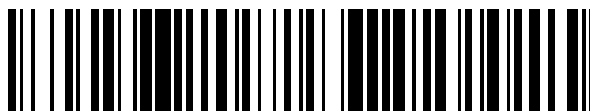


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 505 695**

51 Int. Cl.:

A61K 39/09

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.07.2004** **E 04809528 (5)**

97 Fecha y número de publicación de la concesión europea: **09.07.2014** **EP 1648500**

54 Título: **Composiciones inmunógenas para Streptococcus pyogenes**

30 Prioridad:

31.07.2003 US 491822 P
03.02.2004 US 541565 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
10.10.2014

73 Titular/es:

NOVARTIS VACCINES AND DIAGNOSTICS, INC.
(100.0%)
4560 HORTON STREET
EMERYVILLE, CA 94608, US

72 Inventor/es:

GRANDI, GUIDO;
TELFORD, JOHN y
BENSI, GIULIANO

74 Agente/Representante:

IZQUIERDO FACES, José

ES 2 505 695 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).

COMPOSICIONES INMUNÓGENAS PARA *STREPTOCOCCUS PYOGENES***Descripción****5 CAMPO DE LA INVENCION**

La presente invención se encuentra en los campos de la inmunología y la vacunología. En concreto, se refiere a antígenos derivados de *Streptococcus pyogenes* y a su uso en la inmunización.

10 ANTECEDENTES DE LA INVENCION

El estreptococo del grupo A ("GAS", *S. pyogenes*) es un patógeno humano frecuente, que se estima se encuentra presente en el 5%-15% de los individuos normales sin signos de enfermedad. Sin embargo, cuando las defensas del hospedador se encuentran afectadas, o cuando el organismo es capaz de ejercer su virulencia, o cuando se introduce en hospedadores o tejidos vulnerables, se produce una infección aguda. Las enfermedades relacionadas incluyen la fiebre puerperal, la escarlatina, la erisipela, la faringitis, el impétigo, la fascitis necrotizante, la miositis y el síndrome de shock tóxico estreptocócico.

GAS es una bacteria con forma de coco, grampositiva y no formadora de esporas, que se produce normalmente en cadenas o pares de células. Aunque *S. pyogenes* pueden tratarse mediante antibióticos, resulta deseable una vacuna profiláctica para prevenir la aparición de la enfermedad. Las tentativas para desarrollar una vacuna de este tipo han estado en curso durante muchas décadas. Aunque se han sugerido diversos enfoques de vacunas contra GAS y algunos enfoques se encuentran actualmente en ensayos clínicos, hasta la fecha no existen vacunas contra GAS a disposición del público.

En la SEQ ID NO: 9188 del documento WO 02/034771 se proporciona la secuencia de aminoácidos de GAS40. En el documento WO 02/034771 se analiza el uso de antígenos de GAS en composiciones inmunógenas.

En el documento WO 2004/078907 (que se publicó el 16 de septiembre de 2004 y se encuentra dentro del Artículo 54(3) EPC) también se proporciona la secuencia de aminoácidos de GAS40 y se analiza el uso de antígenos de GAS en composiciones inmunógenas.

Es un objeto de la invención proporcionar composiciones adicionales y mejoradas para proporcionar inmunidad contra la enfermedad y/o la infección por GAS. Las composiciones incluyen GAS 40, un factor de virulencia de GAS identificado por los solicitantes, que es particularmente adecuado para su uso en vacunas. Además, las composiciones se basan en una combinación de dos o más (por ejemplo tres o más) antígenos de GAS.

40 RESUMEN DE LA INVENCION

Los solicitantes han descubierto un grupo de treinta antígenos de GAS que son particularmente adecuados para los fines de inmunización, particularmente cuando se utilizan en combinaciones. Además, los solicitantes han identificado un antígeno de GAS (GAS 40) que es particularmente inmunógeno utilizado solo o en combinación con antígenos de GAS adicionales.

La invención proporciona una composición inmunógena que comprende una combinación de antígenos de GAS, consistiendo dicha combinación en dos a diez antígenos de GAS, en la que dicha combinación incluye GAS40, siendo dicho "GAS40" una proteína que comprende una secuencia de aminoácidos (a) que tiene un 95% o más identidad con la SEQ ID NO: 1 y/o (b) que es un fragmento de al menos 200 aminoácidos consecutivos de la SEQ ID NO: 1 y en la que dicha combinación incluye GAS57, siendo dicho "GAS57" una proteína que comprende una secuencia de aminoácidos (a) que tiene un 95% o más identidad con la SEQ ID NO: 116 y/o (b) que es un fragmento de al menos 200 aminoácidos consecutivos de la SEQ ID NO: 116. La invención también proporciona la presente composición para uso terapéutico. La invención también proporciona un método de fabricación de la presente composición. La invención también proporciona un kit que comprende un primer componente que comprende la presente composición.

Por lo tanto, la invención proporciona una composición inmunógena que comprende GAS 40 (incluidos fragmentos del mismo o un polipéptido con identidad de secuencia con el mismo). Un fragmento preferente de GAS 40 comprende una o más regiones con superenrollamiento. La invención proporciona una composición inmunógena que comprende una combinación de antígenos de GAS, consistiendo dicha combinación en dos a diez antígenos de GAS, en la que dicha combinación incluye GAS 40 o un fragmento del mismo o un polipéptido con identidad de secuencia con el mismo, y GAS 57 o un fragmento del mismo o un polipéptido con identidad de secuencia con el mismo. Preferentemente, la combinación consiste en tres, cuatro, cinco, seis o siete antígenos de GAS. Aún más preferentemente, la combinación consiste en tres, cuatro o cinco antígenos de GAS.

En el presente documento también se describe una composición inmunógena que comprende una

combinación de antígenos de GAS, consistiendo dicha composición en dos a treinta y un antígenos de GAS de un primer grupo de antígenos, consistiendo dicho primer grupo de antígenos en: GAS 117, GAS 130, GAS 277, GAS 236, GAS 40, GAS 389, GAS 504, GAS 509, GAS 366, GAS 159, GAS 217, GAS 309, GAS 372, GAS 039, GAS 042, GAS 058, GAS 290, GAS 511, GAS 533, GAS 527, GAS 294, GAS 253, GAS 529, GAS 045, GAS 095, GAS 193, GAS 137, GAS 084, GAS 384, GAS 202 y GAS 057. Estos antígenos se denominan en el presente documento "primer grupo de antígenos". Preferentemente, la combinación de antígenos de GAS consiste en tres, cuatro, cinco, seis, siete, ocho, nueve o diez antígenos de GAS seleccionados del primer grupo de antígenos. Preferentemente, la combinación de antígenos de GAS consiste en tres, cuatro o cinco antígenos de GAS seleccionados del primer grupo de antígenos.

GAS 39, GAS 40, GAS 57, GAS 117, GAS 202, GAS 294, GAS 527, GAS 533 y GAS 511 resultan antígenos de GAS particularmente preferentes. Preferentemente, la combinación de antígenos de GAS incluye uno o ambos de GAS 40 y GAS 117. Preferentemente, la combinación incluye GAS 40.

Más adelante se analizan los ejemplos representativos de algunas de estas combinaciones de antígenos.

La combinación de antígenos de GAS puede consistir en tres antígenos de GAS seleccionados del primer grupo de antígenos. Por consiguiente, la combinación de antígenos de GAS puede consistir en GAS 40, GAS 117 y un tercer antígeno de GAS seleccionado del primer grupo de antígenos. Las combinaciones preferentes incluyen GAS 40, GAS 117 y un tercer antígeno de GAS seleccionado del grupo que consiste en GAS 39, GAS 57, GAS 202, GAS 294, GAS 527, GAS 533 y GAS 511.

La combinación de antígenos de GAS puede consistir en GAS 40 y dos antígenos de GAS adicionales seleccionados del primer grupo de antígenos. Las combinaciones preferentes incluyen GAS 40 y dos antígenos de GAS seleccionados del grupo que consiste en GAS 39, GAS 57, GAS 117, GAS 202, GAS 294, GAS 527, GAS 533 y GAS 511. La combinación de antígenos de GAS puede consistir en GAS 117 y dos antígenos de GAS adicionales seleccionados del primer grupo de antígenos.

La combinación de antígenos de GAS puede consistir en cuatro antígenos de GAS seleccionados del primer grupo de antígenos. La combinación de antígenos de GAS puede consistir en GAS 40, GAS 117 y dos antígenos de GAS adicionales seleccionados del primer grupo de antígenos. Las combinaciones preferentes incluyen GAS 40, GAS 117 y dos antígenos de GAS seleccionados del grupo que consiste en GAS 39, GAS 57, GAS 202, GAS 294, GAS 527, GAS 533 y GAS 511.

La combinación de antígenos de GAS puede consistir en GAS 40 y tres antígenos de GAS adicionales seleccionados del primer grupo de antígenos. Las combinaciones preferentes incluyen GAS 40 y tres antígenos de GAS adicionales seleccionados del grupo que consiste en GAS 39, GAS 57, GAS 117, GAS 202, GAS 294, GAS 527, GAS 533 y GAS 511. La combinación de antígenos de GAS puede consistir en GAS 117 y tres antígenos adicionales seleccionados del primer grupo de antígenos.

La combinación de antígenos de GAS puede consistir en cinco antígenos de GAS seleccionados del primer grupo de antígenos. La combinación de antígenos de GAS puede consistir en GAS 40, GAS 117 y tres antígenos de GAS adicionales seleccionados del primer grupo de antígenos. Las combinaciones preferentes incluyen GAS 40, GAS 117 y tres antígenos de GAS adicionales seleccionados del grupo que consiste en GAS 39, GAS 57, GAS 202, GAS 294, GAS 527, GAS 533 y GAS 511.

La combinación de antígenos de GAS puede consistir en GAS 40 y cuatro antígenos de GAS adicionales seleccionados del primer grupo de antígenos. Las combinaciones preferentes incluyen GAS 40 y cuatro antígenos de GAS adicionales seleccionados del grupo que consiste en GAS 39, GAS 57, GAS 117, GAS 202, GAS 294, GAS 527, GAS 533 y GAS 511. La combinación de antígenos de GAS puede consistir en GAS 117 y cuatro antígenos de GAS adicionales seleccionados del primer grupo de antígenos.

La combinación de antígenos de GAS puede consistir en ocho antígenos de GAS seleccionados del primer grupo de antígenos. La combinación de antígenos de GAS puede consistir en GAS 40, GAS 117 y seis antígenos de GAS adicionales seleccionados del primer grupo de antígenos. La combinación de antígenos de GAS puede consistir en GAS 40 y siete antígenos de GAS adicionales seleccionados del primer grupo de antígenos. La combinación de antígenos de GAS puede consistir en GAS 117 y siete antígenos de GAS adicionales seleccionados del primer grupo de antígenos.

La combinación de antígenos de GAS puede consistir en diez antígenos de GAS seleccionados del primer grupo de antígenos. La combinación de antígenos de GAS puede consistir en GAS 40, GAS 117 y ocho antígenos de GAS adicionales seleccionados del primer grupo de antígenos. La combinación de antígenos de GAS puede consistir en GAS 40 y nueve antígenos de GAS adicionales seleccionados del primer grupo de antígenos. La combinación de antígenos de GAS puede consistir en GAS 117 y nueve antígenos de GAS adicionales seleccionados del primer grupo de antígenos.

BREVE DESCRIPCIÓN DE LAS FIGURAS

La figura 1 identifica una secuencia de péptido líder, dos secuencias con superenrollamiento, una secuencia con cremallera de leucina y una secuencia transmembrana dentro de una secuencia de aminoácidos de GAS 40.

La figura 2 representa un esquema de GAS 40 que identifica una secuencia de péptido líder, dos secuencias con superenrollamiento, una secuencia con cremallera de leucina y una secuencia transmembrana, así como regiones con superenrollamiento de GAS 40 que tienen bajo nivel de homología con otras proteínas estreptocócicas con función conocida o prevista.

La figura 3 incluye el análisis de alineamiento BLAST de las regiones con superenrollamiento identificadas de GAS 40 con otras bacterias *Streptococcus*.

La figura 4 proporciona la estructura secundaria predicha para una secuencia de aminoácidos de GAS 40.

La figura 5 representa esquemáticamente la ubicación de GAS 40 dentro del genoma de GAS. También incluye la comparación que representa esquemáticamente un mutante de GAS con delección para GAS 40. Los detalles adicionales en estos esquemas demuestran la probabilidad de que GAS 40 se adquiriese por transferencia horizontal a través de un factor transposón.

La figura 6 proporciona la comparación del análisis FACS que representa la exposición en superficie de GAS 40 en una cepa de tipo silvestre (y sin exposición en superficie en el mutante por delección de GAS 40).

La figura 7 presenta los datos de opsonofagocitosis para GAS 40 (en diversos constructos de expresión).

La figura 8 presenta los datos de inmunización y provocación para varios antígenos de GAS de la descripción.

DESCRIPCIÓN DETALLADA DE LA INVENCION

Como se ha analizado anteriormente, en el presente documento se describen composiciones que comprenden una combinación de antígenos de GAS, en las que las combinaciones pueden seleccionarse de entre grupos de antígenos que los solicitantes han identificado como particularmente adecuados para fines de inmunización, particularmente cuando se utilizan en combinación. La invención proporciona composiciones que comprenden GAS 40.

GAS 40 y los demás antígenos de GAS del primer grupo de antígenos se describen con más detalle más adelante. Las secuencias genómicas de al menos tres cepas de GAS están a disposición del público. La secuencia genómica de una cepa M1 de GAS se presenta en Ferretti *et al.*, PNAS (2001) 98(8):4658 - 4663. La secuencia genómica de una cepa M3 de GAS se presenta en Beres *et al.*, PNAS (2002) 99(15):10078 - 10083. La secuencia genómica de una cepa M18 de GAS se presenta en Smoot *et al.*, PNAS (2002) 99(7):4668 - 4673. Preferentemente, los antígenos de GAS de la invención comprenden la secuencia de aminoácidos o de polinucleótidos de una cepa M1, M3 o M18 de GAS. Más preferentemente, los antígenos de GAS de la invención comprenden una secuencia de aminoácidos o de polinucleótidos de una cepa M1.

Como habrá varianza entre los antígenos de GAS identificados entre los tipos de M de GAS y los aislados de la cepa de GAS, las referencias a las secuencias de aminoácidos o de polinucleótidos de GAS descritas en el presente documento incluyen preferentemente secuencias de aminoácidos o de polinucleótidos con identidad de secuencia con las mismas. Las secuencias de aminoácidos o de polinucleótidos preferentes tienen un 50% o más de identidad de secuencia (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95 %, 96%, 97%, 98%, 99%, 99,5% o más). Asimismo, las referencias a las secuencias de aminoácidos o de polinucleótidos de GAS descritas en el presente documento incluyen preferentemente fragmentos de esas secuencias (es decir, fragmentos que conservan o codifican las propiedades inmunitarias del antígeno de GAS). Los fragmentos de aminoácidos preferentes incluyen al menos *n* aminoácidos consecutivos, en los que *n* es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250 o más). Los fragmentos de polinucleótidos preferentes incluyen al menos *n* polinucleótidos consecutivos, en los que *n* es 12 o más (por ejemplo, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250 o más). Los fragmentos de aminoácidos o de polinucleótidos de la invención pueden no ser idénticos a las secuencias de aminoácidos o de polinucleótidos de otras bacterias (no GAS) (por ejemplo, los fragmentos no son idénticos a las secuencias en otras bacterias *Streptococcus*).

(1) GAS 40

GAS 40 corresponde a los números de registro del GenBank de M1 GI:13621545 y GI:15674449, al número de registro del GenBank de M3 GI:21909733, al número de registro del GenBank de M18 GI:19745402, y también se conoce como 'Spy0269' (M1), 'SpyM3_0197' (M3), 'SpyM18_0256' (M18) y 'prgA'. GAS 40 también se ha identificado como una supuesta proteína de exclusión superficial. Las secuencias de aminoácidos y de polinucleótidos de GAS 40 de una cepa M1 se exponen a continuación y en el listado de secuencias como las SEQ ID NO: 1 y 2.

SEQ ID NO: 1

MDLEQTKPNQVKQKIALTSTIALLSASVGVSHQVKADDRASGETKASNTHDDSLPKPETIQEAKATIDAVEKTLSSQKAE
LTELATALTKTTAEINHLKEQQDNEQKALTSQEIIYNTNLASSEETLLAQGAHQRELATETELHNAQADQHSKETALS
EQKASISAETTRAQDLVEQVKTSEQNIAKLNAMISNPDAITKAAQTANDNTKALSSELEKAKADLENQKAKVKKQLTEEL
AAQKAALAEKEAELSRKSSAPSTQDSIVGNNTMKAPQGYPLEELKKLEASGYIGSASYNYYKEHADQIIAKASPGNQL
NQYQDIPADRNRFVDPDNLTPEVQNELAQFAAHMINSVRRQLGLPPVTVTAGSQEFARLLSTSYKKTHGNTRPSFVYQGP
GVSGHYGVGPHDKTIIEDSAGASGLIRNDDNMYENIGAFNDVHTVNGIKRGIYDSIKYMLFTDHLHGNTYGHAINFLRVD
KHNPNAPVYLGFTSTNVGSLNEHFVMPESNIAHQRFNKTPIKAVGSTKDYAQRVGTVDITIAAIKGVSSLENRLSAI
HQEADIMAAQKVSQQLGKGLASTLKQSDSLNLQVRQLNDTKGSLRTELLAAKAKQAQLEATRDQSLAKLASLKAALHQTE
ALAEQAAARVATLVAKKAHLQYLRDFKLPNRLQVIRERIDNTKQDLAKTSSLLNAQEALALQAKQSSLEATIATTEH
QLTLKTLANEKEYRHLDEDIATVPDLQVAPPLTGVPKPLSYSKIDTTPLVQEMVKETKQLEASARLAAENTSLVAEALV
QGTSEMVASNAIVSKITSSITQPSKTSYSGSSTTSLNLSVDDESTQRLKAGVVMLAAVGLTGFRFRKESK

SEQ ID NO: 2

ATGGACTTAGAACAAACGAAGCCAAACCAAGTTAAGCAGAAAAATTGCTTTAACCTCAACAATTGCTTTATTGAGTGCCAG
TGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAATACTACGACGATAGTT
TACCAAAACCAGAAACAATTCAAGAGGCAAAGGCAACTATTGATGCAGTTGAAAAAATCTCAGTCAACAAAAAGCAGAA
CTGACAGAGCTTGCTACCGCTCTGACAAAACTACTGCTGAAATCAACCACTTAAAGAGCAGCAAGATAATGAACAAA
AGCTTTAACCTCTGCACAAGAAATTTACACTAATACTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAAC
ATCAAAGAGAGTTAAACAGCTACTGAAACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAGAGAGCTGCATTGTCA
GAACAAAAAGCTAGCATTTTCAGCAGAACTACTCGAGCTCAGAGATTAGTGGAACAAGTCAAAACGCTGGAACAAAATAT
TGCTAAGCTCAATGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAGCAT
TAAGTTCAGAAATTGGAAGAGCTAAAGCTGACTTAGAAAATCAAAAGCTAAAGTTAAAAAGCAATTGACTGAGAGATTG
GCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCAAGCTTAGTCGCTTAAATCCTCAGCTCCGCTACTCAAGATAG
CATTGTGGGTAATAATACCATGAAAGCACCCGAAGGCTATCCTCTTGAAGAACTTAAAAAATTAGAAGCTAGTGGTTATA
TTGGATCAGCTAGTTACAATAATTATTACAAAGAGCATGCAGATCAAATTATTGCCAAAGCTAGTCCAGGTAATCAATTA
AATCAATACCAAGATATTCCAGCAGATCGTAATCGCTTTGTTGATCCCGATAATTTGACACCAGAAGTGCAAAATGAGCT
AGCGCAGTTTGCAGCTCACATGATTAATAGTGAAGAAGACAATTAGGTCTACCACCACTACTGTTACAGCAGGATCAC
AAGAATTTGCAAGATTACTTAGTACCAGCTATAAGAAAACCTCATGGTAATACAAGACCATCATTGTCTACGGACAGCCA
GGGGTATCAGGGCATTATGGTGTGGGCTCATGATAAACTATTATTGAAGACTCTGCCGGAGCGTCAGGGCTCATTG
AAATGATGATAACATGTACGAGAATATCGGTGCTTTTAACGATGTGCATCTGTGAATGGTATTAAACGTGGTATTATG
ACAGTATCAAGTATATGCTCTTTACAGATCATTTACCGAATATACATACGCCCATGCTATTAACTTTTACGTGTAGAT
AAACATAACCTTAATGCGCCTGTTTACCTTGGATTTCACACGCAATGTAGGATCTTGAATGAACACTTTGTGATGTT
TCCAGAGTCTAACATTGCTAACCATCAACGCTTTAATAAGACCCCTATAAAAGCCGTTGGAAGTACAAAAGATTATGCCC
AAAGAGTAGGCACTGTATCTGATATATTGACGCGATCAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTGCGCTATT
CATCAAGAAGCTGATATTATGGCAGCCCAAGCTAAAGTAAGTCAACTCAAGGTAATTAGCAAGCACACTTAAGCAGTC
AGACAGCTTAAATCTCCAAGTGAGACAATTAATGATACATAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTAAAGCAA
AACAAGCACAACTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACTGCACCAGACAGAA
GCCTTAGCAGAGCAAGCCGAGCCAGAGTGACAGCACTGGTGGCTAAAAAGCTCATTGCAATATCTAAGGGACTTTAA
ATTGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATGATAATACTAAGCAAGATTGGCTAAACTACCTCATCTT
TGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCTAAACAAAGCAGTCTAGAAGCTACTATTGCTACCACAGAACAC
CAGTTGACTTTGCTTAAACCTTAGCTAACGAAAAGGAATATCGCCACTTAGACGAAGATATAGCTACTGTGCTGATT
GCAAGTAGCTCCACCTCTTACGGGCGTAAACCGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAATGG
TTAAAGAAAACGAAACAATATTAGAAGCTTCAGCAAGATTAGCTGCTGAAAATACAAGTCTTTGAGCAGAAGCGCTTGT
GGCCAAACCTCTGAAATGGTAGCAAGTAATGCCATTGTGTCTAAATCACATCTTCGATTACTCAGCCCTCATCTAAGAC
ATCTTATGGCTCAGATCTTCTACACGAGCAATCTCATTCTGATGTTGATGAAAGTACTCAAGAGCTCTTAAAGCAG
GAGTCGTATGTTGGCAGCTGTGCGCTCACAGGATTAGGTTCCGTAAGGAATCTAAGTGA

Las proteínas GAS 40 para su uso con la invención comprenden una secuencia de aminoácidos: (a) que tiene un 95% o más identidad (por ejemplo, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 1; y/o (b) que es un fragmento de al menos *n* aminoácidos consecutivos de la SEQ ID NO: 1, donde *n* es 200 o más (por ejemplo, 250 o más). Estas proteínas GAS 40 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 1. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 1. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 1.

Por ejemplo, en una forma de realización, la secuencia de aminoácidos subrayada en el extremo N-terminal (secuencia líder) de la SEQ ID NO: 1 está eliminada. (Las secuencias de aminoácidos y de polinucleótidos para esta secuencia líder N-terminal se enumeran en el listado de secuencias como las SEQ ID NO: 3 y 4. Las secuencias de aminoácidos y de polinucleótidos para el fragmento de GAS 40 restante se enumeran en el listado de secuencias como las SEQ ID NO: 5 y 6).

Como otro ejemplo, en una forma de realización, la secuencia de aminoácidos subrayada en el extremo C-terminal (región transmembrana) de la SEQ ID NO: 1 está eliminada. (Las secuencias de aminoácidos y de polinucleótidos para esta región transmembrana se enumeran en el listado de secuencias como las SEQ ID NO: 7 y 8. Las secuencias de aminoácidos y de polinucleótidos para el fragmento de GAS 40 restante se enumeran en el listado de secuencias como las SEQ ID NO: 9 y 10).

Otros fragmentos pueden omitir uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

La ilustración adicional de los dominios dentro de GAS 40 se muestra en las figuras 1 y 2. Como se muestra en estas figuras, una secuencia de aminoácidos para GAS 40 (SEQ ID NO: 1) contiene una secuencia de péptido líder dentro de los aminoácidos 1-26 (por ejemplo la SEQ ID NO: 3), una primera región con superenrollamiento dentro de los aminoácidos 58 - 261 (SEQ ID NO: 12), una segunda región con superenrollamiento generalmente dentro de los aminoácidos 556 - 733 (SEQ ID NO: 13), una región con cremallera de leucina dentro de los aminoácidos 673 - 701 (SEQ ID NO: 14) y una región transmembrana dentro de los aminoácidos 855 - 866 (SEQ ID NO: 11). La figura 1 representa estas regiones dentro de una secuencia de aminoácidos para GAS 40, mientras que la figura 2 representa estas regiones esquemáticamente a lo largo de la longitud de la proteína GAS 40.

Es probable que las regiones con superenrollamiento identificadas dentro de GAS 40 formen enrollamientos en hélice alfa. Estas estructuras están con frecuencia involucradas en las interacciones de oligomerización, por ejemplo entre diferentes regiones de la proteína o entre regiones de dos proteínas distintas. El motivo cremallera de leucina dentro de la segunda región con superenrollamiento contiene una serie de residuos de aminoácidos de leucina (o isoleucina), espaciados de manera que permitan a la proteína formar una interacción de oligomerización especializada entre dos hélices alfa. En un motivo cremallera de leucina, preferentemente, hay seis residuos de aminoácidos con espacios intermedios entre los residuos de leucina de repetición. En una estructura oligomérica con cremallera de leucina, se cree que las hélices alfa se mantienen unidas por interacciones hidrófobas entre los residuos de leucina, que están situados en un lado de cada hélice. Los motivos cremallera de leucina están con frecuencia implicados en las interacciones de dimerización. La ubicación del motivo cremallera de leucina dentro de la región con superenrollamiento indica además la probabilidad de que esta región de la proteína GAS 40 esté implicada en una interacción de oligomerización.

La figura 2 también ilustra que existe un bajo nivel de homología entre algunas de las regiones identificadas de GAS 40 y otras proteínas estreptocócicas con localización en la superficie o estructuras bidimensionales conocidas o previstas. Tal nivel bajo de homología puede indicar estructuras secundarias similares o incluso una función. Por ejemplo, los aminoácidos 33 a 324 de GAS 40, incluida la primera región con superenrollamiento, tienen aproximadamente un 22% de identidad de secuencia con una región (aminoácidos 112 a 392) de una proteína de *Streptococcus gordonii* llamada precursor de la proteína A de superficie estreptocócica ("SpA") (referencia del Genbank GI 25990270, SEQ ID NO: 15). Se cree que esta proteína es una proteína de superficie adhesina, implicada en la adhesión de ese *Streptococcus* con las membranas celulares del hospedador mamífero. La SpA de *S. gordonii* es un miembro de la familia de proteínas adhesinas antígeno I/II estreptocócico y reconoce la glicoproteína aglutinina salival (gp-340) y el colágeno de tipo I. Los aminoácidos 33 a 258 de GAS 40 también muestran un bajo nivel de identidad de secuencia (23%) con otra proteína de *S. gordonii*, el precursor de la proteína B de superficie estreptocócica (referencia del GenBank GI 25055226, SEQ ID NO: 16).

Una región similar de GAS 40 que también se superpone con la primera región con superenrollamiento (aminoácidos 43 - 238) demuestra un 23% de identidad de secuencia con una región (aminoácidos 43 - 238) de una proteína de *Streptococcus pneumoniae* llamada precursor de la proteína de superficie pspA (referencia del Genbank GI 282335, SEQ ID NO: 17). Se cree que el dominio aminoterminal de pspA es esencial para la virulencia neumocócica completa, y los anticuerpos monoclonales generados contra ella protegen a los ratones contra infecciones neumocócicas. El dominio de pspA tiene una forma monomérica con una relación de conformación axial de aproximadamente 1:12, típica de las proteínas fibrosas. El análisis de secuencias indica una estructura con superenrollamiento en hélice alfa para esta molécula monomérica con sólo algunas interrupciones de tipo bucle en la helicidad.

La segunda región con superenrollamiento de GAS 40 tiene aproximadamente un 46% de identidad de secuencia con una región (aminoácidos 509-717) de una proteína de *Streptococcus equi* llamada proteína inmunorreactiva Se89.9 (referencia del GenBank GI 2330384, SEQ ID NO: 18) (la secuencia de longitud completa para Se89.9 también está disponible en <http://pedant.gsf.de>). Se prevé que esta proteína de *Streptococcus equi* esté expuesta en la superficie. El alineamiento BLAST de cada una de estas secuencias estreptocócicas con GAS 40 se presenta en la figura 3.

La ilustración adicional de la estructura bidimensional de GAS 40 se muestra en la figura 4. En primer lugar, la figura 4(a) presenta el análisis de la estructura secundaria predicha alineada frente a la secuencia de aminoácidos para GAS 40. Las regiones en hélice alfa predichas en la figura 4 generalmente corresponden a las regiones con superenrollamiento indicadas anteriormente. En la figura 4(b), se utiliza la predicción PairCoil para predecir la ubicación de los supuestos bucles superenrollados. En este caso, se identifican dos regiones con enrollamiento, generalmente correspondientes a las regiones con superenrollamiento primera y segunda. La figura 4(c) destaca la región con cremallera de leucina e ilustra los residuos de aminoácidos de leucina (o isoleucina) de repetición regular que es probable participen en la cremallera de leucina.

Por consiguiente, la primera región con superenrollamiento de GAS 40 comprende una secuencia de

aminoácidos de al menos diez (por ejemplo, al menos 10, 13, 15, 18, 20, 25, 30, 35, 40, 50, 70, 90, 100 o más) residuos de aminoácidos consecutivos, seleccionados de la mitad N-terminal de una secuencia de GAS 40 de longitud completa, y que se predice forman un complejo en hélice alfa en base a las características funcionales de los residuos de aminoácidos en la secuencia. La SEQ ID NO: 12 es una primera región con superenrollamiento preferente de GAS 40.

La segunda región con superenrollamiento de GAS 40 comprende una secuencia de aminoácidos de al menos diez (por ejemplo, al menos 10, 13, 15, 18, 20, 25, 30, 35, 40, 50, 70, 90, 100 o más) residuos de aminoácidos consecutivos, seleccionados de la mitad C-terminal de una secuencia de GAS 40 de longitud completa, y que se predice forman un complejo en hélice alfa en base a las características funcionales de los residuos de aminoácidos en la secuencia. La segunda región con superenrollamiento incluye preferentemente un motivo cremallera de leucina. La SEQ ID NO: 13 es una segunda región con superenrollamiento preferente de GAS 40.

