de la cápsula codifica la sintasa y dos enzimas para la producción de UDP-GlcUA, UDP-glucosa deshidrogenasa y UDP-glucosa pirofosforilasa (~ 4 kilobases de DNA; Crater y van de Rijn, 1995). Análisis topológicos de una serie de proteínas de fusión de spHAS estreptocócica que contenían enzimas informadoras indican que esta sintasa atraviesa la bicapa al menos cuatro veces y está íntimamente asociada con la membrana (Heldermon *et al.*, 2001) (Figura 19). A partir de análisis bioquímicos y biofísicos, parece que un complejo compuesto de un monómero del polipéptido spHAS o seHAS y ~ 16 moléculas lipídicas cataliza la transferencia de ambos UDP- azúcares a la cadena de HA naciente (Tlapak-Simmons *et al.*, 1998). Se especuló que spHAS o seHAS, pequeños polipéptidos integrales de membrana, requerirían la ayuda de los lípidos para facilitar el transporte de la cadena polímera de HA en crecimiento a través del núcleo hidrófobo de la bicapa al crear un poro de proteína/lípido.

Por otra parte, las bacterias Gram-negativas capaces de la biosíntesis de GAG, *Escherichia coli* y *Pasteurella multocida*, poseen dos membranas lipídicas, y los loci de su cápsula codifican muchas proteínas asociadas al transporte además de las glicosiltransferasas y las enzimas formadoras de UDP-GlcUA (~ 10-18 kilobases; Roberts, 1996; Townsend *et al.*, 2001). Aunque muchos detalles no son bien entendidos, en el modelo mejor estudiado, el sistema capsular del Grupo II de *E. coli*, parece que el transporte de la cadena de polímero naciente necesita un aparato compuesto de al menos 7 especies polipeptídicas distintas (Whitfield y Roberts, 1999; Silver *et al.*, 2001). En resumen, un complejo que contiene KpsC,M,S,T se ensambla en la membrana interna e interacciona con el complejo catalítico KfiA,B,C. KpsM y T forman el transportador “casete de unión a ATP” (ABC; del inglés, ATP-binding cassette). Una proteína periplásmica, KpsD, y un dímero de otra proteína de la membrana interna, KpsE, ayudan a transportar el polímero a través del espacio periplásmico (Arrecubieta, 2001). Se reúne un complejo de porinas en la membrana externa para transportar la cadena polisacárida en crecimiento fuera de la célula.

Ciertos mutantes para Kps polimerizan la cadena polisacárida capsular pero poseen una translocación defectuosa, lo que da lugar a la acumulación de polímero en el citoplasma o el periplasma. También se piensa que *P. multocida*, basándose en las similitudes secuenciales y la disposición génica de sus supuestas proteínas de transporte con respecto a las proteínas de *E. coli*, tiene un sistema de transporte de tipo Grupo II.

En el caso de pmHAS y pmCS, es probable que la cola carboxilo-terminal contenga un segmento de acoplamiento que interaccione con el mecanismo de transporte (Jing y DeAngelis, 2000) (Figura 19).

HAS de Streptococcus uberis

En la Figura 21 se ilustra la multiplicación, por PCR, del gen de HAS de cuatro colonias mucoides distintas de *S. uberis*. Para cada muestra, resultó evidente una banda en el esperado tamaño de aproximadamente 1,25 kb, correspondiente al marco de lectura completo de suHAS más sitios de restricción que fueron añadidos por la multiplicación por PCR.

En la Figura 22 se ilustra la actividad HA sintasa de *S. pyogenes*, *S. equisimilis* y *S. uberis*.

Otras secuencias de HAS identificadas

Además de las secuencias de HAS que han sido aquí descritas e ilustradas en la alineación de la Figura 2, se han identificado otras secuencias de HAS que pueden ser usadas en métodos para producción de HA en una especie de *Bacillus*. Por ejemplo, las ID. SEC. números 15 y 16 describen las secuencias de nucleótidos y aminoácidos, respectivamente, de una HA sintasa hallada en la arqueobacteria *Sulfolobus solfataricus*. El aislamiento de esta HA sintasa de un organismo extremófilo proporciona una HAS que tiene una mejor estabilidad y una cinética más rápida que las HA sintasas previamente aquí descritas, a causa de su capacidad para actuar a temperaturas elevadas, es decir, a aproximadamente 75°C.

Un grupo de genes similar al operón estreptocócico *hasABC* ha sido identificado en el plásmido pXO1 de *Bacillus anthracis*, que contiene los genes de la toxina del ántrax. Sin embargo, el orden de los genes en pXO1 es A, C, B. La secuencia completa para el plásmido pXO1 está bajo el N° de Acceso AF065404, y la secuencia similar a *hasA* es el ORF 93 de esta secuencia y comienza en 111374 y se detiene en 112474. Las ID. SEC. números 17 y 18 representan las secuencias de nucleótidos y aminoácidos, respectivamente, del gen similar a *hasA* identificado en pXO1 de *Bacillus anthracis*. No hay comunicados sobre una cápsula polisacárida en *Bacillus anthracis* y, por lo tanto, Okinaka *et al.*, el grupo que identificó estos genes, creen que los ORFs 93, 94 y 95 de pXO1 son ejemplos de genes no funcionales que aún han de deteriorarse [*J. Bacteriol.* 181: 6509 (1999)].

Se ha identificado un supuesto tercer HAS en un virus que infecta las algas pardas *Ectocarpus siliculosus*. En las ID. SEC. números 19 y 20, respectivamente, se pueden hallar las secuencias de aminoácidos y nucleótidos. Este caso es probablemente similar al de la cvHAS del virus PBCV-1.

ES 2 346 196 T3

Un método para demostrar la actividad HA sintasa (nativa o recombinante) de cualquier supuesta HA sintasa implica desarrollar las bacterias en un cultivo líquido, extraer la fracción polisacárida (es decir, precipitación con detergente catiónico/extracción con elevada concentración salina/precipitación con alcohol/redisolución en agua/extracción con disolvente/precipitación con alcohol), y analizar la composición de monosacáridos después de una hidrólisis ácida. Otro análisis incluye una electroforesis, en gel de agarosa, de los polímeros intactos y las muestras enzimáticamente tratadas (HA liasa, condroitinasa, etc.). Además, son útiles los ensayos biológicos en que se utilizan proteínas ligantes de HA específicas en un formato ELISA o de competición. Para ensayar la enzima, se preparan membranas de células, y se proporcionan diversos sustratos de UDP-azúcar y luego se analiza su incorporación al polímero, lo que va seguido de cromatografía y/o electroforesis. Se observa expresión heteróloga al preparar un casete génico usando una PCR con cebadores y DNA genómico que permite la clonación del ORF en un vector de expresión. Se pueden transformar diversos huéspedes con dicho vector y se pueden analizar las células recombinantes resultantes en cuanto al polisacárido y/o la enzima, como aquí se describió previamente.

De este modo, debería ser evidente que aquí se han descrito una célula huésped recombinante que tiene un segmento purificado de ácido nucleico que tiene una región de codificación que codifica una HAS enzimáticamente activa en ella introducida, así como métodos para producir ácido hialurónico a partir de la célula huésped recombinante, que satisfacen completamente los objetivos y ventajas anteriormente expuestos. Aunque el invento ha sido descrito junto con realizaciones específicas del mismo, es evidente que muchas alternativas, modificaciones y variaciones resultarán obvias a los expertos en la técnica. En consecuencia, se pretende abarcar todas las citadas alternativas, modificaciones y variaciones, que caen dentro del alcance de las reivindicaciones adjuntas.

Todas las mediciones numéricas y cuantitativas expuestas en esta solicitud (incluyendo las de los ejemplos y las reivindicaciones) son aproximaciones.

El invento ilustrativamente descrito o reivindicado aquí adecuadamente puede ser llevado a la práctica en ausencia de cualquier elemento que no sea específicamente descrito o reivindicado aquí. De este modo, el invento puede comprender, consistir en, o consistir esencialmente en, los elementos aquí reivindicados.

Las siguientes reivindicaciones dan derecho al más amplio alcance posible consistente con esta solicitud. Las reivindicaciones no se limitarán necesariamente a las realizaciones preferidas ni a las realizaciones mostradas en los ejemplos.

REIVINDICACIONES

1. Una célula huésped recombinante, en que la célula huésped recombinante es una célula de *Bacillus* que comprende:

un vector recombinante que comprende un segmento purificado de ácido nucleico que tiene una región de codificación que codifica una hialuronano sintasa enzimáticamente activa, en que la región de codificación está bajo el control de un promotor, y en que el segmento purificado de ácido nucleico es seleccionado del grupo que consiste en:

(A) un segmento purificado de ácido nucleico de acuerdo con la ID. SEC. nº 11;

(B) un segmento purificado de ácido nucleico que comprende una región de codificación que codifica la hialuronano sintasa enzimáticamente activa de la ID. SEC. nº 12; y

(C) un segmento purificado de ácido nucleico que tiene una identidad de entre 90% y 99% con respecto a la ID. SEC. nº 11 y comprende cambios conservativos o semiconservativos de codones de aminoácidos cuando se compara con la ID. SEC. nº 11, siendo dichos cambios la sustitución de uno o más codones nucleotídicos que codifican un aminoácido de uno o más de los grupos siguientes por un codón que codifica otro aminoácido del mismo grupo:

Grupo A: alanina, valina, leucina, isoleucina, prolina, metionina, fenilalanina y triptófano;

Grupo B: glicocola, serina, treonina, cisteína, asparagina y glutamina;

Grupo C: ácido aspártico y ácido glutámico;

Grupo D: lisina, arginina e histidina; y

un vector recombinante que comprende un segmento purificado de ácido nucleico que tiene una región de codificación que codifica una enzima enzimáticamente activa de la ruta biosintética de UDP-GlcUA, en que la enzima enzimáticamente activa de la ruta biosintética de UDP-GlcUA es seleccionada del grupo que consiste en UDP-glucosa deshidrogenasa, UDP-glucosa pirofosforilasa y combinaciones de las mismas.

2. La célula huésped recombinante de la Reivindicación 1, en que la célula huésped recombinante produce ácido hialurónico.

3. La célula huésped recombinante de la Reivindicación 1, en que la célula huésped recombinante es una célula de *Bacillus subtilis* o *Bacillus licheniformis*.

4. La célula huésped recombinante de la Reivindicación 1, en que la región de codificación que codifica la hialuronano sintasa enzimáticamente activa del segmento purificado de ácido nucleico está bajo el control de un promotor compatible con bacterias Gram-positivas.

5. La célula huésped recombinante de la Reivindicación 4, en que el promotor es un promotor compatible con *Bacillus*.

6. La célula huésped recombinante de cualquiera de las Reivindicaciones 1-5, en que el vector recombinante que comprende un segmento purificado de ácido nucleico que tiene una región de codificación que codifica una hialuronano sintasa enzimáticamente activa comprende además un segmento purificado de ácido nucleico que tiene una región de codificación que codifica UDP-glucosa deshidrogenasa enzimáticamente activa.

7. La célula huésped recombinante de la Reivindicación 6, en que la región de codificación que codifica la HAS enzimáticamente activa y la región de codificación que codifica la UDP-glucosa deshidrogenasa enzimáticamente activa están bajo el control de al menos una copia de al menos un promotor.

8. La célula huésped recombinante de la Reivindicación 6, en que la región de codificación que codifica la HAS enzimáticamente activa y la región de codificación que codifica la UDP-glucosa deshidrogenasa enzimáticamente activa están bajo el control de promotores diferentes.

9. La célula huésped recombinante de cualquiera de las Reivindicaciones 1-8, en que la célula huésped recombinante presenta una producción aumentada de al menos uno de UDP-GlcUA y UDP-GlcNAc.

10. La célula huésped recombinante de la Reivindicación 9, en que la célula huésped recombinante incluye además al menos un promotor de RNA polimerasa modificado, de modo que, cuando el promotor de RNA polimerasa modificado es reconocido por una RNA polimerasa, la RNA polimerasa es capaz de expresar RNA en una cantidad

mayor que en el caso de un promotor de RNA polimerasa endógeno, y en que la modificación en el promotor de RNA polimerasa es seleccionada del grupo que consiste en una mutación y elementos promotores en tándem.

11. La célula huésped recombinante de la Reivindicación 9, en que la célula huésped recombinante incluye además al menos uno de:

- (A) al menos un elemento estabilizador de RNA mensajero adicional con respecto a los hallados en una célula nativa de *Bacillus*; y
- (B) al menos un elemento desestabilizador de RNA mensajero menos con respecto a los hallados en una célula nativa de *Bacillus*.

12. La célula huésped recombinante de la Reivindicación 9, en que la célula huésped recombinante incluye además al menos un segmento de ácido nucleico que tiene una región de codificación que codifica una enzima funcionalmente activa de una ruta para la biosíntesis de precursores de UDP-azúcar, de modo que la célula huésped recombinante tiene una actividad mayor que la de una célula huésped nativa que expresa una enzima endógena de la ruta para la biosíntesis de precursores de UDP-azúcar.

13. La célula huésped recombinante de la Reivindicación 9, en que la célula huésped recombinante incluye además al menos un gen mutado para la biosíntesis de precursores de UDP-azúcar, seleccionado del grupo que consiste en: un gen mutado de precursor de UDP-azúcar que aumenta la semivida de un RNA mensajero transcrito, y un gen mutado de precursor de UDP-azúcar que codifica un RNA mensajero que tiene una eficacia translacional aumentada.

14. La célula huésped recombinante de la Reivindicación 13, en que la mutación en el gen para la biosíntesis de precursores de UDP-azúcar tiene lugar en un sitio de unión al ribosoma del gen para la biosíntesis de precursores de UDP-azúcar, de modo que un ribosoma tiene una afinidad ligante aumentada por el sitio de unión al ribosoma.

15. La célula huésped recombinante de la Reivindicación 1, en que el segmento purificado de ácido nucleico es introducido en la célula de *Bacillus* mediante al menos una de las técnicas: transformación, transfección, transducción y electroporación.

16. La célula huésped recombinante de la Reivindicación 1, en que el segmento purificado de ácido nucleico que tiene una región de codificación que codifica la hialuronano sintasa enzimáticamente activa es el segmento purificado de ácido nucleico de (A).

17. La célula huésped recombinante de la Reivindicación 1, en que el segmento purificado de ácido nucleico que tiene una región de codificación que codifica la hialuronano sintasa enzimáticamente activa es el segmento purificado de ácido nucleico de (B).

18. Un método para producir ácido hialurónico, que comprende las operaciones de:

- proporcionar la célula huésped recombinante de cualquiera de las Reivindicaciones 1-17;
- cultivar la célula huésped recombinante en un medio para que secrete ácido hialurónico; y
- recuperar el ácido hialurónico secretado.

19. El método de acuerdo con la Reivindicación 18, en que la operación de recuperación del ácido hialurónico comprende extraer del medio el ácido hialurónico secretado.

20. El método de acuerdo con la Reivindicación 19, que comprende además la operación de purificar el ácido hialurónico extraído.

21. El método de acuerdo con la Reivindicación 18, en que la operación de cultivar la célula huésped recombinante en un medio para que secrete ácido hialurónico es adicionalmente definida por ser seleccionada del grupo que consiste en:

- (A) cultivar la célula huésped recombinante en un medio químicamente definido a una temperatura de aproximadamente 25°C a aproximadamente 42°C para que secrete ácido hialurónico;
- (B) cultivar la célula huésped recombinante en un medio complejo a una temperatura de aproximadamente 25°C a aproximadamente 42°C para que secrete ácido hialurónico; y
- (C) cultivar la célula huésped recombinante en un medio que contiene glucosa y al menos uno de N-acetil-glucosamina y glucosamina, para que secrete ácido hialurónico.

ES 2 346 196 T3

22. El método de acuerdo con la Reivindicación 18, en que la operación de recuperar el ácido hialurónico secretado es adicionalmente definida como:

- separar el ácido hialurónico de las células y los fragmentos celulares mediante al menos una de las técnicas: filtración, centrifugación y floculación;
- concentrar el ácido hialurónico separado; y
- separar del medio el ácido hialurónico concentrado mediante al menos un método seleccionado del grupo que consiste en: precipitación, ultrafiltración y diálisis.