Es probable que las regiones con superenrollamiento de GAS 40 estén implicadas en la formación de oligómeros tales como dímeros o trímeros. Tales oligómeros pueden ser homómeros (que contienen dos o más proteínas GAS 40 oligomerizadas juntas) o heterómeros (que contienen una o más proteínas GAS adicionales oligomerizadas con GAS 40). Como alternativa, las regiones primera y segunda con superenrollamiento pueden estar interactuando juntas dentro de la proteína GAS 40 para formar reacciones oligoméricas entre las regiones primera y segunda con superenrollamiento.

Por consiguiente, en una forma de realización, las composiciones de la invención incluyen un antígeno GAS 40 en forma de oligómero. El oligómero puede comprender dos antígenos más de GAS 40 o fragmentos de los mismos, o puede comprender GAS 40 o un fragmento del mismo oligomerizado a un segundo antígeno de GAS. Los fragmentos de GAS 40 preferentes comprenden una secuencia de aminoácidos seleccionada del grupo que consiste en la primera región con superenrollamiento y la segunda región con superenrollamiento. Tales fragmentos de GAS 40 preferentes pueden utilizarse en solitario o en las combinaciones de la invención.

Las secuencias de aminoácidos y de polinucleótidos de GAS de la invención pueden manipularse para facilitar u optimizar la expresión recombinante. Por ejemplo, puede sustituirse la secuencia líder N-terminal con una secuencia que codifica una proteína marcadora tal como polihistidina ("HIS") o glutatión S-transferasa ("GST"). Tales proteínas marcadoras pueden utilizarse para facilitar la purificación, la detección y la estabilidad de la proteína expresada. Las variaciones de tales modificaciones para GAS 40 se analizan más adelante. Tales modificaciones pueden aplicarse a cualquiera de las proteínas GAS de la invención.

Un ejemplo de una secuencia de GAS 40 con un marcador GST y un marcador HIS se indica en el presente documento como "GST 40 HIS". Este constructo incluye una secuencia de GAS 40 en la que se elimina la secuencia líder, se añade al extremo N-terminal una secuencia que codifica el marcador GST y se añade al extremo C-terminal una secuencia que codifica el marcador HIS (utilizando, por ejemplo, un vector pGEXNNH con sitios de restricción NdeI y NotI). Las secuencias de polinucleótidos y de aminoácidos para la región fusionada del marcador GST, la secuencia de GAS 40 y el marcador HIS C-terminal de GST 40 HIS se muestran en las SEQ ID NO: 19 y 20.

Como alternativa, puede utilizarse una secuencia de marcador única. Un ejemplo de una secuencia de GAS 40 con sólo un marcador HIS se indica como "40a-HIS". Este constructo incluye una secuencia de GAS 40 en la que se elimina la secuencia líder N-terminal y el extremo C-terminal que contiene la secuencia transmembrana. En este constructo, se añade al extremo C-terminal la secuencia del marcador HIS (utilizando, por ejemplo, un vector de clonación tal como pET21b+ (Novagen) en los sitios de restricción NdeI y NotI). Las secuencias de polinucleótidos y de aminoácidos para 40a-HIS se muestran en las SEQ ID NO: 21 y 22.

Además de la adición de marcadores de purificación, también puede facilitarse la expresión recombinante optimizando las secuencias codificantes a las más abundantes o accesibles para el hospedador recombinante. Por ejemplo, la secuencia de polinucleótidos AGA codifica un residuo de aminoácido arginina. La arginina también puede ser codificada por la secuencia de polinucleótidos CTG. Este codón CTG resulta preferente para las enzimas de traducción en *E. coli*. En la secuencia de polinucleótidos de 40a-HIS SEQ ID NO: 21, se ha sustituido un CTG C-terminal que codifica la arginina con CGT.

Los siguientes codones están generalmente sobrerrepresentados en *E. coli*: AGA, AGG y CGA. Cuando estos codones se dan en una secuencia de polinucleótidos de GAS, pueden sustituirse con uno de los otros dos codones opcionales que codifican el mismo residuo de aminoácido.

Un total de tres codones ATG se optimizan a CTG en el constructo "40a-RR-HIS", SEQ ID NO: 23 y 24. La SEQ ID NO: 23 también se muestra más adelante, con los codones optimizados subrayados. (Aparte de la optimización de codones adicional, 40a-RR-HIS es idéntico a 40a-HIS.)

SEO ID N: 23

ATGAGTGTAGGCGTATCTCACCAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAAATACTCACGACG
ATAGTTTACCAAAACGAGAAACAATTCAAGAGCGAAAGGCCAATTGATGTCAGTTGAAAAACTCTCAGTCAACAAAA
AGCAGAAGTACAGACAGCTTGCTACCGCTCTGACGAAAAACTACTGCTGAAATCAACCCTTAAAGAGCAGCAAGATAAT
GAACAAAAAGCTTTAACCCTGCACAGAATAATTACATAATACTTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAG
GAGCCGAACATCAAAGAGAGTTAACAGCTACTGAAACAGAGCTTCATAATGCTCAGCAGATCAACATTCAAAGAGAC
TGCATTGTCAGAAACAAAAGCTAGCATTTTCAGCAGAAACTCTGAGCTCAAGATTTAGTGGAACAAAGTCAAAAAGCTCT
GAACAAAATATGCTAAGCTCAATGCTATGATTAGCAATCTCTGAGCTATCACTAAAGAGCTCAAAACGGCTAATGATA
ATACAAAAGCATTAAGCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAATCAAAAAGCTAAAGTTAAAAAGCAAT

GACTGAAGAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCAGAACCTTAGTCGTCTTAAATCCTCAGCTCCG
TCTACTCAAGATAGCATTTGGGGTAATAATACCATGAAGCACCAGGCGTATCCTCTTGAAGAACTTAAAAAATTAG
AAGCTAGTGGTTATATTGGATCAGCTAGTTACATAATTATTACAAAGAGCATGCAGATCAAAATTTGCCAAAGCTAG
TCCAGGTAATCAATTAATCAATACCAAGATATTCCAGCAGATCGTAATCGCTTTGTTGATGCCGATAATTGTGACACCA
GAAGTGCAAAATGAGCTAGCGCAGTTTGCAGCTCACATGATTAATAGTGTAG**GtGtGt**CAATTAGGTCTACCACCAGTTA
CTGTTACAGCAGGATCAAGAGAATTTGCAAGGATTAATTTAGTACCAGCTATAAGAAAACTCATGGTAATACAAGACCATA
ATTGTGCTACGGACAGCCAGGGGATCAGGGCATTTGGTGTGGGCCCTCATGATAAAACTATTATTGAAGACTCTGCC
GGAGCGTCAGGGCTCATTTGAAATGATGATAACATGTACAGAAATATCGTGCTCTTTAACAGATGTGCACTACTGTGAATG
GTATTAAACGTGGTATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAATACATACGGCCATGC
TATTAACCTTTTTTACGTGTGATAGTAATACCAACCTTAATGCGCCTGPTTACCTTGGATTTTCAACCGCAATGTAGGATCT
TTGATAAGAACCTTTGTAATTTTCCAGAGTCTAAACTTGACTTAACCTCAACGCTTTTAATAAGACCCCTATAAAAGCCG
TTGGAAGTACAAAGAAATATGCCCAAAGATAGGCATGTATCTGATACATTGACGGATCAAGGAATAAGTAAGTCT
ATTAGAAAAATCGTTTGTGCGGTATTCATCAAGAAGCTGATATTATGGCAGCCCAAGCTAAAGTAAGTCAACTTCAAGGT
AAATTAGCAAGCAGCTTAAGCAGTCAAGACGCTTAATCTCCAAGTGAGACAATTAATGATACTAAAGGTTCTTTGA
GAACAGAATTACTAGCAGCTAAAGCAAAACAAAGCAAACTCGAAGCTACTCGTGATCAATCTTAGCTAAGCTAGCATC
GTTGAAGCCGCTACTGCACAGACAGAAAGCCTTAGCAGAGCAAGCCGAGCCAGAGTGCAGCAGCTGGTGGCTAAAAA
GCTCATTTGCAATATCTAAGGGACTTTAAATTTGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATTTGATAATACTA
AGCAAGATTTGGCTAAAACTACCTCATCTTTGTAAATGCACAGAAGAGCTTTAGCAGCCTTACAGCTAAACAAAGCAG
CTTAGAAGCTACTATTGCTTACCACAGAACCCAGGTGACTTTGCTTAAACCTTAGCTAACGAAAGGAATATCGCCAC
TTAGACGAAGATATAGCTACTGTGCTGATTTGCAAGTAGCTCCACCTCTTACGGGGCTAAACCCGCTATCATATAGTA
AGATAGATACTACTCCGCTTGTTCAGAAATGGTTAAAGAAACGAAACAACATATTAGAAGCTTCAGCAAGATTAGCTGC
TGAAAATACAAGTTCTGTAGCAGAAAGCGCTTGTGGCCAAACCTCGAAATGGTAGGCAAGTAATGCCATTGTGTCTAAA
ATCACATCTTCGATTACTACGCCCCTCATCTAAGACATCTTATGGCTCAGGATCTTCAACACGAGCAATCTCATTTCTG
ATGTTGATGAAAGTACTCA**GtGtGCGGCGCATCGAGCACCACCAACCACCAACCAC**

La optimización de codones también puede utilizarse sin un marcador de purificación. El constructo "40a-RR-Nat", SEQ ID NO: 25 y 26, proporciona un ejemplo de este tipo. Este constructo comprende GAS 40 sin la secuencia líder N-terminal y la secuencia transmembrana C-terminal, con tres optimizaciones de codones (y no incluye una secuencia del marcador HIS).

Pueden utilizarse diferentes vectores de clonación para optimizar la expresión en diferentes células hospedadoras o en diferentes condiciones de cultivo. Los constructos analizados anteriormente utilizaron el vector pET21b+ (Novagen) que incluye un promotor inducible por IPTG. Como alternativa, puede utilizarse un vector lanzadera de expresión de *E. coli/B. subtilis* tal como pSM214gNH. Este vector utiliza un promotor constitutivo en lugar de un promotor inducible por IPTG. Un ejemplo de un constructo de GAS 40 que utiliza este vector se indica como "HIS-40a-NH", SEQ ID NO: 27 y 28 En este constructo, se eliminan la secuencia líder N-terminal y la secuencia transmembrana C-terminal y se añade un marcador HIS al extremo N-terminal. Con la donación se introducen aminoácidos N-terminales adicionales. Además, se indican dos cambios de nucleótidos que muy probablemente se produjeron durante la PCR - ninguno de estos cambios da como resultado cambios de aminoácidos.

Como otra alternativa, puede utilizarse el vector lanzadera pSM214gCH. Un ejemplo de un constructo de GAS 40 que utiliza este vector se indica como "HIS-40a-CH", SEQ ID NO: 29 y 30. En este constructo, se eliminan la secuencia líder N-terminal y la secuencia transmembrana C-terminal y se coloca el marcador HIS en el extremo C-terminal. También se introducen dos aminoácidos adicionales en el extremo amino terminal. En la secuencia de ADN se muestran tres cambios de nucleótidos introducidos con la clonación, indicándose un cambio de aminoácido resultante en la secuencia de la proteína (del aminoácido F al S).

La optimización de codones también puede utilizarse con estos vectores de clonación alternativos. El constructo de GAS 40 "HIS-40a-RR-NH" comprende el constructo "HIS-40a-NH" con tres optimizaciones de codones. HIS-40a-RR-NH se expone en el listado de secuencias como las SEQ ID NO: 31 y 32.

Por consiguiente, los antígenos de GAS utilizados en la invención pueden producirse por recombinación utilizando constructos de expresión que facilitan su producción recombinante. Las modificaciones de secuencia preferentes para facilitar la expresión pueden seleccionarse del grupo que consiste en (1) la adición de una secuencia de marcador de purificación y (2) la optimización de codones.

Como se ha analizado anteriormente, los solicitantes han identificado GAS 40 como particularmente

adecuado para uso en composiciones inmunógenas, ya sea solo o en combinaciones. El uso de GAS 40 como antígeno de GAS particularmente eficaz es respaldado por su asociación con la virulencia, su localización en la superficie, su eficacia en ensayos de opsonofagocitosis bacteriana y en experimentos de provocación mediante inmunización. Además, la potencial adquisición horizontal de este factor de virulencia indica que este antígeno puede ser específico para GAS (con respecto a otras bacterias estreptocócicas). El respaldo adicional para las propiedades antigénicas de GAS 40 también incluye la identificación de las regiones con superenrollamiento dentro de la estructura bidimensional de GAS 40 y el bajo nivel de homología de estas regiones con las proteínas de superficie de otras bacterias estreptocócicas, incluidas algunas proteínas de adhesión.

El análisis de los solicitantes de la ubicación de GAS 40 dentro del genoma de *Streptococcus pyogenes* indica que este factor de virulencia fue probablemente adquirido por GAS durante la evolución como resultado de una transferencia horizontal de genes. La figura 5A representa GAS 40 dentro del genoma de GAS. Está precedido en el extremo 5' por una secuencia denominada "represor del operón de purina" o "PurR". Va seguido en el extremo 3' de dos secuencias que codifican las proteínas ribosómicas denominadas "proteína ribosómica S12" o "rpsL" y "proteína ribosómica S7" o "rpsG". (Las secuencias de aminoácidos y de polinucleótidos para estos genes flanqueantes están a disposición del público en el GenBank. Las secuencias PurR pueden encontrarse por ejemplo con la referencia del Genbank GI:15674250. Las secuencias rpsL pueden encontrarse por ejemplo con la referencia del Genbank GI:15674250. Las secuencias rpsG pueden encontrarse por ejemplo con la referencia del GenBank G:15674250). Cabe destacar que hay dos supuestas secuencias de promotor indicadas al comienzo de la secuencia rpsL. La figura 5B representa un mutante de GAS en el que una gran porción de GAS 40 está eliminada. La única porción de la secuencia de GAS 40 que queda corresponde a los polinucleótidos 1-97 de la SEQ ID NO: 2. La delección incluía uno de los promotores de rpsL, dejando el segundo, P*, intacto. (Las flechas horizontales subrayan el esquema que indica la región delecionada).

La figura 5C proporciona detalles adicionales sobre la secuencia de GAS de tipo silvestre. En este caso, las secuencias de repetición directa, denominadas "DR", se muestran flanqueando los extremos 5' y 3' de GAS 40. (Las secuencias correspondientes en el mutante por delección de GAS 40 se identifican en la figura 5D). Estas secuencias de repetición directa son aproximadamente 8 pares de bases. Un ejemplo de una repetición directa de pares de bases de este tipo comprende la SEQ ID NO: 136. Tales motivos de secuencia dentro de un genoma bacteriano indican con frecuencia una transferencia horizontal de genes. Los experimentos de infección *in vivo* muestran que el mutante por delección de GAS 40 es varias unidades logarítmicas menos virulento que la cepa de tipo silvestre. (Los detalles de este experimento se proporcionan en el Ejemplo 2).

La combinación de la presencia de las secuencias de repetición directa flanqueantes y la virulencia asociada con GAS 40 sugiere claramente que la secuencia de GAS 40 fue adquirida horizontalmente por *Streptococcus pyogenes* durante la evolución. Cabe destacar que aunque los purR y rpsL relacionados están presentes en las bacterias estreptocócicas relacionadas *Streptococcus agalactiae* y *Streptococcus mutans*, ninguna de estas bacterias son conocidas por tener un homólogo de GAS 40. (La figura 5E representa esquemáticamente la ubicación de purR, rpsL y los homólogos rpsG dentro de *S. agalactiae* (estreptococo del grupo B) y muestra el porcentaje de homología de los homólogos GBS con los equivalentes de GAS. Cabe destacar que los genomas de GBS generalmente sólo poseen una de las secuencias de repetición directa - y no contienen un par de las secuencias de repetición directa que flanquean la secuencia de GAS 40).

La localización en superficie de GAS 40 se ilustra mediante el diagrama FACS presentado en la figura 6. (El análisis de los protocolos referentes al análisis FACS se presenta en el Ejemplo 1). La figura 6 incluye los diagramas FACS para el GAS de tipo silvestre (denominado DSM 2071, un tipo M23 de GAS) y el mutante por delección (denominado DSM 2071Δ40). El desplazamiento de la absorbancia para la cepa de tipo silvestre indica que el GAS 40 es reconocido en la superficie de las bacterias por anticuerpos contra GAS 40 (y que no es reconocido en la superficie del mutante por delección).

La exposición en superficie de GAS 40 se demuestra adicionalmente mediante un ensayo de opsonofagocitosis bacteriana ilustrado en la figura 7 y en el Ejemplo 3. En este ensayo, las cepas de GAS se incuban con inmunosuero y pre-inmunosuero, polimorfonucleares y complemento. (El inmunosuero se genera mediante inmunización de ratones con la proteína GAS indicada). La fagocitosis o el crecimiento de las bacterias se miden logarítmicamente. Las barras positivas del histograma representan la fagocitosis (o la muerte bacteriana). Las barras negativas del histograma representan el crecimiento bacteriano. Como se muestra en la figura 7, los inmunosueños generados por cada una de las proteínas expresadas GAS40 dieron como resultado una reducción de las bacterias (barras positivas del histograma).

Los estudios de provocación mediante inmunización con GAS 40 se analizan detalladamente en el Ejemplo 4. Como se muestra en este ejemplo, GAS 40, tal como se produce utilizando diversos constructos, proporciona una protección considerable en ratones adultos. Cabe destacar que la mayoría de los constructos de GAS40 proporcionan casi tanta protección como la proteína M de GAS. (La proteína M de GAS se utiliza para la comparación, ya que se sabe que es muy inmunógena. Sin embargo, la proteína M no se considera generalmente como un candidato a vacuna contra GAS adecuado, ya que varía ampliamente entre las cepas de GAS y tiene epítopos con reactividad cruzada potencial con tejidos humanos). Además, un fragmento N-terminal de GAS 40

también proporcionó una protección significativa en este modelo. El fragmento N-terminal comprende aproximadamente 292 aminoácidos del extremo N-terminal de GAS 40 que se superpone con la primera región con superenrollamiento. "40N-HIS" (SEQ ID NO: 33 y 34) es un ejemplo de este fragmento de GAS 40 que comprende la región con superenrollamiento de GAS 40 y un marcador HIS C-terminal.

(2) GAS 117

GAS 117 corresponde a los números de registro del GenBank de M1 GI:13621679 y GI:15674571, al número de registro del GenBank de M3 GI:21909852, al número de registro del GenBank de M18 GI:19745578, y también se conoce como 'Spy0448' (M1), 'SpyM3_0316' (M3) y 'SpyM18_0491' (M18). Los ejemplos de secuencias de aminoácidos y de polinucleótidos de GAS 117 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 35 y 36.

Las proteínas GAS 117 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 35; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 35, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas GAS 117 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 35. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 1. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 35. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 35 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 37 comprende la secuencia de aminoácidos N-terminal eliminada. La SEQ ID NO: 38 comprende un fragmento de GAS 117 sin la secuencia de aminoácidos N-terminal). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 35

MTLKHHYLLSLALVTVGAAFNSTQSVSAQVYSNEGYHQHLLTDEKSHLQYSKDNAQLQLRNILDGYQNDLGRHYSSYYY
YNLRTVMGLSSEQDIEKHYEELKNKLHDMYNHY

(3) GAS 130

GAS 130 corresponde a los números de registro del GenBank de M1 GI:13621794 y GI:15674677, al número de registro del GenBank de M3 GI:21909954, al número de registro del GenBank de M18 GI:19745704, y también se conoce como 'Spy0591' (M1), 'SpyM3_0418' (M3) y 'SpyM18_0660' (M18). GAS 130 se ha identificado potencialmente como una supuesta proteasa. Los ejemplos de secuencias de aminoácidos y de polinucleótidos de GAS 130 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 39 y 40.

Las proteínas GAS 130 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 39; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 39, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas GAS 130 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 39. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 39. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 39. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(4) GAS 277

GAS 277 corresponde a los números de registro del GenBank de M1 GI:13622962 y GI:15675742, al número de registro del GenBank de M3 GI:21911206, al número de registro del GenBank de M18 GI:19746852, y también se conoce como 'Spy1939' (M1), 'SpyM3_1670' (M3) y 'SpyM18_2006' (M18). Las secuencias de aminoácidos y de polinucleótidos de GAS 277 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 41 y 42.

Las proteínas GAS 277 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 41; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 41, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas GAS 277 incluyen variantes (por ejemplo, variantes

alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 41. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 41. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 41. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 41 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 43 comprende el aminoácido N-terminal subrayado. La SEQ ID NO: 44 comprende un fragmento de GAS 277 con la secuencia de aminoácidos N-terminal eliminada). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 41

MTTMQKTIISLLSLALLIGLLGTSGKAISVYAQDQHTDNVIAESTISQVSVEASMRGTEPYIDATVTTDQFVRQPTQATIT
LKDASDNTINSWVYTMAAQQRRTAFWDLTGQKSGDYHVTVTVHTQEKAVTGQSGTVHFDQNKARKTPTNMQQKDTSKAM
TNSVDVDTKAQTNQSANQEIDSTSNPFRSATNHRSTSLKRSTKNEKLTFTASNSQKNGSNKTKMLVDKEEVKPTSKRGFP
WVLLGLVVSIAAGLFIAIQVSRK

(5) GAS 236

GAS 236 corresponde a los números de registro del GenBank de M1 GI:13622264 y GI:15675106, al número de registro del GenBank de M3 GI:21910321 y al número de registro del GenBank de M18 GI:19746075, y también se conoce como 'Spy1126' (M1), 'SpyM3_0785' (M3) y 'SpyM18_1087' (M18). Las secuencias de aminoácidos y de polinucleótidos de GAS 236 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 45 y 46.

Las proteínas GAS 236 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 45; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 45, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas GAS 236 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 45. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 45. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 45. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 45 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 47 comprende la secuencia de aminoácidos N-terminal. La SEQ ID NO: 48 comprende un fragmento de GAS 236 con la secuencia N-terminal eliminada). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 45

MTQMNYTGKVKRVIAIANGKYQSKRVASKLFSVFKDDPDFYLSKKNPDIVISIGGDGMLLSAFHMYEKELDKVRFVGIHT
GHLGFYTDYRDFEVDKLIDNLRKDKGEQISYPILKVAITLDDGRVVKARALNEATVKRIEKTADVIVNHVKFESFRGD

GISVSTPTGSTAYNKSLGGAVLHPTIEALQLTEISSLNRRVFRITLGSSIIIPKKDKIELVPKRLGIYITISIDNKTYQLKN
VTKVEYFIDDEKIHVSSPSHTSFWERVKDAFIGEIDS

(6) GAS 389

GAS 389 corresponde a los números de registro del GenBank de M1 GI:13622996 y GI:15675772, al número de registro del GenBank de M3 GI:21911237, al número de registro del GenBank de M18 GI:19746884, y también se conoce como 'Spy1981' (M1), 'SpyM3_1701' (M3), 'SpyM18_2045' (M18) y 'relA'. GAS 389 también se ha identificado como una (p)ppGpp sintetasa. Las secuencias de aminoácidos y de polinucleótidos de GAS 389 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 49 y 50.

Las proteínas GAS 389 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 49; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 49, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250 o más). Estas proteínas GAS 389 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 49. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 49. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 49.

Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(7) GAS 504

GAS 504 corresponde a los números de registro del GenBank de M1 GI:13622806 y GI:15675600, al número de registro del GenBank de M3 GI:21911061, al número de registro del GenBank de M18 GI:19746708, y también se conoce como 'Spy1751' (M1), 'SpyM3_1525', 'SpyM18_1823' (M18) y 'fabK'. GAS 504 también se ha identificado como una supuesta trans-2-enoil-ACP reductasa II. Las secuencias de aminoácidos y de polinucleótidos de GAS 504 de una cepa M1 se exponen más adelante y en el listado de secuencias como las SEQ ID NO: 51 y 52.

Las proteínas GAS 504 preferentes para uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 51; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 51, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas GAS 504 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 51. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 51. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 51. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(8) GAS 509

GAS 509 corresponde a los números de registro del GenBank de M1 GI:13622692 y GI:15675496, al número de registro del GenBank de M3 GI:21910899, al número de registro del GenBank de M18 GI:19746544, y también se conoce como 'Spy1618' (M1), 'SpyM3_1363' (M3), 'SpyM18_1627' (M18) y 'cysM'. GAS 509 también se ha identificado como una supuesta O-acetil-serina liasa. Las secuencias de aminoácidos y de polinucleótidos de GAS 509 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 53 y 54.

Las proteínas GAS 509 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 53; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 53, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas GAS 509 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 53. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 53. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 53. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo C-terminal de la SEQ ID NO: 53 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 55 comprende la secuencia de aminoácidos C-terminal. La SEQ ID NO: 56 comprende un fragmento de GAS 509 con la secuencia C-terminal eliminada). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 53

MTKIYKTITELVGQTPIIKLNRLIPNEAADVYVKLEAFNPGSSVKDRIALSMIEAAEAEGGLISPGDVIIIE
PTSGNTGIGLAWGAAGYRVIIIVPETMSLERRQIIQAYGAELVLTPGAEGMKGAIAKAETLAIELGAW
MEMQFNNPANPSIHEKTTAQEILEAFKEISLDAFVSGVGTGGTSLGVSHVLKKNPETVIYAVEAEESAV
LSGQEPGPHKIQGISAGFIPNTLDTKAYDQIRVKSKDALETARLTGAKEGFLVGISSGAALYAAIEVAK
QLGKGKHLVLTILPDNGERYLSTELYDVPVIKTK

(9) GAS 366

GAS 366 corresponde a los números de registro del GenBank de M1 GI:13622612, GI:15675424 y GI:30315979, al número de registro del GenBank de M3 GI:21910712, al número de registro del GenBank de M18 GI:19746474, y también se conoce como 'Spy1525' (M1), 'SpyM3_1176' (M3), 'SpyM18_1542' (M18) y 'murD'. GAS 366 también se ha identificado como una UDP-N-acetilmuramoylalanina-D-glutamato ligasa o una enzima de adición de ácido D-glutámico. Las secuencias de aminoácidos y de polinucleótidos de GAS 366 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 57 y 58.

Las proteínas GAS 366 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%,

92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 57; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 57, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas GAS 366 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 57. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 57. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 57. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 57 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 59 comprende la secuencia líder N-terminal. La SEQ ID NO: 60 comprende un fragmento de GAS 366 en el que la secuencia N-terminal está eliminada). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 57

MKVISNFFQNKILILGLAKSGEAAAKLLTKLGLVTVNDSKPFDPNPAQALLEEGIKVICGSHPEVLLDENFEYMKVKNP
GIPYDNFPMVRKALAKEIPILTEVELAYFVSEAPIIGITGSNGKTTTMTIADVLNAGGQSALLSGNIGYPASKVVQKAI

GDTLVMESSFLVGVNAFRPHIAVITNLMPTHLDYHGSFEDYVAAKWMIIQAQMTESDYLIINANQEISATLAKTTKATV
IPFSTQKVVDGAYLKDGILYFKEQAIIAATDLGVPGSHNIENALATIAVAKLSGIADDIIAQCLSHFGGVKHLRQVRVQGI
KDITFYNDKSTNIIATQKALSGFDNSRLIIAGGLDRGNEFDLVPDLGLKQMIILGESAERMKRAANKAEVSYLEAR
NVAEATELAFKLAQTGDTILLSPANASWDMYPNFEVRGDEFATFDCLRGDA

(10) GAS 159

GAS 159 corresponde a los números de registro del GenBank de M1 GI:13622244 y GI:15675088, al número de registro del GenBank de M3 GI:21910303, al número de registro del GenBank de M18 GI:19746056, y también se conoce como 'Spy1105' (M1), 'SpyM3_0767' (M3), 'SpyM18_1067' (M18) y 'potD'. GAS 159 también se ha identificado como un supuesto transportador ABC de espermidina/putrescina (una proteína de transporte periplásmico). Las secuencias de aminoácidos y de polinucleótidos de GAS 159 de una cepa M1 se exponen a continuación y en el listado de secuencias como las SEQ ID NO: 61 y 62.

SEQ ID NO: 61

MRKLYSFLAGVLGVIVILTSLSFILQKKSGSGSQSDKLVIVNWGDYIDPALLKKFTKETGIEVQYETFDSDNEAMYTKIKQ
GGTTYDIAVPSDYIDKMIKENLLNKLDKSKLVGMDNIGKEFLGKSFDPQNDYSLPYFWGTGVIYNDQLVDKAPMHWED
LWRPEYKNSIMLIDGAREMLGVGLTTFGYSVNSKNLEQLQAAERKLQQLTPNVKAIVADEMKGMYIQGDAAGITFSGEA
SEMILDSNEHLHYIVPSEGSNLWFDNLVLPKTMKHEKEAYAFNFINRPNENAAQNAAYIGYATPNKKAKALLPDEIKNDPA
FYPTDDI IKKLEVDNLGSRWNLGIYNDLYLQFKMYRK

SEQ ID NO: 62

ATGCGTAAACTTTTATCCTTTCTAGCAGGAGTTTTGGGTGTTATTGTTATTTTAAACAAGTCTTTCTTCTATCTTGCAGAA
AAAATCGGGTTCTGGTAGTCAATCGGATAAATTAGTTATTTTAACTGGGGAGATTACATTGATCCAGCTTGGCTCAAAA
AATTACCAAAGAAACGGGCATTGAAGTGCAGTATGAACTTTTCGATTCCAATGAAGCCATGTACACTAAAATCAAGCAG
GGCGGAACCACTTACGACATTGCTGTTCTAGTGATTACACCATGTATAAAATGATCAAGAAAAACCTACTCAATAAGCT
TGATAAGTCAAAATAGTTGGCATGGATAAATATCGGAAAGAAATTTTAGGGAAGAGCTTTGACCCACAAAACGACTATT
CTTTGCCCTTATTTCTGGGGAACCGTTGGGATTGTTTATAATGATCAATTAGTTGATAAGGCGCCTTGCCTGGGAAGAT
CTGTGGCGTCCAGAAATATAAAATAGTATTATGCTGATTGATGGAGCGCGTGAAATGCTAGGGGTTGGTTTAACTTT
TGGTTATAGTGTGAATTTCAAAAATCTAGAGCAGTTGCAGGCAGCCGAGAGAAAACTGCAGCAGTTGACGCCGAATGTGA
AAGCCATTGTAGCAGATGAGATGAAAGGCTACATGATTCAAGGTGACGCTGCTATTGGAATTACCTTTTCTGGTGAAGCC
AGTGAGATGTAGATAGTAACGAACACCTTCACTACATCGTGCCTTCAGAAAGGCTTAACCTTTGGTTTGATAATTGGT
ACTACCAAAAACCATGAAACAGGAAAAAGAGCTTATGCTTTTGAACCTTTATCAATCGTCTGAAAAATGCTGCGCAAA
ATGCTGCATATATTGGTTATGCGACACCAATAAAAAAGCCAAAGGCTTACTTCCAGATGAGATAAAAAATGATCCTGCT
TTTTATCCAACAGATGACATTATCAAAAATTTGAAGTTTATGACAATTTAGGGTCAAGATGGTTGGGGATTTATAATGA
TTTATACCTCCAATTAAATGTATCGCAAATAA

Las proteínas GAS 159 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 61; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 61, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas GAS 159 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 61. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 61. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 61. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 61 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 63 comprende la secuencia de aminoácidos líder N-terminal. La SEQ ID NO: 64 comprende un fragmento de GAS 159 en el que la secuencia de aminoácidos líder N-terminal está eliminada). En otro ejemplo, la secuencia de aminoácidos subrayada en el extremo C-terminal de la SEQ ID NO: 61 está eliminada. (La SEQ ID NO: 65 comprende la región hidrófoba C-terminal. La SEQ ID NO: 66 comprende un

fragmento de GAS 159 en el que la región hidrófoba C-terminal está eliminada. La SEQ ID NO: 67 comprende un fragmento de GAS 159 en el que la secuencia líder N-terminal y la región hidrófoba C-terminal están eliminadas). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 61

```
MRKLYSFLAGVLGVIVILTSLSFILQKKSGSGSQSDKLVINWGDYIDPALLKKFTKETGIEVQYETFDSEAMYTKIKQ
GGTTYDIAVPSDYITDKMIKENLLNKLDKSKLVGMDNIGKEFLGKSFDPQNDYSLPYFWGTGIVYNDQLVDKAPMHWED
LWRPEYKNSIMLIDGAREMLGVGLTTFGYSVNSKNLEQLQAAERKQLQTPNVKAIVADEMKGYMIQGDAAIGITFSGEA
SEMLDSNEHLHYIVPSEGSNLWFDNLVLPKTMKHEKEAYAFNLNFINRPENAAQNAAYIGYATPNKKAKALLPDEIKNDPA
FYPTDDI IKKLEVYDNLGSRWLGIYNDLYLQFKMYRK
```

(11) GAS 217

GAS 217 corresponde a los números de registro del GenBank de M1 GI:13622089 y GI:15674945, al número de registro del GenBank de M3 GI:21910174, al número de registro del GenBank de M18 GI:19745987, y también se conoce como 'Spy0925' (M1), 'SpyM3_0638' (M3) y 'SpyM18_0982' (M18). GAS 217 también se ha identificado como una supuesta oxidoreductasa. Las secuencias de aminoácidos y de polinucleótidos de GAS 217 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 68 y 69.

Las proteínas GAS 217 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 68; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 68, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas GAS 217 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 68. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 68. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 68. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(12) GAS 309

GAS 309 corresponde a los números de registro del GenBank de M1 GI:13621426 y GI:15674341, al número de registro del GenBank de M3 GI:21909633, al número de registro del GenBank de M18 GI:19745363, y también se conoce como 'Spy0124' (M1), 'SpyM3_0097' (M3), 'SpyM18_0205' (M18), 'nra' y 'rofA'. GAS 309 también se ha identificado como una proteína reguladora y un regulador negativo de la transcripción. Las secuencias de aminoácidos y de polinucleótidos de GAS 309 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 70 y 71.

Las proteínas GAS 309 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 70; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 70, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas GAS 309 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 70. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 70. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 70. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(13) GAS 372

GAS 372 corresponde a los números de registro del GenBank de M1 GI:13622698 y GI:15675501, al número de registro del GenBank de M3 GI:21910905, al número de registro del GenBank de M18 GI:19746500, y también se conoce como 'Spy1625' (M1), 'SpyM3_1369' (M3) y 'SpyM18_1634' (M18). GAS 372 también se ha identificado como una supuesta proteína quinasa o una supuesta serina/treonina quinasa de tipo eucariota. Las secuencias de aminoácidos y de polinucleótidos de GAS 372 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 72 y 73.

Las proteínas GAS 372 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 72; y/o (b) que es un fragmento de al

menos n aminoácidos consecutivos de la SEQ ID NO: 72, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250 o más). Estas proteínas GAS 372 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 72. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 72. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 72. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

14) GAS 039

GAS 039 corresponde a los números de registro del GenBank de M1 GI:13621542 y GI:15674446, al número de registro del GenBank de M3 GI:21909730, al número de registro del GenBank de M18 GI:19745398, y también se conoce como 'Spy0266' (M1), 'SpyM3_0194' (M3) y 'SpyM18_0250' (M18). Las secuencias de aminoácidos y de polinucleótidos de GAS 039 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 74 y 75.

Las proteínas GAS 039 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 74; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 74, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas GAS 039 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 74. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 74. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 74. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(15) GAS 042

GAS 042 corresponde a los números de registro del GenBank de M1 GI:13621559 y GI:15674461, al número de registro del GenBank de M3 GI:21909745, al número de registro del GenBank de M18 GI:19745415, y también se conoce como 'Spy0287' (M1), 'SpyM3_0209' (M3) y 'SpyM18_0275' (M18). Las secuencias de aminoácidos y de polinucleótidos de GAS 042 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 76 y 77.

Las proteínas GAS 042 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 76; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 76, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas GAS 042 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 76. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 76. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 76. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(16) GAS 058

GAS 058 corresponde a los números de registro del GenBank de M1 GI:13621663 y GI:15674556, al número de registro del GenBank de M3 GI:21909841, al número de registro del GenBank de M18 GI:19745567, y también se conoce como 'Spy0430' (M1), 'SpyM3_0305' (M3) y 'SpyM18_0477' (M18). Las secuencias de aminoácidos y de polinucleótidos de GAS 058 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 78 y 79.

Las proteínas GAS 058 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 78; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 78, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas GAS 058 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 78. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 78. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 78.

Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 78 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 80 comprende la secuencia líder N-terminal. La SEQ ID NO: 81 comprende un fragmento de GAS 58 en el que la secuencia líder N-terminal está eliminada). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de una señal péptido, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 78

MKWSGFMKTKSKRFLNLATLCLALLGTTLLMAHPVQAEVISKRDYMTFRFGLGDLEDDSSANYPSNLEARYKGYLEGYKGL
KGDDIPERPKIQVPEDVQPSDHGDYRDGYEEGFGEGQHKRDPLETEAEDDSQGGRQEGRQGHQEGADSSDLNVEESDGLS
VIDEVGVIIYQAFSTIWTYLSGLF

(17) GAS 290

GAS 290 corresponde a los números de registro del GenBank de M1 GI:13622978 y GI:15675757, al número de registro del GenBank de M3 GI:21911221, al número de registro del GenBank de M18 GI:19746869, y también se conoce como 'Spy1959' (M1), 'SpyM3_1685' (M3) y 'SpyM18_2026' (M18). Las secuencias de aminoácidos y de polinucleótidos de GAS 290 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 82 y 83.

Las proteínas GAS 290 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 82; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 82, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas GAS 290 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 82. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 82. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 82. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(18) GAS 511

GAS 511 corresponde a los números de registro del GenBank de M1 GI:13622798 y GI:15675592, al número de registro del GenBank de M3 GI:21911053, al número de registro del GenBank de M18 GI:19746700, y también se conoce como 'Spy1743' (M1), 'SpyM3_1517' (M3), 'SpyM18_1815' (M18) y 'accA'. Las secuencias de aminoácidos y de polinucleótidos de GAS 511 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 84 y 85.

Las proteínas GAS 511 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 84; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 84, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas GAS 511 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 84. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 84. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 84. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(19) GAS 533

GAS 533 corresponde a los números de registro del GenBank de M1 GI:13622912 y GI:15675696, al número de registro del GenBank de M3 GI:21911157, al número de registro del GenBank de M18 GI:19746804, y también se conoce como 'Spy1877' (M1), 'SpyM3_1621' (M3), 'SpyM18_1942' (M18) y 'glnA'. GAS 533 también se ha identificado como una supuesta glutamina sintetasa. Las secuencias de aminoácidos y de polinucleótidos de GAS 533 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 86 y 87.

Las proteínas GAS 533 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 86; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 86, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 533 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 86. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 86. Otros fragmentos preferentes carecen de uno o

más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 86. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(20) GAS 527

GAS 527 corresponde a los números de registro del GenBank de M1 GI:13622332, GI:15675169 y GI:24211764, al número de registro del GenBank de M3 GI:21910381, al número de registro del GenBank de M18 GI:19746136, y también se conoce como 'Spy1204' (M1), 'SpyM3_0845' (M3), 'SpyM18_1155' (M18) y 'guaA'. GAS 527 también se ha identificado como una supuesta GMP sintetasa (hidrolítica del glutamato) (glutamato amidotransferasa). Las secuencias de aminoácidos y de polinucleótidos de GAS 527 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 88 y 89.

Las proteínas GAS 527 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 88; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 88, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 527 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 88. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 88. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 88. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(21) GAS 294

GAS 294 corresponde a los números de registro del GenBank de M1 GI:13622306, GI:15675145 y GI:26006773, al número de registro del GenBank de M3 GI:21910357, al número de registro del GenBank de M18 GI:19746111, y también se conoce como 'Spy1173' (M1), 'SpyM3_0821' (M3), 'SpyM18_1125' (M18) y 'gid'. GAS 294 también se ha identificado como una supuesta proteína de división inhibida por glucosa. Las secuencias de aminoácidos y de polinucleótidos de GAS 294 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 90 y 91.

Las proteínas GAS 294 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 90; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 90, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 294 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 90. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 90. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 90. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(22) GAS 253

GAS 253 corresponde a los números de registro del GenBank de M1 GI:13622611, GI:15675423 y GI:21362716, al número de registro del GenBank de M3 GI:21910711, al número de registro del GenBank de M18 GI:19746473, y también se conoce como 'Spy1524' (M1), 'SpyM3_1175' (M3), 'SpyM18_1541' (M18) y 'murG'. GAS 253 también se ha identificado como una supuesta undecaprenil-PP-MurNAc-pentapéptido-UDPGlcNAc GlcNAc transferasa. Las secuencias de aminoácidos y de polinucleótidos de GAS 253 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 92 y 93.

Las proteínas GAS 253 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 92; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 92, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 253 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 92. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 92. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 92. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un

dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(23) GAS 529

GAS 529 corresponde a los números de registro del GenBank de M1 GI:13622403, GI:15675233 y GI:21759132, al número de registro del GenBank de M3 GI:21910446, al número de registro del GenBank de M18 GI:19746203, y también se conoce como 'Spy1280' (M1), 'SpyM3_0910' (M3), 'SpyM18_1228' (M18) y 'glmS'. GAS 529 también se ha identificado como una supuesta L-glutamina-D-fructosa-6-fosfato aminotransferasa (glucosamina-6-fosfato sintasa). Las secuencias de aminoácidos y de polinucleótidos de GAS 529 de una cepa M1 se exponen más adelante y en el listado de secuencias como las SEQ ID NO: 94 y 95.

Las proteínas GAS 529 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 94; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 94, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 529 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 94. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 94. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 94. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(24) GAS 045

GAS 045 corresponde al número de registro del GenBank de M3 GI:21909751, al número de registro del GenBank de M18 GI:19745421, y se conoce como 'SpyM3_0215' (M3), 'SpyM18_oppA' (M18) y 'oppA'. GAS 045 se ha identificado como una oligopéptido permeasa. Las secuencias de aminoácidos y de polinucleótidos de GAS 045 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 96 y 97.

Las proteínas GAS 045 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 96; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 96, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 045 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 96. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 96. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 96. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 96 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 98 comprende la secuencia líder N-terminal subrayada. La SEQ ID NO: 99 comprende un fragmento de GAS 45 en el que la secuencia líder N-terminal está eliminada). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 96

VTFMKSKWLAASVSAIVSALAACGNKNASGGSEATKTYKYVFNNDPKSLDYILTNGGGTTDVITQMVDGLENDEYGNLVPISLAKDWKSKDGLTYTYTLRDGVSWYTADGEEYAPVTAEDFVTGLKHAVDDKSDALYVVEDSIKNLKAYQNGEVDFKEVGVKALDDKTQYTLNKPESYWNSTKTYSVLFPVNAKFLKSKGKDFGTTDPSSILVNGAYFLSAFTSKSSMEFHKNENYWDKNAVIGIESVKLTYS DGSDPGSFYKNFDKGEFSVARLYPNDDPTYS AKKNYADNITYGMLTGDIRHLTWNLNRTSFKN TKKDPAQDQAGKKALNNKDFRQAIQFAFDRASFQAQTAGQDAKTALRNMLVPPTFTVTIGESDFGSEVEKEMAKLGDEWK DVLNLAQDQGFYNPEKAKAEFAKAKEALTAEGVTFFVQLDYPVDQANAATVQEAQSFQSVESLKGENVINVLETETSTHEAQGFYAETPEQQDYDISSWGWPDYQDPRTYLDIMSPVGGGSGVIQKLGKAGQNKDVAAAGLDTYQTLTDEAAAIT DDNDARYKAYAKAQAYLTDNAVDIPVVALGGTPRVTKAVPFSGGFSWAGSKGPLAYKGMKLQDKPVTVKQYEKAKEKWMK AKAKSNAKYAELADHVEK

(25) GAS 095

GAS 095 corresponde a los números de registro del GenBank de M1 GI:13622787 y GI:15675582, al número de registro del GenBank de M3 GI:21911042, al número de registro del GenBank de M18 GI:19746634, y también se conoce como 'Spy1733' (M1), 'SpyM3_1506' (M3), 'SpyM18_1741' (M18). GAS 095 también se ha identificado como un supuesto regulador de la transcripción. Las secuencias de aminoácidos y de polinucleótidos de GAS 095 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 100 y 101.

Las proteínas GAS 095 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 100; y/o (b) que es un fragmento de

al menos n aminoácidos consecutivos de la SEQ ID NO: 100, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 095 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 100. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 100. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 100. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 100 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 102 comprende la secuencia de aminoácidos de la secuencia líder N-terminal subrayada. La SEQ ID NO: 103 comprende un fragmento de GAS 95 en el que la secuencia líder N-terminal está eliminada). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 100

```

MKIGKKIVLMFTAIVLTTVLALGVYLTSAITFSTGELSKTFKDFSTSSNKSDAIKQTRAFSILLMGVDTGSSERASKWEG
NSDSMILVTVPKTKKTTMTSLERDTLTTLTGPKNNEMNGVEAKLNAAYAAGGAQMAIMTVQDLLNITIDNYVQINMQGL
IDLNAVGGITVTNEFDFFISIAENEPEYQATVPAGTHKINGEQALVYARMRYDDPEGDYGRQKRQREVIQVLKKILAL
DSISSYRKILSAVSSNMQTNIEISSRTIPSLGLGYRDLRTIKTYQLKGEDATLSDGGSYQIVTSNHLLEIQNRIRTELGL
HKVNQLKTNATVYENLYGSKTSQTVNNNYDSSGQAPSYSDDSHSSYANYSSGVDTGQSASTDQDSTASSHRPATPSSSSDA
LADESSSSSGSGLVPPANINPQT

```

(26) GAS 193

GAS 193 corresponde a los números de registro del GenBank de M1 GI:13623029 y GI:15675802, al número de registro del GenBank de M3 GI:21911267, al número de registro del GenBank de M18 GI:19746914, y también se conoce como 'Spy2025' (M1), 'SpyM3_1731' (M3), 'SpyM18_2082' (M18) e 'isp'. GAS 193 también se ha identificado como un precursor de la proteína secretada inmunógena. Las secuencias de aminoácidos y de polinucleótidos de GAS 193 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 104 y 105.

Las proteínas GAS 193 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 104; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 104, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 193 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 104. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 104. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 104. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(27) GAS 137

GAS 137 corresponde a los números de registro del GenBank de M1 GI:13621842, GI:15674720 y GI:30173478, al número de registro del GenBank de M3 GI:21909998, al número de registro del GenBank de M18 GI:19745749, y también se conoce como 'Spy0652' (M1), 'SpyM3_0462' y 'SpyM18_0713' (M18). Las secuencias de aminoácidos y de polinucleótidos de GAS 137 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 106 y 107.

Las proteínas GAS 137 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 106; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 106, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 137 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 106. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 106. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 106. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(28) GAS 084

GAS 084 corresponde a los números de registro del GenBank de M1 GI:13622398 y GI:15675229, al número de registro del GenBank de M3 GI:21910442, al número de registro del GenBank de M18 GI:19746199, y

también se conoce como 'Spy1274' (M1), 'SpyM3_0906' y 'SpyM18_1223' (M18). GAS 084 también se ha identificado como un supuesto transportador ABC de aminoácidos/proteína periplásmica de unión de aminoácidos. Las secuencias de aminoácidos y de polinucleótidos de GAS 084 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 108 y 109.

Las proteínas GAS 084 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 108; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 108, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 084 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 108. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 108. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 108. Por ejemplo, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 108 (que se muestra a continuación) puede eliminarse. (La SEQ ID NO: 110 comprende una secuencia de aminoácidos que comprende la secuencia líder N-terminal subrayada de GAS 84. La SEQ ID NO: 111 comprende un fragmento de GAS 84 en el que la secuencia líder N-terminal está eliminada). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 108

MIKKRTVAIIAIASSFFLVACQATKSLKSGDAWGVYQKQKSITVGFNTFVPMGYKDESGRCKGFDIDLAKVEVFHQYGL
KVNFAQINWDMKEAELNNGKIDVIWNGYSITKERQDKVAFTDSYMRNEQIIVVKRSDIKTISDMKHKVLGAQSASSGYD
SLLRTPKLLKDFIKNKDANQYETFTQAFIDLKSDRIDGILIDKVYANYLAKEGQLENYRMIPTTFENEAFSVGLRKEDK
TLQAKINRAFRVLYQNGKFQAISEKWFQDDVATANIKS

(29) GAS 384

GAS 384 corresponde a los números de registro del GenBank de M1 GI:13622908 y GI:15675693, al número de registro del GenBank de M3 GI:21911154, al número de registro del GenBank de M18 GI:19746801, y también se conoce como 'Spy1874' (M1), 'SpyM3_1618' (M3) y 'SpyM18_1939' (M18). GAS 384 también se ha identificado como una supuesta glicoproteína endopeptidasa. Las secuencias de aminoácidos y de polinucleótidos de GAS 384 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 112 y 113.

Las proteínas GAS 384 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 112; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 112, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 384 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 112. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 112. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 112. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(30) GAS 202

GAS 202 corresponde a los números de registro del GenBank de M1 GI:13622431 y GI:15675258, al número de registro del GenBank de M3 GI:21910527, al número de registro del GenBank de M18 GI:19746290, y también se conoce como 'Spy1309' (M1), 'SpyM3_0991' (M3), 'SpyM18_1321' (M18) y 'dltD'. GAS 202 también se ha identificado como una supuesta proteína de membrana externa. Las secuencias de aminoácidos y de polinucleótidos de GAS 202 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 114 y 115.

Las proteínas GAS 202 preferentes para su uso con la descripción comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 114; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 114, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más). Estas proteínas GAS 202 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 114. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 114. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO:

114. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(31) GAS 057

GAS 057 corresponde a los números de registro del GenBank de M1 GI:13621655 y GI:15674549, al número de registro del GenBank de M3 GI:21909834, al número de registro del GenBank de M18 GI:19745560, y también se conoce como 'Spy0416' (M1), 'SpyM3_0298' (M3), 'SpyM18_0464' (M18) y 'prtS'. GAS 057 también se ha identificado como una supuesta proteinasa de la envoltura celular. Las secuencias de aminoácidos y de polinucleótidos de GAS 057 de una cepa M1 se exponen en el listado de secuencias como las SEQ ID NO: 116 y 117.

Las proteínas GAS 057 preferentes para su uso con la invención comprenden una secuencia de aminoácidos: (a) que tiene un 95% o más identidad (por ejemplo, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 116; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 116, donde n es 200 o más. Estas proteínas GAS 057 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 116. Los fragmentos preferentes de (b) comprenden un epitopo de la SEQ ID NO: 116. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 116. Por ejemplo, en una forma de realización, la secuencia de aminoácidos subrayada en el extremo N-terminal de la SEQ ID NO: 116 (que se muestra a continuación) está eliminada. (La SEQ ID NO: 118 comprende la secuencia líder N-terminal subrayada. La SEQ ID NO: 119 comprende un fragmento de GAS 57 en el que la secuencia líder N-terminal está eliminada). En otro ejemplo, la secuencia de aminoácidos subrayada en el extremo C-terminal de la SEQ ID NO: 116 está eliminada. (La SEQ ID NO: 120 comprende la región hidrófoba C-terminal subrayada. La SEQ ID NO: 121 comprende un fragmento de GAS 57 en el que la región hidrófoba C-terminal está eliminada. La SEQ ID NO: 122 comprende un fragmento de GAS 57 en el que la secuencia líder N-terminal y la región hidrófoba C-terminal están eliminadas). Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

SEQ ID NO: 116

MEKKQRFSLRKYKSGTFSVLIGSVFLVMTTVADELSTMSSEPTITNHAQQQAQHLTNTLSSAESKSQDTSQITLKTNR
EKEQSQDLVSEPTTTELADTDAASMAN TGSDATQKSASLPVNTDVHDVWTKGAWDKGYKGQGVVAVIDTGIDPAHQ
MRISDVSTAKVSKEDMLARQKAAGINYGSWINDKVVFAHNYVENSNDNIKENQFEDFDEWENFEFDAEAEPKAIKKHKI
YRPQSTQAPKETVIKTEETDGSHDIDWTQTD DDTKYESHGMHVTGIVAGNSKEAAATGERFLGIAPEAQVMFMRV FANDI
MGSAESLFKAIKEDAVAGADVINLSLGTANGAQLSGSKPLMEAIKAKKAGVSVVVAAGNERVYGS DHDDEPLATNP DYG
LVGSPSTGRTPTSVAAINS KWIQRLMTVKEL ENRADLNHGKAIYSESVD FKDIKDSLGYDKSHQFAYVKEST DAGYNAQ
DVKGKIALIERDPNKTYDEMI ALAKKHGALGVLI FNNKPGQSNRSMRLTANGMGIPSAFISHEFGKAMSQ LINGNGTGSLE
FDSVVS KAPSQKGNEMNHF SNWGLTSDGYL KPDITAPGGDIYSTYNDNH YGSQTGTSMAS PQIAGASLLVKQYLEKTQPN
LPKEKIADIVKNLLMSNAQIHVNPETKTTSPRQQGAGLLNIDGAVTSGLYVTGKDN YGSI SLGNITDTMTFDVTVHNLS
NKDKTLRYDTELLTDHVD PQGRFTLTSHSLKTYQGGEVTPANGKVTVRVTMDVSQFTKELTKQMPNGYYLEGFVRFRD
SQDDQLNRVNIPFVGFKGQFENLAVAEESIYRLKSQKGTGFYFDES GPKDDIYVGKHFTGLVTLGSETNVSTKTI SDNGL
HTLGT FKNADGKFI LEKNAQGNPVLAI SPNGDNNQDFAAFKGVFLRKYQGLKASVYHASDKEHKNPLWVS PESFKGDKNF
NSDIRFAKSTLLGTAFSGKSLTGAELPDGHYHYVVSYPDVVGAKRQEMTFDMILDRQKPVLSQATFDPETNRFKPEPL
KDRGLAGVRKDSVFYLERKDNKPYTVTINDSYKYVSVEDNKT FVERQADGSFILPLDKAKLGDFYVMVEDFAGNVAIAKL
GDHLPQTLGKTPIKLKLTDGNYQTKETLKNLEMTQSDTGLVTNQAQLAVVHRNQPSQLTKMNQDFFISPNE DGNKDFV
AFKGLKNNVYNDLT VNVYAKDDHQKQTP IWSQAGASVSAIESTAWYGITARGSKVMPGDYQYVVYTYRDEHGKEHQKYT
ISVNDKKPMITQGRFDTINGVDHFTPDKTKALDSSGIVREEVFYLA KKNRKFVDVTEGKDGITVSDNKVYIPKNPDG SYT
ISKRDGVTLSDYLYLVEDRAGNVSFATL RDLKAVGKDKAVNFGLDLPVPEDKQIVNFTYLVRDADGKPIENLEY YNNSG
NSLILPYGKYTVELLTYDTNAAKLES DKIVSFTLSADNNFQQVTFKITMLATSQITAHFDHLLPEGSRVSLKTAQDQLIP
LEQSLYVPKAYGKTVQEGTYEVVVS LPKGYRIEGNTKVNLTLPNEVHELRLRVKVGDA SDSTGDHKVMSKNNSQALTASA
TPTKSTTSATAKALPSTGEKMGLKLRIVGLVLLGLTQVFSRKKSTKD

Los ejemplos representativos de inmunización con antígenos de GAS de la descripción en el modelo murino (en ratones) analizado anteriormente se resumen en la figura 8. La primera columna identifica el antígeno de GAS utilizado en el experimento. En algunos casos se hace referencia en esta lista a aspectos de la purificación. Además, las modificaciones en la secuencia de polinucleótidos que se han realizado para facilitar la expresión recombinante del antígeno se indican en el gráfico con las siguientes anotaciones: "a" indica que las regiones hidrófobas N o C-terminales se han eliminado; RR indica la optimización de codones; "NH" y "CH" corresponden a los vectores de expresión similares a los indicados en los ejemplos de constructo de GAS 40. Cuando se proporciona un valor p, éste se calcula en base a los valores de terminación de HIS de control en la parte inferior del gráfico.

Los ratones inmunizados con GAS 40 produjeron tasas de supervivencia sustancialmente mejoradas en la provocación - en un grupo de más de 100 inmunizaciones de ratones, la inmunización con GAS 40 produjo más de un 50% de supervivencia. Los demás antígenos de GAS en el gráfico ofrecieron una cantidad de protección que, por ejemplo si se combina con GAS 40, podría ofrecer una protección mejorada.

La inmunogenicidad de otros antígenos de GAS conocidos puede mejorarse mediante la combinación con dos o más GAS del primer grupo de antígenos. Tales otros antígenos de GAS conocidos incluyen un segundo grupo de antígenos que consiste en (1) una o más variantes de la proteína M de superficie o fragmentos de la misma, (2) proteína de unión a fibronectina, (3) proteína estreptocócica asociada al grupo hemo, o (4) SagA. Estos antígenos se denominan en el presente documento "segundo grupo de antígenos".

En el presente documento se describe una composición inmunógena que comprende una combinación de antígenos de GAS, consistiendo dicha combinación en dos a treinta y un antígenos de GAS del primer grupo de antígenos y uno, dos, tres o cuatro antígenos de GAS del segundo grupo de antígenos. Preferentemente, la combinación consiste en tres, cuatro, cinco, seis, siete, ocho, nueve o diez antígenos de GAS del primer grupo de antígenos. Aún más preferentemente, la combinación consiste en tres, cuatro o cinco antígenos de GAS del primer grupo de antígenos. Preferentemente, la combinación de antígenos de GAS incluye uno o ambos de GAS 40 y GAS 117. Preferentemente, la combinación de antígenos de GAS incluye una o más variantes de la proteína M de superficie.

Cada uno de los antígenos de GAS del segundo grupo de antígenos se describe con más detalle a continuación.

(1) Proteína M de superficie

La proteína M es un factor de virulencia de GAS que se ha asociado con la colonización y la resistencia a la fagocitosis. Se han identificado más de 100 variantes de tipo diferente de la proteína M en base a la especificidad antigénica y se cree que la proteína M es la causa principal del cambio antigénico y de la deriva antigénica en GAS. La proteína M también se une al fibrinógeno del suero y bloquea la unión del complemento a los peptidoglicanos subyacentes. Se cree que esta acción aumenta la supervivencia de GAS dentro de un hospedador mamífero inhibiendo la fagocitosis.

Lamentablemente, la proteína M de GAS contiene algunos epítomos que imitan a los de los músculos y el tejido conectivo de mamíferos. Determinadas proteínas M de GAS pueden ser reumatogénicas ya que contienen epítomos relacionados con el músculo cardíaco, y puede conducir a la carditis reumática autoinmunitaria (fiebre reumática) después de una infección aguda.

Se han identificado en la región amino terminal epítomos con mayor actividad bactericida y con menor probabilidad de reacción cruzada con los tejidos humanos y se han combinado en proteínas de fusión que contienen aproximadamente seis, siete u ocho fragmentos de proteína M unidos en tándem. Véanse Hu *et al.*, *Infection & Immunity* (2002) 70(4):2171-2177; Dale, *Vaccine* (1999) 17:193-200; Dale *et al.*, *Vaccine* 14(10):944-948; los documentos WO 02/094851 y WO 94/06465.

Por consiguiente, las composiciones de la invención pueden comprender adicionalmente una proteína M de superficie de GAS o un fragmento o derivado de la misma. Pueden combinarse juntos, en una proteína de fusión, uno o más fragmentos de proteína M de superficie de GAS. Como alternativa, uno o más fragmentos de proteína M de superficie de GAS se combinan con un antígeno de GAS o fragmento del mismo del primer grupo de antígenos. Un ejemplo de una proteína M de GAS se expone en el listado de secuencias como la SEQ ID NO: 123.

Las proteínas M de GAS preferentes para su uso con la invención comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con una proteína M conocida tal como la SEQ ID NO: 123; y/o (b) que es un fragmento de al menos *n* aminoácidos consecutivos de una proteína M conocida tal como la SEQ ID NO: 123, donde *n* es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150 o más). Estas proteínas M de GAS incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 123. Los fragmentos preferentes de (b) comprenden un epítomo de una proteína M conocida, tal como la SEQ ID NO: 123. Preferentemente, el fragmento es uno de los descritos en las referencias anteriores. Preferentemente, el fragmento se construye en una proteína de fusión con uno o más fragmentos adicionales de proteína M. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de una proteína M conocida, tal como la SEQ ID NO: 123. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(2) Proteína de unión a fibronectina

La proteína de unión a fibronectina ('Sfbl') de GAS es una proteína bacteriana multifuncional que se cree interviene en la fijación de las bacterias a las células hospedadoras, facilita la internalización bacteriana en las células y se une al fragmento Fc de la IgG humana, interfiriendo así con la fagocitosis mediada por el receptor Fc y la citotoxicidad celular dependiente de anticuerpos. La inmunización de ratones con Sfbl y un "fragmento H12" (codificado por las posiciones 1240 - 1854 del gen Sfbl) se analiza en Schulze *et al.*, *Vaccine* (2003) 21:1958 -1964;

Schulze *et al.*, Infection and Immunity (2001) 69(1): 622-625 y Guzmán *et al.*, Journal of Infectious Diseases (1999) 179:901 – 906. Un ejemplo de una secuencia de aminoácidos para SfbI de GAS se muestra en el listado de secuencias como la SEQ ID NO: 124.