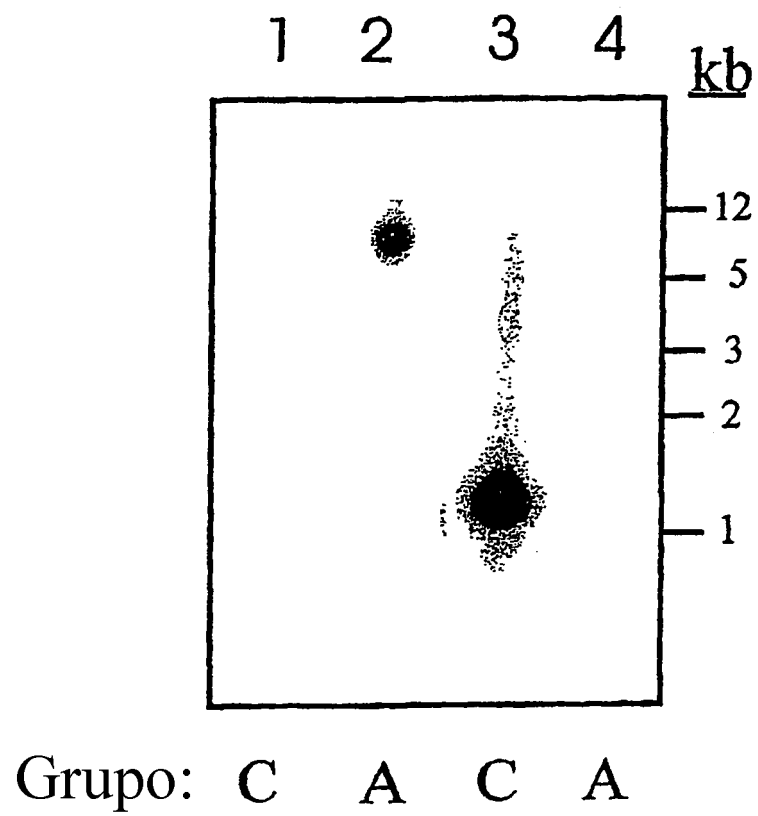


FIGURA 1

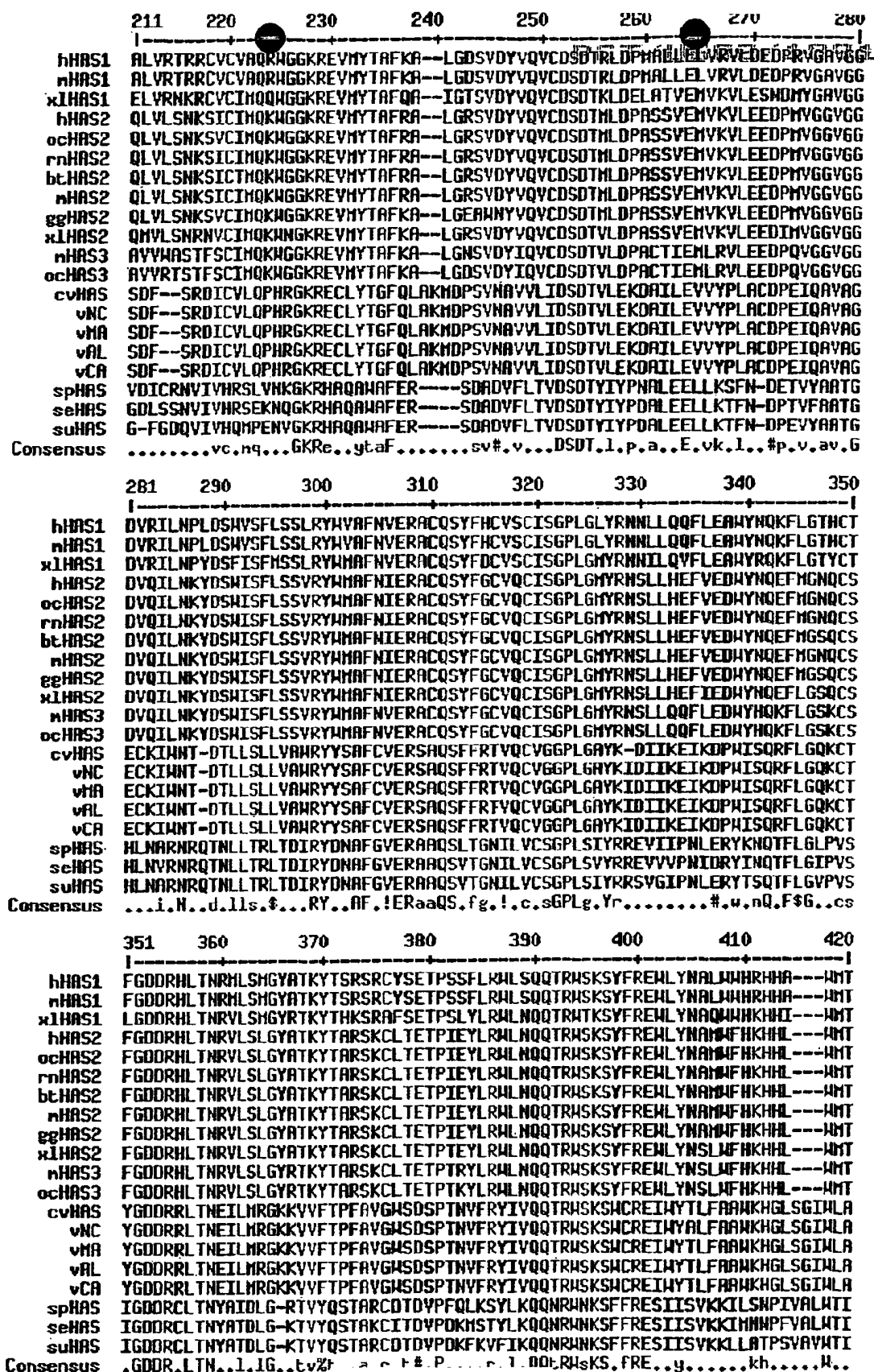


FIGURA 2

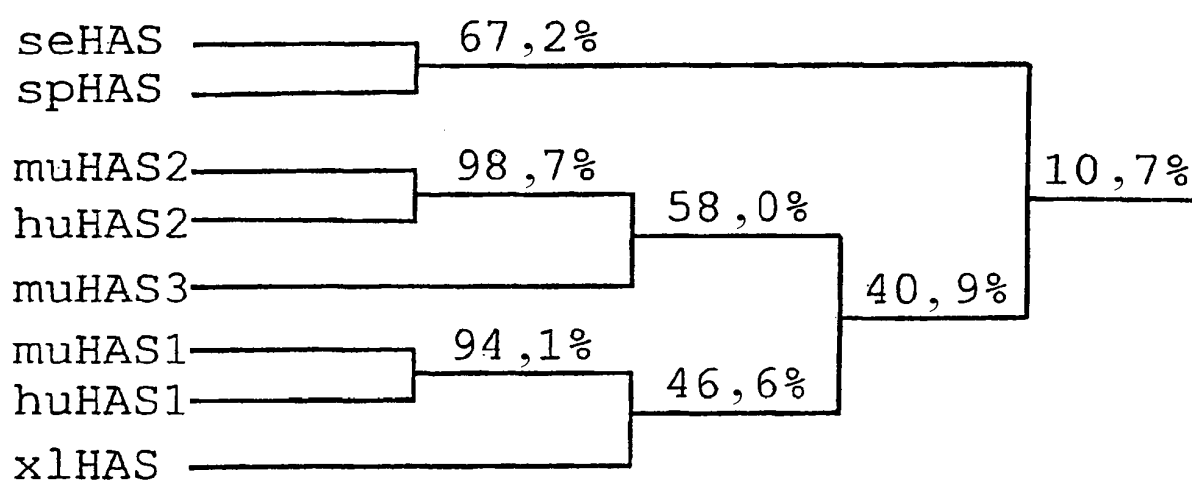


FIGURA 3

DISTRIBUCIÓN DE TAMAÑOS DEL HIALURONANO
PRODUCIDO POR DIFERENTES ENZIMAS HAS
ESTREPTOCÓCICAS GENÉTICAMENTE MODIFICADAS

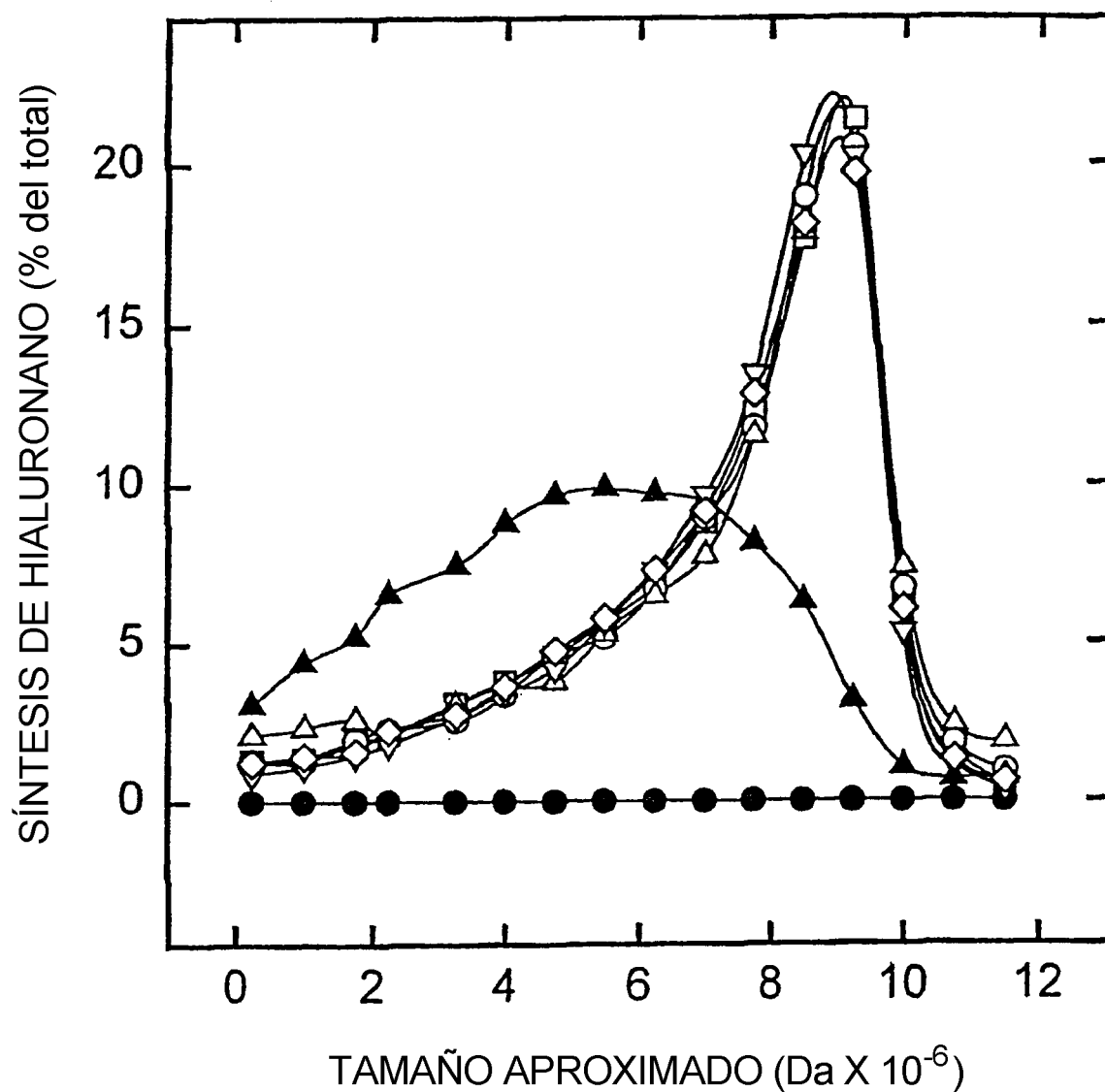


FIGURA 4

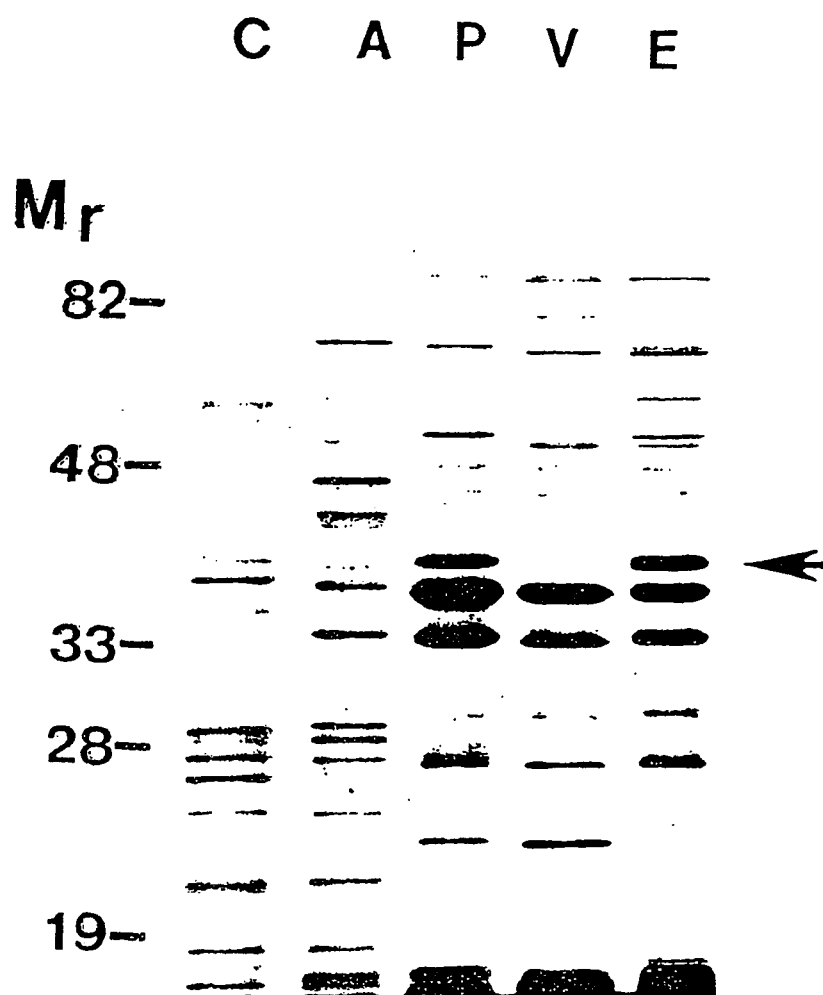


FIGURA 5

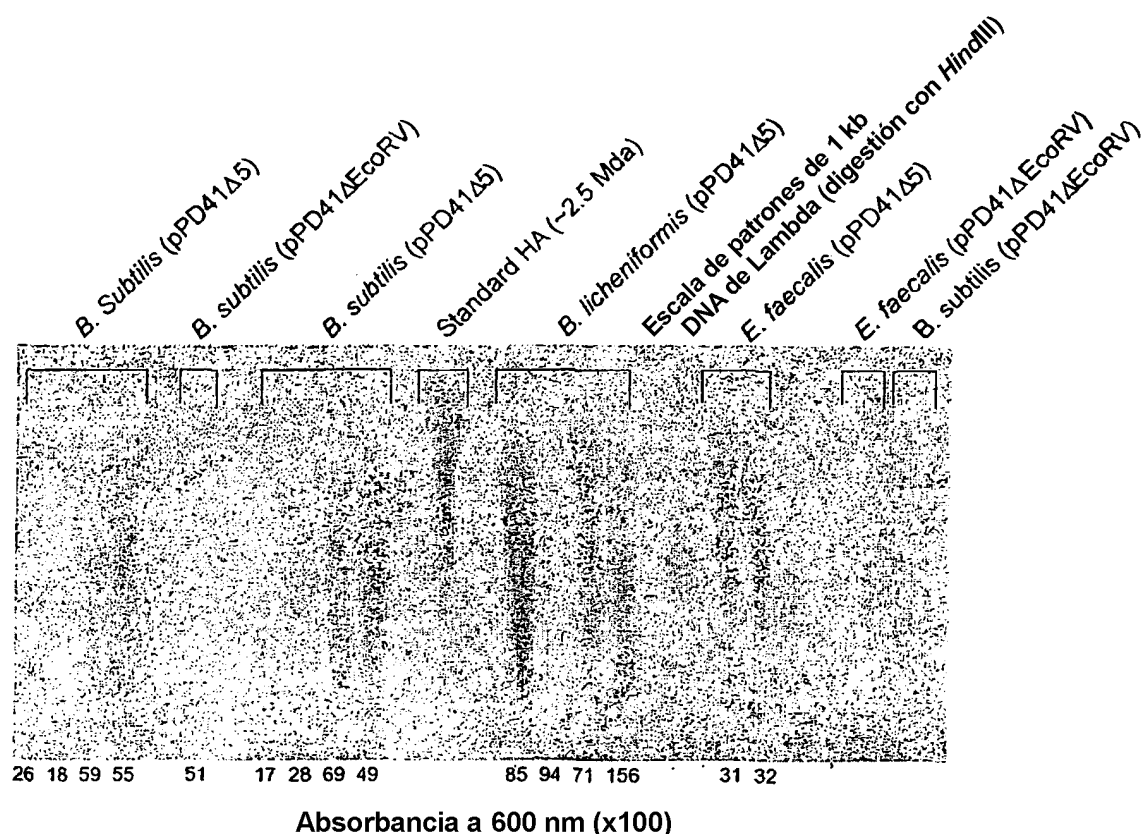
Producción de hialuronano recombinante en *Bacillus* y *Enterococcus*

Imagen digital de un gel de agarosa teñido para hialuronano. Se transformaron diferentes especies bacterianas con los plásmidos indicados que contenían el gen *hasB* y o bien un gen *hasA* funcional (pPD41Δ5) o bien un gen *hasA* no funcional (pPD41ΔEcoRV). Cada carril representa una colonia independiente sembrada en medio fresco de cultivo, y todos los cultivos se recolectaron al mismo tiempo. Se indica en el fondo la absorbancia a 600 nm cuando se recolectaron los cultivos. Se separó el medio después de la centrifugación de las células y se analizó mediante electroforesis en gel de agarosa. Se demostró la presencia de HA por la tinción oscura con StainsAll. En un experimento separado, casi todo el material teñido producido por las bacterias transformadas fue eliminado por tratamiento con la ácido hialurónico liasa específica, lo que confirma que el material es HA. Las cepas bacterianas que contienen pPD41ΔEcoRV no deberían poder producir HA, y, en realidad, todas fueron negativas. Los cultivos de *B. licheniformis* también demostraron la producción de HA.