Las proteínas SfbI preferentes para su uso con la invención comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 124; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 124, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas SfbI incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 124. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 124. Preferentemente, el fragmento es uno de los descritos en las referencias anteriores. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 124. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(3) Proteína estreptocócica asociada al grupo hemo

Se ha identificado la proteína estreptocócica asociada al grupo hemo ('Shp') de GAS como una proteína de superficie celular de GAS. Se cree que se cotranscribe con genes que codifican homólogos de un transportador ABC implicado en la absorción de hierro en las bacterias gramnegativas. La proteína Shp se describe adicionalmente en Lei *et al.*, "Identification and Characterization of a Novel Heme-Associated Cell Surface Protein Made by Streptococcus pyogenes", Infection and Immunity (2002) 70(8):4494 – 4500. Un ejemplo de una proteína Shp se muestra en el listado de secuencias como la SEQ ID NO: 125.

Las proteínas Shp preferentes para su uso con la invención comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 125; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 125, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 o más). Estas proteínas Shp incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 125. Los fragmentos preferentes de (b) comprenden un epítipo de la SEQ ID NO: 125. Otros fragmentos preferentes carecen de uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo C-terminal y/o uno o más aminoácidos (por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25 o más) del extremo N-terminal de la SEQ ID NO: 125. Otros fragmentos omiten uno o más dominios de la proteína (por ejemplo, omisión de un péptido señal, de un dominio citoplasmático, de un dominio transmembrana o de un dominio extracelular).

(4) SagA

Se cree que la estreptolisina S (SLS), también conocida como 'SagA', es producida por casi todas las colonias de GAS. Esta toxina citolítica es responsable de la beta-hemólisis alrededor de las colonias de GAS cultivadas en agar sangre y se cree que está asociada con la virulencia. Aunque el péptido SagA completo no ha demostrado ser inmunógeno, se ha utilizado un fragmento de los aminoácidos 10 - 30 (SagA 10 - 30) para producir anticuerpos neutralizantes. Véase Dale *et al.*, "Antibodies against a Synthetic Peptide of SagA Neutralize the Cytolytic Activity of Streptolysin S from Group A Streptococci", Infection and Immunity (2002) 70(4):2166 – 2170. La secuencia de aminoácidos de SagA 10 - 30 se muestra en el listado de secuencias como la SEQ ID NO: 126.

Las proteínas SagA 10-30 preferentes para su uso con la invención comprenden una secuencia de aminoácidos: (a) que tiene un 50% o más de identidad (por ejemplo, 60%, 65%, 70%, 75%, 80%, 85%, 90 %, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o más) con la SEQ ID NO: 126; y/o (b) que es un fragmento de al menos n aminoácidos consecutivos de la SEQ ID NO: 126, donde n es 7 o más (por ejemplo, 8, 10, 12, 14, 16, 18, o 20). Estas proteínas SagA 10 - 30 incluyen variantes (por ejemplo, variantes alélicas, homólogos, ortólogos, parálogos, mutantes, etc.) de la SEQ ID NO: 126.

Existe un límite superior al número de antígenos de GAS que estarán en las composiciones de la invención. Preferentemente, el número de antígenos de GAS en una composición de la invención es inferior a 20, inferior a 19, inferior a 18, inferior a 17, inferior a 16, inferior a 15, inferior a 14, inferior a 13, inferior a 12, inferior a 11, inferior a 10, inferior a 9, inferior a 8, inferior a 7, inferior a 6, inferior a 5, inferior a 4 o inferior a 3. Aún más preferentemente, el número de antígenos de GAS en una composición de la invención es inferior a 6, inferior a 5 o inferior a 4. Aún más preferentemente, el número de antígenos de GAS en una composición de la invención es 3. Los antígenos de GAS utilizados en la invención están preferentemente aislados, es decir, están separados y diferenciados del organismo entero con el que la molécula se encuentra en la naturaleza o, cuando el polinucleótido o polipéptido no se encuentra en la naturaleza, está suficientemente libre de otras macromoléculas biológicas de manera que el polinucleótido o polipéptido pueda utilizarse para su propósito.

Proteínas de fusión

Los antígenos de GAS utilizados en la invención pueden estar presentes en la composición como polipéptidos individuales separados, pero resulta preferente que al menos dos (es decir, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ó 20) de los antígenos se expresen como una única cadena polipeptídica (un polipéptido "híbrido"). Los polipéptidos híbridos ofrecen dos ventajas principales: en primer lugar, un polipéptido que puede ser inestable o expresarse mal por sí mismo puede ser ayudado por la adición de una pareja de híbrido adecuada que supere el problema; en segundo lugar, se simplifica la fabricación comercial, ya que sólo es necesario emplear una expresión y purificación con el fin de producir dos polipéptidos que son útiles antigénicamente.

El polipéptido híbrido puede comprender dos o más secuencias de polipéptidos del primer grupo de antígenos. Por consiguiente, en el presente documento se describe una composición que comprende una primera secuencia de aminoácidos y una segunda secuencia de aminoácidos, en la que dichas secuencias de aminoácidos primera y segunda se seleccionan de entre un antígeno de GAS o un fragmento del mismo del primer grupo de antígenos. Preferentemente, las secuencias de aminoácidos primera y segunda en el polipéptido híbrido comprenden diferentes epítomos.

El polipéptido híbrido puede comprender una o más secuencias de polipéptidos del primer grupo de antígenos y una o más secuencias de polipéptidos del segundo grupo de antígenos. Por consiguiente, en el presente documento se describe una composición que comprende una primera secuencia de aminoácidos y una segunda secuencia de aminoácidos, dicha primera secuencia de aminoácidos seleccionada de entre un antígeno de GAS o un fragmento del mismo del primer grupo de antígenos y dicha segunda secuencia de aminoácidos seleccionada de entre un antígeno de GAS o un fragmento del mismo del segundo grupo de antígenos. Preferentemente, la secuencias de aminoácidos primera y segunda en el polipéptido híbrido comprenden diferentes epítomos.

Resultan preferentes los híbridos que consisten en secuencias de aminoácidos de dos, tres, cuatro, cinco, seis, siete, ocho, nueve o diez antígenos de GAS. En concreto, resultan preferentes los híbridos que consisten en secuencias de aminoácidos de dos, tres, cuatro o cinco antígenos de GAS.

Pueden mezclarse en una única formulación diferentes polipéptidos híbridos. Dentro de tales combinaciones, un antígeno de GAS puede estar presente en más de un polipéptido híbrido y/o como un polipéptido no híbrido. Sin embargo, resulta preferente que un antígeno esté presente como híbrido o como no híbrido, pero no como ambos.

Los polipéptidos híbridos pueden representarse mediante la fórmula $\text{NH}_2\text{-A-}\{-\text{X-L-}\}_n\text{-B-COOH}$, en la que: X es una secuencia de aminoácidos de un antígeno de GAS o un fragmento del mismo del primer grupo de antígenos o del segundo grupo de antígenos; L es una secuencia de aminoácidos conectora opcional; A es una secuencia de aminoácidos N-terminal opcional; B es una secuencia de aminoácidos C-terminal opcional; y n es 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 ó 15.

Si un resto -X- tiene una secuencia de péptido líder en su forma de tipo silvestre, ésta puede incluirse u omitirse en la proteína híbrida. En algunas formas de realización, los péptidos líder estarán delecionados a excepción de aquel del resto -X- situado en el extremo N-terminal de la proteína híbrida, es decir, se conservará el péptido líder de X_1 , pero se omitirán los péptidos líder de $X_2 \dots X_n$. Esto es equivalente a deleccionar todos los péptidos líder y utilizar el péptido líder de X_1 como resto -A-.

Para cada caso n de $\{-\text{X-L-}\}$, la secuencia de aminoácidos conectora -L- puede estar presente o ausente. Por ejemplo, cuando $n=2$ el híbrido puede ser $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-L}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-X}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-X}_2\text{-L}_2\text{-COOH}$, etc. La(s) secuencia(s) de aminoácidos conectora(s) -L- será(n) por lo general corta(s) (por ejemplo, 20 o menos aminoácidos, es decir, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Los ejemplos comprenden secuencias peptídicas cortas que facilitan la clonación, los enlaces de poliglicina (es decir, que comprenden Gly_n donde $n = 2, 3, 4, 5, 6, 7, 8, 9, 10$ o más) y los marcadores de histidina (es decir, His_n donde $n = 3, 4, 5, 6, 7, 8, 9, 10$ o más). Otras secuencias de aminoácidos conectoras adecuadas resultarán evidentes para los expertos en la materia. Un conector útil es GSGGGG, formándose el dipéptido Gly-Ser a partir de un sitio de restricción *Bam*HI, ayudando así a la clonación y la manipulación, y siendo el tetrapéptido $(\text{Gly})_4$ un conector de poliglicina típico.

-A- es una secuencia de aminoácidos N-terminal opcional. Esta será por lo general corta (por ejemplo, 40 o menos aminoácidos, es decir, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Los ejemplos incluyen secuencias líder para dirigir el tráfico de proteínas, o secuencias peptídicas cortas que facilitan la clonación o la purificación (por ejemplo, marcadores de histidina, es decir, His_n donde $n = 3, 4, 5, 6, 7, 8, 9, 10$ o más). Otras secuencias de aminoácidos N-terminales adecuadas resultarán evidentes para los expertos en la materia. Si X_1 carece de su propia metionina N-terminal, -A- es preferentemente un oligopéptido (por ejemplo, con 1, 2, 3, 4, 5, 6, 7 u 8 aminoácidos) que proporciona una metionina N-terminal.

-B- es una secuencia de aminoácidos C-terminal opcional. Esta será por lo general corta (por ejemplo, 40 o menos aminoácidos, es decir, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Los ejemplos incluyen secuencias para dirigir el tráfico de proteínas, secuencias peptídicas cortas que facilitan la clonación o la purificación (por ejemplo, que comprenden marcadores de histidina, es decir, His_n, donde $n = 3, 4, 5, 6, 7, 8, 9, 10$ o más), o secuencias que potencian la estabilidad de las proteínas. Otras secuencias de aminoácidos C-terminales adecuadas resultarán evidentes para los expertos en la materia.

Lo más preferentemente, n es 2 ó 3.

Los constructos de fusión descritos en el presente documento pueden incluir una combinación de dos o más antígenos de GAS, en los que dicha combinación incluye GAS 40 o un fragmento del mismo o un polipéptido con identidad de secuencia con el mismo.

Los constructos de fusión descritos en el presente documento pueden incluir una combinación de antígenos de GAS, consistiendo dicha combinación en dos a treinta y un antígenos de GAS del primer grupo de antígenos, consistiendo dicho primer grupo de antígenos en: GAS 117, GAS 130, GAS 277, GAS 236, GAS 40, GAS 389, GAS 504, GAS 509, GAS 366, GAS 159, GAS 217, GAS 309, GAS 372, GAS 039, GAS 042, GAS 058, GAS 290, GAS 511, GAS 533, GAS 527, GAS 294, GAS 253, GAS 529, GAS 045, GAS 095, GAS 193, GAS 137, GAS 084, GAS 384, GAS 202 y GAS 057. Preferentemente, la combinación de antígenos de GAS consiste en tres, cuatro, cinco, seis, siete, ocho, nueve o diez antígenos de GAS seleccionados del primer grupo de antígenos. Preferentemente, la combinación de antígenos de GAS consiste en tres, cuatro o cinco antígenos de GAS seleccionados del primer grupo de antígenos.

GAS 39, GAS 40, GAS 57, GAS 117, GAS 202, GAS 294, GAS 527, GAS 533 y GAS 511 resultan antígenos de GAS particularmente preferentes para su uso en los constructos de fusión descritos en el presente documento. Preferentemente, la combinación de antígenos de GAS incluye uno o ambos de GAS 40 y GAS 117. Preferentemente, la combinación incluye GAS 40.

La expresión recombinante de los constructos de fusión descritos en el presente documento puede mejorarse u optimizarse mediante los mismos métodos descritos para la expresión de los antígenos de GAS en solitario (analizados anteriormente). Los constructos de fusión de GAS 40 y GAS 117 se ejemplifican más adelante.

En el primer ejemplo, GAS 117 se une a GAS 40a-RR. (Como se ha analizado anteriormente, GAS 40a-RR es una secuencia de GAS 40 con codones optimizados en la que se eliminan la secuencia líder N-terminal y la secuencia transmembrana C-terminal). En este constructo, se coloca un fragmento de GAS 117 (en el que la secuencia líder N-terminal está eliminada) en el extremo N-terminal de la secuencia de GAS 40 y se añade un marcador HIS al extremo C-terminal de la secuencia de GAS 40. Este constructo se denomina "117-40a-RR". Las secuencias de aminoácidos y de polinucleótidos para este constructo se muestran en el listado de secuencias como las SEQ ID NO: 127 y 128.

Las secuencias de GAS 117 y GAS 40 se unen preferentemente mediante una secuencia conectora que comprende múltiples residuos de glicina. Por ejemplo, se utiliza el conector utilizado en el constructo de fusión 117-40a-RR, una secuencia conectora de la SEQ ID NO: 129 (YASGGGS).

En un segundo ejemplo, pueden intercambiarse las ubicaciones relativas de las secuencias de GAS 40 y GAS 117. En este constructo, denominado "40a-RR-117", la secuencia 40a-RR de GAS se coloca en el extremo N-terminal de la secuencia de GAS 117 y se añade el marcador HIS al extremo C-terminal de la secuencia de GAS 117. Las secuencias de aminoácidos y de polinucleótidos para este constructo de fusión se muestran en el listado de secuencias como las SEQ ID NO: 130 y 131.

Como alternativa, los constructos de fusión pueden diseñarse sin optimizaciones de codones. Por ejemplo, las secuencias de polinucleótidos y de aminoácidos para el constructo de fusión "117-40a" se muestran en el listado de secuencias como las SEQ ID NO: 132 y 133. (Aunque no se utilizaron optimizaciones de codones, aparentemente se produjeron tres mutaciones puntuales durante la clonación, sólo una de las cuales implicaba un cambio de aminoácidos conservador (glucina a glicina). En el modelo de inmunización murino (ya analizado anteriormente), la inmunización con "117-40a" produjo hasta un 80% de supervivencia tras la provocación.

Una secuencia de fusión de GAS40 preferente comprende un fragmento de GAS 40 que comprende una o más de las regiones con superenrollamiento. Por ejemplo, el constructo de fusión puede comprender una secuencia de GAS 40 que comprende la primera región con superenrollamiento. "117-40N" es un ejemplo de este tipo de constructo. Las secuencias de aminoácidos y de polinucleótidos para este constructo se muestran en el listado de secuencias como las SEQ ID NO: 132 y 133.

En el presente documento también se describen ácidos nucleicos que codifican los polipéptidos híbridos de la descripción. Además, en el presente documento se describe un ácido nucleico que puede hibridar con este ácido

nucleico, preferentemente en condiciones de "alta rigurosidad" (por ejemplo, 65°C en una solución 0,1xSSC, SDS al 0,5%).

Los antígenos de GAS descritos en el presente documento también pueden utilizarse para preparar anticuerpos específicos contra los antígenos de GAS. Los anticuerpos son preferentemente específicos contra las regiones con superenrollamiento primera o segunda de GAS 40. En el presente documento también se describe el uso de una combinación de dos o más tipos de anticuerpos seleccionados del grupo que consiste en anticuerpos específicos contra GBS 80, GAS 117, GAS 130, GAS 277, GAS 236, GAS 40, GAS 389, GAS 504, GAS 509, GAS 366, GAS 159, GAS 217, GAS 309, GAS 372, GAS 039, GAS 042, GAS 058, GAS 290, GAS 511, GAS 533, GAS 527, GAS 294, GAS 253, GAS 529, GAS 045, GAS 095, GAS 193, GAS 137, GAS 084, GAS 384, GAS 202 y GAS 057. Preferentemente, la combinación incluye un anticuerpo específico contra GAS 40, o un fragmento del mismo.

Los anticuerpos específicos contra GAS descritos en el presente documento incluyen uno o más restos biológicos que, a través de medios químicos o físicos, pueden unirse a o asociarse con un epítipo de un polipéptido GAS. Los anticuerpos descritos en el presente documento incluyen anticuerpos que se unen específicamente a un antígeno de GAS, preferentemente GAS 80. En el presente documento se describen anticuerpos obtenidos a partir de preparaciones policlonales y monoclonales, así como los siguientes: moléculas de anticuerpo híbridas (quiméricas) (véanse, por ejemplo, Winter *et al.* (1991) *Nature* 349: 293-299 y la patente de EE.UU. N° 4.816.567; fragmentos F(ab')₂ y F(ab); moléculas Fv (heterodímeros no covalentes, véanse, por ejemplo, Inbar *et al.* (1972) *Proc Natl Acad Sci USA* 69:2659-2662 y Ehrlich *et al.* (1980) *Biochem* 19:4091-4096); moléculas Fv de cadena sencilla (sFv) (véase, por ejemplo, Huston *et al.* (1988) *Proc Natl Acad Sci EE.UU.* 85:5897-5883); constructos de fragmentos de anticuerpos diméricos y triméricos; minicuerpos (véanse, por ejemplo, Pack *et al.* (1992) *Biochem* 31: 1579-1584; Cumber *et al.* (1992) *J Immunology* 149B: 120-126); moléculas de anticuerpos humanizados (véanse, por ejemplo, Riechmann *et al.* (1988) *Nature* 332:323-327; Verhoeven *et al.* (1988) *Science* 239:1534-1536 y la publicación de patente británica N° GB 2.276.169, publicada el 21 de septiembre de 1994); y cualquier fragmento funcional obtenido a partir de tales moléculas, en los que tales fragmentos conservan las propiedades de unión inmunitarias de la molécula de anticuerpo parental. En el presente documento se describen adicionalmente anticuerpos obtenidos a través de procesos no convencionales, tales como la presentación en fagos.

Preferentemente, los anticuerpos específicos contra GAS descritos en el presente documento son anticuerpos monoclonales. Los anticuerpos monoclonales descritos en el presente documento incluyen una composición de anticuerpo que tiene una población de anticuerpos homogénea. Los anticuerpos monoclonales descritos en el presente documento pueden obtenerse a partir de hibridomas murinos, así como anticuerpos monoclonales humanos obtenidos utilizando hibridomas humanos en lugar de murinos. Véase, por ejemplo, Cote, *et al.* *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, 1985, p. 77.

Los polipéptidos de la invención pueden prepararse mediante diversos medios (por ejemplo, expresión recombinante, purificación a partir de cultivo celular, síntesis química, etc.) y en diversas formas (por ejemplo, natural, fusiones, no glicosilada, lipidada, etc.). Se preparan preferentemente en forma sustancialmente pura (es decir, sustancialmente libre de otras proteínas GAS o de la célula hospedadora).

El ácido nucleico descrito en el presente documento puede prepararse de muchas maneras (por ejemplo, mediante síntesis química, a partir de genotecas de ADN_g o de ADN_c, a partir del propio organismo, etc.) y puede adoptar diversas formas (por ejemplo, monocatenario, bicatenario, vectores, sondas, etc.). Se preparan preferentemente en forma sustancialmente pura (es decir, sustancialmente libres de otros ácidos nucleicos de GAS o de la célula hospedadora).

La expresión "ácido nucleico" incluye ADN y ARN, y también sus análogos, tales como los que contienen cadenas principales modificadas (por ejemplo, fosforotioatos, etc.) y también ácidos nucleicos peptídicos (PNA), etc. En el presente documento se describe un ácido nucleico que comprende secuencias complementarias a las descritas anteriormente (por ejemplo, para fines de hibridación o antisentido).

La invención también proporciona un procedimiento para producir un polipéptido de la invención, que comprende la etapa de cultivar una célula hospedadora transformada con el ácido nucleico descrito en el presente documento en condiciones que induzcan la expresión del polipéptido.

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

En el presente documento se describe un proceso para producir el ácido nucleico descrito en el presente documento, que comprende la etapa de amplificar el ácido nucleico utilizando un método de amplificación basado en cebadores (por ejemplo, PCR).

En el presente documento se describe un proceso para producir el ácido nucleico descrito en el presente documento, que comprende la etapa de sintetizar al menos parte del ácido nucleico por medios químicos.

Cepas

Los polipéptidos preferentes de la invención comprenden una secuencia de aminoácidos que se encuentra en una cepa M1, M3 o M18 de GAS. La secuencia genómica de una cepa M1 de GAS se presenta en Ferretti *et al.*, PNAS (2001) 98(8):4658 - 4663. La secuencia genómica de una cepa M3 de GAS se presenta en Beres *et al.*, PNAS (2002) 99(15):10078 - 10083. La secuencia genómica de una cepa M18 de GAS se presenta en Smoot *et al.*, PNAS (2002) 99(7): 4668-4673.

Cuando se utilizan polipéptidos híbridos, los antígenos individuales dentro del híbrido (es decir, los restos -X- individuales) pueden ser de una o más cepas. Cuando $n = 2$, por ejemplo, X_2 puede ser de la misma cepa que X_1 o de una cepa diferente. Cuando $n = 3$, las cepas pueden ser (i) $X_1=X_2=X_3$ (ii) $X_1=X_2 \neq X_3$ (iii) $X_1 \neq X_2=X_3$ (iv) $X_1 \neq X_2 \neq X_3$ o (v) $X_1=X_3 \neq X_2$, etc.

Purificación y expresión recombinante

Los antígenos de GAS de la invención pueden aislarse a partir de un *Streptococcus pyogenes*, o pueden producirse por recombinación, por ejemplo, en un hospedador heterólogo. Preferentemente, los antígenos de GAS se preparan utilizando un hospedador heterólogo. El hospedador heterólogo puede ser procariota (por ejemplo, una bacteria) o eucariota. Es preferentemente *E. coli*, pero otros hospedadores adecuados incluyen *Bacillus subtilis*, *Vibrio cholerae*, *Salmonella typhi*, *Salmonella typhimurium*, *Neisseria lactamica*, *Neisseria cinerea*, micobacterias (por ejemplo *M. tuberculosis*), levaduras, etc.

La producción recombinante de polipéptidos se facilita añadiendo una proteína marcadora al antígeno de GAS para que se exprese como una proteína de fusión que comprende la proteína marcadora y el antígeno de GAS. Tales proteínas marcadoras pueden facilitar la purificación, la detección y la estabilidad de la proteína expresada. Las proteínas marcadoras adecuadas para uso en la invención incluyen un marcador poliarginina (Arg-tag), un marcador polihistidina (His-tag), FLAG-tag, Strep-tag, c-myc-tag, S-tag, péptido de unión a calmodulina, dominio de unión a celulosa, SBP-tag, dominio de unión a quitina, marcador glutatión S-transferasa (GST), proteína de unión a maltosa, factor antiterminación de terminación de la transcripción (NusA), tiorredoxina de *E. coli* (TrxA) y proteína disulfuro isomerasa I (DsbA). Las proteínas marcadoras preferentes incluyen His-tag y GST. Puede encontrarse un análisis completo del uso de proteínas marcadoras en Terpe *et al.*, Appl Microbiol Biotechnol (2003) 60:523-533.

Después de la purificación, las proteínas marcadoras pueden eliminarse opcionalmente de la proteína de fusión expresada, es decir, mediante tratamientos enzimáticos adaptados específicamente conocidos en la técnica. Las proteasas comúnmente utilizadas incluyen enteroquinasa, virus del mosaico del tabaco (TEV), trombina y factor X_a .

Composiciones inmunógenas y medicamentos

Las composiciones de la invención son composiciones inmunógenas, y son más preferentemente composiciones de vacuna. El pH de la composición se encuentra preferentemente de entre 6 y 8, preferentemente aproximadamente 7. El pH puede mantenerse mediante el uso de un tampón. La composición puede ser estéril y/o apirógena. La composición puede ser isotónica con respecto a los seres humanos.

Las vacunas según la invención pueden ser profilácticas (es decir, para prevenir la infección) o terapéuticas (es decir, para tratar la infección), pero serán por lo general profilácticas. Por consiguiente, en el presente documento se describe un método para el tratamiento terapéutico o profiláctico de una infección por *Streptococcus pyogenes* en un animal susceptible a la infección por estreptococos, que comprende administrar a dicho animal una cantidad terapéutica o profiláctica de las composiciones inmunógenas de la invención. Preferentemente, la composición inmunógena comprende una combinación de antígenos de GAS, consistiendo dicha composición en dos a treinta y un antígenos de GAS del primer grupo de antígenos. Preferentemente, la combinación de antígenos de GAS consiste en tres, cuatro, cinco, seis, siete, ocho, nueve o diez antígenos de GAS seleccionados del primer grupo de antígenos. Preferentemente, la combinación de antígenos de GAS consiste en tres, cuatro o cinco antígenos de GAS seleccionados del primer grupo de antígenos. Preferentemente, la combinación de antígenos de GAS incluye uno o ambos de GAS 40 y GAS 117.

Como alternativa, en el presente documento se describe una composición inmunógena que comprende una combinación de antígenos de GAS, consistiendo dicha combinación en dos a treinta y un antígenos de GAS del primer grupo de antígenos y uno, dos, tres o cuatro antígenos de GAS del segundo grupo de antígenos. Preferentemente, la combinación consiste en tres, cuatro, cinco, seis, siete, ocho, nueve o diez antígenos de GAS del primer grupo de antígenos. Aún más preferentemente, la combinación consiste en tres, cuatro o cinco antígenos de GAS del primer grupo de antígenos. Preferentemente, la combinación de antígenos de GAS incluye uno o ambos de GAS 40 y GAS 117. Preferentemente, la combinación de antígenos de GAS incluye una o más variantes de la proteína M de superficie.

La invención también proporciona una composición de la invención para su uso como medicamento. El

medicamento es preferentemente capaz de provocar una respuesta inmunitaria en un mamífero (es decir, es una composición inmunógena) y es más preferentemente una vacuna.

La invención también proporciona el uso de las composiciones de la invención en la fabricación de un medicamento para provocar una respuesta inmunitaria en un mamífero. El medicamento es preferentemente una vacuna. La invención también proporciona un kit que comprende un primer componente que comprende una combinación de antígenos de GAS. La combinación de antígenos de GAS puede consistir en una mezcla de dos a treinta y un antígenos de GAS seleccionados del primer grupo de antígenos. Preferentemente, la combinación consiste en tres, cuatro, cinco, seis, siete, ocho, nueve o diez antígenos de GAS del primer grupo de antígenos. Preferentemente, la combinación consiste en tres, cuatro o cinco antígenos de GAS del primer grupo de antígenos. Preferentemente, la combinación incluye uno o ambos de GAS 117 y GAS 040.

El kit puede comprender un primer componente que comprende una combinación de antígenos de GAS que consiste en una mezcla de dos a treinta y un antígenos de GAS del primer grupo de antígenos y uno, dos, tres o cuatro antígenos de GAS del segundo grupo de antígenos. Preferentemente, la combinación consiste en tres, cuatro, cinco, seis, siete, ocho, nueve o diez antígenos de GAS del primer grupo de antígenos. Aún más preferentemente, la combinación consiste en tres, cuatro o cinco antígenos de GAS del primer grupo de antígenos. Preferentemente, la combinación de antígenos de GAS incluye uno o ambos de GAS 40 y GAS 117. Preferentemente, la combinación de antígenos de GAS incluye una o más variantes de la proteína M de superficie.

La invención también proporciona un dispositivo de administración precargado con las composiciones inmunógenas de la invención.

La invención también proporciona un método para provocar una respuesta inmunitaria en un mamífero, que comprende la etapa de administrar una cantidad eficaz de una composición de la invención. La respuesta inmunitaria es preferentemente protectora y preferentemente implica anticuerpos y/o inmunidad celular. El método puede provocar una respuesta de refuerzo.

El mamífero es preferentemente un ser humano. Cuando la vacuna es para uso profiláctico, el ser humano es preferentemente un niño (por ejemplo, un niño en edad de empezar a caminar o un bebé) o un adolescente; cuando la vacuna es para uso terapéutico, el ser humano es preferentemente un adolescente o un adulto. Una vacuna destinada a los niños también puede administrarse a adultos, por ejemplo, para evaluar la seguridad, la dosificación, la inmunogenicidad, etc.

Estos usos y métodos son preferentemente para la prevención y/o el tratamiento de una enfermedad provocada por *Streptococcus pyogenes* (por ejemplo, la faringitis (tal como el dolor de garganta por estreptococos), la escarlatina, el impétigo, las erisipelas, la celulitis, la septicemia, el síndrome de choque tóxico, la fascitis necrotizante (enfermedad devoradora de carne) y sus secuelas (tal como la fiebre reumática y la glomerulonefritis aguda). Las composiciones también pueden ser eficaces contra otras bacterias estreptocócicas.

Una forma de comprobar la eficacia del tratamiento terapéutico implica supervisar la infección por GAS después de la administración de la composición de la invención. Una forma de comprobar la eficacia del tratamiento profiláctico implica supervisar las respuestas inmunitarias contra los antígenos de GAS en las composiciones de la invención después de la administración de la composición.

Las composiciones de la invención se administrarán por lo general directamente a un paciente. La administración directa puede llevarse a cabo mediante inyección parenteral (por ejemplo, por vía subcutánea, intraperitoneal, intravenosa, intramuscular o al espacio intersticial de un tejido), o por vía rectal, oral (por ejemplo, comprimido, pulverización), vaginal, tópica, transdérmica (por ejemplo, véase el documento WO99/27961) o transcutánea (por ejemplo, véanse los documentos WO02/074244 y WO02/064162), intranasal (por ejemplo, véase el documento WO03/028760), ocular, ótica, pulmonar u otra administración a las mucosas. La invención puede utilizarse para provocar la inmunidad sistémica y/o de las mucosas.

El tratamiento de dosificación puede ser un régimen monodosis o un régimen multidosis. Las dosis múltiples pueden utilizarse en un programa de vacunación primaria y/o en un programa de vacunación de refuerzo. En un régimen multidosis las diversas dosis pueden administrarse por ejemplo a través de la misma o de diferentes vías, una sensibilización parenteral y un refuerzo por vía mucosa, una sensibilización por vía mucosa y un refuerzo parenteral, etc.

Las composiciones de la invención pueden prepararse en diversas formas. Por ejemplo, las composiciones pueden prepararse como inyectables, bien como soluciones líquidas o como suspensiones. También pueden prepararse formas sólidas adecuadas para la solución en, o suspensión en, vehículos líquidos antes de la inyección (por ejemplo, una composición liofilizada). La composición puede prepararse para la administración tópica, por ejemplo como ungüento, crema o polvo. La composición puede prepararse para la administración oral, por ejemplo como comprimido o cápsula, como pulverización o como jarabe (opcionalmente aromatizada). La composición puede prepararse para la administración pulmonar, por ejemplo como inhalador, utilizando un polvo fino o un aerosol. La

composición puede prepararse como supositorio o pesario. La composición puede prepararse para la administración nasal, ótica u ocular, por ejemplo, como gotas. La composición puede estar en forma de kit, diseñado de manera que una composición combinada se reconstituya inmediatamente antes de la administración a un paciente. Tales kits pueden comprender uno o más antígenos en forma líquida y uno o más antígenos liofilizados. Las composiciones inmunógenas utilizadas como vacunas comprenden una cantidad inmunitariamente eficaz de antígeno(s), así como cualquier otro componente, según sea necesario. "Cantidad inmunitariamente eficaz" se refiere a que la administración de esa cantidad a un individuo, ya sea en una sola dosis o como parte de una serie, es eficaz para el tratamiento o la prevención. Esta cantidad varía en función de la salud y la condición física del individuo a tratar, la edad, el grupo taxonómico del individuo a tratar (por ejemplo, primate no humano, primate, etc.), la capacidad del sistema inmunitario del individuo para sintetizar anticuerpos, el grado de protección deseado, la formulación de la vacuna, la evaluación por parte del médico tratante de la situación médica y otros factores pertinentes. Se espera que la cantidad se encuentre dentro de un intervalo relativamente amplio que puede determinarse mediante ensayos rutinarios.

Componentes adicionales de la composición

La composición de la invención comprenderá por lo general, además de los componentes mencionados anteriormente, uno o más "vehículos farmacéuticamente aceptables", que incluyen cualquier vehículo que no induzca por sí mismo la producción de anticuerpos perjudiciales para el individuo que recibe la composición. Los vehículos adecuados son por lo general macromoléculas grandes de metabolización lenta, tales como proteínas, polisacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos y agregados lipídicos (tales como gotitas de aceite o liposomas). Tales vehículos son bien conocidos por los expertos habituales en la materia. Las vacunas también pueden contener diluyentes, tales como agua, solución salina, glicerol, etc. Además, pueden estar presentes sustancias auxiliares, tales como agentes humectantes o emulsionantes, sustancias amortiguadoras del pH, y similares. Se dispone de un análisis minucioso de excipientes farmacéuticamente aceptables en Gennaro (2000) Remington: The Science and Practice of Pharmacy. 20^a ed., ISBN: 0683306472.

Las vacunas de la invención pueden administrarse junto con otros agentes inmunorreguladores. En concreto, las composiciones incluirán normalmente un adyuvante.

Los adyuvantes adicionales preferentes incluyen, pero no se limitan a, uno o más de los siguientes, que se exponen a continuación:

A. Composiciones que contienen minerales

Las composiciones que contienen minerales adecuadas para su uso como adyuvantes en la invención incluyen sales minerales, tales como sales de aluminio y sales de calcio. La invención puede utilizarse con sales minerales tales como hidróxidos (por ejemplo, oxihidróxidos), fosfatos (por ejemplo, hidroxifosfatos, ortofosfatos), sulfatos, etc. {por ejemplo, véanse los capítulos 8 y 9 del documento Vaccine design: the subunit and adjuvant approach (1995) Powell y Newman. ISBN 0-306-44867-X}, o mezclas de diferentes compuestos minerales, adoptando los compuestos cualquier forma adecuada (por ejemplo, gel, cristalina, amorfa, etc.) y siendo preferente la adsorción. Las composiciones que contienen minerales también pueden formularse en forma de partícula de sal metálica. Véase el documento WO00/23105.

B. Emulsiones de aceite

Las composiciones de emulsión de aceite adecuadas para su uso como adyuvantes en la invención incluyen emulsiones de escualeno-agua, tal como MF59 (escualeno al 5%, Tween 80 al 0,5% y Span 85 al 0,5%, formulado en partículas submicrométricas utilizando un microfluidizador). Véase el documento WO90/14837. Véanse también, Podda, "The adjuvanted influenza vaccines with novel adjuvants: experience with the MF59-adjuvanted vaccine", Vaccine (2001) 19:2673-2680; Frey *et al.*, "Comparison of the safety, tolerability, and immunogenicity of a MF59-adjuvanted influenza vaccine and a non-adjuvanted influenza vaccine in non-elderly adults", Vaccine (2003) 21:4234-4237. MF59 se utiliza como adyuvante en la vacuna de subunidades trivalente contra el virus de la gripe, FLUADTM.

Los adyuvantes particularmente preferentes para uso en las composiciones son emulsiones submicrométricas de aceite en agua. Las emulsiones submicrométricas de aceite en agua preferentes para uso en el presente documento son emulsiones de escualeno/agua que contienen opcionalmente distintas cantidades de MTP-PE, tal como una emulsión submicrométrica de aceite en agua que contiene escualeno al 4%-5% p/v, Tween 80TM (monooleato de polioxietilensorbitán) al 0,25%-1,0% p/v y/o Span 85TM (trioleato de sorbitán) al 0,25%-1,0% y, opcionalmente, N-acetilmuramil-L-alanil-D-isogluatminil-L-alanina-2-(1'-2'-dipalmitoil-sn-glicero-3-hidroxfosforiloxi)-etilamina (MTP-PE), por ejemplo, la emulsión submicrométrica de aceite en agua conocida como "MF59" (publicación internacional N° WO90/14837; patentes de EE.UU. N° 6.299.884 y 6.451.325; y Ott *et al.*, "MF59 - Design and Evaluation of a Safe and Potent Adjuvant for Human Vaccines" en Vaccine Design: The Subunit and Adjuvant Approach (Powell, M.F. y Newman, M.J. eds.) Plenum Press, Nueva York, 1995, págs. 277-296). MF59

contiene escualeno al 4%-5% p/v (por ejemplo, 4,3%), Tween 80™ al 0,25%-0,5% p/v y Span 85™ al 0,5% p/v y, opcionalmente, contiene distintas cantidades de MTP-PE, formulado en partículas submicrométricas utilizando un microfluidizador tal como el microfluidizador Modelo 110Y (Microfluidics, Newton, MA). Por ejemplo, MTP-PE puede estar presente en una cantidad de aproximadamente 0-500 µg/dosis, más preferentemente 0-250 µg/dosis y, lo más preferentemente, 0-100 µg/dosis. Tal como se utiliza en el presente documento, el término "MF59-0" se refiere a la anterior emulsión submicrométrica de aceite en agua sin MTP-PE, mientras que el término MF59-MTP indica una formulación que contiene MTP-PE. Por ejemplo, "MF59-100" contiene 100 µg de MTP-PE por dosis, y así sucesivamente. MF69, otra emulsión submicrométrica de aceite en agua para uso en el presente documento, contiene escualeno al 4,3% p/v, Tween 80™ al 0,25% p/v y Span 85™ al 0,75% p/v y, opcionalmente, MTP-PE. Todavía otra emulsión submicrométrica de aceite en agua es MF75, también conocida como SAF, que contiene escualeno al 10%, Tween 80 al 0,4%™, polímero de bloque pluronic L121 al 5% y thr-MDP, también microfluidizado en una emulsión submicrométrica. MF75-MTP indica una formulación de MF75 que incluye MTP, tal como de 100 µg-400 µg de MTP-PE por dosis.

En la publicación internacional N° WO90/14837 y en las patentes de EE.UU. N° 6.299.884 y 6.451.325 se describen detalladamente emulsiones submicrométricas de aceite en agua, métodos para fabricarlas y agentes inmunoestimuladores, tales como péptidos muramilo, para su uso en las composiciones.

También pueden utilizarse el adyuvante completo de Freund (CFA) y el adyuvante incompleto de Freund (IFA) como adyuvantes en la invención.

C. Formulaciones de saponina

También pueden utilizarse formulaciones de saponina como adyuvantes en la invención. Las saponinas son un grupo heterólogo de glucósidos de esterol y glucósidos triterpenoides que se encuentra en la corteza, las hojas, los tallos, las raíces e incluso las flores de una gran variedad de especies vegetales. La saponina de la corteza del árbol *Quillaia saponaria* Molina se ha estudiado ampliamente como adyuvante. La saponina también puede obtenerse en el mercado de *Smilax ornata* (zarzaparrilla), *Gypsophilla paniculata* (velo de novia) y *Saponaria officinalis* (jabonera). Las formulaciones de adyuvantes de saponina incluyen formulaciones purificadas, tales como QS21, así como formulaciones lipídicas, tales como los ISCOM. Las composiciones de saponina se han purificado mediante cromatografía de capa fina de alto rendimiento (HP-LC) y cromatografía de líquidos de alto rendimiento en fase inversa (RP-HPLC). Se han identificado fracciones purificadas específicas utilizando estas técnicas, incluidas QS7, QS17, QS18, QS21, QH-A, QH-B y QH-C. Preferentemente, la saponina es QS21. Un método de producción de QS21 se describe en la patente de EE.UU. N° 5.057.540. Las formulaciones de saponina también pueden comprender un esterol, tal como colesterol (véase el documento WO 96/33739).

Las combinaciones de saponinas y colesterol pueden utilizarse para formar partículas únicas denominadas complejos inmunoestimuladores (ISCOM). Los ISCOM también incluyen por lo general un fosfolípido tal como fosfatidiletanolamina o fosfatidilcolina. Puede utilizarse cualquier saponina conocida en los ISCOM. Preferentemente, el ISCOM incluye una o más de entre Quil A, QHA y QHC. Los ISCOM se describen adicionalmente en los documentos EP 0 109 942, WO 96/11711 y WO 96/33739. Opcionalmente, los ISCOM pueden estar desprovistos de detergente adicional. Véase el documento WO00/07621.

Puede encontrarse una revisión del desarrollo de adyuvantes basados en saponina en Barr, *et al.*, Advanced Drug Delivery Reviews (1998) 32:247 – 271. Véase también Sjolander, *et al.*, Advanced Drug Delivery Reviews (1998) 32:321-338.

C. Virosomas y partículas pseudovirales (VLP)

También pueden utilizarse virosomas y partículas pseudovirales (VLP) como adyuvantes en la invención. Estas estructuras contienen generalmente una o más proteínas de un virus opcionalmente combinadas o formuladas con un fosfolípido. Son generalmente no patógenas, no replicantes y generalmente no contienen nada del genoma viral natural. Las proteínas virales pueden producirse por recombinación o aislarse a partir de virus enteros. Estas proteínas virales adecuadas para su uso en virosomas o VLP incluyen proteínas derivadas del virus de la gripe (tal como HA o NA), virus de la hepatitis B (tal como las proteínas del core o de la cápside), virus de la hepatitis E, virus del sarampión, virus Sindbis, Rotavirus, virus de la fiebre aftosa, retrovirus, virus de Norwalk, virus del papiloma humano, VIH, fagos de ARN, fago Qβ (tal como proteínas de la cubierta), fago GA, fago fr, fago AP205 y Ty (tal como la proteína pI del retrotransposón Ty). Las VLP se analizan adicionalmente en los documentos WO 03/024480, WO 03/024481, y en Niikura *et al.*, Virology (2002) 293:273-280, Lenz *et al.*, Journal of Immunology (2001) 5246 - 5355; Pinto *et al.*, Journal of Infectious Diseases (2003) 188:327-338 y Gerber *et al.*, Journal of Virology (2001) 75(10): 4752 - 4760. Los virosomas se analizan adicionalmente por ejemplo en Gluck *et al.*, Vaccine (2002) 20: B10-B16.

D. Derivados bacterianos o microbianos

Los adyuvantes adecuados para su uso en la invención incluyen derivados bacterianos o microbianos tales

como:

(1) *Derivados no tóxicos de lipopolisacárido enterobacteriano (LPS)*

- 5 Tales derivados incluyen monofosforil lípido A (MPL) y MPL 3-O-desacilado (3dMPL). El 3dMPL es una mezcla de monofosforil lípido A 3 De-O-acilado con 4, 5 ó 6 cadenas aciladas. En el documento EP 0 689 454 se describe una forma de "partícula pequeña" preferente de monofosforil lípido A 3 De-O-acilado. Tales "partículas pequeñas" de 3dMPL son lo suficientemente pequeñas para esterilizarse por filtración a través de una membrana de 0,22 micras (véase el documento EP 0 689 454). Otros derivados de LPS no tóxicos incluyen miméticos de monofosforil lípido A, tales como derivados de aminoalquil glucosaminida fosfato, por ejemplo, RC-529. Véase Johnson *et al.* (1999) Bioorg Med Chem Lett 9: 2273-2278.

(2) *Derivados de lípido A*

- 15 Los derivados de lípido A incluyen derivados de lípido A de *Escherichia coli* tal como OM-174. OM-174 se describe, por ejemplo, en Meraldi *et al.*, Vaccine (2003) 21:2485-2491 y Pajak, *et al.*, Vaccine (2003) 21:836-842.

(3) *Oligonucleótidos inmunoestimuladores*

- 20 Los oligonucleótidos inmunoestimuladores adecuados para su uso como adyuvantes en la invención incluyen secuencias de nucleótidos que contienen un motivo CpG (una secuencia que contiene una citosina no metilada seguida de guanósina y unida por un enlace fosfato). El ARN bicatenario bacteriano o los oligonucleótidos que contienen secuencias palindrómicas o de poli(dG) también han demostrado ser inmunoestimuladores.

- 25 Los CpG pueden incluir modificaciones/análogos de nucleótidos tales como modificaciones fosforotioato y pueden ser bicatenarios o monocatenarios. Opcionalmente, la guanósina puede sustituirse con un análogo tal como 2'-desoxi-7-desazaguanósina. Véase Kandimalla, *et al.*, Nucleic Acids Research (2003) 31(9):2393-2400; los documentos WO 99/62923 y WO 02/26757 para ejemplos de posibles sustituciones análogas. El efecto adyuvante de los oligonucleótidos CpG se analiza adicionalmente en Krieg, Nature Medicine (2003) 9(7):831-835; McCluskie, *et al.*, FEMS Immunology and Medical Microbiology (2002) 32:179-185; en el documento WO 98/40100, la patente de EE.UU. N° 6.207.646, la patente de EE.UU. N° 6.239.116 y la patente de EE.UU. N° 6.429.199.

- 30 La secuencia CpG puede dirigirse a TLR9, como el motivo GTCGTT o TTCGTT. Véase Kandimalla, *et al.*, Biochemical Society Transactions (2003) 31 (parte 3): 654 - 658. La secuencia CpG puede ser específica para inducir una respuesta inmunitaria de Th1, tal como un ODN CpG-A, o puede ser más específica para inducir una respuesta de linfocitos B, tal como un ODN CpG-B. Los ODN CpG-A y CpG-B se analizan en Blackwell, *et al.*, J. Immunol. (2003) 170(8):4061-4068; Krieg, TRENDS in Immunology (2002) 23(2): 64 - 65 y en el documento WO 01/95935. Preferentemente, el CpG es un ODN CpG-A.

- 40 Preferentemente, el oligonucleótido CpG se construye de manera que el extremo 5' sea accesible para el reconocimiento del receptor. Opcionalmente, pueden fijarse dos secuencias de oligonucleótidos CpG en sus extremos 3' para formar "inmunómeros". Véanse, por ejemplo, Kandimalla, *et al.*, BBRC (2003) 306:948-953; Kandimalla, *et al.*, Biochemical Society Transactions (2003) 31 (parte 3): 664-658; Bhagat *et al.*, BBRC (2003) 300:853-861 y el documento WO 03/035836.

- 45 (4) *Toxinas de ribosilación de ADP y derivados detoxificados de las mismas.*

- 50 Pueden utilizarse toxinas de ribosilación de ADP bacterianas y derivados detoxificados de las mismas como adyuvantes en la invención. Preferentemente, la proteína se deriva de *E. coli* (es decir, enterotoxina termolábil de *E. coli* "LT"), del cólera ("CT") o de pertussis ("PT"). El uso de toxinas de ribosilación de ADP detoxificadas como adyuvantes para administración en mucosas se describe en el documento WO 95/17211 y como adyuvantes para administración parenteral en el documento WO 98/42375. Preferentemente, el adyuvante es un mutante de LT detoxificado, tal como LT-K63.

55 E. Inmunomoduladores humanos

- Los inmunomoduladores humanos adecuados para su uso como adyuvantes en la invención incluyen citocinas, tales como interleucinas (por ejemplo IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc.), interferones (por ejemplo interferón-γ), factor estimulador de colonias de macrófagos y factor de necrosis tumoral.

60 F. Bioadhesivos y mucoadhesivos

- También pueden utilizarse bioadhesivos y mucoadhesivos como adyuvantes en la invención. Los bioadhesivos adecuados incluyen microesferas de ácido hialurónico esterificado (Singh *et al.* (2001) J. Cont Rele 70:267-276) o mucoadhesivos tales como derivados reticulados de poli(ácido acrílico), alcohol polivinílico, polivinilpirrolidona, polisacáridos y carboximetilcelulosa. También pueden utilizarse como adyuvantes en la invención

el quitosano y derivados del mismo. Por ejemplo, documento WO99/27960.

G. Micropartículas

5 También pueden utilizarse micropartículas como adyuvantes en la invención. Resultan preferentes las micropartículas (es decir, una partícula de ~100 nm a ~150 µm de diámetro, más preferentemente ~200 nm a ~30 µm de diámetro, y lo más preferentemente ~500 nm a ~10 µm de diámetro) formadas a partir de materiales biodegradables y no tóxicos (por ejemplo, un poli(α-hidroxiácido), un ácido polihidroxibutírico, un poliortoéster, un polianhídrido, una policaprolactona, etc.), con poli(lactida-co-glicólido), opcionalmente tratadas para que tengan una superficie cargada negativamente (por ejemplo, con SDS) o una superficie cargada positivamente (por ejemplo, con un detergente catiónico, tal como CTAB).

H. Liposomas

15 Los ejemplos de formulaciones de liposomas adecuadas para su uso como adyuvantes se describen en la patente de EE.UU. N° 6.090.406, la patente de EE.UU. N° 5.916.588 y el documento EP 0 626 169.

I. Formulaciones de éter de polioxietileno y de éster de polioxietileno

20 Los adyuvantes adecuados para su uso en la invención incluyen éteres de polioxietileno y ésteres de polioxietileno. Documento WO99/52549. Tales formulaciones incluyen adicionalmente tensioactivos de éster de polioxietilensorbitán en combinación con un octoxinol (documento WO01/21207), así como tensioactivos de éster o éteres de polioxietileno alquilo en combinación con al menos un tensioactivo no iónico adicional tal como un octoxinol (WO01/21152).

25 Los éteres de polioxietileno preferentes se seleccionan del siguiente grupo: éter de polioxietileno-9-laurilo (laureth 9), éter de polioxietileno-9-esteorilo, éter de polioxietileno-8-esteorilo, éter de polioxietileno-4-laurilo, éter de polioxietileno-35-laurilo y éter de polioxietileno-23-laurilo.

J. Polifosfaceno (PCPP)

30 Se describen formulaciones de PCPP por ejemplo en Andrianov *et al.*, "Preparation of hydrogel microspheres by coacervation of aqueous polyphosphazene solutions", *Biomaterials* (1998) 19(1-3):109-115 y Payne *et al.*, "Protein Release from Polyphosphazene Matrices", *Adv. Drug. Delivery Review* (1998) 31(3):185-196.

K. Péptidos muramilo

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

L. Compuestos de imidazoquinolona.

45 Los ejemplos de compuestos de imidazoquinolona adecuados para su uso como adyuvantes en la invención incluyen Imiquamod y sus homólogos, que se describen adicionalmente en Stanley, "Imiquimod and the imidazoquinolones: mechanism of action and therapeutic potential", *Clin Exp Dermatol* (2002) 27(7):571- 577 y Jones, "Resiquimod 3M", *Curr Opin Investig Drugs* (2003) 4(2):214 - 218. La invención también puede comprender combinaciones de aspectos de uno o más de los adyuvantes identificados anteriormente. Por ejemplo, pueden utilizarse en la invención las siguientes composiciones de adyuvantes:

- 50 (1) una saponina y una emulsión de aceite en agua (WO99/11241);
 (2) una saponina (por ejemplo, QS21) + un derivado no tóxico de LPS (por ejemplo, 3dMPL) (véase el documento WO 94/00153);
 (3) una saponina (por ejemplo, QS21) + un derivado no tóxico de LPS (por ejemplo, 3dMPL) + un colesterol;
 55 (4) una saponina (por ejemplo QS21) + 3dMPL + IL-12 (opcionalmente + un esteroide) (WO98/57659);
 (5) combinaciones de 3dMPL con, por ejemplo, QS21 y/o emulsiones de aceite en agua (solicitudes de patente europea 0835318, 0735898 y 0761231);
 (6) SAF, que contiene escualano al 10%, Tween 80 al 0,4%, polímero de bloque pluronic L121 al 5% y thr-MDP, microfluidizado en una emulsión submicrométrica o agitado en vórtex para generar una emulsión de mayor tamaño de partícula.
 60 (7) sistema adyuvante RIBI™ (RAS), (Ribi Immunochem) que contiene escualano al 2%, Tween 80 al 0,2% y uno o más componentes de la pared celular bacteriana del grupo que consiste en monofosforil lípido A (MPL), dimicolato de trehalosa (TDM) y el esqueleto de la pared celular (CWS), preferentemente MPL + CWS (Detox™);
 y
 65 (8) una o más sales minerales (tales como una sal de aluminio) + un derivado no tóxico de LPS (tal como 3dPML).

(9) una o más sales minerales (tal como una sal de aluminio) + un oligonucleótido inmunoestimulador (tal como una secuencia de nucleótidos que incluya un motivo CpG).

Las sales de aluminio y MF59 son adyuvantes preferentes para la inmunización por vía parenteral. Las toxinas bacterianas mutantes son adyuvantes para administración en mucosas preferentes.

La composición puede incluir un antibiótico.

Antígenos adicionales

Las composiciones de la invención pueden comprender adicionalmente uno o más antígenos no GAS adicionales, incluidos antígenos bacterianos, virales o parasitarios.

En una forma de realización, las combinaciones de antígenos de GAS de la invención se combinan con uno o más antígenos no GAS adicionales adecuados para uso en una vacuna pediátrica. Por ejemplo, las combinaciones de antígenos de GAS pueden combinarse con uno o más antígenos derivados de una bacteria o virus seleccionados del grupo que consiste en *N. meningitidis* (incluidos los serogrupos A, B, C, W135 y/o Y), *Streptococcus pneumoniae*, *Bordetella pertussis*, *Moraxella catarrhalis*, tétanos, difteria, virus respiratorio sincicial ('RSV'), poliomielitis, sarampión, paperas, rubéola y rotavirus.

En otra forma de realización, las combinaciones de antígenos de GAS de la invención se combinan con uno o más antígenos no GAS adicionales adecuados para su uso en una vacuna diseñada para proteger a individuos de edad avanzada o inmunocomprometidos. Por ejemplo, las combinaciones de antígenos de GAS pueden combinarse con un antígeno derivado del grupo que consiste en *Enterococcus faecalis*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Listeria monocytogenes*, gripe y virus de la parainfluenza ('PIV').

Cuando se utiliza un antígeno sacárido o carbohidrato, éste se conjuga preferentemente con una proteína transportadora con el fin de potenciar la inmunogenicidad {por ejemplo, Ramsay *et al.* (2001) Lancet 357(9251): 195-196; Lindberg (1999) Vaccine 17 suppl. 2:S28-36; Buttery y Moxon (2000) J R Coll Physician Lond 34:163-168; Ahmad y Chapnick (1999) Infect Dis Clin North Am 13:113-133, vii. Goldblatt (1998) J. Med. Microbiol. 47:563-567; la patente europea 0 477 508; la patente de EE.UU. Nº 5.306.492; el documento WO98/42721; Conjugate Vaccines (Eds. Cruse *et al.*) ISBN 3805549326, en concreto el vol. 10:48-114; Hermanson (1996) Bioconjugate Techniques ISBN: 0123423368 ó 012342335X}. Las proteínas transportadoras preferentes son toxinas o toxoides bacterianos, tales como los toxoides diftérico o tetánico. El toxoide diftérico CRM₁₉₇ resulta particularmente preferente {Research Disclosure, 453077 (enero 2002)}. Otros polipéptidos transportadores incluyen la proteína de membrana externa de *N. meningitidis* {EP-A-0372501}, péptidos sintéticos {EP-A-0378881 y EP-A-0427347}, proteínas de choque térmico {WO93/17712 y WO94/03208}, proteínas de pertussis {WO98/58668 y EP-A-0471177}, proteína D de *H. influenzae* {WO00/56360}, citocinas {WO91/01146}, linfocinas, hormonas, factores de crecimiento, toxina A o B de *C. difficile* {WO00/61761}, proteínas de captación de hierro {WO01/72337}, etc. Cuando una mezcla comprende sacáridos capsulares de ambos serogrupos A y C, puede resultar preferente que la relación (p/p) de sacárido MenA:sacárido MenC sea superior a 1 (por ejemplo, 2:1, 3:1, 4:1, 5:1, 10:1 o superior). Pueden conjugarse diferentes sacáridos con el mismo o diferente tipo de proteína transportadora. Puede utilizarse cualquier reacción de conjugación adecuada, con cualquier conector adecuado cuando sea necesario.

Los antígenos proteicos tóxicos pueden detoxificarse cuando sea necesario, por ejemplo detoxificación de la toxina de pertussis, por medios químicos y/o genéticos.

Cuando se incluya un antígeno diftérico en la composición, resulta preferente incluir también un antígeno tetánico y antígenos de pertussis. De manera similar, cuando se incluya un antígeno tetánico resulta preferente incluir también antígenos diftéricos y de pertussis. De manera similar, cuando se incluya un antígeno de pertussis resulta preferente incluir también antígenos diftéricos y tetánicos.

Los antígenos en la composición estarán presentes por lo general a una concentración de al menos 1 µg/ml cada uno. En general, la concentración de cualquier antígeno determinado será suficiente para provocar una respuesta inmunitaria contra ese antígeno.

Como alternativa al uso de antígenos proteicos en la composición de la invención, puede utilizarse el ácido nucleico que codifica el antígeno {por ejemplo, Robinson y Torres (1997) Seminars in Immunology 9:271-283; Donnelly *et al.* (1997) Annu Rev Immunol 15:617-648; Scott-Taylor y Dalglish (2000) Expert Opin Investig Drugs 9:471-480; Apostolopoulos y Plebanski (2000) Curr Opin Mol Ther 2:441-447; Ilan (1999) Curr Opin Mol Ther 1:116-120; Dubensky *et al.* (2000) Mol Med 6:723-732; Robinson y Pertmer (2000) Adv Virus Res 55:1-74; Donnelly *et al.* (2000) Am J Respir Crit Care Med 162(4 Pt 2):S190-193; Davis (1999) Mt. Sinai J. Med. 66:84-90}. Los componentes proteicos de las composiciones de la invención pueden sustituirse por el ácido nucleico (preferentemente ADN, por ejemplo en forma de plásmido) que codifica la proteína.

Definiciones

La expresión "que comprende" significa "que incluye", así como "que consiste". Por ejemplo, una composición "que comprende" X puede consistir exclusivamente en X o puede incluir algo adicional, por ejemplo X + Y.

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

Las referencias a un porcentaje de identidad de secuencia entre dos secuencias de aminoácidos significan que, cuando se alinean, ese porcentaje de aminoácidos es el mismo al comparar las dos secuencias. Este alineamiento y el porcentaje de homología o identidad de secuencia pueden determinarse utilizando programas de software conocidos en la técnica, por ejemplo los descritos en la sección 7.7.18 de Current Protocols in Molecular Biology (F.M. Ausubel *et al.*, Eds., 1987), suplemento 30. Un alineamiento preferente se determina mediante el algoritmo de búsqueda de homología de Smith-Waterman utilizando una búsqueda de huecos afines con una penalización por apertura de hueco de 12 y una penalización por extensión de hueco de 2, matriz BLOSUM de 62. El algoritmo de búsqueda de homología de Smith-Waterman se describe en Smith y Waterman (1981) Adv. Appl. Math. 2: 482-489. Pueden utilizarse métodos similares de identidad de secuencia para determinar la homología de secuencias entre dos secuencias de polinucleótidos.

El siguiente ejemplo muestra una forma de preparar los antígenos de GAS recombinantes de la invención y ensayar su eficacia en un modelo murino.

EJEMPLO 1: Preparación de antígenos de GAS recombinantes y demostración de la eficacia en un modelo murino.

Las proteínas GAS recombinantes correspondientes a dos o más de los antígenos de GAS del primer grupo de antígenos se expresan de la siguiente manera.

1. Clonación de antígenos de GAS para la expresión en *E. coli*

Los antígenos de GAS seleccionados se clonaron de manera que se obtuviesen dos clases diferentes de proteínas recombinantes: (1) proteínas con un marcador hexa-histidina en el extremo carboxilo terminal (Gas-His) y (2) proteínas con el marcador hexa-histidina en el extremo carboxilo terminal y GST en el extremo amino terminal (GST-Gas-His). Las proteínas de tipo (1) se obtuvieron por clonación en un vector pET21b+ (disponible en Novagen). Las proteínas de tipo (2) se obtuvieron por clonación en un vector pGEX-NNH. Esta estrategia de clonación permitió utilizar el ADN genómico de GAS para amplificar los genes seleccionados mediante PCR, para realizar una sola digestión con enzimas de restricción de los productos de PCR y a continuación clonarlos simultáneamente en ambos vectores.