FIGURA 6

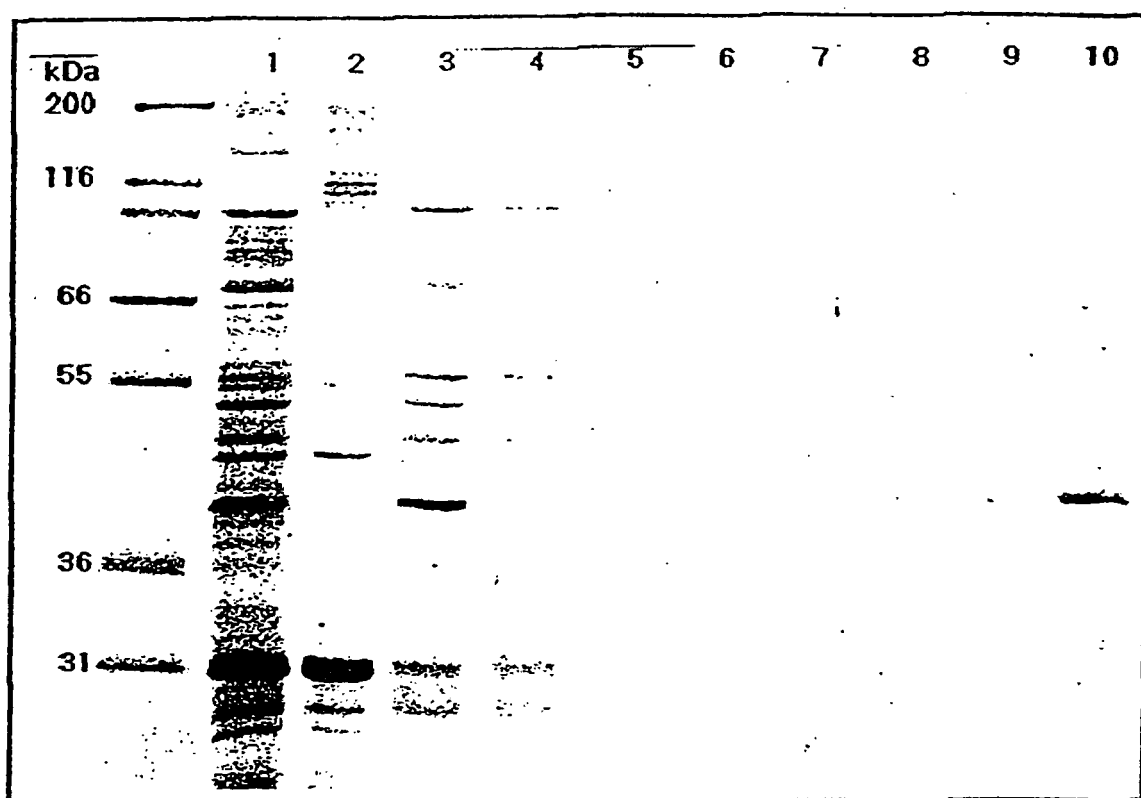


FIGURA 7

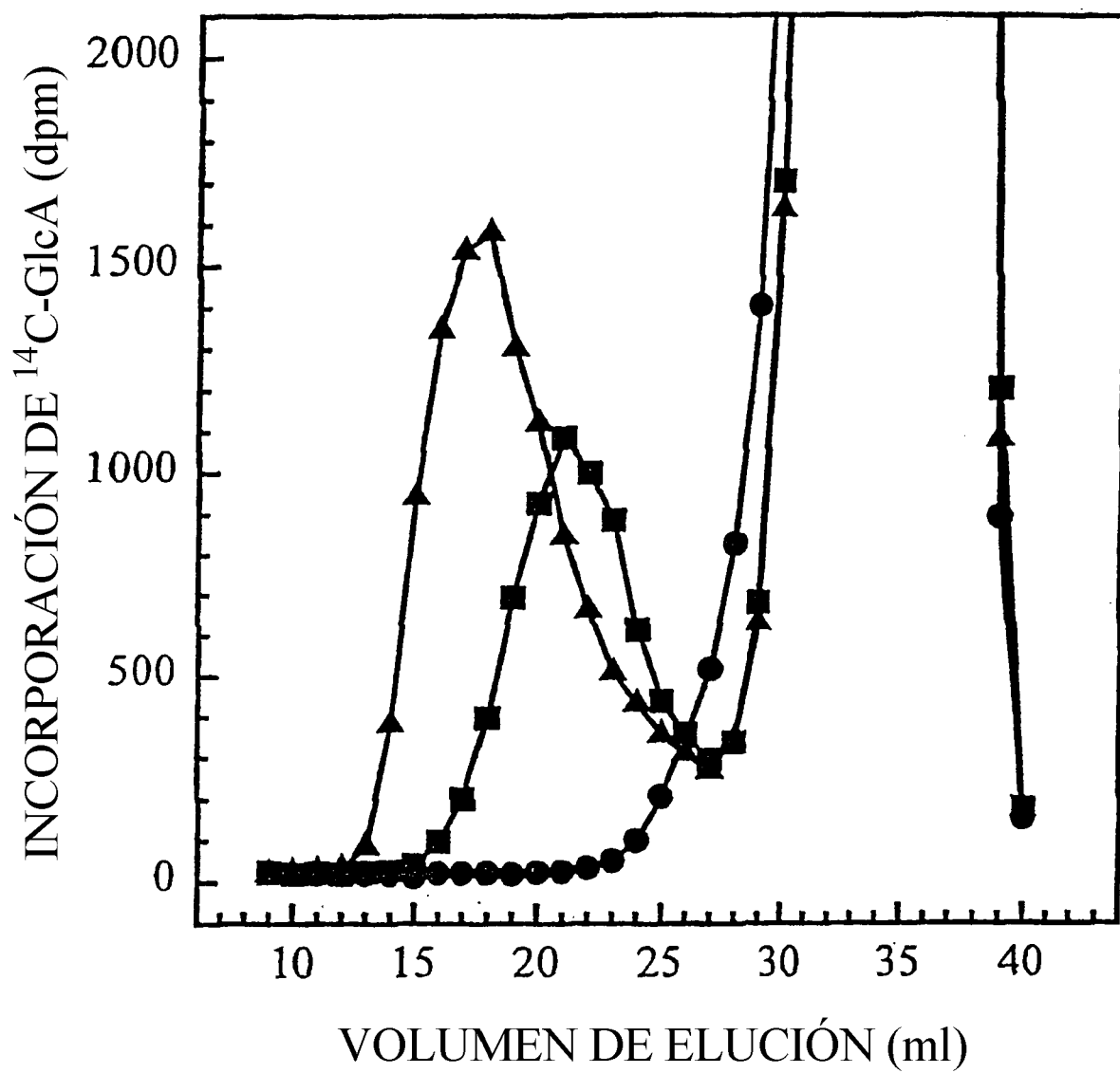


FIGURA 8

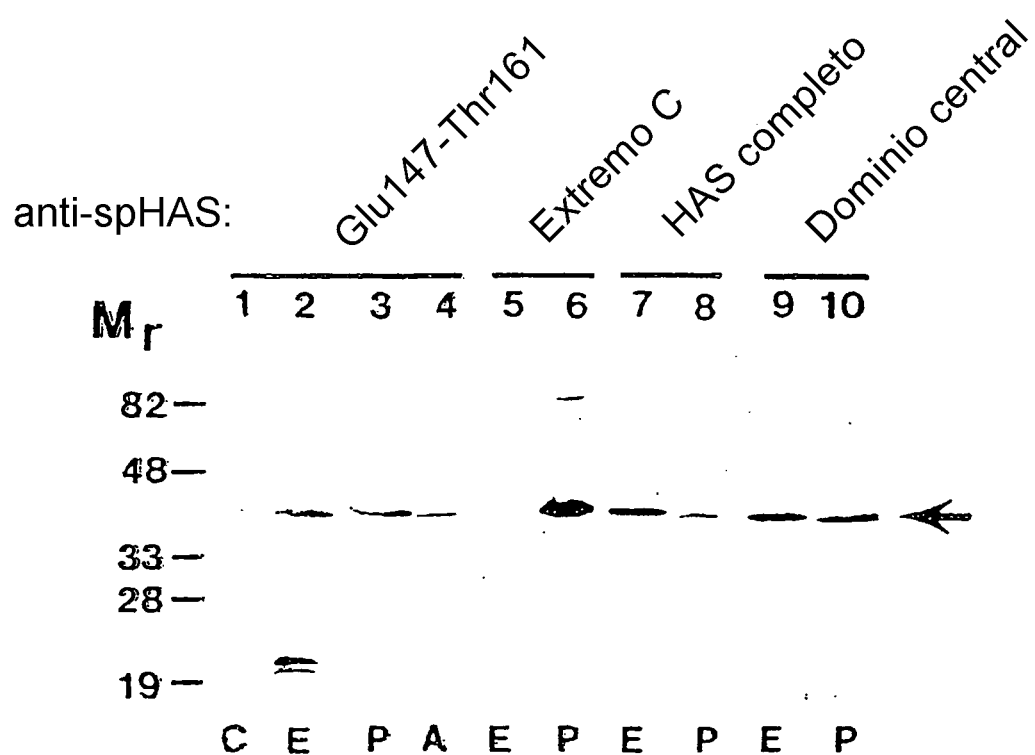


FIGURA 9

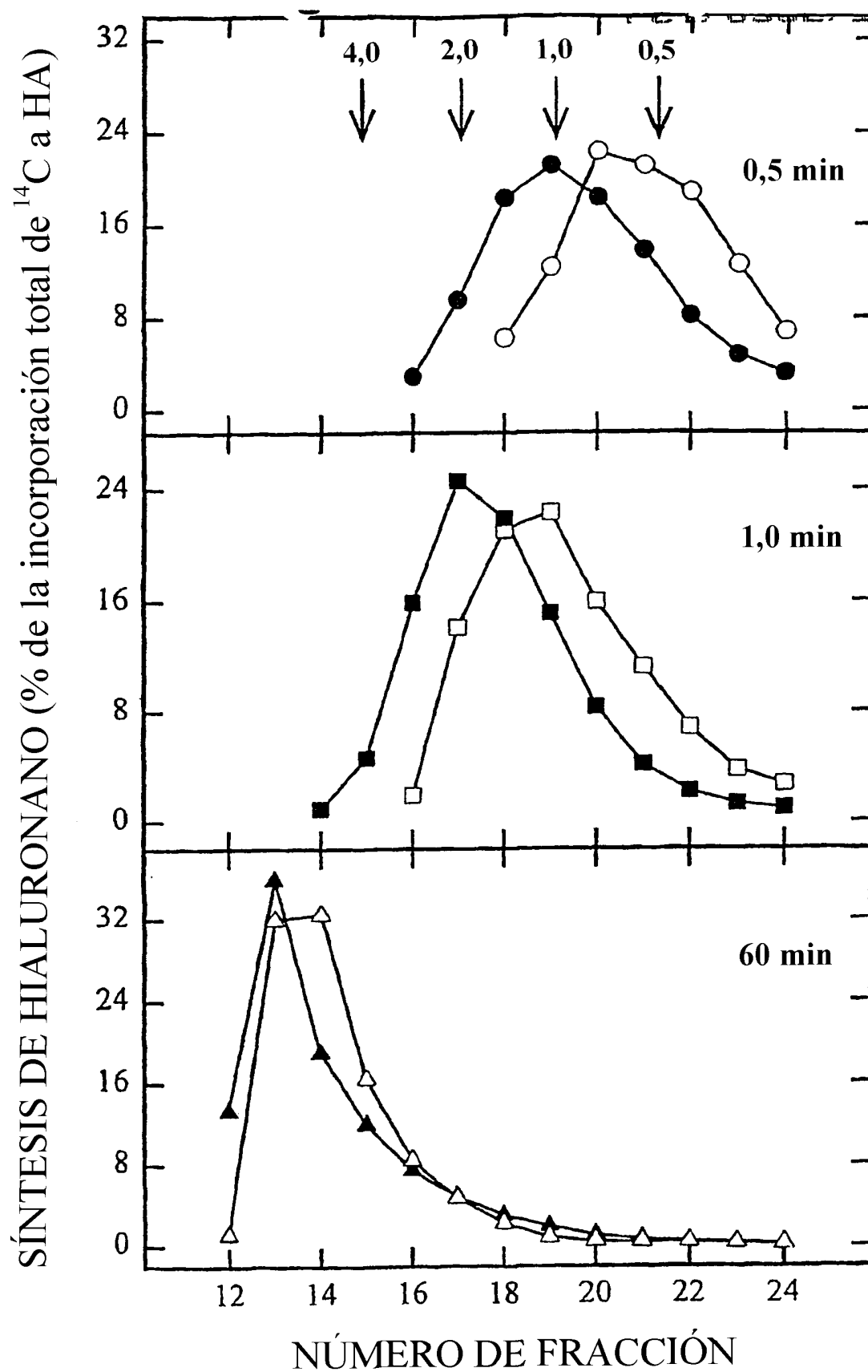


FIGURA 10

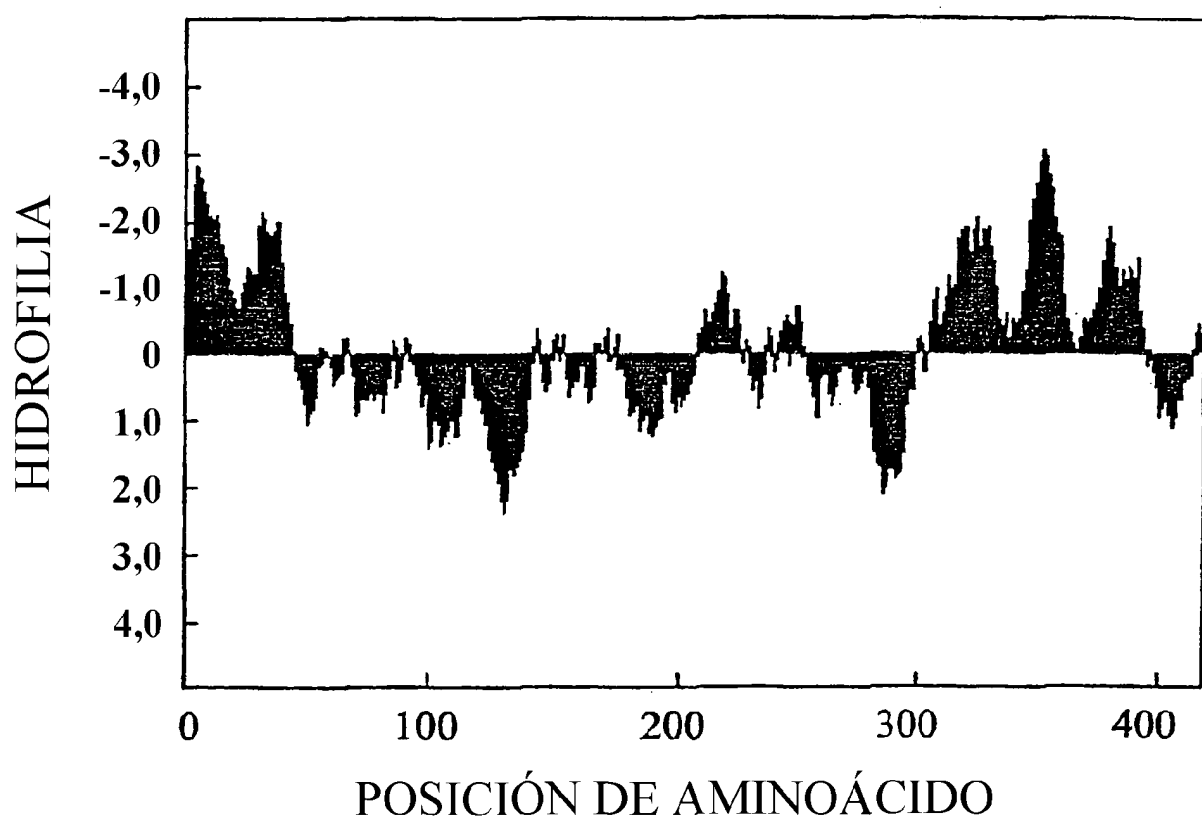


FIGURA 11

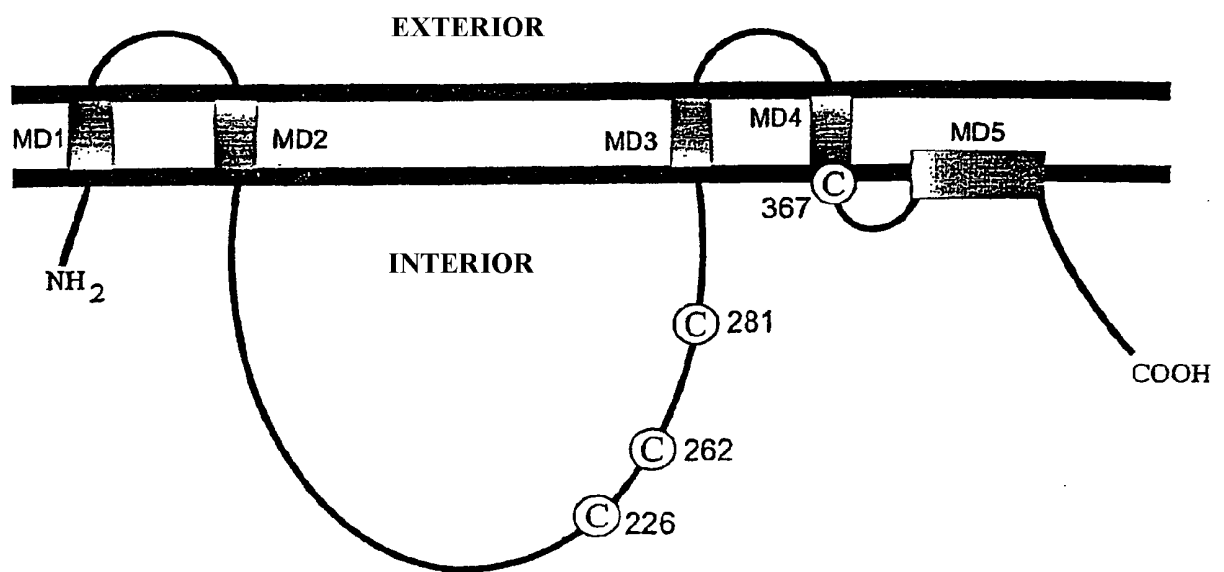


FIGURA 12

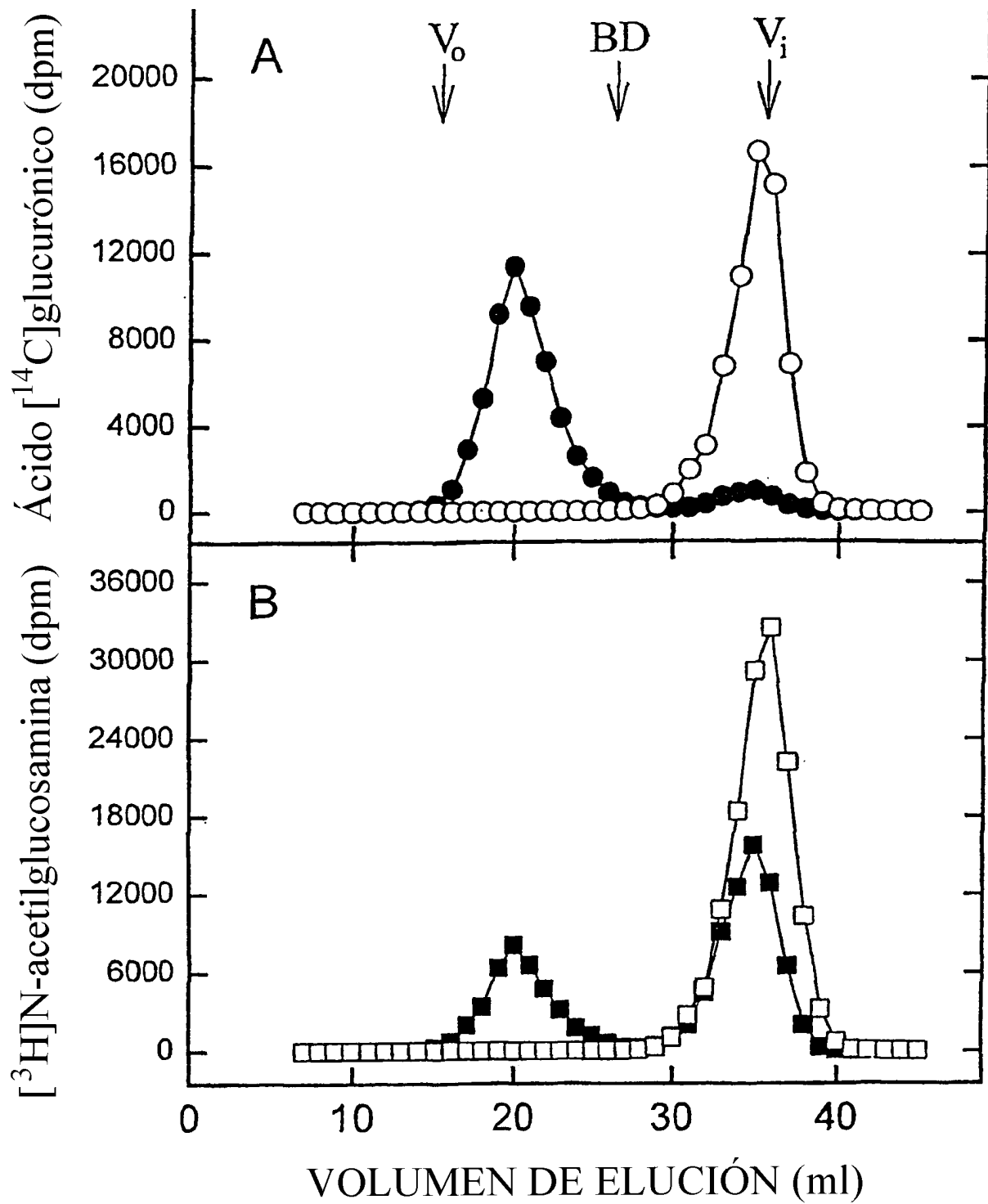


FIGURA 13

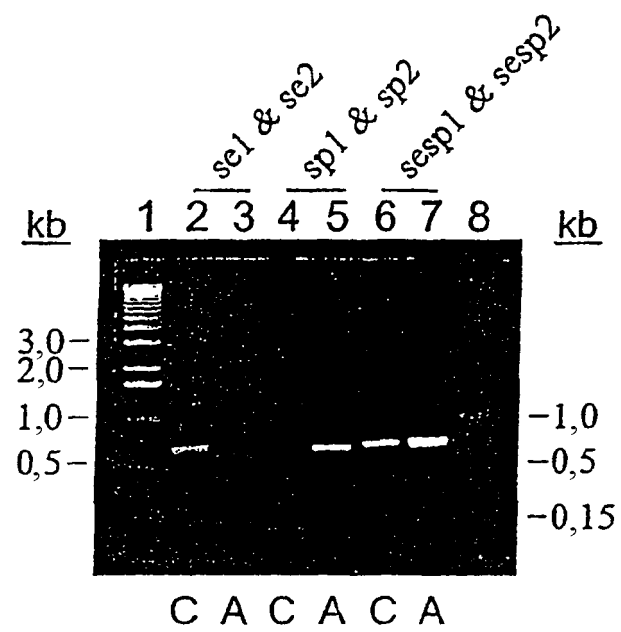


FIGURA 14

IDENTIFICACIÓN DE SECUENCIA N° 3

5' - G C T G A T G A G A C A G G T A T T A A G C

cebador: se1 (sentido, para los nucleótidos G³¹⁶ - C³³⁷)

IDENTIFICACIÓN DE SECUENCIA N° 4

5' - A T C A A A T T C T C T G A C A T T G C

cebador: se2 (antisentido, para los nucleótidos sentido G¹⁰³¹ - T¹⁰⁵⁰)

IDENTIFICACIÓN DE SECUENCIA N° 5

5' - G A C T C A G A T A C T T A T A T C T A

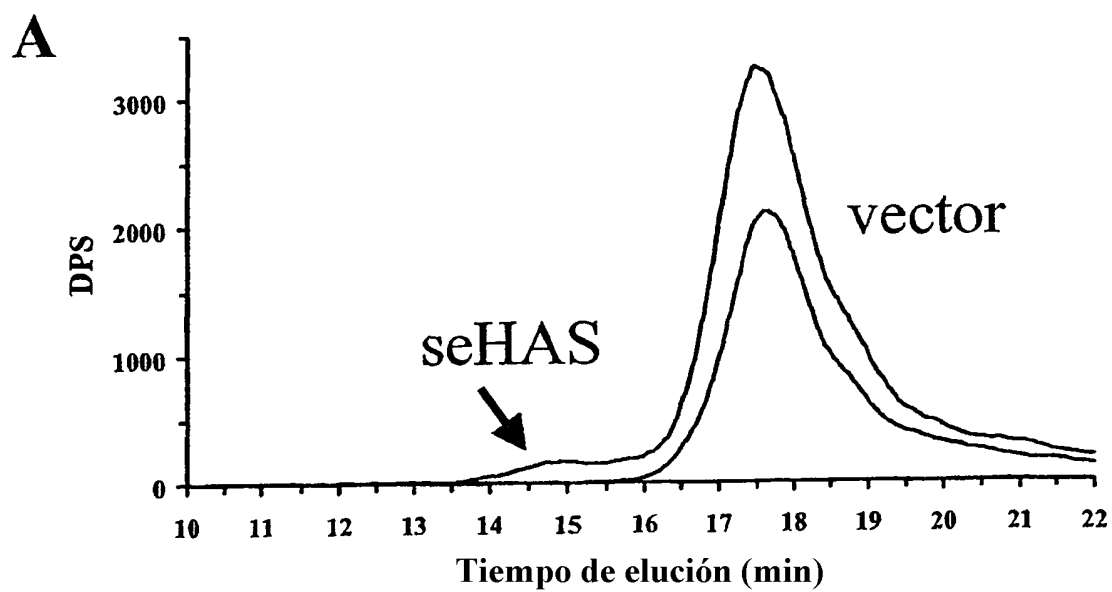
cebador: sesp1 (sentido, para los nucleótidos G⁴⁷⁵ - A⁴⁹⁴)

IDENTIFICACIÓN DE SECUENCIA N° 6

5' - T T T T T A C G T G T T C C C C A

cebador: sesp2 (antisentido, para los nucleótidos sentido T¹²²⁸ - A¹²⁴⁴)

FIGURA 15



“B”

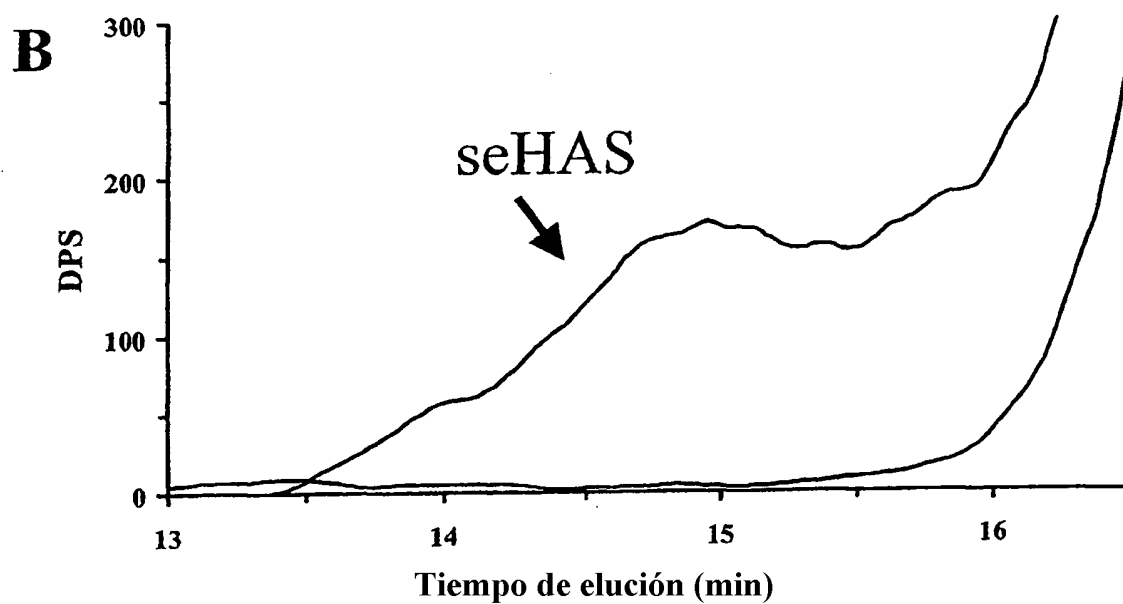


FIGURA 16

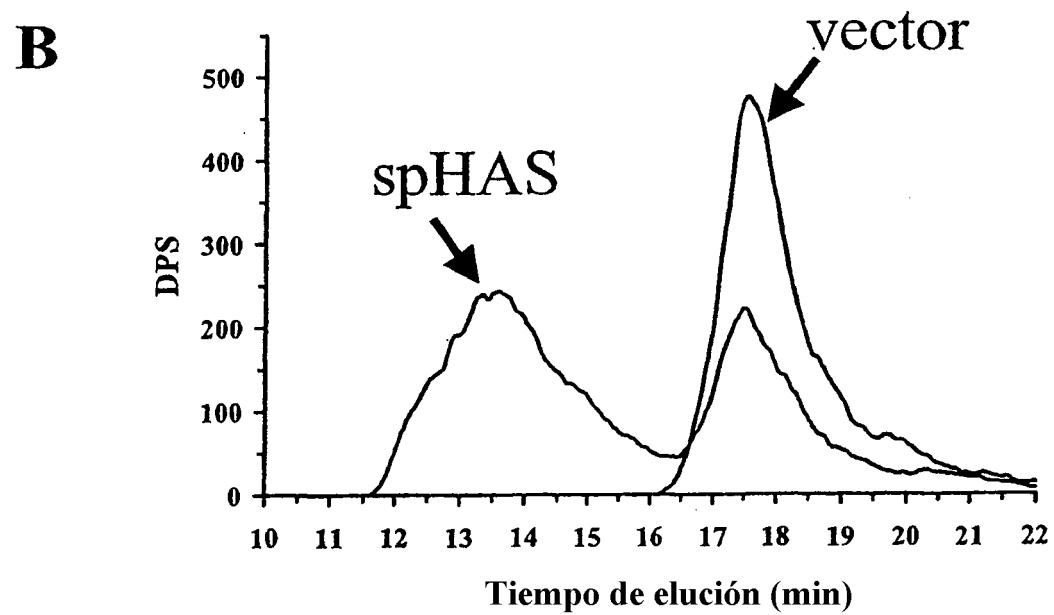
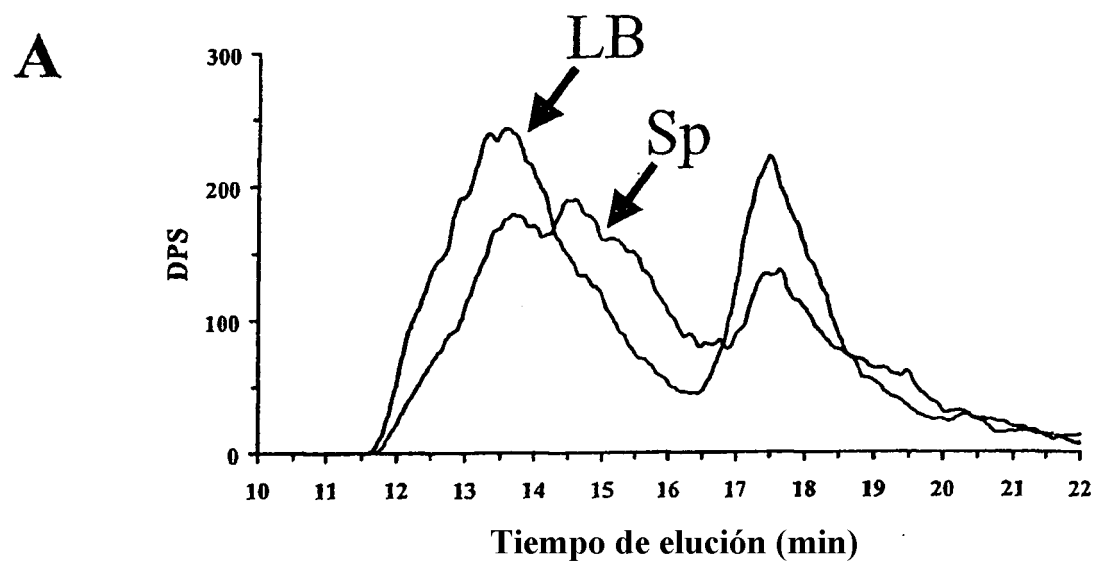