(a) Construcción de vectores de expresión pGEX-NNH

Se sintetizan dos parejas de oligodesoxirribonucleótidos complementarios utilizando el sintetizador de ADN ABI394 (Perkin Elmer) y los reactivos de Cruachem (Glasgow, Escocia). Se hibridan cantidades equimolares de los pares de oligonucleótidos (50 ng de cada oligonucleótido) en tampón con ADN ligasa de T4 (New England Biolabs) durante 10 minutos en un volumen final de 50 μ l y a continuación se deja enfriar lentamente a temperatura ambiente. Con el procedimiento descrito se obtienen los siguientes conectores de ADN:

Conector gexNN

```

NdeI  NheI  XmaI  EcoRI  NcoI      SalI      XhoI      SacI
GATCCCATATGGCTAGCCCGGGAATTCGTCCATGGAGTGAGTCGACTGACTCGAGTGATCGAGCTC
GGTATACCGATCGGGCCCCTTAAGCAGGTACCTCACTCAGCTGACTGAGCTCACTAGCTCGAG

NotI
CTGAGCGGCCGCATGAA
GACTCGCCGGCGTACTTTCGA

```

Conector gexNNH

```

HindIII  NotI  XhoI  Hexa-Histidina
TCGACAAGCTTGC GGCGCACTCGAGCATCACCATCACCATCACTGAT
GTTCGAACGCGCGGTGAGCACGTAGAGGTAGTGGTAGTGACTATCGA

```

El plásmido pGEX-KG [K. L. Guan y J. E. Dixon, Canal Biochem. 192, 262 (1991)] se digiere con BamHI y HindIII y se ligan 100 ng al conector gexNN durante toda la noche a 16°C, con una relación molar de 3:1 conector/plásmido utilizando 200 unidades de ADN ligasa de T4 (New England Biolabs). Después de la transformación del producto de ligación en *E. coli* DH5, se selecciona un clon que contiene el plásmido pGEX-NN, que tiene el conector correcto, por medio del análisis con enzimas de restricción y secuenciación de ADN.

El nuevo plásmido pGEX-NN se digiere con Sall y HindIII y se liga al conector gexNNH. Después de la transformación del producto de ligación en *E. coli* DH5, se selecciona un clon que contiene el plásmido pGEX-NNH, que tiene el conector correcto, por medio del análisis con enzimas de restricción y secuenciación de ADN.

(b) Reparación del ADN cromosómico

Se cultiva la cepa SF370 de GAS en medio THY hasta que la DO_{600} es 0,6-0,8. A continuación, las bacterias se centrifugan, se suspenden en tampón TES con lisozima (10 mg/ml) y mutanolisina (10 U/μl) y se incuban 1 hora a 37°C. Después del tratamiento de la suspensión bacteriana con ARNasa, proteinasa K y sarcosil/EDTA al 10%, se lleva a cabo la extracción de proteínas con fenol saturado y fenol/cloroformo. El sobrenadante resultante se hace precipitar con acetato sódico/etanol y el ADN extraído se sedimenta por centrifugación, se suspende en tampón Tris y se mantiene a -20°C.

(c) Diseño de oligonucleótidos

Se diseñan cebadores de oligonucleótidos sintéticos en base a la secuencia codificante de cada antígeno de GAS utilizando la secuencia de la cepa M1 de *Streptococcus pyogenes* SF370. Se omite cualquier péptido señal predicho, deduciendo la secuencia de cebador de amplificación del extremo 5' inmediatamente cadena abajo de la secuencia líder predicha. Para la mayoría de los antígenos de GAS, la cola 5' de los cebadores (véase la Tabla 1, a continuación) incluye sólo un sitio de reconocimiento de la enzima de restricción (NdeI, o NheI, o SpeI, dependiendo del patrón de restricción propio del gen); las colas 3' del cebador (véase la Tabla 1) incluyen un sitio de restricción XhoI o NotI o HindIII.

Colas 5'	Colas 3'
NdeI 5' GTGCGT CATATG 3'	XhoI 5' GCGT CTCGAG 3'
NheI 5' GTGCGT GCTAGC 3'	NotI 5' ACTCGCTAG CGGCCGC 3'
SpeI 5' GTGCGT ACTAGT 3'	HindIII 5' GCGT AAGCTT 3'

Tabla 1. Colas de oligonucleótidos de los cebadores utilizados para amplificar los genes que codifican los antígenos de GAS seleccionados.

Además de contener las secuencias de reconocimiento de enzimas de restricción, los cebadores incluyen nucleótidos que hibridan con la secuencia a amplificar. El número de nucleótidos que hibridan depende de la temperatura de fusión de los cebadores que pueden determinarse como se describe [(Breslauer *et al.*, Proc. Nat. Acad. Sci. 83, 3746-50 (1986)]. La temperatura media de fusión de los oligonucleótidos seleccionados es de 50°C-55°C para la región de hibridación sola y 65°C-75°C para el oligonucleótido entero. Los oligonucleótidos pueden adquirirse de MWG-Biotech S.p.A. (Firenze, Italia).

(d) Amplificación mediante PCR

El protocolo estándar de PCR es de la siguiente manera: se utilizan 50 ng de ADN genómico como molde en presencia de 0,2 μM de cada cebador, 200 μM de cada dNTP, MgCl₂ 1,5 mM, 1x tampón de PCR sin Mg (Gibco-BRL) y 2 unidades de Taq ADN polimerasa (Platinum Taq, Gibco-BRL) en un volumen final de 100 μl. Se somete cada muestra a una amplificación en dos etapas: los primeros 5 ciclos se realizan utilizando como temperatura de hibridación la de uno de los oligonucleótidos excluyendo la cola de la enzima de restricción, seguido de 25 ciclos realizados según la temperatura de hibridación de los cebadores de longitud completa. Los ciclos estándar son los siguientes: un ciclo:

desnaturalización: 94°C, 2 minutos,

5 ciclos:

desnaturalización: 94°C, 30 segundos,

hibridación: 51°C, 50 segundos,

elongación: 72°C, 1 minuto ó 2 minutos y 40 segundos,

25 ciclos:

desnaturalización: 94°C, 30 segundos,
hibridación: 70°C, 50 segundos,
elongación: 72°C, 1 minuto ó 2 minutos y 40 segundos,

**72°C, 7 minutos,
4°C**

El tiempo de elongación es de 1 minuto para los antígenos de GAS codificados por ORF de menos de 2.000 pb, y 2 minutos y 40 segundos para los ORF de más de 2.000 pb. Las amplificaciones se realizaron utilizando un sistema de PCR Gene Amp 9600 (Perkin Elmer).

Para comprobar los resultados de la amplificación, se cargan 4 µl de cada producto de PCR en gel de agarosa al 1%-1,5% y se compara el tamaño de los fragmentos amplificados con los patrones de peso molecular de ADN (marcadores de ADN III o IX, Roche). Los productos de PCR se cargan en gel de agarosa y después de la electroforesis se cortan del gel las bandas del tamaño correcto. El ADN se purifica a partir de la agarosa utilizando el Gel Extraction Kit (Qiagen) siguiendo las instrucciones del fabricante. El volumen de elución final del ADN es 50 µl de TE (Tris-HCl 10 mM, EDTA 1 mM, pH 8). Se carga un µl de cada ADN purificado en gel de agarosa para evaluar el rendimiento.

(e) *Digestión de los fragmentos de PCR*

Se someten a doble digestión uno-dos µg de productos de PCR purificados, durante toda la noche a 37°C con las enzimas de restricción apropiadas (60 unidades de cada enzima) utilizando el tampón de restricción apropiado en un volumen final de 100 µl. Las enzimas de restricción y los tampones de digestión son de New England Biolabs. Después de la purificación del ADN digerido (Kit de purificación de PCR, Qiagen) y la elución con 30 µl de TE, se somete 1 µl a electroforesis en gel de agarosa para evaluar el rendimiento en comparación con los patrones de peso molecular titulados (marcadores de ADN III o IX, Roche).

(f) Digestión de los vectores de cierre (pET21b+ y pGEX-NNH)

Se someten a doble digestión 10 µg de plásmido con 100 unidades de cada enzima de restricción en 400 µl de volumen de reacción en presencia del tampón apropiado mediante incubación durante toda la noche a 37°C. Después de la electroforesis en un gel de agarosa al 1%, la banda correspondiente al vector digerido se purifica a partir del gel utilizando el Qiaex II Gel Extraction Kit de Qiagen y se eluye el ADN con 50 µl de TE. La concentración de ADN se evalúa midiendo la DO₂₆₀ de la muestra.

(g) *Clonación de los productos de PCR*

Setenta y cinco ng de los vectores apropiadamente digeridos y purificados y los fragmentos digeridos y purificados correspondientes a cada antígeno de GAS seleccionado se ligan en volúmenes finales de 10 µl-20 µl con una relación molar de 1:1 fragmento/vector, utilizando 400 unidades de ADN ligasa de T4 (New England Biolabs) en presencia del tampón suministrado por el fabricante. Las reacciones se incuban durante toda la noche a 16°C.

Se realiza la transformación de células electrocompetentes de *E. coli* BL21 (Novagen) y *E. coli* BL21-DE3 (Novagen) utilizando ligaciones con pGEX-NNH y ligaciones con pET21b+ respectivamente. El procedimiento de transformación es la siguiente: se mezclan 1 μ l-2 μ l de la reacción de ligación con 50 μ l de células competentes enfriadas en hielo, a continuación se vierten las células en una cubeta con electrodos de 0,1 cm de un Gene Pulser (Biorad). Después de pulsar las células en un electroporador MicroPulser (Biorad) siguiendo las instrucciones del fabricante, se suspenden las células en 0,95 ml de medio SOC y se incuban durante 45 minutos a 37°C con agitación. Se siembran 100 μ l y 900 μ l de suspensiones de células en placas distintas de agar LB con 100 μ g/ml de ampicilina y se incuban las placas durante toda la noche a 37°C. El cribado de los transformantes se realiza mediante PCR: los transformantes seleccionados al azar se recogen y se suspenden en 30 μ l de mezcla de reacción de PCR que contiene el tampón de PCR, los 4 dNTP, MgCl₂ 1,5 mM, Taq polimerasa y cebadores de oligonucleótidos directos e inversos apropiados que son capaces de hibridar cadena arriba y cadena abajo del policonector de los vectores pET21b+ o pGEX-NNH. Después de 30 ciclos de PCR, se someten 5 μ l de los productos resultantes a electroforesis en gel de agarosa con el fin de seleccionar clones positivos a partir de los que se obtiene la banda de PCR esperada. Los clones positivos de PCR se eligen en base al tamaño correcto del producto de PCR, tal como se evalúa por comparación con marcadores de peso molecular apropiados (marcadores de ADN III o IX, Roche).

2. Expresión de proteínas

Las colonias positivas de la PCR se inoculan en 3 ml de LB con 100 µg/ml de ampicilina y se cultivan a 37°C durante toda la noche. Se inoculan 70 µl del cultivo durante toda la noche en 2 ml de LB/Amp y se cultivan a 37°C hasta que la DO₆₀₀ de los clones pEt alcanza el valor 0,4-0,8 o hasta que la DO₆₀₀ de los clones pGEX alcanza el valor 0,8-1. A continuación, se induce la expresión de proteínas añadiendo a los mini-cultivos IPTG 1 mM (isopropil β-D tio-galacto-piranosido). Al cabo de 3 horas de incubación a 37°C se comprueba la DO₆₀₀ final y se enfrían los cultivos en hielo. Después de la centrifugación de 0,5 ml de cultivo, se suspende el sedimento celular en 50 µl de tampón de muestra de carga de proteínas (Tris-HCl mM, pH 6,8, 60, SDS al 5% p/v, glicerina al 10% v/v, azul de bromofenol al 0,1% p/v, DTT 100 mM) y se incuba a 100°C durante 5 minutos. Se analiza, mediante SDS-PAGE, un volumen de muestra hervida correspondiente a un cultivo con una DO₆₀₀ de 0,1 y se tiñe con azul de Coomassie para verificar la presencia de la banda de proteína inducida.

3. Purificación de las proteínas recombinantes

Se inoculan las colonias individuales en 25 ml de LB con 100 µg/ml de ampicilina y se cultivan a 37°C durante toda la noche. El cultivo de una noche se inocula en 500 ml de LB/Amp y se cultiva con agitación a 25°C hasta una DO₆₀₀ de 0,4-0,7. A continuación, se induce la expresión de proteínas añadiendo a los cultivos IPTG 1 mM. Al cabo de 3,5 horas de incubación a 25°C se comprueba la DO₆₀₀ final y se enfrían los cultivos en hielo. Después de centrifugación a 6.000 rpm (rotor JA10, Beckman), se procesa el sedimento celular para la purificación o se congela a -20°C.

(a) Procedimiento para la purificación de proteínas solubles marcadas con His de E. coli

- (1) Transferir los sedimentos de -20°C a baño de hielo y reconstituir con 10 ml de tampón NaHPO₄ 50 mM, NaCl 300 mM, pH 8,0, pasar a tubos de centrifugación de 40 ml-50 ml y romper las células de acuerdo al siguiente resumen.
- (2) Romper los sedimentos en la prensa francesa realizando tres pases con lavado en línea.
- (3) Centrifugar a aproximadamente 30-40.000 x g durante 15-20 minutos. Si es posible, utilizar el rotor JA 25.50 (21.000 rpm, 15 minutos) o JA-20 (18.000 rpm, 15 minutos).
- (4) Equilibrar las columnas Poly-Prep con 1 ml de resina Fast Flow Chelating Sepharose con tampón fosfato 50 mM, NaCl 300 mM, pH 8,0.
- (5) Almacenar el sedimento de centrifugación a -20°C, y cargar el sobrenadante en las columnas.
- (6) Recoger el flujo que atraviesa la columna.
- (7) Lavar las columnas con 10 ml (2 ml + 2 ml + 4 ml) de tampón fosfato 50 mM, NaCl 300 mM, pH 8,0.
- (8) Lavar de nuevo con 10 ml de tampón imidazol 20 mM, fosfato 50 mM, NaCl 300 mM, pH 8,0.
- (9) Eluir las proteínas unidas a las columnas con 4,5 ml (1,5 ml + 1,5 ml + 1,5 ml) de tampón imidazol 250 mM, fosfato 50 mM, NaCl 300 mM, pH 8,0 y recoger las 3 fracciones correspondientes de -1,5 ml cada una. Añadir a cada tubo 15 µl de DTT 200 mM (concentración final 2 mM)
- (10) Medir la concentración de proteína de las dos primeras fracciones con el método de Bradford, recoger una alícuota de 10 µg de proteínas de cada muestra y analizar mediante SDS-PAGE. (N.B.: si la muestra está demasiado diluida, cargar 21 µl + 7 µl de tampón de carga).
- (11) Almacenar las fracciones recogidas a +4°C a la espera de los resultados del análisis de SDS-PAGE.
- (12) Para la inmunización, preparar 4-5 alícuotas de 100 µg cada una en 0,5 ml de glicerol al 40%. El tampón de dilución es el tampón de elución anteriormente indicado, más DTT 2 mM. Almacenar las alícuotas a -20°C hasta la inmunización.

(b) Purificación de proteínas marcadas con His a partir de cuerpos de inclusión

Las purificaciones se llevan a cabo esencialmente según el siguiente protocolo:

- (1) Las bacterias se recogen por centrifugación de cultivos de 500 ml. De ser necesario, almacenar los sedimentos bacterianos a -20°C. Para la extracción, resuspender cada sedimento bacteriano en 10 ml de tampón TRIS-HCl 50 mM, pH 8,5 en un baño de hielo.
- (2) Romper las bacterias resuspendidas con una prensa francesa, realizando dos pases.
- (3) Centrifugar a 35.000 x g durante 15 minutos y recoger los sedimentos. Utilizar un rotor Beckman JA 25.50 (21.000 rpm, 15 minutos) o JA-20 (18.000 rpm, 15 minutos).
- (4) Disolver los sedimentos de centrifugación con TRIS-HCl 50 mM, TCEP 1 mM {clorhidrato de Tris(2-carboxietil)-fosfina, Pierce}, cloruro de guanidinio 6M, pH 8,5. Agitar durante ~ 10 minutos con una barra magnética.
- (5) Centrifugar como se ha descrito anteriormente y recoger el sobrenadante.
- (6) Preparar un número adecuado de columnas Poly-Prep (Bio-Rad) que contienen 1 ml de Fast Flow Chelating Sepharose (Pharmacia) saturado con níquel según las recomendaciones del fabricante. Lavar las columnas dos veces con 5 ml de H₂O y equilibrar con TRIS-HCl 50 mM, TCEP 1 mM, cloruro de guanidinio 6M, pH 8,5.
- (7) Cargar los sobrenadantes de la etapa 5 en las columnas y lavar con 5 ml de tampón Tris-HCl 50 mM, TCEP 1 mM, urea 6M, pH 8,5.
- (8) Lavar las columnas con 10 ml de imidazol 20 mM, Tris-HCl 50 mM, urea 6M, TCEP 1 mM, pH 8,5. Recoger y

reservar los primeros 5 ml para posibles controles adicionales.

(9) Eluir las proteínas unidas a las columnas con 4,5 ml de un tampón que contiene imidazol 250 mM, Tris-HCl 50 mM, urea 6M, TCEP 1 mM, pH 8,5. Añadir el tampón de elución en tres alícuotas de 1,5 ml y recoger las 3 fracciones correspondientes. Añadir a cada fracción 15 µl de DTT (concentración final 2 mM).

(10) Medir la concentración de proteína eluida con el método de Bradford y analizar alícuotas de aproximadamente 10 µg de proteína mediante SDS-PAGE.

(11) Almacenar las proteínas a -20°C en glicerol al 40% (v/v), Tris-HCl 50 mM, urea 2M, arginina 0,5 M, DTT 2 mM, TCEP 0,3 mM, imidazol 83,3 mM, pH 8,5.

(c) Procedimiento para la purificación de GST-proteínas de fusión a partir de *E. coli*

(1) Transferir los sedimentos bacterianos de -20°C a un baño de hielo y suspender con 7,5 ml de PBS, pH 7,4 al que se ha añadido una mezcla de inhibidores de proteasa (CØMPLETE™ - Boehringer Mannheim, 1 comprimido cada 25 ml de tampón).

(2) Transferir a tubos de centrifugación de 40 ml-50 ml y exponer a ultrasonidos según el siguiente procedimiento:

a. Colocar la sonda a aproximadamente 0,5 cm de la parte inferior del tubo

b. Bloquear el tubo con la abrazadera

c. Sumergir el tubo en un baño de hielo

d. Ajustar el sonicador de la siguiente manera: Temporizador → Espera, Ciclo de trabajo → 55, Control de salida → 6.

e. Realizar 5 ciclos de 10 impulsos en un intervalo de tiempo de 1 minuto (es decir, un ciclo = 10 impulsos + ~45" de espera; b. 10 impulsos + ~45" de espera; c. 10 impulsos + ~45" de espera; d. 10 impulsos + ~45" de espera; e. 10 impulsos + ~45" de espera).

(3) Se centrifuga a aproximadamente 30-40.000 x g durante 15-20 minutos. Por ejemplo: utilizar un rotor Beckman JA 25.50 a 21.000 rpm, durante 15 minutos.

(4) Almacenar los residuos de centrifugación a -20°C y cargar los sobrenadantes en las columnas de cromatografía, de la siguiente manera:

(5) Equilibrar las columnas de Poly-Prep (Bio-Rad) con 0,5 ml (\approx 1 ml de suspensión) de resina glutatión-sefarosa 4B, lavar con 2 ml (1 + 1) de H₂O y a continuación con 10 ml (2 + 4 + 4) de PBS, pH 7,4.

(6) Cargar los sobrenadantes en las columnas y desechar el flujo que atraviesa la columna.

(7) Lavar las columnas con 10 ml (2 + 4 + 4) de PBS, pH 7,4.

(8) Eluir las proteínas unidas a las columnas con 4,5 ml de tampón TRIS 50 mM, glutatión reducido 10 mM, pH 8,0, añadiendo 1,5 ml + 1,5 ml + 1,5 ml y recoger las 3 fracciones respectivas de ~1,5 ml cada una.

(9) Medir la concentración de proteína de las dos primeras fracciones con el método de Bradford, analizar una alícuota de 10 µg de proteínas de cada muestra mediante SDS-PAGE. (N.B.: si la muestra está demasiado diluida, cargar 21 µl + 7 µl de tampón de carga).

(10) Almacenar las fracciones recogidas a +4°C a la espera de los resultados del análisis de SDS-PAGE.

(11) Para cada proteína destinada a la inmunización, preparar 4-5 alícuotas de 100 µg cada una en 0,5 ml de glicerol al 40%. El tampón de dilución es TRIS.HCl 50 mM, DTT 2 mM, pH 8,0. Almacenar las alícuotas a -20°C hasta la inmunización.

4. Modelo murino de protección frente a la infección por GAS

(a) Protocolo de inmunización

Se inmunizan grupos de 10 ratones hembra CD1 de edades comprendidas entre 6 y 7 semanas con dos o más antígenos de GAS, (20 µg de cada antígeno de GAS recombinante), suspendidos en 100 µl de una solución adecuada. Cada grupo recibe 3 dosis los días 0, 21 y 45. La inmunización se realiza a través de inyección intraperitoneal de la proteína con un volumen igual de adyuvante completo de Freund (CFA) para la primera dosis y adyuvante incompleto de Freund (IFA) para las siguientes dos dosis. En cada esquema de inmunización se utilizan grupos de control negativo y positivo.

Para el grupo de control negativo, se inmuniza a los ratones con proteínas de *E. coli* eluidas a partir de las columnas de purificación después del procesamiento de extracto total bacteriano de una cepa de *E. coli* que contenía el vector pET21b o el pGEX-NNH (que expresaba por tanto GST solamente) sin nada de ORF de GAS clonado (los grupos pueden indicarse como HisStop o GSTStop respectivamente).

Para los grupos de control positivo, se inmuniza a los ratones con M de GAS purificada clonada a partir de la cepa SF370 de GAS o la cepa DSM 2071 de GAS (los grupos se indican como 192SF y 192DSM respectivamente).

Se recogen sueros combinados de cada grupo antes de la primera inmunización y dos semanas después de la última. Se infecta a los ratones con GAS aproximadamente una semana después.

Se infecta a los ratones inmunizados utilizando una cepa de GAS diferente de la utilizada para la clonación de las proteínas seleccionadas. Por ejemplo, la cepa de GAS puede ser el tipo M23 de DSM 2071, que puede obtenerse a partir de la Colección Alemana de Microorganismos y Cultivos Celulares (DSMZ).

Para los experimentos de infección, se cultiva DSM 2071 a 37°C en caldo THY hasta una DO_{600} de 0,4. Las bacterias se sedimentan por centrifugación, se lavan una vez con PBS, se suspenden y se diluyen con PBS para obtener la concentración apropiada de bacterias/ml y se administra a los ratones mediante inyección intraperitoneal. Se dan a cada ratón entre 50 y 100 bacterias, determinado mediante siembra de alícuotas de la suspensión bacteriana en 5 placas THY. Se observa diariamente a los animales y se comprueban la supervivencia.

5. Análisis de inmunosueros

(a) Preparación de extractos de proteínas totales de GAS

Se preparan extractos de proteínas totales incubando un cultivo bacteriano desarrollado hasta una DO_{600} de 0,4-0,5 en Tris 50 mM pH 6,8/mutanolisina (20 unidades/ml) durante 2 horas a 37°C, seguido de incubación durante diez minutos en hielo en NaOH 0,24 N y β -mercaptoetanol al 0,96%. Las proteínas extraídas se hacen precipitar mediante adición de ácido tricloroacético, se lavan con acetona enfriada en hielo y se suspenden en tampón de carga de proteínas.

(b) Análisis de transferencia de western

Se cargaron alícuotas de extracto de proteínas totales mezcladas con tampón de carga SDS (1x: Tris-HCl 60 mM pH 6,8, SDS al 5% p/v, glicerina al 10% v/v, azul de bromofenol al 0,1%, DTT 100 mM) y hervidas durante 5 minutos a 95°C, en un gel prefabricado de SDS-PAGE al 12,5% (Biorad). Se corre el gel utilizando un tampón de electroforesis de SDS-PAGE que contiene TRIS 250 mM, glicina 2,5 mM y SDS al 0,1%. Se somete el gel a electrotransferencia sobre una membrana de nitrocelulosa a 200 mA durante 60 minutos. La membrana se bloquea durante 60 minutos con PBS/Tween-20 al 0,05% (Sigma), leche desnatada en polvo al 10% y se incuba durante toda la noche a 4°C con PBS/Tween 20 al 0,05%, leche desnatada en polvo al 1%, con la dilución apropiada de los sueros. Después de lavar dos veces con PBS/Tween al 0,05%, se incuba la membrana durante 2 horas con anticuerpo secundario anti-ratón conjugado con peroxidasa (Amersham) diluido 1:4.000. Se lava tres veces la nitrocelulosa durante 10 minutos con PBS/Tween al 0,05% y una vez con PBS y se desarrolla posteriormente mediante el kit Opti-4CN Substrate (Biorad).

(c) Preparación de cultivos de GAS tratados con paraformaldehído

Se lava una vez con PBS un cultivo bacteriano desarrollado hasta una DO_{600} de 0,4-0,5 y se concentra cuatro veces en PBS/paraformaldehído al 0,05%. Después de 1 hora de incubación a 37°C con agitación, se mantiene el cultivo tratado durante toda la noche a 4°C y a continuación se controla la inactivación completa de las bacterias sembrando alícuotas sobre placas de agar sangre THY.

(d) Análisis FACS de cultivos de GAS creados con paraformaldehído con inmunosueros de ratón

Se lavan aproximadamente 10^5 bacterias inactivadas con paraformaldehído con 200 μ l de PBS en una placa con fondo en U de 96 pocillos y se centrifugan durante 10 minutos a 3.000 g, a 4°C. Se desecha el sobrenadante y se suspenden las bacterias en 20 μ l de PBS-BSA al 0,1%. Se añaden a la suspensión bacteriana ochenta μ l de pre-inmunosero o inmunosero de ratón diluido en PBS-BSA al 0,1%, a una dilución final de 1:100, 1:250 ó 1:500, y se incuba en hielo durante 30 minutos. Se lavan una vez las bacterias añadiendo 100 μ l de PBS-BSA al 0,1%, se centrifugan durante 10 minutos a 3.000 g, 4°C, se suspenden en 200 μ l de PBS-BSA al 0,1%, se centrifugan de nuevo y se suspenden en 10 μ l de IgG de cabra anti-ratón conjugada con R-ficoeritrina específica para el fragmento F(ab')₂, (Jackson ImmunoResearch Laboratories Inc., N° de catálogo 115-116-072) en PBS-BSA al 0,1% a una dilución final de 1:100, y se incuban en hielo durante 30 minutos en la oscuridad. Se lavan una vez las bacterias añadiendo 180 μ l de PBS-BSA al 0,1% y se centrifugan durante 10 minutos a 3.000 g, 4°C. Se desecha el sobrenadante y se suspenden las bacterias en 200 μ l de PBS. La suspensión bacteriana se pasa a través de una cámara de citometría de un FACS Calibur (Becton Dickinson, Mountain View, CA, EE.UU.) y se adquieren 10.000 eventos. Los datos se analizan utilizando el software Cell Quest (Becton Dickinson, Mountain View, CA, EE.UU.) dibujando un gráfico de puntos morfológico (utilizando parámetros de dispersión hacia adelante y lateral) sobre las señales bacterianas. A continuación se crea un gráfico de histogramas sobre la intensidad FL2 de la escala logarítmica de fluorescencia que recuerda la región morfológica de las bacterias.

Ejemplo 2: Comparación de la virulencia de la cepa de GAS de tipo silvestre (incluida GAS 40) y el mutante por delección de GAS 40.

El siguiente ejemplo proporciona una comparación entre la virulencia de una cepa de GAS de tipo silvestre y un mutante por delección de GAS 40. Se prepararon mediante métodos convencionales cepas de GAS mutantes en las que la mayoría de la secuencia de GAS 40 está eliminada. Se inocularon grupos de inmunización de diez ratones

por grupo con las cepas de GAS de tipo silvestre o mutante. Como se muestra a continuación, la inoculación de un intervalo de concentraciones del aislado de tipo silvestre dio como resultado la muerte de los ratones, mientras que la inoculación con el mutante GAS Δ40 no lo hizo.

Cepa de GAS	concentración	número de muertes
de tipo silvestre	2×10^5	10
de tipo silvestre	2×10^6	9
de tipo silvestre	2×10^7	5
GAS Δ40	2×10^2	0
GAS Δ40	2×10^3	0
GAS Δ40	2×10^4	0
GAS Δ40	2×10^5	0
GAS Δ40	2×10^6	0
GAS Δ40	2×10^7	0

Ejemplo 3: Ensayo de opsonofagocitosis bacteriana de constructos de GAS 40

El siguiente ejemplo demuestra la exposición en superficie de GAS 40 mediante el uso en un ensayo de opsonofagocitosis bacteriana. En el ensayo se utilizaron los siguientes constructos de GAS, cada uno de los cuales se ha descrito detalladamente anteriormente: 40a-CH, 40a-RR-NH, 40a-RR, GST-40, 40a, 40a y 40a-NH. (Las dos referencias a "40a" en la figura 7 se refieren a sueros preparados en diferentes días).

El ensayo se realizó de la siguiente manera.

1. Preparación del inóculo bacteriano. Se cultivan bacterias GAS en medio THY hasta que alcanzan la fase exponencial media (DO_{600} de 0,4) a 37°C. Se lavan dos veces las bacterias en solución salina enfriada y se suspenden en medio HBSS ajustándose el volumen para cada cepa en función de la cantidad de bacterias que se utilizará. Las células bacterianas se mantienen en hielo hasta su uso.

2. Preparación de PMN. Se preparan PMN a partir de capas leucocitarias de sangre heparinizada de voluntarios sanos. La capa leucocitaria se incuba durante 30 minutos en una solución que contiene dextrano, NaCl y heparina (cantidad 1:1). Después de la incubación se elimina el sobrenadante, rico en leucocitos, se transfiere a un tubo limpio y se centrifuga a 700 x g durante 20 minutos. Se realiza un lavado en agua corto para romper los glóbulos rojos de la sangre y a continuación se añade una solución de NaCl para restaurar la concentración salina apropiada. Después de esta etapa las células se centrifugan, se lavan y se suspenden en MEM a una concentración adecuada.

3. Ensayo de opsonofagocitosis. Se incuban cepas de GAS (preparadas como se ha descrito) con inmunosuero de ratón inactivado por calor proveniente de la inmunización con el antígeno de GAS indicado (o pre-inmunosuero para el control), PMN humanos y complemento de cría de conejo. 1 hora de incubación a 37°C. Las muestras tomadas inmediatamente antes y después de la incubación se siembran en placas de agar sangre THY. Se evalúa la fagocitosis comparando la diferencia en el número de colonias en los dos instantes para el pre-inmunosuero y el inmunosuero. Los datos se presentan como el número de logaritmo de colonias cultivadas en $t = 0$ - el número de logaritmo de las colonias desarrolladas en $t = 60$.