FIGURA 17

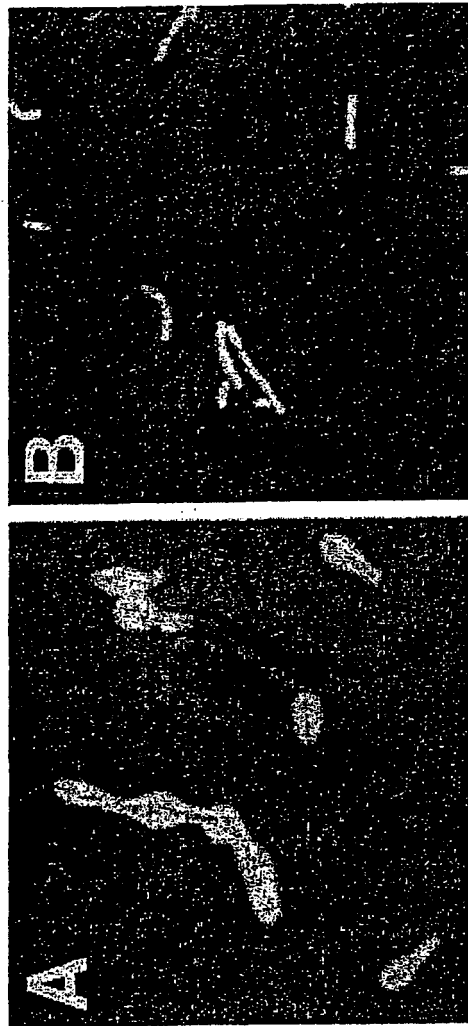


FIGURA 18

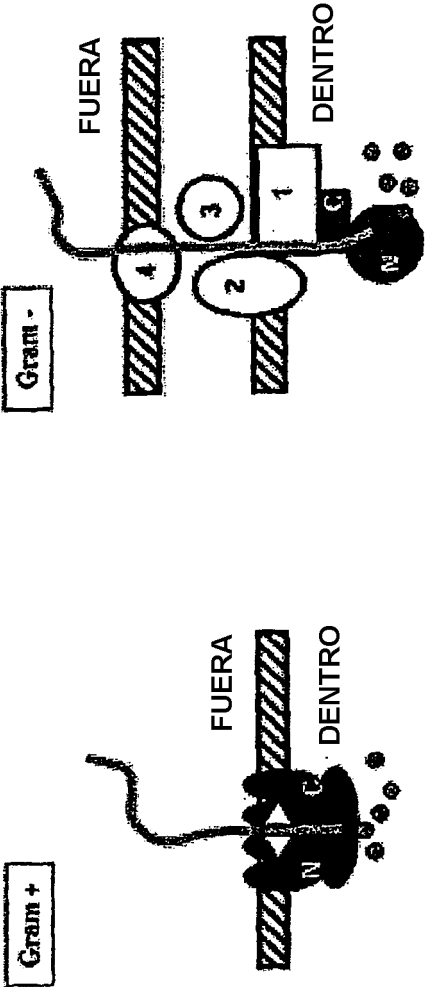
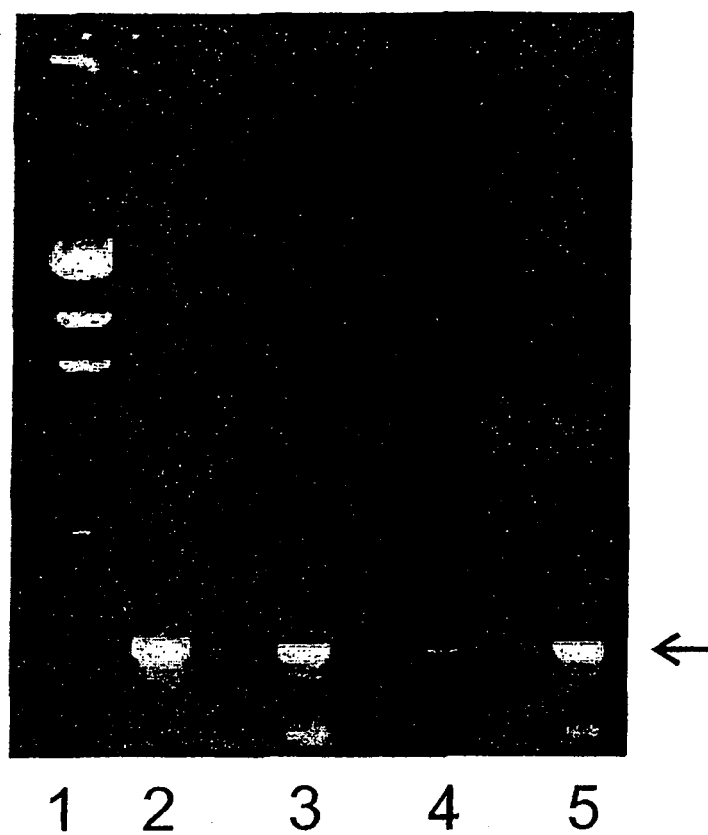
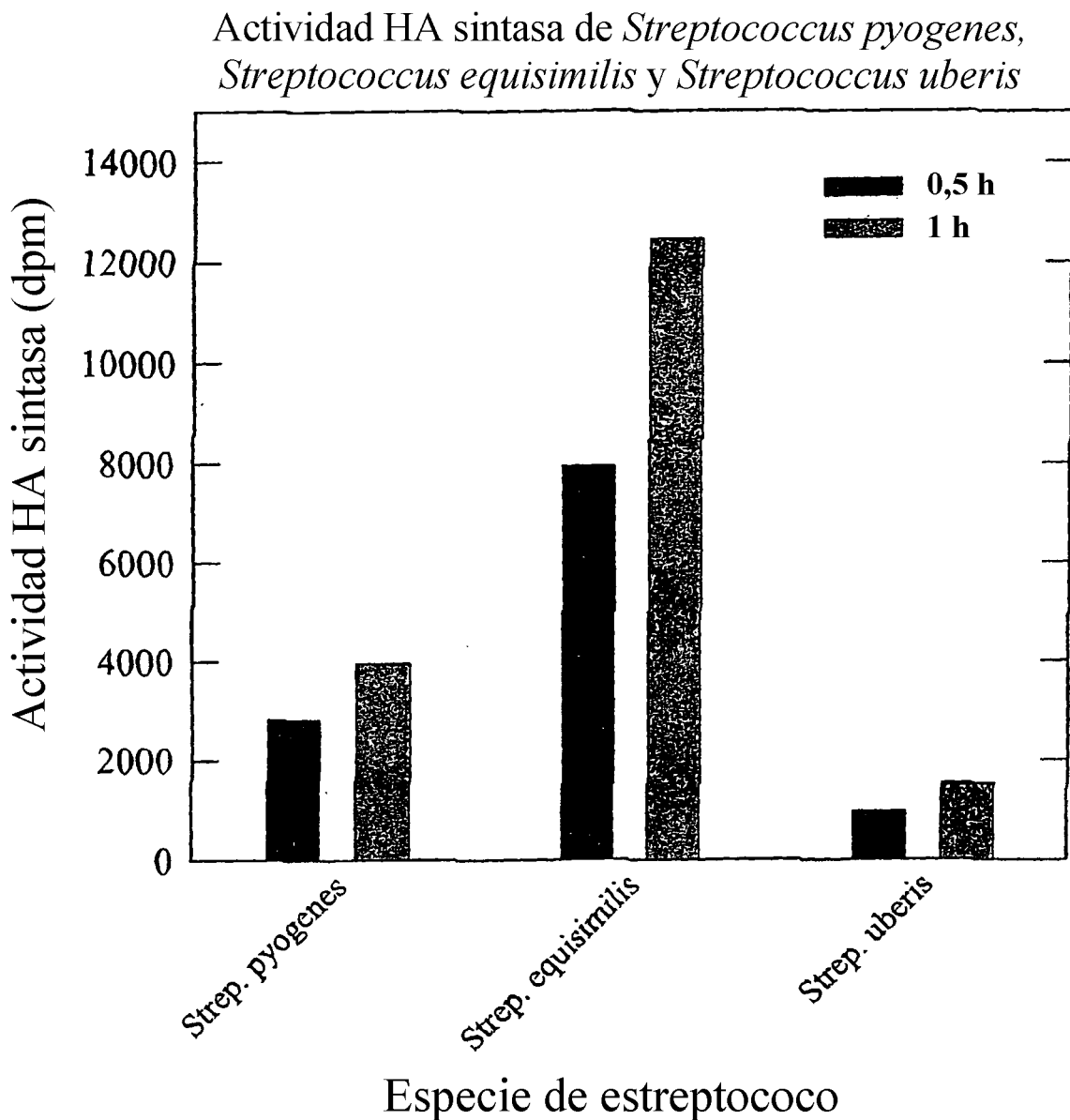


FIGURA 19



Se preparó DNA genómico a partir de cuatro colonias mucoides distintas de *S. uberis* (carriles 2-5) y se sometió a PCR usando cebadores diseñados para multiplicar la región de codificación completa. El cebador 5' (directo), que contiene sitios para enzimas de restricción 5' *Bam*HI y *Pst*II (cursiva), era: AGGATCCGAATTC ATG GAA AAA CTA AAA AAT CTC. El cebador 3' (inverso), que contiene sitios 5' *Eco*RI y *Pst*II (cursiva), era: AGAATTCTGCAG TTA TTT ACT TGT CTT TTT ACG. Los productos de reacción fueron analizados mediante electroforesis en gel de agarosa y teñidos con bromuro de etidio, y se preparó una imagen digital. En cada muestra, era evidente una banda en el tamaño esperado de ~ 1,25 kb (indicada por la flecha) que correspondía al marco de lectura completo de suHAS más los sitios de restricción añadidos. El carril 1 contiene un producto de digestión de DNA de Lambda con *Hind*III, como un conjunto de patrones o marcadores.

FIGURA 20



La actividad de síntesis de HA fue medida (para ~ 5-12 µg de proteína de membrana) mediante la incorporación de [^{14}C]UDP-GlcUA a HA utilizando un ensayo por cromatografía en papel.

FIGURA 21

ES 2 346 196 T3

LISTA DE SECUENCIAS

<110> Weigel, Paul H.
Kumari, Kshama
DeAngelis, Paul

5

<120> GENES DE HIALURONANO SINTASA Y EXPRESIÓN DE LOS MISMOS EN *BACILLUS SUBTILIS*

10 <130> 3554.078wo

<150> 60/297.744

<151> 2001-06-13

15

<150> 60/297.788

<151> 2001-06-13

20 <150> 09/469.200

<151> 1999-12-21

<150> 09/178.851

25 <151> 1998-10-26

<160> 20

30 <170> PatentIn, versión 3.1

<210> 1

<211> 1254

35 <212> DNA

<213> *Streptococcus equisimilis*

<400> 1

40

atgagaacat taaaaaacct cataactggt gtggccttta gtatTTTTTtg ggtactgttg
60

45

atttacgtca atgtttatct ctttggtgct aaaggaagct tgtcaattta tggcTTTTTtg
120

50

ctgatagctt acctattagt caaaatgtcc ttatcctttt ttacaagcc atttaaggga
180

55

agggctgggc aatataaggt tgcagccatt attcctctt ataacgaaga tgctgagtca
240

ttgctagaga ccttaaaaag tggtcagcag caaacctatc ccctagcaga aatttatgtt
300

60

65

ES 2 346 196 T3

gttgacgatg gaagtgctga tgagacaggt attaagcgca ttgaagacta tgtgcgtgac
 360
 5 actggtgacc tatcaagcaa tgtcattggt catcggtcag agaaaaatca aggaaagcgt
 420
 catgcacagg cctgggcctt tgaaagatca gacgctgatg tctttttgac cgttgactca
 10 480
 gatacttata tctaccctga tgctttagag gagttgttaa aaacctttaa tgaccaact
 540
 15 gtttttgctg cgacgggtca ccttaatgtc agaaatagac aaaccaatct cttaacacgc
 600
 20 ttgacagata ttcgctatga taatgctttt ggcgttgaac gagctgcca atccgttaca
 660
 ggtaatatcc ttgtttgctc aggtccgctt agcgtttaca gacgcgaggt ggttgttctt
 25 720
 aacatagata gatacatcaa ccagacctc ctgggtattc ctgtaagtat tggatgatgac
 780
 30 aggtgcttga ccaactatgc aactgattta ggaaagactg tttatcaatc cactgctaaa
 840
 35 tgtattacag atgttcctga caagatgtct acttacttga agcagcaaaa ccgctggaac
 900
 aagtccttct ttagagagtc cattatttct gttaagaaaa tcatgaacaa tccttttgta
 960
 40 gccctatgga ccatacttga ggtgtctatg tttatgatgc ttgtttattc tgtggtggat
 1020
 45 ttctttgtag gcaatgtcag agaatttgat tggtcaggg ttttagcctt tctggtgatt
 1080
 atcttcattg ttgccctgtg tcggaacatt cattacatgc ttaagcaccg gctgtccttc
 50 1140
 ttgttatctc cgttttatgg ggtgctgcat ttgtttgtcc tacagccctt gaaattatat
 1200
 55 tctcttttta ctattagaaa tgctgactgg ggaacacgta aaaaattatt ataa
 1254
 60 <210> 2
 <211> 417
 <212> PRT
 65 <213> *Streptococcus equisimilis*

ES 2 346 196 T3

<400> 2

```

5  Met Arg Thr Leu Lys Asn Leu Ile Thr Val Val Ala Phe Ser Ile Phe
   1           5           10           15

10 Trp Val Leu Leu Ile Tyr Val Asn Val Tyr Leu Phe Gly Ala Lys Gly
   20           25           30

15 Ser Leu Ser Ile Tyr Gly Phe Leu Leu Ile Ala Tyr Leu Leu Val Lys
   35           40           45

20 Met Ser Leu Ser Phe Phe Tyr Lys Pro Phe Lys Gly Arg Ala Gly Gln
   50           55           60

25 Tyr Lys Val Ala Ala Ile Ile Pro Ser Tyr Asn Glu Asp Ala Glu Ser
   65           70           75           80

30 Leu Leu Glu Thr Leu Lys Ser Val Gln Gln Gln Thr Tyr Pro Leu Ala
   85           90           95

35 Glu Ile Tyr Val Val Asp Asp Gly Ser Ala Asp Glu Thr Gly Ile Lys
   100          105          110

40 Arg Ile Glu Asp Tyr Val Arg Asp Thr Gly Asp Leu Ser Ser Asn Val
   115          120          125

45 Ile Val His Arg Ser Glu Lys Asn Gln Gly Lys Arg His Ala Gln Ala
   130          135          140

50 Trp Ala Phe Glu Arg Ser Asp Ala Asp Val Phe Leu Thr Val Asp Ser
   145          150          155          160

```

55

60

65

ES 2 346 196 T3

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

ES 2 346 196 T3

Arg Val Leu Ala Phe Leu Val Ile Ile Phe Ile Val Ala Leu Cys Arg
355 360 365

5

Asn Ile His Tyr Met Leu Lys His Pro Leu Ser Phe Leu Leu Ser Pro
370 375 380

10

Phe Tyr Gly Val Leu His Leu Phe Val Leu Gln Pro Leu Lys Leu Tyr
385 390 395 400

15

Ser Leu Phe Thr Ile Arg Asn Ala Asp Trp Gly Thr Arg Lys Lys Leu
405 410 415

20

Leu

25

<210> 3

<211> 22

<212> DNA

30

<213> *Streptococcus equisimilis*

<400> 3

35

gctgatgaga caggtattaa gc
22

<210> 4

40

<211> 20

<212> DNA

<213> *Streptococcus equisimilis*

45

<400> 4

atcaaattct ctgacattgc
20

50

<210> 5

<211> 20

55

<212> DNA

<213> *Streptococcus equisimilis*

<400> 5

60

gactcagata cttatatcta
20

65

<210> 6

<211> 17

ES 2 346 196 T3

<212> DNA

<213> *Streptococcus equisimilis*

5 <400> 6

tttttacgtg ttcccca
17

10

<210> 7

<211> 567

15 <212> PRT

<213> Virus PBCV-1 de Chlorella

<400> 7

20

Met Gly Lys Asn Ile Ile Ile Met Val Ser Trp Tyr Thr Ile Ile Thr
1 5 10 15

25

Ser Asn Leu Ile Ala Val Gly Gly Ala Ser Leu Ile Leu Ala Pro Ala
20 25 30

30

Ile Thr Gly Tyr Val Leu His Trp Asn Ile Ala Leu Ser Thr Ile Trp
35 40 45

35

Gly Val Ser Ala Tyr Gly Ile Phe Val Phe Gly Phe Phe Leu Ala Gln
50 55 60

40

Val Leu Phe Ser Glu Leu Asn Arg Lys Arg Leu Arg Lys Trp Ile Ser
65 70 75 80

45

Leu Arg Pro Lys Gly Trp Asn Asp Val Arg Leu Ala Val Ile Ile Ala
85 90 95

50

Gly Tyr Arg Glu Asp Pro Tyr Met Phe Gln Lys Cys Leu Glu Ser Val
100 105 110

55

Arg Asp Ser Asp Tyr Gly Asn Val Ala Arg Leu Ile Cys Val Ile Asp

60

65

ES 2 346 196 T3

115

120

125

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

65

ES 2 346 196 T3

	305		310				315		320							
5	Lys	Val	Val	Phe	Thr	Pro	Phe	Ala	Val	Gly	Trp	Ser	Asp	Ser	Pro	Thr
					325					330					335	
10	Asn	Val	Phe	Arg	Tyr	Ile	Val	Gln	Gln	Thr	Arg	Trp	Ser	Lys	Ser	Trp
				340					345					350		
15	Cys	Arg	Glu	Ile	Trp	Tyr	Thr	Leu	Phe	Ala	Ala	Trp	Lys	His	Gly	Leu
			355					360					365			
20	Ser	Gly	Ile	Trp	Leu	Ala	Phe	Glu	Cys	Leu	Tyr	Gln	Ile	Thr	Tyr	Phe
		370					375					380				
25	Phe	Leu	Val	Ile	Tyr	Leu	Phe	Ser	Arg	Leu	Ala	Val	Glu	Ala	Asp	Pro
	385					390					395					400
30	Arg	Ala	Gln	Thr	Ala	Thr	Val	Ile	Val	Ser	Thr	Thr	Val	Ala	Leu	Ile
					405					410					415	
35	Lys	Cys	Gly	Tyr	Phe	Ser	Phe	Arg	Ala	Lys	Asp	Ile	Arg	Ala	Phe	Tyr
			420						425					430		
40	Phe	Val	Leu	Tyr	Thr	Phe	Val	Tyr	Phe	Phe	Cys	Met	Ile	Pro	Ala	Arg
			435					440					445			
45	Ile	Thr	Ala	Met	Met	Thr	Leu	Trp	Asp	Ile	Gly	Trp	Asp	Thr	Arg	Gly
	450						455					460				
50	Gly	Asn	Glu	Lys	Pro	Ser	Val	Gly	Thr	Arg	Val	Ala	Leu	Trp	Ala	Lys
	465					470					475				480	
55	Gln	Tyr	Leu	Ile	Ala	Tyr	Met	Trp	Trp	Ala	Ala	Val	Val	Gly	Ala	Gly
					485					490					495	
60	Val	Tyr	Ser	Ile	Val	His	Asn	Trp	Met	Phe	Asp	Trp	Asn	Ser	Leu	Ser

65

ES 2 346 196 T3

500

505

510

5 Tyr Arg Phe Ala Leu Val Gly Ile Cys Ser Tyr Ile Val Phe Ile Val
515 520 525

10 Ile Val Leu Val Val Tyr Phe Thr Gly Lys Ile Thr Thr Trp Asn Phe
530 535 540

15 Thr Lys Leu Gln Lys Glu Leu Ile Glu Asp Arg Val Leu Tyr Asp Ala
545 550 555 560

20 Thr Thr Asn Ala Gln Ser Val
565

25 <210> 8
<211> 1740
<212> DNA
30 <213> Virus PBCV-1 de Chlorella

<400> 8

35 aagacttctt gaaagttaca atgggtaaaa atataatcat aatgggtttcg tggtagacca
60

40 tcataacttc aaatctaate gcgggttgag gagcctctct aatcttggct ccggcaatta
120

ctgggtatgt tctacattgg aatattgctc tctcgacaat ctggggagta tcagcttatg
180

45 gtattttcgt ttttgggttt ttccttgcac aagttttatt ttcagaactg aacaggaaac
240

50 gtcttcgcaa gtggatttct ctcagacctt aggggttgaa tgatgttcgt ttggctgtga
300

55 tcattgctgg atatcgcgag gatccttata tgttccagaa gtgcctcgag tctgtacgtg
360

actctgatta tggcaacgtt gcccgctctga tttgtgtgat tgacgggtgat gaggacgatg
420

60 atatgaggat ggctgccgtt tacaaggcga tctacaatga taatatcaag aagcccgagt
480

65

ES 2 346 196 T3

ttgttctgtg tgagtcagac gacaaggaag gtgaacgcat cgactctgat ttctctcgcg
 540
 5 acatttgtgt cctccagcct catcgtggaa aacgggagtg tctttatact gggtttcaac
 600
 10 ttgcaaagat ggaccccagt gtcaatgctg tcgttctgat tgacagcgat accgttctcg
 660
 agaaggatgc tattctggaa gttgtatacc cacttgcatt cgatcccagag atccaagccg
 720
 15 ttgcaggtga gtgtaagatt tggaaacacag acactctttt gactcttctc gtcgcttggc
 780
 20 ggtactattc tgcgttttgt gtggagagga gtgccagtc ttttttcagg actgttcagt
 840
 gcgttggggg gccactgggt gcctacaaga ttgatattcat taaggagatt aaggaccct
 900
 25 ggatttccca gcgctttctt ggtcagaagt gtacttacgg tgacgaccgc cggctaacca
 960
 30 acgagatctt gatgcgtggt aaaaagggtg tgttactcc atttgctgtt ggttggctcg
 1020
 acagtccgac caatgtgttt cggtacatcg ttcagcagac ccgctggagt aagtcgtggt
 1080
 35 gccgcgaaat ttggtacacc ctcttcgccg cgtggaagca cggtttgtct ggaatttggc
 1140
 40 tggcctttga atgtttgtat caaattacat acttcttctt cgtgatttac ctcttttctc
 1200
 45 gcctagccgt tgaggccgac cctcgcgcc agacagccac ggtgattgtg agcaccacgg
 1260
 ttgcattgat taagtgtggg tatttttcat tccgagccaa ggatattcgg gcgttttact
 1320
 50 ttgtgcttta tacatttgtt tactttttct gtatgattcc ggccaggatt actgcaatga
 1380
 55 tgacgctttg ggacattggc tgggatactc gcggtggaaa cgagaagcct tccgttggca
 1440
 60
 65