Los resultados del ensayo se muestran en la figura 7. El eje Y presenta la diferencia entre el logaritmo del recuento de colonias en el instante 0 y el logaritmo de los recuentos de colonias después de 60 segundos: $\log(CFU @ T_0) - \log(CFU @ T_{60})$. Si ha habido crecimiento (es decir, las bacterias no están muertas activamente), se obtienen números negativos (barras negativas). Si las bacterias están muertas, se obtienen números positivos (barras positivas del histograma). Como se muestra en la figura 7, se presentan barras positivas del histograma para cada uno de los constructos de GAS. Las últimas cuatro barras amarillas en la figura 7 representan los controles: **B** = bacterias solas, **B PMN** = bacterias + polimorfonucleares, **B C** = Bacterias + complemento, **P PMN C** = bacterias + polimorfonucleares + complemento (sin suero).

Ejemplo 4: Experimentos de provocación mediante inmunización con GAS 40 en un modelo de protección murino (en ratones)

A continuación se enumera una muestra de los resultados de porcentaje de supervivencia de numerosos experimentos en modelo murino (en ratones) utilizando el antígeno GAS 40. Las anotaciones indican cuándo se modificó el constructo utilizado para que expresase el antígeno GAS 40 recombinante para facilitar la expresión.

Antígeno de GAS	% de supervivencia en un modelo de provocación en ratones
40a	55
40a-RR	70
40a-RR-NH	60

LISTADO DE SECUENCIAS

SEQ ID NO: 1 secuencia de aminoácidos que comprende GAS 40

MDLEQTKPNQVKQKIALTSTIALLSASVGVSHQVKADDRASGETKASNTTHDSSLPKPETIQEAKATIDAVE
 KTLSQQKAELTELATATKTAEINHLKEQDNEQKALTSAQEIYTNLTASSEETLLAQGAHQREL/TATE
 TELHNAQADQHSKETALSEQKASISAEETRAQDLVEQVKTSEQNIAKLNAMISNPDAITKAAQTANDNTKA
 LSSELEKAKADLENQKAKVKQLTEELAAQKAALAEKEAELSRKSSAPSTQDSIVGNNTMKAPQGYPLEE
 LKKLEASGYIGSASYNMYKEHADQIIAKASPGNQLNQYQDIPADNRNRFVDPDNLTPVQNELAQFAAHMI
 NSVRRQLGLPPVTVTAGSQEFARLLSTSYKKTGNTTRPSFVYGQPGVSGHYGVGPHDKTIIEDSAGASGLI
 RNDNMYENIGAFNDVHTVNGIKRGIYDSIKYMLFTDHLHGNTYGHAINFLRVDKHNPNAPVYLGFSTSNV
 GSLNEHFVMPESNIAHQRFNKTPIKAVGSTKDYAQRVGTVSDTIAAIKGVSSLENRLSAIHQADIMA
 AQAKVSQLQGLASTLKQSDSLNLQVRQLNDTKGSLRTELLAAKAKQAQLEATRDQSLAKLASLKAALHQT
 EALAEQAAARVTALVAKKAHLQYLRDPKLNPNRLQVIRERIDNTKQDLAKTTSSLLNAQBALAALQAKQSS
 LEATIAETHEQLTLKTLANEKEYRHLDEDIATVPDLQVAPPLTGVPKPLSYSKIDTTPLVQEMVKETKQLL
 EASARLAAENTSLVAELVGQTSSEMVASNAIVSKITSSITQPSKTSYSGSGSSTTSNLI SDVDESTQRLK
 AGVVM LAAVGLTGFRFRKESK

SEQ ID NO: 2 secuencia de polinucleótidos que codifica GAS 40

ATGGACTTAGAACAAACGAAGCCAAACCAAGTTAAGCAGAAAATTGCTTTAACCTCAACAATTGCTTTATT
 GAGTGCCAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCCGAGTA
 ATACTCAGCAGTAGTATTACCAAAACAGAAAACAATTCAAGAGGCCAAAGGCAACTATTGATGCGATTGAA
 AAAACTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAAAATCTAGCTGAAAT
 CAACCACTTAAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAATTTACACTAATA
 CTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTAACAGCTACTGAA
 ACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCATTGTCAGAACAAAAAGCTAGCAT
 TTCAGCAGAACTACTCGAGCTCAAGATTTAGTGGAACAAGTCAAACGCTCTGAACAAAAATATTGCTAAGC
 TCAATGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAGCA
 TTAAGCTCAGAAATTGGAGAAGGCTAAAGCTGACTTAGAAAAATCAAAAAGCTAAAGTTAAAAAGCAATTGAC
 TGAAGAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCCAGAACTTAGTCGCTCTTAAATCCTCAG
 CTCCGCTCTACTCAAGATAGCATTTGTGGGTAATAATACCATGAAAGCACCGCAAGGCTATCCTCTTGAAGAA
 CTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAGAGCATGCAGA
 TCAAATTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATATTCCAGCAGATCGTAATC
 GCTTTGTTGATCCCGATAATTTGACACCAGAAAGTGCAAAATGAGCTAGCGCAGTTTGACGCTCACATGATT
 AATAGTGTAAGAAGACAATTAGGTCTACCACAGTTACTGTTACAGCAGGATCACAGAATTTGCAAGATT
 ACTTAGTACCAGCTATAAGAAAATCATGGTAATACAAGACCATCATTTGTCTACGGCAGCCAGGGGTAT
 CAGGGCATTGCTGTTGGGCCCTCATGATAAACTATTATTGAAGACTCTGCCGGAGCGTCAGGCTCATT
 CGAAATGATGATAACATGTACGAGAAATATCGGTGCTTTTAAACGATGTGCATACTGTGAATGGTATTAAACG
 TGGTATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAATACATACGGCCATGCTA
 TTAACCTTTTACGTGTAGATAAAACATAACCCCTAATGCGCCTGTTTACCTTGGATTTTCAACCAGCAATGTA
 GGATCTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATTGCTAACCATCAACGCTTTAATAAGAC
 CCCTATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAAGAGTAGGCACTGTATCTGATACTATTGCGA
 CGATCAAAGGAAAAGTAAGCTCATTAGAAAACTGTTTGTGCGCTATTTCATCAAGAAGCTGATATTATGGCA
 GCCCAAGCTAAAGTAAGTCAACTTCAAGGTAAATATGCAAGCACACTTAAGCAGTCAGACAGCTTAAATCT
 CCAAGTGAGACAATTAATGATACTAAAGGTTCTTTGAGAACAGAAATTACTAGCAGCTAAAGCAAAAACAG
 CACAACCTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACTGCACCAGACA
 GAAGCCTTAGCAGAGCAAGCCGACAGAGTGACAGCACTGGTGGCTAAAAAAGCTCATTGCAATATCT
 AAGGGACTTTAAATTTGAATCCTAACCCGCTTCAAGTGATACGTGAGCGCATTGATAATACTAAGCAAGATT
 TGGCTAAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCTAAACAAAGCAGT
 CTAGAAGCTACTATTGCTACACAGAACACCGTTGACTTTGCTTAAAACCTTAGCTAACGAAAAGGAATA
 TCGCCACTTAGCAAGATATAGCTACTGTGCTGATTGCAAGTAGCTCCACCTCTTACGGGCGTAAAAAC
 CGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAATGGTTAAAGAAACGAAACAACTATTA
 GAAGCTTCAGCAAGATTAGCTGCTGAAAATACAAGTCTTTGTAGCAGAAGCGCTTGTGGCCAAACCTCTGA
 AATGGTAGCAAGTAATGCCATTGTGTCTAAAATCACATCTTCGATTACTCAGCCCTCATCTAAGACATCTT
 ATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAAAGAGCTCTTAAA
 GCAGGAGTCGTATGTTGGCAGCTGTGCGCCTCACAGGATTTAGGTTCCGTAAGGAATCTAAGTGA

SEQ ID NO: 3 secuencia de aminoácidos que comprende una secuencia líder N-terminal de GAS 40

MDLEQTKPNQVKQKIALTSTIALLSA

SEQ ID NO: 4 secuencia de polinucleótidos que codifica una secuencia líder N-terminal de GAS 40

5 ATGGACTTAGAACAAACGAAGCCAAACCAAGTTAAGCAGAAAATTGCTTTAACCCTCAACAATTGCTTTATT
GAGTGCC

SEQ ID NO: 5 secuencia de aminoácidos que comprende un fragmento de GAS 40 con la secuencia líder N-terminal eliminada

10 SVGVSHQVKADDRASGETKASNTHDDSLPKPETIQEAKATIDAVEKTLSSQKAELELATALTKTTAEINH
LKEQQDNEQKALTSAQEIYTNLTLASSEETLLAQGAHQRELATETELHNAQADQHSKETALSEQKASISA
15 ETTRAQDLVEQVKTSEQNIKLNAMISNPDAITKAAQTANDNTKALSSLEKAKADLENQKAKVKKQLTEE
LAAQKAALAEKEALSRLKSSAPSTQDSIVGNNTMKAPQGYPLEELKKLEASGYIGSASYNYYKEHADQI
IAKASPGNQLNQYQDI PADRNRFVDPDNLTPVQNELAQFAAHMINSVRRQLGLPPVTVTAGSQEFARLLS
TSYKKTGHGNTSPFVYQPGVSGHYGVGPHDKTIIEDSAGASGLIRNDDNMYENIGAFNDVHTVNGIKRGI
YDSIKYMLFTDHLHGNTYGHAINFLRVDKHNPNAPVYLGFTSNVGSLSNEHFVMPESNIAHQRFNKPTPI
20 KAVGSTKDYAQRVGTVDSTIAAIKGVSSLENRLSAIHQEADIMAAQAKVSQLQGKLASTLKQSDSLNLQV
RQLNDTKGSLRTELLAAKAKQAQLEATRDQSLAKLASLKAALHQTEALAEQAAARVATLVAKKAHLQYLRD
FKLNPRLQVIRERIDNTKQDLAKTTSSLLNAQEALALQAKQSSLEATITTEHQLTLLKTLANEKEYRH
LDEDIATVPDLQVAPPLTGKPLSYSKIDTTPLVQEMVKETKQLEASARLAAENTSLVAELVGQTSSEMV
ASNAIVSKITSSITQPSKTSYSGSSSTTSNLSIDVDESTQRLKAGVVMALAAVGLTGFRFRKESK

SEQ ID NO: 6 secuencia de polinucleótidos que codifica un fragmento de GAS 40 con la secuencia líder N-terminal eliminada

30 AGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAATACTCA
CGACGATAGTTTACCAAAACAGAAACAATTCAAGAGGCAAAGGCAACTATTGATGCAGTTGAAAAAAGCTC
TCAGTCAACAAAAAGCAGAAGTACAGAGCTTGCTACCGCTCTGACAAAAACTACTGCTGAAATCAACCAC
TTAAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAATTTACACTAATACTCTTGC
AAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTAACAGCTACTGAAACAGAGC
TTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCATTGTCAGAACAAAAAGCTAGCATTTTCAGCA
35 GAAACTACTCGAGCTCAAGATTTAGTGAACAAGTCAAAACGCTGAAACAAAAATATTGCTAAGCTCAATGC
TATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAGCATTAAAGCT
CAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAATCAAAAGCTAAAGTTAAAAAGCAATTGACTGAAGAG
TTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCGAGAAGTATGCTCGTCTTAAATCCTCAGCTCCGTC
TACTCAAGATAGCATTTGTGGGTAATAATACCATGAAAGCACCAGGCTATCCTCTTGAAGAAGCTTAAAA
40 AATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAGAGCATGCAGATCAAAAT
ATTGCCAAAGCTAGTCCAGGTAATCAATTAATCAATACCAAGATATTCCAGCAGATCGTAATCGCTTTGT
TGATCCCGATAAATTGACACCAGAAGTGCAAAATGAGCTAGCGCAGTTTGAGCTCACATGATTAAATAGTG
TAAGAAGACAATTAGGTCTACCACAGTTACTGTTACAGCAGGATCACAAGAAATTGCAAGATTACTTAGT
ACCAGCTATAAGAAAAGCTCATGGTAATACAAGACCATCATTTGTCTACGGACAGCCAGGGGTATCAGGGCA
45 TTATGGTGTGGGCCCTCATGATAAAACTATTATTGAAGACTCTGCCGAGCGTCAGGGCTCATTCGAAATG
ATGATAACATGTACGAGAATATCGGTGCTTTTAAAGATGTGCATCTGTGAATGGTATTAAACAGTGTATT
TATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAAATACATACGGCCATGCTATTAACTT
TTTACGTGTAGATAAACATAACCCTAATGCGCCTGTTTACCTTGGATTTTCAACCAGCAATGTAGGATCTT
TGAATGAACACTTTGTAAATGTTTCCAGAGTCTAACAATTGCTAACCATCAACGCTTTAATAAGACCCCTATA
50 AAAGCCGTTGGAAGTACAAAAGATTATGCCCCAAGAGTAGGCACTGTATCTGATACTATTGACGCGATCAA
AGGAAAAGTAAGCTCATTAGAAAAATCGTTTGTGCGCTATTTCATCAAGAAGCTGATATTATGGCAGCCCCAAG
CTAAAGTAAGTCAACTTCAAGGTAAATTAGCAAGCACTTAAGCAGTCAGACAGCTTAAATCTCCAAGTG
AGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTAAAGCAAAACAAGCACAACT
CGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACTGCACCAGACAGAAGCCT
TAGCAGAGCAAGCCGAGCCAGAGTGACAGCAGTGGTGGCTAAAAAAGCTCATTTGCAATATCTAAGGGAC
55 TTTAAATTGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATTTGATAATACTAAGCAAGATTGGCTAA
AACTACCTCATCTTTGTTAAATGCAAGAAAGCTTTAGCAGCCTTACAAGCTAAACAAAGCAGTCTAGAAG
CTACTATTGCTACCCAGAACACCAAGTTGACTTTGCTTAAACCTTAGCTAACGAAAAGGAATATCGCCAC
TTAGACGAAGATATAGCTACTGTGCTGATTGTCAGTAGCTCCACCTCTTACGGGCGTAAACCGCTATC
ATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAATGGTTAAAGAAACGAAACAACTATTAGAAGCTT
60 CAGCAAGATTAGCTGCTGAAAATACAAGTCTTGTAGCAGAAGCGCTTGTGGCCAAACCTCTGAAATGGTA
GCAAGTAATGCCATTGTGTCTAAAATCACATCTTCGATTACTCAGCCCTCATCTAAGACATCTTATGGCTC

AGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAAAGAGCTCTTAAAGCAGGAG
TCGTATGTTGGCAGCTGTGCGCCTCACAGGATTTAGGTTCCGTAAGGAATCTAAGTGA

65

SEQ ID NO: 7 secuencia de aminoácidos que comprende una región transmembrana C-terminal de GAS 40
ALKAGVVMLAAVGLTGFRFRKESK

SEQ ID NO: 8 secuencia de polinucleótidos que codifica una región transmembrana C-terminal de GAS 40

GCTCTTAAAGCAGGAGTCGTCATGTTGGCAGCTGTGCGCCTCACAGGATTTAGGTTCCGTAAGGAATCTAA
GTGA

SEQ ID NO: 9 secuencia de aminoácidos que comprende un fragmento de GAS 40 con una secuencia transmembrana C-terminal eliminada

MDLEQTKPNQVKQKIALSTIALLSASVGVSHQVKADDRASGETKASNTHDDSLPKPETIQEAKATIDAVE
KTLSSQKAELTELATALTKTTAEINHLKEQDNEQKALTSAQEIYNTNLASSEETLLAQGAHQRELTA
TELHNAQADQHSKETALSEQKASISAETTRAQDLVEQVKTSEQNIAKLNAMISNPDAITKAAQTANDNTKA
LSSELEKAKADLENQKAKVKKQLTEELAAQKAALAEKAEKLSRLKSSAPSTQDSIVGNNTMKAPQGYPLEE
LKKLEASGYIGSASYNMYKEHADQIIAKASPGNQLNQYQDI PADRNRFVDPDNLTPFVQNELAQFAAAMI
NSVRRQLGLPPVTVTAGSQEFARLLSTSYYKTHGNTRPSFVYGGQPGVSGHYGVGPHDKTIIEDSAGASGLI
RNDNMYENIGAFNDVHTVNGIKRGIYDSIKYMLFTDHLHGNTYGHAINFLRVDKHNPAPVYLGFTSNV
GSLNEHFVMPFESNIAHQRFNKTPIKAVGSTKDYAQRVGTVDITAAIKGVSSLENRLSAIHQEADIMA
AQAKVSQLOGLASTLKQSDSLNLQVRQLNDTKGSLRTELLAAKAKQAQLEATRDQSLAKLASLKAALHQT
EALAEQAAARVLTALVAKKAHLQYLRFKLPNRLQVIRERIDNTKQDLAKTSSLLNAQELAAALQAKQSS
LEATTATTEHQTLTKLANEKEYRHLDEDIATVPDLQVAPPLTGKPLSYSKIDTTPLVQEMVKETKQLL
EASARLAAENTSLVAEALVGQTSSEMVASNAIVSKITSSITQPSKTSYSGSSSTSNLISDVDESTQR

SEQ ID NO: 10 secuencia de polinucleótidos que codifica un fragmento de GAS 40 con una secuencia transmembrana C-terminal eliminada

ATGGACTTAGAACAAACGAAGCCAAACCAAGTTAAGCAGAAAATTGCTTTAACCTCAACAATTGCTTTATT
GAGTGCCAGTGTAGGCGTATCTCACCAGTCAAAGCAGATGATAGAGCTCAGGAGAAAACGAGGCGAGTA
ATACTCAGCAGTATAGTTTACCAAAACAGAAAATCAAGAGGCAAAGGCACTATTGATGCAGTTGAA
AAAACCTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAAAATCTGCTGAAAT
CAACCACTTAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAAGAAATTTACACTAATA
CTCTTGCAAGTAGTGAGGAGACGCTATAGCCCAAGGAGCCGAACATCAAGAGAGTTAACAGCTACTGAA
ACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCAATTTGTCAGAAACAAAAGCTAGCAT
TTCAGCAGAAACTACTCGAGCTCAAGATTTAGTGGAACAAGTCAAAACGCTCTGAACAAAATATTGCTAAGC
TCAATGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAGCA
TTAAGCTCAGAAATGGAGAAGGCTAAAGCTGACTTAGAAAATCAAAAAGCTAAAGTTAAAAAGCAATTGAC
TGAAGAGTTGGCAGCTCAGAAAGCTGCTTAGCAGAAAAAGAGGCAGAACTTAGTCGTCTTAAATCCTCAG
CTCCGCTACTCAAGATAGCATTGTGGGTAATAATACCATGAAAGCACCAGGCTATCCTCTTGAAGAA
CTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAAGAGCATGCAGA
TCAAATTATTGCCAAAGCTAGTCCAGGTAATCAATTAATCAATACCAAGATATTCCAGCAGATCGTAATC
GCTTTGTGATCCCGATAATTTGACACCAGAAGTGCAAAATGAGCTAGCGCAGTTTGCAGCTCACATGATT
AATAGTGTAAAGAGACAATTAGGCTTACCACCAAGTTACTGTTACAGCAGGATCACAGAAGATTGCAAGATT
ACTTAGTACCAGCTATAAGAAACCTCATGGTAATACAAGACCATCATTTGCTTACGGAGCAGGCGGTAT
CAGGGCATTATGGTGTGGGCTCATGATAAACTATTATTGAAGACTCTGCCGGAGCGCTCAGGGCTCATT
CGAAATGATGATAACATGTACGAGAATATCGGTGCTTTTAAACGATGTGCATCTGTGAATGGTATTAAACG
TGGTATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAATACATACGCCCATGCTA
TTAATCTTTTACGTGTAGATAAACATAACCCTAATGCGCCTGTTTACCTTGGATTTCACCCAGCAATGTA
GGATCTTTTGAACACTTTGTAATGTTTCCAGAGTCTAACATTTGCTAACCATCAACGCTTAAATAAGAC
CCCTATAAAAGCCGTTTGAAGTACAAAAGATTATGCCCAAAGAGTAGGCACTGTATCTGATACTATTGCAG
CGATCAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTGCGCTATTTCATCAAGAGCTGATATTATGGCA
GCCCAAGCTAAAGTAAGTCAACTTCAAGGTAATTTAGCAAGCACACTTAAGCAGTCAGACAGCTTAAATCT
CCAAGTGAGACAATTAAATGATCTAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTAAAGCAAAACAG
CACAACCTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACTGCACCCAGACA
GAAGCCTTAGCAGAGCAAGCCGAGCCAGAGTGACAGCACTGGTGGCTAAAAAGCTCATTGCAATATCT
AAGGACTTTTAAATTGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATTGATAATACTAAGCAAGATT

TGGCTAAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCTAAACAAAGCAGT
CTAGAAGCTACTATTGCTACCACAGAACACCAAGTTGACTTTGCTTAAACCTTAGCTAACGAAAAGGAATA
TCGCCACTTAGACGAAGATATAGCTACTGTGCGCTGATTGCAAGTAGCTCCACCTCTTACGGGCGTAAAC
CGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAATGGTTAAAGAAACGAAACAATATTA
GAAGCTTCAGCAAGATTAGCTGCTGAAAATACAGTCTTGTAGCAGAAGCGCTTGTGGCCAAACCTCTGA
AATGGTAGCAAGTAATGCCATTGTGTCTAAAATCACATCTTCGATTACTCAGCCCTCATCTAAGACATCTT
ATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAAAGA

SEQ ID NO: 11 secuencia de aminoácidos que comprende una región transmembrana de GAS 40 como se muestra en las figuras 1 y 2. ALKAGVVMLAAVGLTG

SEQ ID NO: 12 secuencia de aminoácidos que comprende una primera región con superenrollamiento de GAS 40

ETIQEAKATIDAVEKTLSSQKAELETALATLTKTTAEINHLKEQQDNEQKALTSAQEIYTNLTASSEETLL
AQGAHQRELATETELHNAQADQHSKETALSEQKASI SAETTRAQDLVEQVKTSEQNIAKLNAMISNPDA
ITKAAQTANDNTKALSSELEKAKADLENQKAKVKKQLTEELAAQKALAEKEAELSRLKSSA

SEQ ID NO: 13 secuencia de aminoácidos que comprende una segunda región con superenrollamiento de GAS 40

RLSAIHQEADIMAAQKVSQQLQGKLASTLKQSDSLNLQVRQLNDTKGSLRTELLAAKAKQAQLEATRDQSL
AKLASLKAALHQTEALAEQAAARVTALVAKKAHLQYLRDFKLNPNRLQVIRERIDNTKQDLAKTTSSLLNA
QEALAAQAKQSSLEATITATTEHQLTLLKTLANEKE

SEQ ID NO: 14 secuencia de aminoácidos que comprende un motivo cremallera de leucina dentro de la segunda región con superenrollamiento de GAS 40. QVIRFRIDNTKQDL2kKTTSSLLNAQEALAA

SEQ ID NO: 15 secuencia de aminoácidos que comprende SpA de *Streptococcus gordonii*, referencia del Genbank GI 25990270

MNKRKEVFGFRKSKVAKTLCGAVLGAALIAIADQQVLADDEVTTETNSTANVAVTTTGNPATNLPEAQGEATE
AASQSQAGSKEGALPVEVSADDLNQAVTDAKAAGVNVVQDQTSKGTATTAAENAQKQAEIKSDYAKQA
EEIKKTTEAYKKEVEAHQAETDKINAENKAAEDKYQEDLKAHQAEVEKINTANATAKAEYEAKLAQYQKDL
AAVQKANEDSQLDYQNKLSAYQAEELARVQKANAEEAEYKAVKENTAKNAALQAENEAIKQRNETAKANY
DAAMQYEAADLAAIKKAKEDNDADYQAKLAAYQAEELARVQKANADAKAAYEKAVEENTAKNTAIQAENEAI
KQRNAAAKATYEALQYEAADLAAKAKANEDSDADYQAKLAAYQTELARVQKANADAKAAYEKAVEDNKAK
NAALQAENEAIKQRNAAAKTDYEAELAKYEAADLAKYKELAEYPAKLKAYEDEQAQIKAALEVELEKNKNQD
GYLSKPSAQSLVYDSEPNAQLSLTNGKMLKASAVDEAFSHDTAQYSKKILQPDNLNVSYLQQAADDVTSSM
ELYGNFGDKAGWTTTVGNNTVEVKFASVLLERQSVTATYTNLEKSYNGKKISKAVFKYSLSDSKFKNV
KAWLGVLDPDTLGVFASAYTGQEEKDTSIFIKNEFTFYDENDQPINFDNALLSVAASLNRENNNSIEMAKDYS
GTFVKISGSSVGEKDGKIYATETLNFQGGQSSRWMTMYKNSQPGSGWSSDAPNSWYGAGALSMSGPTNHV
TVGAISATQVVPSPDPMVAATGKRPNWIWYSLNGKIRAVNVPKITKEKPTPPVAPTEPQAPTYEVEKPLEPA
PVAPTYENEPTPPVKTPDQPEPSKPEEPTYETEKPLEPAPVVPVPTYENEPTPPVKTPDQPEPSKPEEPTYET
EKPLEPAPVAPTYENEPTPPVKTPDQPEPSKPEEPTYDPLPTPPVAPTPKQLPTPPVPTVHFHYSSLLAQ
PQINKEIKNEDGVDIDRTLVAKQSIKVFELKTEALTAGRPKTTSFVLVDPLPTGYKFDLATAKASTGFD
TYDEASHTVTFKATDETATYNADLTKEVETLHPTVGRVLNDGATYINNFTLTVNDAYGIKSNVVRVTT
GKPNPDPNPNNNYIKPTKVNKNKEGLNIDGKEVLGASTNYEELTWDLQYKGDKSSKEAIQNGFYVDDYP
EEALDVRPDLVKVADEKGNQVSGVSVQYDLSLEAPKKVQDLLKKNITVKGAFLFSADNPEEFYKQYVS
TGTSLVITDPMTVKSEFGKTGGKYENKAYQIDFGNGYATEVVVNVNPKITPKKDVTVSLDPTSENLDGQTV
QLYQTFNYRLIGGFIPQNHSELEEDYSFVDDYDQAGDQYTGNYKTFSSNLNLMKDGSVIKAGTDLTSQTTA
ETDAANGIVTVRSKEDSLQKISLDSPPQAETYLQMRRIAIGTFENTYVNTVNVKAYASNTVTRTTTPIRTP
DKPTPIPTPKPDKPKETPKPKVPSPKVEDPSAPIPVSVGKELTLPKTGTNDSSYPYLGALVGV
GLGQLKRKEDESN

SEQ ID NO: 16 Secuencia de aminoácidos que comprende el precursor de la proteína B de superficie estreptocócica de *Streptococcus gordonii*, de referencia del Genbank GI 25055226 AAC44102.3

5
10
15
20
25
30
35
40
45
50
55
60
65

MQKREVFGRKSKVAKTLCGAVLGAALIAIADQQVLADDEVTTETNSTANVAVTTTGNPATNLPEAQGEATEA
ASQSQAQAGSKDQALPVEVSADDLNKAVTDAKAAGVNVVQDQTSKGTATTAAENAQKQAEIKSDYAKQAE
EIKKTTEAYKKEVEAHQAETDKINAENKAAEDKYQEDLKAHQAEVEKINTANATAKAAYEAKLAQYQKDLA
AVQKANEDSQLDYQNKLSAYQAEELARVQKANAEAKEAYEKAVKENTAKNAALQAENEAIKQRNETAKANYD
AAMKQYEAADLAAIKKAKEDNDADYQAKLAAYQAEELARVQKANADAKAAYEKAVEENTAKNTAIQAENEAIK
QRNETAKATYEAAYKQYEAADLAAVQKANATNEADYQAKLAAYQTELARVQKANADAKATYKAVEDNKA
AALQAENEAIKQRNAAAKTDYEAELAKYEADLAKYKDDFAAYTAALAEAESKKKQDGYLSEPRSQSLNPKS
EPNAIRTIDSSVHQYGGQELDALVKSWSGISPTNPDRKKSTAYSIFYNAINSNNTYAKLVLEKDKPVDVITYTG
LKNSSFNKKISKVYTYTLKETGFDGKMTMFASDDPTVTAWYNDYFTSTNINVKVYDEEGQLMNL
GGLVNFSSLNRGNGSGAIDKDAIESVRNFNGRYIPISGSSIKIHENNSAYADSSNAEKSRGARWDTSEWDT
TSSPNNWYGAIVGEBITQSEISPNMASSKSGNIWFAPNSNINAIGVPTKPVAPTAPTQPMYETEKPLEPAPV
VPTYENEPTPPVKTPDQPEPSKPEEPTYETEKPLEPAPVAPTYPPEPTPPVKIPDQPEPSKPEEPTYETEK
PLEPAPVAPTYPPEPTPPVKTPDQPEPSKPEEPTYDPLPTPLAPTQKQLPTPPVPTVHFHYSSLLAQFQ
INKEIKNEDGVDIDRTLVAQSIGKPFELKTEALTAGRPKTTSFVLVDPLPTGYKFDLDATKAASTGFDITY
DEASHTVTFKATDETLATYNADLTTPVETLHPTVVGRLNDGATYTNNFTLTVDAYGKSNVVRVTPGK
PNDPDNPNNNYIKPTKVKNKEGLNIDGKEVLGAGSTNYEELTWDLQYKGDSSKEAIQNGFYVDDYPPE
ALDVRPDLVKVADEKGNQVSGVSVQYDSLEAPKKVQDLLKKANTVKGAPQLFSADNPPEEFYKQVYSTG
TSLVITDPMPTVKSEFGKTGGKYENKAYQIDFGNGYATEVVVNVNPKITPKKDVTVSLDPTSENLDGQTVQL
YQTFNYRLIGGFIPQNHSEELDYSFVDDYDQAGDQYTGNYKTFSSNLMTMKDGSVIKAGTDLTSQTAE
DATNGIVTVRFKEDFLQKISLSDSPFQAETYLQMRRIAIGTFENTYVNTVNVKAYASNTVTTTTPIPTPDK
PTPIPTPKPKDPDKPETPKPKVPSPKVEDPSAPIPVSVGKELTTLPKTGTNDATYMPYLGLAALVGFLGL
GLAKRKED

SEQ ID NO: 17 secuencia de aminoácidos que comprende PspA de *Streptococcus pneumoniae*, referencia del Genbank GI 282335