ES 2 346 196 T3

cccggttcgc tctgtgggca aagcaatatc tcattgcata tatgtggtgg gccgcggttg
1500

5 ttggcgctgg agtttacagc atcgctccata actggatggt cgattggaat tctctttctt
1560

10 atcgttttgc tttggttggg atttggtctt acattgtttt tattgttatt gtgctggtgg
1620

tttatttcac cggcaaaatt acgacttggg atttcacgaa gcttcagaag gagctaatac
1680

15 aggatcgctg tctgtacgat gcaactacca atgctcagtc tgtgtgattt ttcttgcaag
1740

20

<210> 9

<211> 2937

25 <212> DNA

<213> *Pasteurella multocida*

<400> 9

30

attttttaag gacagaaaat gaatacatta tcacaagcaa taaaagcata taacagcaat
60

35

gactatcaat tagcactcaa attatttgaa aagtcggcgg aaatctatgg acgaaaaatt
120

40 gttgaatttc aaattaccaa atgccaagaa aaactctcag cacatccttc tgtaattca
180

45 gcacatcttt ctgtaaataa agaagaaaaa gtcaatgttt gcgatagtcc gttagatatt
240

gcaacacaac tgttactttc caacgtaaaa aaattagtagc tttctgactc ggaaaaaac
300

50

acgttaaaaa ataaatggaa attgctcact gagaagaaat ctgaaaatgc ggaggtaaga
360

55 gcggtcgccc ttgtaccaa agattttccc aaagatctgg ttttagcgcc tttacctgat
420

catgttaatg attttacatg gtacaaaaag cgaaagaaaa gacttggcat aaaacctgaa
480

60

catcaacatg ttggtctttc tattatcggt acaacattca atcgaccagc aattttatcg

65

ES 2 346 196 T3

540

5 attacattag cctgtttagt aaaccaaaaa acacattacc cgtttgaagt tatcgtgaca
600

10 gatgatggta gtcaggaaga tctatcaccg atcattcgcc aatatgaaaa taaattggat
660

15 attcgctacg tcagacaaaa agataacggt tttcaagcca gtgccgctcg gaatatggga
720

20 ttacgcttag caaaatatga ctttattggc ttactcgact gtgatatggc gccaaatcca
780

25 ttatgggttc attcttatgt tgcagagcta ttagaagatg atgatttaac aatcattggg
840

30 ccaagaaaat acatcgatac acaacatatt gacccaaaag acttcttaaa taacgcgagt
900

35 ttgcttgaat cattaccaga agtgaaaacc aataatagtg ttgccgcaaa aggggaagga
960

40 acagtttctc tggattggcg cttagaacaa ttcgaaaaaa cagaaaatct ccgcttatcc
1020

45 gattcgccctt tccgtttttt tgcggcgggt aatgttgctt tcgctaaaaa atggctaaat
1080

50 aaatccgggt tctttgatga ggaatttaat cactgggggtg gagaagatgt ggaatttgga
1140

55 tatcgcttat tccgttacgg tagtttcttt aaaactattg atggcattat ggcctaccat
1200

60 caagagccac caggtaaaga aaatgaaacc gatcgtgaag cgggaaaaaa tattacgctc
1260

65 gatattatga gagaaaaggt cccttatatc tatagaaaac ttttaccaat agaagattcg
1320

70 catatcaata gactaccttt agtttcaatt tatatcccag cttataactg tgcaaactat
1380

75 attcaacggt gcgtagatag tgcactgaat cagactgttg ttgatctcga ggtttgtatt
1440

80 tgtaacgatg gttcaacaga taatacctta gaagtgatca ataagcttta tggtataaat

ES 2 346 196 T3

1500

5 cctagggtac gcatcatgtc taaaccaa at ggcggaatag cctcagcatc aaatgcagcc
1560

10 gtttcttttg ctaaagggtta ttacattggg cagtttagatt cagatgatta tcttgagcct
1620

15 gatgcagttg aactgtgttt aaaagaattt ttaaaagata aaacgctagc ttgtgtttat
1680

20 accactaata gaaacgtcaa tccggatggg agcttaatcg ctaatgggtta caattggcca
1740

25 gaattttcac gagaaaaact cacaacggct atgattgctc accactttag aatgttcacg
1800

30 attagagctt ggcatttaac tgatggattc aatgaaaaaa ttgaaatgc cgtagactat
1860

35 gacatgttcc tcaaactcag tgaagttgga aaatttaaac atcttaataa aatctgctat
1920

40 aaccgtgtat tacatgggtga taacacatca attaa aac ttggcattca aaagaaaaac
1980

45 cattttgttg tagtcaatca gtcattaaat agacaaggca taacttatta taattatgac
2040

50 gaatttgatg atttagatga aagtagaaag tatattttca ataaaaccgc tgaatatcaa
2100

55 gaagagattg atatcttaaa agatattaaa atcatccaga ataaagatgc caaatcgca
2160

60 gtcagtattt tttatcccaa tacattaaac ggcttagtga aaaaactaaa caatattatt
2220

65 gaatataata aaaatatatt cgttattggt ctacatgttg ataagaatca tcttacacca
2280

70 gatatcaaaa aagaaatact agccttctat cataaacatc aagtgaatat ttactaaat
2340

75 aatgatattt catattacac gagtaataga ttaataaaaa ctgaggcgca ttttaagtaat
2400

80 attaataaat taagtcagtt aaatctaaat tgtgaataca tcatttttga taatcatgac

ES 2 346 196 T3

2460

5 agcctattcg ttaaaaatga cagctatgct tatatgaaaa aatatgatgt cggcatgaat
2520

ttctcagcat taacacatga ttggatcgag aaaatcaatg cgcattccacc atttaaaaag
2580

10 ctcatataaaa cttatttttaa tgacaatgac ttaaaaagta tgaatgtgaa aggggcatca
2640

15 caaggtatgt ttatgacgta tgcgctagcg catgagcttc tgacgattat taaagaagtc
2700

20 atcacatctt gccagtcaat tgatagtgtg ccagaatata aactgagga tatttggttc
2760

caatttgac ttttaatctt agaaaagaaa accggccatg tatttaataa aacatcgacc
2820

25 ctgacttata tgccttggga acgaaaatta caatggacaa atgaacaaat tgaaagtgca
2880

30 aaaagaggag aaaatatacc tgtaacaag ttcattatta atagtataac tctataa
2937

35 <210> 10

<211> 972

<212> PRT

<213> *Pasteurella multocida*

40 <400> 10

45 Met Asn Thr Leu Ser Gln Ala Ile Lys Ala Tyr Asn Ser Asn Asp Tyr
1 5 10 15

50 Gln Leu Ala Leu Lys Leu Phe Glu Lys Ser Ala Glu Ile Tyr Gly Arg
20 25 30

55 Lys Ile Val Glu Phe Gln Ile Thr Lys Cys Gln Glu Lys Leu Ser Ala
35 40 45

60 His Pro Ser Val Asn Ser Ala His Leu Ser Val Asn Lys Glu Glu Lys
50 55 60

65

ES 2 346 196 T3

Val Asn Val Cys Asp Ser Pro Leu Asp Ile Ala Thr Gln Leu Leu Leu
65 70 75 80

5

Ser Asn Val Lys Lys Leu Val Leu Ser Asp Ser Glu Lys Asn Thr Leu
85 90 95

10

Lys Asn Lys Trp Lys Leu Leu Thr Glu Lys Lys Ser Glu Asn Ala Glu
100 105 110

15

Val Arg Ala Val Ala Leu Val Pro Lys Asp Phe Pro Lys Asp Leu Val
115 120 125

20

Leu Ala Pro Leu Pro Asp His Val Asn Asp Phe Thr Trp Tyr Lys Lys
130 135 140

25

Arg Lys Lys Arg Leu Gly Ile Lys Pro Glu His Gln His Val Gly Leu
145 150 155 160

30

Ser Ile Ile Val Thr Thr Phe Asn Arg Pro Ala Ile Leu Ser Ile Thr
165 170 175

35

Leu Ala Cys Leu Val Asn Gln Lys Thr His Tyr Pro Phe Glu Val Ile
180 185 190

40

Val Thr Asp Asp Gly Ser Gln Glu Asp Leu Ser Pro Ile Ile Arg Gln
195 200 205

45

Tyr Glu Asn Lys Leu Asp Ile Arg Tyr Val Arg Gln Lys Asp Asn Gly
210 215 220

50

Phe Gln Ala Ser Ala Ala Arg Asn Met Gly Leu Arg Leu Ala Lys Tyr
225 230 235 240

55

Asp Phe Ile Gly Leu Leu Asp Cys Asp Met Ala Pro Asn Pro Leu Trp
245 250 255

60

65

ES 2 346 196 T3

Val His Ser Tyr Val Ala Glu Leu Leu Glu Asp Asp Asp Leu Thr Ile
260 265 270

5

Ile Gly Pro Arg Lys Tyr Ile Asp Thr Gln His Ile Asp Pro Lys Asp
275 280 285

10

Phe Leu Asn Asn Ala Ser Leu Leu Glu Ser Leu Pro Glu Val Lys Thr
290 295 300

15

Asn Asn Ser Val Ala Ala Lys Gly Glu Gly Thr Val Ser Leu Asp Trp
305 310 315 320

20

Arg Leu Glu Gln Phe Glu Lys Thr Glu Asn Leu Arg Leu Ser Asp Ser
325 330 335

25

Pro Phe Arg Phe Phe Ala Ala Gly Asn Val Ala Phe Ala Lys Lys Trp
340 345 350

30

Leu Asn Lys Ser Gly Phe Phe Asp Glu Glu Phe Asn His Trp Gly Gly
355 360 365

35

Glu Asp Val Glu Phe Gly Tyr Arg Leu Phe Arg Tyr Gly Ser Phe Phe
370 375 380

40

Lys Thr Ile Asp Gly Ile Met Ala Tyr His Gln Glu Pro Pro Gly Lys
385 390 395 400

45

Glu Asn Glu Thr Asp Arg Glu Ala Gly Lys Asn Ile Thr Leu Asp Ile
405 410 415

50

Met Arg Glu Lys Val Pro Tyr Ile Tyr Arg Lys Leu Leu Pro Ile Glu
420 425 430

55

Asp Ser His Ile Asn Arg Val Pro Leu Val Ser Ile Tyr Ile Pro Ala
435 440 445

60

65

ES 2 346 196 T3

Tyr Asn Cys Ala Asn Tyr Ile Gln Arg Cys Val Asp Ser Ala Leu Asn
 450 455 460
 5
 Gln Thr Val Val Asp Leu Glu Val Cys Ile Cys Asn Asp Gly Ser Thr
 465 470 475 480
 10
 Asp Asn Thr Leu Glu Val Ile Asn Lys Leu Tyr Gly Asn Asn Pro Arg
 485 490 495
 15
 Val Arg Ile Met Ser Lys Pro Asn Gly Gly Ile Ala Ser Ala Ser Asn
 500 505 510
 20
 Ala Ala Val Ser Phe Ala Lys Gly Tyr Tyr Ile Gly Gln Leu Asp Ser
 515 520 525
 25
 Asp Asp Tyr Leu Glu Pro Asp Ala Val Glu Leu Cys Leu Lys Glu Phe
 530 535 540
 30
 Leu Lys Asp Lys Thr Leu Ala Cys Val Tyr Thr Thr Asn Arg Asn Val
 545 550 555 560
 35
 Asn Pro Asp Gly Ser Leu Ile Ala Asn Gly Tyr Asn Trp Pro Glu Phe
 565 570 575
 40
 Ser Arg Glu Lys Leu Thr Thr Ala Met Ile Ala His His Phe Arg Met
 580 585 590
 45
 Phe Thr Ile Arg Ala Trp His Leu Thr Asp Gly Phe Asn Glu Lys Ile
 595 600 605
 50
 Glu Asn Ala Val Asp Tyr Asp Met Phe Leu Lys Leu Ser Glu Val Gly
 610 615 620
 55
 Lys Phe Lys His Leu Asn Lys Ile Cys Tyr Asn Arg Val Leu His Gly
 625 630 635 640
 60
 65

ES 2 346 196 T3

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

ES 2 346 196 T3

Met Asn Phe Ser Ala Leu Thr His Asp Trp Ile Glu Lys Ile Asn Ala
835 840 845

5

His Pro Pro Phe Lys Lys Leu Ile Lys Thr Tyr Phe Asn Asp Asn Asp
850 855 860

10

Leu Lys Ser Met Asn Val Lys Gly Ala Ser Gln Gly Met Phe Met Thr
865 870 875 880

15

Tyr Ala Leu Ala His Glu Leu Leu Thr Ile Ile Lys Glu Val Ile Thr
885 890 895

20

Ser Cys Gln Ser Ile Asp Ser Val Pro Glu Tyr Asn Thr Glu Asp Ile
900 905 910

25

Trp Phe Gln Phe Ala Leu Leu Ile Leu Glu Lys Lys Thr Gly His Val
915 920 925

30

Phe Asn Lys Thr Ser Thr Leu Thr Tyr Met Pro Trp Glu Arg Lys Leu
930 935 940

35

Gln Trp Thr Asn Glu Gln Ile Glu Ser Ala Lys Arg Gly Glu Asn Ile
945 950 955 960

40

Pro Val Asn Lys Phe Ile Ile Asn Ser Ile Thr Leu
965 970

45

<210> 11

<211> 3466

<212> DNA

50

<213> *Streptococcus uberis*

<400> 11

55

gcaaaagttt aaaggaggaa ttatggaaaa actaaaaaat ctcattacat ttatgacttt
60

60

tattttcctg tggctcataa ttattgggct taatgttttt gtatttggaa ctaaaggaag

65

ES 2 346 196 T3

120

5 tctaacagtg tatgggatta ttctattaac ctatttgtcg ataaaaatgg gattatcttt
180

10 tttttatcgt ccctataaag gaagtgtagg tcaatataag gtagcagcta ttatcccatc
240

15 ttataatgag gatgggtgtcg gtttactaga aactctaaag agtggttcaaa aacaaacata
300

20 tccaattgca gaaatttttcg taattgacga tgggtcagta gataaaacag gtataaaatt
360

25 ggtcgaagac tatgtgaagt taaatggctt tggagaccaa gttatcggtc atcagatgcc
420

30 tgaaaatggt ggtaaaagac atgctcaggc ttgggcattt gaaaggctctg atgctgatgt
480

35 tttcttaaca gtggattcag atacctacat ctatcctgat gctcttgaag aattattaaa
540

40 gacattttaat gatccagagg tctacgctgc aactgggtcat ttaaattgcaa gaaatagaca
600

45 aactaatctc ttaactagac tgactgatat tcgttacgat aatgcatttg gtgtagaacg
660

50 tgctgctcag tctgttacgg gaaatatttt ggtttggtcc ggacctttaa gtatttatag
720

55 acgttccgtc ggtattccaa atcttgaacg ctatacctca caaacatttc ttggtgtccc
780

60 tgtaagcata ggggatgacc gttgtttgac aaattatgca actgatttgg gaaaaacggt
840

65 ttatcagtca actgcaagat gtgatactga cgttccagat aagtttaagg ttttcatcaa
900

70 acaacaaaat cgttggaata agtcattttt tagggagtct attatctctg ttaagaagtt
960

75 attagccaca ccaagtgttg ctgtttggac tattacagaa gtttccatgt tcatcatgct
1020

80 agttttattct atcttttagct tattgatagg agaggctcaa gaatttaatc tcataaaact

ES 2 346 196 T3

1080

5 ggttgctttt ttagttatta ttttcatagt agctctttgt agaaatgttc attacatggg
1140

taagcatcca tttgcttttt tattgtcacc gttttatgga ttgatacatc tattcgtttt
1200

10 gcaacctctt aagatatatt cgttatttac tataagaaat gctacatggg gaactcgtaa
1260

15 aaagacaagt aaataattca attagagaaa ggacaaaata gtgaaaattg cagttgcagg
1320

ttctggctat gttggcctat cattaagtgt attattagca cagaaaaatc ctgttacagt
20 1380

tgtagatatt attgagaaga aagtaaactc cataaatcaa aaacaatcac caatccagga
1440

25 tgttgatatt gaaaactatt taaaagaaaa aaagttacaa ttaagagcta ctctagacgc
1500

cgatcaagca tttagggatg cagatatact aattattgct acaccaacca attatgatgt
30 1560

ggagaagaat ttttttgata ctagtcatgt tgagactgta attgagaaag ctttagcttt
1620

35 aaatagtcag gctttgttag ttattaaatc aacgatacca cttgggtttta ttaaaaagat
1680

40 gcgtaaaaaa tatcagacag accgtattat ttttagtccc gaatttctta gagagtctaa
1740

agctttaaaa gataatcttt atcctagtcg aataattggt tcctttgaag atgatgattc
45 1800

tatggaagta atagaagcag caaagacttt tgctcaattg ttaaaagatg gttctttgga
1860

50 taaagatggt cctgtacttt ttatgggttc agcagaggct gaagcagtaa aattatttgc
1920

caatacctat ttagctatgc gtgtctccta ttttaatgag ttagatacat atgctgaaaa
55 1980

gaatggttta cgtgtggata atattattga gggcgtttgc catgatcgac gcataggaat

60

65

ES 2 346 196 T3

2040

5 tcattataat aacccttctt ttggctatgg aggatactgc ttacctaaag ataccaaaca
2100

10 gttgctagca ggctatgatg gtattcctca atcgcttata aaagcaattg ttgattctaa
2160

15 taaaattcgt aaagagtata tcgcatcaca aattttacaa caattgagtg atattaatgt
2220

20 agatcctaaa gatgcaacga ttggtattta ccgccttata atgaaaagta actctgataa
2280

25 tttcagagag agtgcaataa aagatattat tgatcatatt aagagctata aaattaatat
2340

30 agtcttgtat gagccaatga tgaatgaaga ttttgattta ccaatcattg atgatttata
2400

35 tgacttcaaa gccatgtcac atattatcgt ttcaaataga tatgatttag ccttagaaga
2460

40 tgttaaagaa aaagtttaca ccagagatat ttacggtgtg gattaagttt gatttttaac
2520

45 aaatctccaa aaaatagata aaaaaaacag actctgataa aagagtctgt tttttaaaag
2580

50 tgtgagcatc ctattgctag gatgctcagg aaatttatga aaagggagat aagagggaac
2640

55 ttatcttccc caacggtttg ggagaccatt atttaggata gtcttatcat aagctatcaa
2700

60 ccttaaagat ttcttaactc gttttcgttt gggctctgtc tttttaattt ttgatgaga
2760

65 attaaacttg atggaatgag aatcaggaca ctgcctatcc aggctgctgg attagctgaa
2820

70 gccacaccaa caaaaccaa gtataacagg ccaataatgg cgactcctgc tctcatgact
2880

75 aattccataa tgccagctaa agtaggaaca aaaccgtatc cgagaccttg aatgaaactt
2940

80 cttagtataa ataggatggc taaaatccaa taaagagagc cattaatcag ataatagaga

ES 2 346 196 T3

3000

5 taggctaaat ggaaaacagc tggatcagcc ttactaatga aaatgccaga gaaaaagcgg
3060

tgttggaaaa ttaacagaat agcaaaaaga acagaccaa taatacagat aatgagtga
3120

10 tctttaagac cctcaaggat tcttttataa gcttttagcgc catagttctg agctgtaaag
3180

15 gttgacaagg ctaagcccag atttaacatc ggtagcatgg ccagttgggc tgttttactg
3240

20 gcaatagcaa tagctgcgat agcttcgggc cccaacttat taatgggttac ctgcagtga
3300

atggctccaa tggctataat actagcctga aatgccatgg gaaaaccaag gcgagcatga
3360

25 tttctgagat tttccctatc aagagtcaaa tcgtctttct tcagtcggaa atgggggac
3420

30 tttttgttga tgtaaaggac caaatagagt acggagaaaag cttgca
3466

35 <210> 12

<211> 417

<212> PRT

40 <213> *Streptococcus uberis*

<400> 12

45 Met Glu Lys Leu Lys Asn Leu Ile Thr Phe Met Thr Phe Ile Phe Leu
1 5 10 15

50 Trp Leu Ile Ile Ile Gly Leu Asn Val Phe Val Phe Gly Thr Lys Gly
20 25 30

55 Ser Leu Thr Val Tyr Gly Ile Ile Leu Leu Thr Tyr Leu Ser Ile Lys
35 40 45

60 Met Gly Leu Ser Phe Phe Tyr Arg Pro Tyr Lys Gly Ser Val Gly Gln
50 55 60

65

ES 2 346 196 T3

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

ES 2 346 196 T3

Gly Asp Asp Arg Cys Leu Thr Asn Tyr Ala Thr Asp Leu Gly Lys Thr
260 265 270

5

Val Tyr Gln Ser Thr Ala Arg Cys Asp Thr Asp Val Pro Asp Lys Phe
275 280 285

10

Lys Val Phe Ile Lys Gln Gln Asn Arg Trp Asn Lys Ser Phe Phe Arg
290 295 300

15

Glu Ser Ile Ile Ser Val Lys Lys Leu Leu Ala Thr Pro Ser Val Ala
305 310 315 320

20

Val Trp Thr Ile Thr Glu Val Ser Met Phe Ile Met Leu Val Tyr Ser
325 330 335

25

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

30

Leu Val Ala Phe Leu Val Ile Ile Phe Ile Val Ala Leu Cys Arg Asn
355 360 365

35

Val His Tyr Met Val Lys His Pro Phe Ala Phe Leu Leu Ser Pro Phe
370 375 380

40

Tyr Gly Leu Ile His Leu Phe Val Leu Gln Pro Leu Lys Ile Tyr Ser
385 390 395 400

45

Leu Phe Thr Ile Arg Asn Ala Thr Trp Gly Thr Arg Lys Lys Thr Ser
405 410 415

50

Lys

55

<210> 13

<211> 1440

<212> DNA

60

<213> *Streptococcus pyogenes*

65

ES 2 346 196 T3

<400> 13

5 tttaatggaa acacaatttt attaaaaata tctctatatc tagttgacat tatttcttat
60

10 ttatattata atattgaggt cctttctttc aaggaaatta aaaaagaaag aggtgtaatt
120

15 gtgcctattt ttaaaaaaac ttttaattggt ttatccttta tttttttgat atctatcttg
180

20 atttatctaa atatgtatct atttggaaca tcaactgtag gaatttatgg agtaatatta
240

25 ataacctatc tagttatcaa acttggaata tctttccttt atgagccatt taaaggaaat
300

30 ccacatgact ataaagttgc tgcgtgtaatt ccttcttata atgaagatgc cgagtcatta
360

35 ttagaaacac ttaaaagtgt gttagcacag acctatccgt tatcagaaat ttatattggt
420

40 gatgatggga gttcaaacac agatgcaata caattaattg aagagtatgt aaatagagaa
480

45 gtggatattt gtcgaaacgt tatcggtcac cgttcccttg tcaataaagg aaaacgccat
540

50 gctcaagcgt gggcatttga aagatctgac gctgacgttt ttttaaccgt agactcagat
600

55 acttatatct atccaaatgc cttagaagaa ctcttaaaaa gcttcaatga tgagacagtt
660

60 tatgctgcaa caggacattt gaatgctaga aacagacaaa ctaatctatt aacgcgactt
720

65 acagatatcc gttacgataa tgcctttggg gtggagcgtg ctgctcaatc attaacaggt
780

70 aatatttttag ttgctcagg accattgagt atttatcgac gtgaagtgat tattcctaac
840

75 ttagagcgct ataaaaatca aacattccta ggtttacctg ttagcattgg ggatgatcga
900

ES 2 346 196 T3

tgtttaacaa attatgctat tgatttagga cgcactgtct accaatcaac agctagatgt
960

5 gatactgatg tacctttcca attaaaaagt tatttaaagc aacaaaatcg atggaataaa
1020

10 tcttttttta gagaatctat tatttctgtt aaaaaaattc tttctaatacc catcggtgcc
1080

15 ttatggacta ttttcgaagt cggtatgttt atgatgttga ttgtcgcaat tgggaatctt
1140

20 ttgtttaatc aagctattca attagacctt attaaacttt ttgccttttt atccatcatc
1200

25 tttatcggtg ctttatgtcg taatgttcat tatatgggtca aacatcctgc tagttttttg
1260

30 ttatctcttc tgtatggaat attacaettg tttgtcttac agcccctaaa actttattct
1320

35 ttatgcacca ttaaaaatac ggaatgggga acacgtaaaa aggtcactat ttttaaataa
1380

40 tatatgcata gagtagttag agaaggagta attttatgaa aatagcagtt gctggatcag
1440

<210> 14

<211> 419

40 <212> PRT

<213> *Streptococcus pyogenes*

<400> 14

45

Met Pro Ile Phe Lys Lys Thr Leu Ile Val Leu Ser Phe Ile Phe Leu
1 5 10 15

50

Ile Ser Ile Leu Ile Tyr Leu Asn Met Tyr Leu Phe Gly Thr Ser Thr
20 25 30

55

Val Gly Ile Tyr Gly Val Ile Leu Ile Thr Tyr Leu Val Ile Lys Leu
35 40 45

60

65

ES 2 346 196 T3

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

ES 2 346 196 T3

Leu Glu Arg Tyr Lys Asn Gln Thr Phe Leu Gly Leu Pro Val Ser Ile
245 250 255

5

Gly Asp Asp Arg Cys Leu Thr Asn Tyr Ala Ile Asp Leu Gly Arg Thr
260 265 270

10

Val Tyr Gln Ser Thr Ala Arg Cys Asp Thr Asp Val Pro Phe Gln Leu
275 280 285

15

Lys Ser Tyr Leu Lys Gln Gln Asn Arg Trp Asn Lys Ser Phe Phe Arg
290 295 300

20

Glu Ser Ile Ile Ser Val Lys Lys Ile Leu Ser Asn Pro Ile Val Ala
305 310 315 320

25

Leu Trp Thr Ile Phe Glu Val Val Met Phe Met Met Leu Ile Val Ala
325 330 335

30

Ile Gly Asn Leu Leu Phe Asn Gln Ala Ile Gln Leu Asp Leu Ile Lys
340 345 350

35

Leu Phe Ala Phe Leu Ser Ile Ile Phe Ile Val Ala Leu Cys Arg Asn
355 360 365

40

Val His Tyr Met Val Lys His Pro Ala Ser Phe Leu Leu Ser Pro Leu
370 375 380

45

Tyr Gly Ile Leu His Leu Phe Val Leu Gln Pro Leu Lys Leu Tyr Ser
385 390 395 400

50

Leu Cys Thr Ile Lys Asn Thr Glu Trp Gly Thr Arg Lys Lys Val Thr
405 410 415

55

Ile Phe Lys

60

<210> 15

<211> 1380

<212> DNA

65

<213> *Sulfolobus solfataricus*

ES 2 346 196 T3

<400> 15

5 tgccctcatg cggttaagct tgaaatggcc ttttttagaa agagaaagga gttatctagt
 60

 taaccaatct tgccaccaga aagttatcat agcatatagg gaagcaatat attgcactgc
10 120

 tagtgcgata gaaccgtata ttagagtttt caatttatct tctggaatta aataaatcat
 180

15 ggtcattaca aaggggatta tagctataaa tcctccgtaa tgcaagaata aataaaagat
 240

20 aaagaagtcg gaatggtaac gggttggcag atacaatagt aataagagct tagtattaat
 300

 tacattaatt gaggaataga tcttaagaat tctagtgaag ctcataataa ggaacaagag
25 360

 cgtaaataat ggtaacagat tgggtgtatat catattaaaa acgtataatg aacccttttt
 420

30 acttatacta ccgtcagcta tctcccttat aaaattaaga taatttgctc ttgtccatct
 480

 agttacttgt ttcgtaaaca tttttatgtc tctagggggg tttgtatatg ccactgcac
35 540

 aaagactttt acagccctat accctttttt tataacaaaa tcggttaaat ctctatcatc
 600

40 ggaaatttta attgggtcttc caaacatttt cggtctctaa aactcttttag ataatatata
 660

45 tggtttttac agttcgggtc tatatattac acattgtcca cttaatatata tagcacttcc
 720

 aaaatagttt accgccctgt ttactatctc acttattctc tcaaagaatt caccataata
50 780

 atatgcatat ttatttttct cgtcatacat aattctaata tttggcccta ctccacctac
 840

55

60

65

ES 2 346 196 T3

tgactcatca aaaacactta acatcttttag tatagagtct ttataaataa tcgtatcact
900

5 atctagaaac atcactagag gagatcttac atacttaact ccctcggcta acgcgtatct
960

10 tttcccttta tgttcacgca tataaataaa ttaccacca tatctttccg taattgattt
1020

gtatggttct agaacactat cccctacaac aataaattct aaccttgtgt catataaagt
1080

15 ccttatcact ttttcaaaaa tatctatttc ctccttataa actggtatca caactgtaag
1140

20 atcagagaga ttataaaaac ttgagtgttg agtttttcta ttattactta ttactgcaaa
1200

25 aaatgaattc aaaaagaaat aaagaatagt tataattgtg aatgaaagag aataaatgaa
1260

atatgagact ccgtgaaata agtgaaacat aatcaccact ataatgctcg atatcgaaat
1320

30 atataacgat ttttcctaat tcaccattcg aattctccgt tcaaaaaggg gttagttaac
1380

35 <210> 16

<211> 415

<212> PRT

40 <213> *Sulfolobus solfataricus*

<400> 16

45 Met Val Ile Met Phe His Leu Phe His Gly Val Ser Tyr Phe Ile Tyr
1 5 10 15

50 Ser Leu Ser Phe Thr Ile Ile Thr Ile Leu Tyr Phe Phe Leu Asn Ser
20 25 30

55 Phe Phe Ala Val Ile Ser Asn Asn Arg Lys Thr Gln His Ser Ser Phe
35 40 45

60 Tyr Asn Leu Ser Asp Leu Thr Val Val Ile Pro Val Tyr Lys Glu Glu

65

ES 2 346 196 T3

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

65

ES 2 346 196 T3

245

250

255

5 Ala Tyr Thr Lys Pro Pro Arg Asp Ile Lys Met Phe Thr Lys Gln Val
 260 265 270

10 Thr Arg Trp Thr Arg Ala Asn Tyr Leu Asn Phe Ile Arg Glu Ile Ala
 275 280 285

15 Asp Gly Ser Ile Ser Lys Arg Gly Ser Leu Tyr Val Phe Asn Met Ile
 290 295 300

20 Tyr Thr Asn Leu Leu Pro Leu Phe Thr Leu Leu Phe Leu Tyr Met Ser
305 310 315 320

25 Phe Thr Arg Ile Leu Lys Ile Tyr Ser Ser Ile Asn Val Ile Asn Thr
 325 330 335

30 Lys Leu Leu Leu Leu Tyr Leu Pro Thr Arg Tyr His Ser Asp Phe
 340 345 350

35 Phe Ile Phe Tyr Leu Phe Leu His Tyr Gly Gly Phe Ile Ala Ile Ile
 355 360 365

Pro Phe Val Met Thr Met Ile Tyr Leu Ile Pro Glu Asp Lys Leu Lys
370 375 380

45 Thr Leu Ile Tyr Gly Ser Ile Ala Leu Ala Val Gln Tyr Ile Ala Ser
385 390 395 400

Leu Tyr Ala Met Ile Thr Phe Trp Trp Gln Asp Trp Leu Thr Arg
405 410 415

55 $\langle 210 \rangle$ 17

$\langle 211 \rangle$ 1200

<212> DNA

60 <213> pXO1 de *Bacillus anthracis*

65

ES 2 346 196 T3

<400> 17

5 taaggttcca aatatgaaag tggcagtagt ggtccctca tataatgaaa gtgcaagtgc
60
tattgttaat acaattaata gcgttttagc tcaagattat ccaattcatg aaattttctt
10 120
tgttgatgat ggtagtaagg ataaatcggc ttatgaagta gcacttaaaa tgagggagga
180
15 acttcttaga actcaacgag aaattgctgc tacaactaag aatatttggt ctgaaatatt
240
20 aggtattcct gacttaatcg tacatcgttt acctaagaat tgcgggaaaa gacatgctca
300
attatgggct tttaaacgga caacagcaga tgcatttggt accattgatt cagatgggtga
25 360
tttgttccca aatgctgtta gagagttatt gaaacccttt aatgatgaaa aagtaatggc
420
30 cacaactggg cagctgaaca ttcgtaatag aaatgataat ttattaacaa aactaattga
480
tattgcgttat gacaatgcgt tccgtgtgga gcgtgcagca cagtccgtaa caggaaatgt
35 540
tcttgtttgt agtgggccgt taagttgtta tcgtagagaa gtaataactg aaaatttaga
600
40 acattatgga agtcagatgt tccttggtga ggaggtgcag tttggagatg atagatgtct
660
45 aactaattat gctattttga aagggaanaac agtttatcaa tccactgctc gatgtattac
720
tgatgctcca actacattaa aacaatttct taaacagcaa ctacgttgga acaagtcatt
50 780
ttttagagaa agtttaattt cacttggcat tggatatgaaa aaaccaaattg ttcttgtttg
840
55 gacaattttc gaaatatcgt tatggatttt atttgggctt tccctacttc taagtattat
900
60 tctcaaggca agtcatgtag ggtaatttt ggctgtttat tatttgggtt atatttcatt
960

65

ES 2 346 196 T3

agctgtatat gctagaaatg tatttttatct attaaaacat ccccttactt tcttactggc
1020

5 gccattatat ggaattctcc atgtatttagc actattacct atacgctttt atgcttttact
1080

10 aactattaaa tctaattgggtt ggggaacacg ttaattacag taatttttatg tatttttttta
1140

ggaggatatt attaagtga gattagaaaa gcgattatcc cagcagcggg attaggcaca
1200
15

<210> 18

20 <211> 366

<212> PRT

<213> pXO1 de *Bacillus anthracis*

25 <400> 18

30 Met Lys Val Ala Val Val Val Pro Ser Tyr Asn Glu Ser Ala Ser Ala
1 5 10 15

35 Ile Val Asn Thr Ile Asn Ser Val Leu Ala Gln Asp Tyr Pro Ile His
20 25 30

40 Glu Ile Phe Phe Val Asp Asp Gly Ser Lys Asp Lys Ser Ala Tyr Glu
35 40 45

45 Val Ala Leu Lys Met Arg Glu Glu Leu Leu Arg Thr Gln Arg Glu Ile
50 55 60

50 Ala Ala Thr Thr Lys Asn Ile Cys Ser Glu Ile Leu Gly Ile Pro Asp
65 70 75 80

55 Leu Ile Val His Arg Leu Pro Lys Asn Cys Gly Lys Arg His Ala Gln
85 90 95

60 Leu Trp Ala Phe Lys Arg Thr Thr Ala Asp Ala Ile Val Thr Ile Asp
100 105 110

65

ES 2 346 196 T3

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

ES 2 346 196 T3

Ile Leu Ala Val Tyr Tyr Leu Gly Tyr Ile Ser Leu Ala Val Tyr Ala
305 310 315 320

5

Arg Asn Val Phe Tyr Leu Leu Lys His Pro Leu Thr Phe Leu Leu Ala
325 330 335

10

Pro Leu Tyr Gly Ile Leu His Val Leu Ala Leu Leu Pro Ile Arg Phe
340 345 350

15

Tyr Ala Leu Leu Thr Ile Lys Ser Asn Gly Trp Gly Thr Arg
355 360 365

20

<210> 19

<211> 1680

25 <212> DNA

<213> Virus de *Ectocarpus siliculosus*

<400> 19

30

gccatttaca tcttggcgcg atcgttgcaa agtaaaaata atgttttctg tgacggaaac
60

35

aggggcatga catgcttctc ggcctgaaga aaaacctctt gttcattgcg tatctgttct
120

40

tcaatgccct tttaacgctt ttgttgcttg gattcgatta cggctacatc atcgatatga
180

45

tattcgttgt cggcggacac ttcagagatg tgatcaacgt agcataccaa ttacttcaca
240

tgacacaggat actcaggcgt tgtgccgaca tcccgggaaga cgatgccaag atagttattt
300

50

gttgcttggt tccggtgtac aacgagaagc cttcgatggt aaagaagaat cttgatgctc
360

55

tgacgacaca gaagctatcg gaaaacacca agttgggtggt gatgcttctc ttgacggac
420

tgaacaacca caacgcagat ctcttcaatg ccgtcgtcga tgccattggc cttgacaccg
480

60

gatgcggaga agagcaatgg tttccgaatt ggaagagcaa gctgctgaag aagttggtgt

65

ES 2 346 196 T3

540

5 acaaaatcgg catatacaat gacacgtcgg tcatcctgtc gtacaaggag aacaattcgg
600

10 gaaaaaaaga ctctctcatc atcggggaaa acttcatcgt gtcgggcac ccgaggatcg
660

15 aatctctgga cgtacgacaa gtggatttca tctatcacac ggacggcgac accatttccg
720

20 acgaaaactg tttgaacgag atggtgaagt ctctcgtgga tgatccagac ctcgacggcg
780

25 tctctggcct cctgagaaca tacctcaagg acgacgcgac ttgctcggaa agtgcgttcg
840

30 tagcgatgca agactttcag tacttcttct ccattgttgt ccgtaggatg acggagagca
900

35 taatgaattc aactacctgc ctcccggggg gctccaacat gatcaggata agcgaaaaga
960

40 ctcacgctgc gattgaaaaa tacggaaacc ttccgggtcaa gaagagcggg ctggtgcaga
1020

45 cagtcacgcg gatgcaagga accgaccgac gatacaccac gctcttgctg agacaggggt
1080

50 ccaagctaca gatgaattgg cgtgcgtttg ttcacacgga gccaccgctg aacgcgacgg
1140

55 cgtttgtgaa tcaacgcaga cgatgggtctt caaactcctt cttcaattcc atgatcacgc
1200

60 tgtactccaa caacatcccg atgtacatca agctatcgaa cdtgtgcgac atcgcccgag
1260

65 tcttcaccac gatctttcgc gtgatatcgt acttgtgctt ttgggtttac gtcaagaatt
1320

70 tttcccttgt caacatcggt tttttctcta tatttatcgc ccttccctac ctctatgcct
1380

75 tcgcctggat attctgtatc gttccagagt ggaagcagat gatagccggg ttttttttga
1440

80 acaaaatcct cagccctttt ttatctgtga tcgcggtcac aaagatgttc ttcacttcaa

ES 2 346 196 T3

1500

5 ccgatttcgc ttggggcagt acgcggttga caccaccgga tgcagcgtct taaactcaac
1560

10 caacaatcgt ctctattaat aaacctgatt cgtgattttg ggttcgattc tcctgtttcc
1620

15 tttttttcct ttttttcctt ttttttcctt ttttctatt ttcctttttg attttgtttt
1680

<210> 20

<211> 493

20 <212> PRT

<213> Virus de *Ectocarpus siliculosus*

<400> 20

30 Met Leu Leu Gly Leu Lys Lys Asn Leu Leu Phe Ile Ala Tyr Leu Phe
1 5 10 15

35 Phe Asn Ala Leu Leu Thr Leu Leu Leu Leu Gly Phe Asp Tyr Gly Tyr
20 25 30

40 Ile Ile Val Ser Ile Phe Val Val Gly Gly His Phe Arg Asp Val Ile
35 40 45

45 Asn Val Ala Tyr Gln Leu Leu His Met His Arg Ile Leu Arg Arg Cys
50 55 60

50 Ala Asp Ile Pro Glu Asp Asp Ala Lys Ile Val Ile Cys Cys Leu Val
65 70 75 80

55 Pro Val Tyr Asn Glu Lys Pro Ser Met Leu Lys Lys Asn Leu Asp Ala
85 90 95

60 Leu Thr Thr Gln Lys Leu Ser Glu Asn Thr Lys Leu Val Val Met Leu
100 105 110

65 Leu Phe Asp Gly Leu Asn Asn His Asn Ala Asp Leu Phe Asn Ala Val

ES 2 346 196 T3

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

65

ES 2 346 196 T3

305 310 315 320

5 Met Gln Gly Thr Asp Arg Arg Tyr Thr Thr Leu Leu Leu Arg Gln Gly
325 330 335

10 Ser Lys Leu Gln Met Asn Trp Arg Ala Phe Val His Thr Glu Pro Pro
340 345 350

15 Leu Asn Ala Thr Ala Phe Val Asn Gln Arg Arg Arg Trp Ser Ser Asn
355 360 365

20 Ser Phe Phe Asn Ser Met Ile Thr Leu Tyr Ser Asn Asn Ile Pro Met
370 375 380

25 Tyr Ile Lys Leu Ser Asn Leu Val Asp Ile Ala Arg Val Phe Thr Thr
385 390 395 400

30 Ile Phe Arg Val Ile Ser Tyr Leu Cys Phe Trp Val Tyr Val Lys Asn
405 410 415

35 Phe Ser Leu Val Asn Ile Val Phe Phe Ser Ile Phe Ile Ala Leu Pro
420 425 430

40 Tyr Leu Tyr Ala Phe Ala Trp Ile Phe Cys Ile Val Pro Glu Trp Lys
435 440 445

45 Gln Met Ile Ala Gly Phe Phe Leu Asn Lys Ile Phe Thr Pro Phe Leu
450 455 460

50 Ser Val Ile Ala Val Thr Lys Met Phe Phe Thr Ser Thr Asp Phe Ala
465 470 475 480

55 Trp Gly Ser Thr Arg Leu Thr Pro Pro Asp Ala Ala Ser
485 490

60

65