MNKKKMILTSLASVAILGAGFVASQPTTVRAEESFVASQSKAEKDYDAKKDAKNAKKAVEDAQKALDDAK
AAQKKYDEDQKKTEKAALKAASEMDKAVAAVQQAYLAYQQATDKAAKDAADKMIDEAKKREERAKTKF
NTVRAMVVPEPEQLAETKKKSEEAKQKAPELTKKLEBAKAKLEEAEEKKATEAKQKVDAEEVAPQAKIAELE
NQVHRLEQELKEIDSESEEDYAKEGFRAPLQSKLDAKKAKLSKLEELSDKIDELDAETAKLEDQLKAAEEN
NNVEDYFKEGLEKTIAAKKAELEKTEADLKKAVNEPEKPAPAPETPAPEAPAEQPKPAPAPQAPAPKPEK
PAEQPKPEKTDQQAEDYARRSEEEYNRLTQQQPPKAEPAPAPKTGWKQENGWYFYNTDGSMTGWLQ
NNGSWYYLNSNGAMATGWLQYNGSWYYLNGAMATGWAKVNGSWYYLNGAMATGWLQYNGSWYYLNGAN
GAMATGWAKVNGSWYYLNGAMATGWLQYNGSWYYLNGAMATGWAKVNGSWYYLNGAMATGWVKG
DTWYYLEASGAMKASQWFKVSDKWYVYVNGLGALAVNTTVDGYKVNANGWV

SEQ ID NO: 18 secuencia de aminoácidos que comprende una porción de Se89.9 de *Streptococcus equi*, referencia del Genbank GI 2330384

ESDIVDATRFSTTEIPKSGQVIDRSASIQALTNDIASIKGKIASLESRLADPSSEAEVTAQAQKISQLQH
QLEAAQAKSHKLDQQVEQLANTKDSLRTQLLAAKEEQALKANLDKALALLASSKATLHKLEAAMEEAKA
RVAGLASQKAQLEDLLAFEKNPNRIELAQEKVAAKKALADTEDKLLAAQASLSDLQAQRARLQLSIATT

SEQ ID NO: 19 secuencia de polinucleótidos que comprende GST-40-HIS

CTGGTTCCGCGTGGATCCCATATGAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGG
 5 AGAAACGAAGGCGAGTAATACTCACGACGATAGTTTACCAAAACCAGAAACAATTCAAGAGGCAAGGCAA
 CTATTGATGCAGTTGAAAAAAGCTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGCACA
 AAACTACTGCTGAAATCAACCACCTAAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACTCTGCACA
 AGAAATTTACTACTACTCTTGCAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAAGAG
 AGTTAACAGCTACTGAAACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCATTSTCA
 10 GAACAAAAAGCTAGCATTTTCAGCAGAACTACTCGAGCTCAAGATTTAGTGGAACAGTCAAAACGCTCTGA
 ACAAATATTGCTAAGCTCAATGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAAACGGCTA
 ATGATAATACAAAAGCATTAAGCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAATCAAAAAGCTAAA
 GTTAAAAAGCAATTGACTGAAGAGTTGGCAGCTCAGAAAAGCTGCTCTAGCAGAAAAAGAGGCAGAACTTAG
 TCGTCTTAAATCCTCAGCTCCGCTACTCAAGATAGCATTTGTGGGTAAATAACCATGAAAGCACCAGCAAG
 15 GCATCTCTTGAAGAACTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAAATAATTAT
 TACAAAGAGCATGCAGATCAAAATTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATAT
 TCCAGCAGATCGTAATCGCTTTGTTGATCCCGATAATTTGACACCAGAGTGCAAAATGAGCTAGCGCAGT

TTGCAGCTCACATGATTAAATAGTGTAAAGACAAATAGGTCTACCACCAGTTACTGTTACAGCAGGATCA
 20 CAAGAATTTGCAAGATTACTTAGTACCAGCTATAAGAAAAGCTCATGGTAATACAAGACCATCATTTGTCTA
 CGGACAGCCAGGGGTATCAGGGCATTTATGGTGTGGGCTCATGATAAACTATTATTGAAGACTTGCCTG
 GAGCGTCAGGGCTCATTGCAATGATGATAACATGTACGAGAAATATCGGTGCTTTTAAAGATGTGCATACT
 GTGAATGGTATTAAACGTGGTATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAAA
 TACATACGGCCATGCTATTAACTTTTACGTGTAGATAAACATAACCTTAATGCGCCTGTTTACCTTGGAT
 25 TTTCAACCAGCAATGTAGGATCTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATTGCTAACCAT
 CAACGCTTTAATAAGACCCCTATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAGAGTAGGCACTGT
 ATCTGATACTATTGCAGCGATCAAAAGGAAAAGTAAAGCTCATTAGAAAATCGTTTGTGCGCTATTCATCAAG
 AAGCTGATATTATGGCAGCCCAAGCTAAAGTAAGTCAACTTCAAGGTAAATTAGCAAGCACACTTAAGCAG
 TCAGACAGCTTAAATCTCCAAGTGAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGAAATTACTAGC
 30 AGCTAAAGCAAAACAAGCAAACTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAG
 CCGCACTGCACAGACAGAAAGCCTTAGCAGAGCAAGCCGACGACAGAGTACAGCACTGGTGGCTAAAAAA
 GCTCATTTGCAATATCTAAGGGACTTTAAATTGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATTGA
 TAATACTAAGCAAGATTGGCTAAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTAC
 AAGCTAAACAAGCAGTCTAGAAGCTACTATTGCTACACAGAACACCAGTTGACTTTGCTTAAACCTTA
 35 GCTAACGAAAAGGAATATCGCCACTTAGACGAAGATATAGCTACTGTGCCTGATTGCAAGTAGCTCCACC
 TCTTACGGGCGTAAACCGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAATGGTTAAAG
 AAACGAAACAACATTATAGAAGCTTCAGCAAGATTAGCTGCTGAAAATACAAGTCTTTGTAGCAGAAAGCGCTT
 GTTGGCCAAACCTCTGAAATGGTAGCAAGTAATGCCATTGTGTCTAAAATCACATCTTCGATTACTCAGCC
 CTCATCTAAGACATCTTATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTA
 40 CTCAAAGAGCTCTTAAAGCAGGAGTCGTCTATGTTGGCAGCTGTGCGCCTCAGAGGATTTAGGTTCCGTAAG
 GAATCTAAGGCGCGCACTCGAGCACCACCACCACCACCACCAC

SEQ ID NO: 20 secuencia de aminoácidos que comprende GST-40-HIS

L V P R G S H Met S V G V S H Q V K A D D R A S G E T K A S N T H D D S
 L P K P E T I Q E A K A T I D A V E K T L S Q Q K A E L T E L A T A L T
 45 K T T A E I N H L K E Q Q D N E Q K A L T S A Q E I Y T N T L A S S E E
 T L L A Q G A E H Q R E L T A T E T E L H N A Q A D Q H S K E T A L S E
 Q K A S I S A E T T R A Q D L V E Q V K T S E Q N I A K L N A Met I S N
 P D A I T K A A Q T A N D N T K A L S S E L E K A K A D L E N Q K A K V
 K K Q L T E E L A A Q K A A L A E K E A E L S R L K S A P S T Q D S I
 50 V G N N T Met K A P Q G Y P L E E L K K L E A S G Y I G S A S Y N N Y Y
 K E H A D Q I I A K A S P G N Q L N Q Y Q D I P A D R N R F V D P D N L
 T P E V Q N E L A Q F A A H Met I N S V R R Q L G L P P V T V T A G S Q
 E F A R L L S T S Y K K T H G N T R P S F V Y G Q P G V S G H Y G V G P
 H D K T I I E D S A G A S G L I R N D D N Met Y E N I G A F N D T V
 N G I K R G I Y D S I K Y Met L F T D H L H G N T Y G H A I N F L R V D
 55 K H N P N A P V Y L G F S T S N V G S L N E H F V Met F P E S N I A N H
 Q R F N K T P I K A V G S T K D Y A Q R V G T V S D T I A A I K G K V S
 S L E N R L S A I H Q E A D I Met A A Q A K V S Q L Q G K L A S T L K Q
 S D S L N L Q V R Q L K G S L R T E L L A A K A K Q A Q L E A T R
 D Q S L A K L A S L K A A L H Q T E A L A E Q A A A R V T A L V A K K A
 60 H L Q Y L R D F K L N P N R L Q V I R E R I D N T K Q D L A K T T S S L
 L N A Q E A L A A L Q A K Q S S L E A T I A T T E H Q L T L L K T L A N
 E K E Y R H L D E D I A T V P D L Q V A P P L T G V K P L S Y S K I D T
 T P L V Q E Met V K E T K Q L L E A S A R L A A E N T S L V A E A L V G
 Q T S E Met V A S N A I V S K I T S S I T Q P S S K T S Y G S G S S T T
 65 S N L I S D V D E S T Q R A L K A G V V Met L A A V G L T G F R F R K E
 S K A A A L E H H H H H H

SEQ ID NO: 21 secuencia de polinucleótidos que comprende 40a-HIS

5 **ATGAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAATAC**
TCACGACGATAGTTTACCAAACAGAAACAATTCAAGAGGCAAAGGCAACTATTGATGCAGTTGAAAAAA
CTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAACTACTGCTGAAATCAAC
CACTTAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAATTTACACTAATACTCT

10 **TGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTAACAGCTACTGAAACAG**
AGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCATTGTCAGAACAAAAAGCTAGCATTTCA
GCAGAACTACTCGAGCTCAAGATTTAGTGGAACAAGTCAAACGCTCTGAACAAAATATTGCTAAGCTCAA
TGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAGCATTA
15 **GCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAATAACAAAAGCTAAAGTTAAAAAGCAATTGACTGAA**
GAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCAGAACTTAGTCGTCTTAAATCCTCAGCTCC
GTCTACTCAAGATAGCATTGTGGGTAATAATACCATGAAAGCACCAGCAAGGCTATCCTCTTGAAGAACTTA
AAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAGAGCATGCAGATCAA
ATTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATATTCCAGCAGATCGTAATCGCTT
20 **TGTTGATCCCGATAATTGACACCAGAAAGTGCAAAATGAGCTAGCGCAGTTTGCAGCTCACATGATTAATA**
GTGTAAGAAGACAATTAGGTCTACCACAGTTACTGTTACAGCAGGATCACAAGAATTTGCAAGATTACTT
AGTACCAGCTATAAGAAAACTCATGGTAATACAAGACCATCATTTGTCTACGGACAGCCAGGGGTATCAGG
GCATTATGGTGTGGGCCCTCATGATAAACTATTATTGAAGACTCTGCCGGAGCGTCAGGGCTCATTCGAA
ATGATGATAACATGTACGAGAATATCGGTGCTTTTAAACGATGTGCATACTGTGAATGGTATTAAACGTGGT
25 **ATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAAATACATACGGCCATGCTATTAA**
CTTTTTACGTGTAGATAAACATAACCCTAATGCGCCTGTTTACCTTGGATTTTCAACCAGCAATGTAGGAT
CTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATTGCTAACCATCAACGCTTTAATAAGACCCCT
ATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAAGAGTAGGCACGTGTATCTGATACTATTGCAGCGAT
CAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTGCGCTATTTCATCAAGAAGCTGATATTATGGCAGCCC
30 **AAGCTAAAGTAAGTCAACTTCAAGGTAAATTAGCAAGCACACTTAAGCAGTCAGACAGCTTAAATCTCCAA**
GTGAGACAAATTAAATGATACTAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTAAAGCAAAACAAGCACA
ACTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACCTGCACCAGACAGAAG
CCTTAGCAGAGCAAGCCGCAGCCAGAGTGACAGCACTGGTGGCTAAAAAGCTCATTTGCAATATCTAAGG
GACTTTAAATTGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATTGATAATACTAAGCAAGATTTGGC
35 **TAAAACCTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCTAAACAAAGCAGTCTAG**
AAGCTACTATTGCTACCACAGAACACCAGTTGACTTTGCTTAAAACCTTAGCTAACGAAAAGGAATATCGC
CACTTAGACGAAGATATAGCTACTGTGCTGATTTGCAAGTAGCTCCACCTCTTACGGGCGTAAAACCGCT
ATCATATAGTAAGATAGATACTACTCCGCTTGTTCAAGAAATGGTTAAAGAAACGAAACAACATATTAGAAG
CTTCAGCAAGATTAGCTGCTGAAAATACAGTCTTGTTAGCAGAAGCGCTTGTTGGCCAAACCTCTGAAATG
40 **GTAGCAAGTAATGCCATTGTGTCTAAATCACATCTTCGATTACTCAGCCCTCATCTAAGACATCTTATGG**
CTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAAAGTGGCGGCCGCACTCG
AGCACCACCACCACCACCACCAC

SEQ ID NO: 22 secuencia de aminoácidos que comprende 40a-HIS

5 MSVGVSHQVKADDDRASGETKASNTHDDSLPKPETIQ
 EAKATIDAVEKTLSSQQKAELTELATALTKTTAETINH
 LKEQQDNEQKALTSAQEIYTNNTLASSSEETLLAQGAEE
 HQRELTAETEELHNAQADQHSKETALSEQKASISAAE
 TTRAQDQDLVEQVKTSSEQNIAKLNAMetISNPDAITKAA
 10 QTANDNTKALSSELEKAKADLENQKAKVKKQLTEEL
 AAQKAAALAEKEAEELSRLEKSSAPSTQDSIVGNNTMKAA
 PQGYPLEELKKLEASGYIGSSASYNMYK'EHADQIIIA
 KASPGNQNLNQYQDIPADRNRPFVDPDNLTPENVQNELA
 QFAAHMINSVRRQLGLPPVTVTAGSQEFARLLSTSY
 15 KKTTHGNTTRPSFVYGPQPGVSGHYGVGPHDKTIIEDSA
 GASGLIRNDDDNMYENIGAFNDVHTVNGIKRGIIYDSI
 KYMLFTDHLHGNTYGHAINFLRVQKHNPNAFVYLGFI
 STSNVGSLSNEHFVMFPESNIAHQRFNKTPIKAVGS
 TKDYAQRVGTVSDTIAAIIKGVSSLENRLSAIHQEA
 20 DIMAAQAKVSQQLQGKLASTLKQSDSLNLQVRQLNDT
 KGSRLRTELLAAKAKQAQLEATRQDQSLAKLASLKAAL
 HQTEALAEQAAARVTAALVAKKAHLQYLRDFKLNPNR
 LQVIRERIDNTKQDLAKTTSSLLNAQEAALAAALQAKQ
 SSLEATIATTEHQLTLLKTLANEKEYRHLDEDIATV
 25 PDLQVAPPPLTGVKPLSYSKIDTTPPLVQEMVKETKQL
 LEASARLAENTSLVAEALVGGQTSEMVASNAIVSKI

 TSSITQPSSSKTSYGS GSSTTSNLI SDVDESTQRAAA
 30 LEHHHHHHHH

SEQ ID NO: 23 secuencia de polinucleótidos que comprende 40a-RR-HIS

5
10
15
20
25
30
35
40
45
50
55
60
65

ATGAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAATAC
TCACGACGATAGTTTACCAAAACCAGAAACAATTCAAGAGGCAAAGGCAACTATTGATGCGAGTTGAAAAA
CTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAAACTACTGCTGAAATCAAC
CACTTAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAATTTACACTAATACTCT
TGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTAACAGCTACTGAAACAG
AGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCATTGTGAGAACAAAAAGCTAGCATTTCA
GCAGAACTACTCGAGCTCAAGATTTAGTGGAAACAAGTCAAACGCTCTGAACAAAAATATTGCTAAGCTCAA
TGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAAGCAATTAA
GCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAAATCAAAAAGCTAAAGTTAAAAAGCAATTGACTGAA
GAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCAGAACTTAGTCGCTCTTAAATCCTCAGCTCC
GTCTACTCAAGATAGCATTGTGGGTAATAATACCATGAAAGCACCGCAAGGCTATCCTCTTGAAGAACCTTA
AAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAAGAGCATGCAGATCAA
ATTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATATTCCAGCAGATCGTAATCGCTT
TGTTGATCCCGATAATTTGACACCAGAAAGTGCAAAATGAGCTAGCGCAGTTTGCAGCTCACATGATTAAATA
GTGTAGCtGtGtCAATTAGGTCTACCACCAGTTACTGTTACAGCAGGATCACAAGAAATTTGCAAGATTACTT
AGTACCAGCTATAAGAAACTCATGGTAATACAAGACCATTATTGTCTACGGACAGCCAGGGGTATCAGG
GCATTATGGTGTGGGCTCATGATAAACTATTATTGAAGACTCTGCCGGAGCTCAGGGCTCATTCGAA
ATGATGATAACATGTACGAGAATATCGGTGCTTTTAAACGATGTGCATACTGTGAATGGTATTAAACGTGGT
ATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTACACGGAAATACATACGGCCATGCTATTAA
CTTTTACGTGTAGATAAAACATAACCCTAATGCGCTGTTTACCTTGGATTTTCAACCAGCAATGTAGGAT
CTTTGAATGAACACTTTGTAATGTTCCAGAGTCTAACATTGCTAACCATCAACGCTTTAATAAGACCCCT
ATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAAGAGTAGGCACTGTATCTGATCTATTGCAGCGAT
CAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTGCGCTATTATCAAGAAGCTGATATTATGGCAGCCC
AAGCTAAAGTAAGTCAACTTCAAGGTAAATTAGCAAGCACACTTAAGCAGTCAGACAGCTTAAATCTCCAA
GTGAGACAATTAATGATACTAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTAAAGCAAAACAGCACA
ACTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACTGCACCAGACAGAAG
CCTTAGCAGAGCAAGCCGACGCCAGAGTGACAGCACTGGTGGCTAAAAAGCTCATTTGCAATATCTAAGG
GACTTTAAATTGAATCCTAACCGCTTCAAGTGATACGTGAGCGCATTGATAATACTAAGCAAGATTTGGC
TAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCTAAACAAAGCAGTCTAG
AAGCTACTATTGCTACCACAGAACACCAGTTGACTTTGCTTAAACCTTAGCTAACGAAAAGGAATATCGC
CACTTAGACGAAGATATAGTACTGTGCTGATTGTCAAGTAGCTCCACCTCTTACGGGCGTAAACCGCT
ATCATATAGTAAGATAGATACTACTCCGCTTGTTCAAGAAATGGTTAAAGAAACGAAACAACTTAGAAG
CTTCAGCAAGATTAGCTGCTGAAAATACAAAGTCTTGTAGCAGAAGCGCTTGTGGCCAAACCTTGAAATG
GTAGCAAGTAATGCCATTGTGTCTAAAAATCACATCTTCGATTACTCAGCCCTCATCTAAGACATCTTATGG
CTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAAGCtGtGCGGCCGCACTCG
AGCACCACCACCACCACCACCAC

SEQ ID NO: 24 secuencia de aminoácidos que comprende 40a-RR-HIS

45
50
55
60
65

M S V G V S H Q V K A D D R A S G E T K A S N T H D D S L P K P E T I Q
E A K A T I D A V E K T L S Q Q K A E L T E L A T A L T K T T A E I N H
L K E Q Q D N E Q K A L T S A Q E I Y T N T L A S S E E T L L A Q G A E
H Q R E L T A T E T E L H N A Q A D Q H S K E T A L S E Q K A S I S A E
T T R A Q D L V E Q V K T S E Q N I A K L N A Met I S N P D A I T K A A
Q T A N D N T K A L S S E L E K A K A D L E N Q K A K V K K Q L T E E L
A A Q K A A L A E K E A E L S R L K S S A P S T Q D S I V G N N T M K A
P Q G Y P L E E L K K L E A S G Y I G S A S Y N N Y Y K E H A D Q I I A
K A S P G N Q L N Q Y Q D I P A D R N R F V D P D N L T P E V Q N E L A
Q F A A H M I N S V R R Q L G L P P V T V T A G S Q E F A R L L S T S Y
K T H G N T R P S F V Y G Q P G V S G H Y G V G P H D K T I E D S A
G A S G L I R N D D N M Y E N I G A F N D V H T V N G I K R G I Y D S I
K Y M L F T D H L H G N T Y G H A I N F L R V D K H N P N A P V Y L G F
S T S N V G S L N E H F V M F P E S N I A N H Q R F N K T P I K A V G S

5
10
T K D Y A Q R V G T V S D T I A A I K G K V S S L E N R L S A I H Q E A
D I M A A Q A K V S Q L Q G K L A S T L K Q S D S L N L Q V R Q L N D T
K G S L R T E L L A A K A K Q A Q L E A T R D Q S L A K L A S L K A A L
H Q T E A L A E Q A A A R V T A L V A K K A H L Q Y L R D F K L N P N R
L Q V I R E R I D N T K Q D L A K T T S S L L N A Q E A L A A L Q A K Q
S S L E A T I A T T E H Q L T L L K T L A N E K E Y R H L D E D I A T V
P D L Q V A P P L T G V K P L S Y S K I D T T P L V Q E M V K E T K Q L
L E A S A R L A A E N T S L V A E A L V G Q T S E M V A S N A I V S K I
T S S I T Q P S S K T S Y G S G S S T T S N L I S D V D E S T Q R A A A
L E H H H H H H H

SEQ ID NO: 25 secuencia de polinucleótidos que comprende 40a-RR (nat)

15
20
25
30
35
40
45
50
ATGAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAATAC
TCACGACGATAGTTTACCAAAACCAGAAAACAATTCAAGAGGCCAAGGCAACTATTGATGCAGTTGAAAAAA
CTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAACTACTGCTGAAATCAAC
CACTTAAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACTCTGCACAAGAAATTTACACTAATACTCT
TGCAAGTAGTGGAGAGCGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTAAACAGCTACTGAAACAG
AGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCATTGTCAGAACAAAAAGCTAGCATTTCA
GCAGAACTACTCGAGCTCAAGATTTAGTGGAACAAGTCAAAACGCTCTGAACAAAATATTGCTAAGCTCAA
TGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAGCATTAA
GCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAATCAAAAAGCTAAAGTTAAAAAGCAATTGACTGAA
GAGTTGGCAGCTCAGAAAAGCTGCTCTAGCAGAAAAAGAGGCGAGAACTTAGTCGCTCTAAATCCTCAGCTCC
GTCTACTCAAGATAGCATTTGTGGGTAAATAATACCATGAAAGCACCGCAAGGCTATCCTCTTGAAGAATT
AAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAAGAGCATGCAGATCAA
ATTATTTGCCAAAGCTAGTCCAGGTAAATCAATTAATCAATACCAAGATATTCCAGCAGATCGTAATCGCTT
TGTTGATCCCGATAATTTGACACCAGAAAGTGCAAAATGAGCTAGCGCAGTTTGCAGCTCACATGATTAATA
GTGTACGtGtCAATTAGGTCTACCAAGTTACTGTTACAGCAGGATCACAAGAATTTGCAAGATTACTT
AGTACAGCTATAAGAAAACTCATGTAATACAGAACCATCATTTGTCTACGGACAGCCAGGGGTATCAGG
GCATTATGGTGTGGGCCCTCATGATAAACTATTATTGAAGACTCTGCCGGAGCGTCAGGGCTCATTGCAA
ATGATGATAACATGTACGAGAATATCGGTGCTTTTAAACGATGTGCATACTGTGAATGGTATTAAACGTGGT
ATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAAATACATACGGCCATGCTATTAA
CTTTTACGTGTAGATAAAACATAACCCCTAATGCCCTGTTTACCTTGGATTTTCAACCAGCAATGTAGGAT
CTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATTGCTAACCATCAACGCTTTAATAAGACCCCT
ATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAGAGTAGGCACTGTATCTGATACTATTGCAGCGAT
CAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTGGCTATTTCATCAAGAAGCTGATATTATGGCAGCCC
AAGCTAAAAGTAAGTCAACTTCAAGGTAAATTAGCAAGCACACTTAAGCAGTCAGACAGCTTAAATCTCCAA
GTGAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGAAATTACTAGCAGCTAAAGCAAAACAAGCACA
ACTCGAAGCTACTCGTGCATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACTGCACAGACAGAAG
CCTTAGCAGCAAGCCGACCCAGAGTACAGCACTGGTGGCTAAAAAGCTCATTGCAATATCTAAGG
GACTTTAAATTGAATCCTAACCGCTTCAAGTGATACGTGAGCGCATTGATAATACTAAGCAAGATTGCGC
TAAACTACCTCATCTTTGTTAAATGCACAAGAGCTTTAGCAGCCTTACAAGCTAAACAAAGCAGTCTAG
AAGCTACTATTGCTACCAACAGAACACAGTTGACTTTGCTTAAACCTTAGCTAACGAAAAGGAATATCGC
CACTTAGACGAAGATATAGCTACTGTGCCTGATTGCAAGTAGCTCCACCTCTTACGGGCGTAAACCCGCT
ATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAAATGGTTAAAGAAACGAAACAACCTATTAGAAG
CTTCAGCAAGATTAGCTGCTGAAAATACAAGTCTGTAGCAGAAAGCGCTTGTGGCCAAACCTCTGAAATG
GTAGCAAGTAATGCCATTGTGTCTAAAATCACATCTTCGATTACTCAGCCCTCATCTAAGACATCTTATGG
CTCAGGATCTTCTACAACGAGCAATCTCATTCTGATGTTGATGAAAGTACTCAAAGT

SEQ ID NO: 26 secuencia de aminoácidos que comprende 40a-RR (nat)

55
60
M S V G V S H Q V K A D D R A S G E T K A S N T H D D S L P K P E T I Q
E A K A T I D A V E K T L S Q Q K A E L T E L A T A L T K T T A E I N H
L K E Q Q D N E Q K A L T S A Q E I Y T N T L A S S E E T L L A Q G A E
H Q R E L T A T E T E L H N A Q A D Q H S K E T A L S E Q K A S I S A E
T T R A Q D L V E Q V K T S E Q N I A K L N A M I S N P D A I T K A A Q
T A N D N T K A L S S E L E K A K A D L E N Q K A K V K K Q L T E E L A
A Q K A A L A E K E A E L S R L K S S A P S T Q D S I V G N N T M K A P

5 Q G Y P L E E L K K L E A S G Y I G S A S Y N N Y Y K E H A D Q I I A K
 A S P G N Q L N Q Y Q D I P A D R N R F V D P D N L T P E V Q N E L A Q
 F A A H M I N S V R R Q L G L P P V T V T A G S Q E F A R L L S T S Y K
 K T H G N T R P S F V Y G Q P G V S G H Y G V G P H D K T I I E D S A G
 A S G L I R N D D N M Y E N I G A F N D V H T V N G I K R G I Y D S I K
 Y M L F T D H L H G N T Y G H A I N F L R V D K H N P N A P V Y L G F S
 10 T S N V G S L N E H F V M F P E S N I A N H Q R F N K T P I K A V G S T
 K D Y A Q R V G T V S D T I A A I K G K V S S L E N R L S A I H Q E A D
 I M A A Q A K V S Q L Q G K L A S T L K Q S D S L N L Q V R Q L N D T K
 G S L R T E L L A A K A K Q A Q L E A T R D Q S L A K L A S L K A A L H
 Q T E A L A E Q A A A R V T A L V A K K A H L Q Y L R D F K L N P N R L
 15 Q V I R E R I D N T K Q D L A K T T S S L L N A Q E A L A A L Q A K Q S
 S L E A T I A T T E H Q L T L L K T L A N E K E Y R H L D E D I A T V P
 D L Q V A P P L T G V K P L S Y S K I D T T P L V Q E M V K E T K Q L L
 E A S A R L A A E N T S L V A E A L V G Q T S E M V A S N A I V S K I T
 20 S S I T Q P S S K T S Y G S G S S T T S N L I S D V D E S T Q R

SEQ ID NO: 27 secuencia de polinucleótidos que comprende HIS-40a NH

25 **ATGGGATCGCATCACCATCACCATCAGCTAGTAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAG**
 AGCCTCAGGAGAAACGAAGGCGAGTAATACTCAGCAGCATAGTTTACCAAAACCAGAAACAATTCAAGAGG
 CAAAGGCAACTATTGATGCAGTTGAAAAAAGCTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACC
 GCTCTGACAAAAACTACTGCTGAAATCAACCACTTAAAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAAC
 CTCTGCACAAGAAATTTACACTAATACTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAAC
 30 **A**CAAAGAGAGTTAACAGCTACTGAAACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACT
 GCATTGTCAGAACAAAAAGCTAGCATTTCAGCAGAACTACTCGAGCTCAAGATTTAGTGGAAACAAGTCAA
 AACGCTCGAACAAAATATTGCTAAGCTCAATGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTC
 AAACGGCTAATGATAATACAAAAGCATTAAAGCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAATCAA
 AAAGCTAAAGTTAAAAAGCAATTGACTGAAGAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGC
 35 AGAAGTTAGTCGCTTTAAATCCTCAGCTCCGCTCTACTCAAGATAGCATTTGTGGGTAATAATACCATGAAAG
 CACCGCAAGGCTATCCTCTTGAAGAACTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTAC
 AATAATTATTACAAAGAGCATGCAGATCAAATTTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATA
 CCAAGATATTCAGCAGATCGTAATCGCTTTGTTGATCCCGATAATTTGACACCAGAAGTGCAAAATGAGC
 TAGCGCAGTTTGCAGCTCAGATGATTAAATAGTGTAGAAGACAATTAGGTCTACCACCAGTTACTGTTACA
 40 GCAGGATCACAAGAATTGCAAGATTACTTAGTACCAGCTATAAGAAAACCATGGTAATACAAGACCATC
 ATTTGTCTACGGACAGCCAGGGGTATCAGGGCATTATGGTGTGGGCCTCATGATAAACTATTATTGAAG
 ACTCTGCCGAGCGTCAGGGCTCATTCGAAATGATGATAACATGTACGAGAATATCGGTGCTTTTAACGAT
 GTGCATCTGTGAATGGTATTAAACGTGGTATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATT
 ACACGGAAATACATACGGCCATGCTATTAACTTTTACGTGTAGATAAAACATAACCCTAATGCGCCTGTTT
 45 ACCTTGGATTTTCAACCAGCAATGTAGGATCTTTGAAATGAACACTTTGTAAATGTTTCCAGAGTCTAACATT
 GCTAACCATCAACGCTTTAATAAGACCCCTATAAAAGCCGTTGGAAGTACAAAAGAGTATGCCCAAAGAGT
 AGGCACTGTATCTGATACTATTGCAGCGATCAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTCCGGCTA
 TTCATCAAGAAGCTGATATTATGGCAGCCCAAGCTAAAGTAAGTCAACTTCAAGGTAATTAGCAAGCACA
 CTTAAGCAGTCAGACAGCTTAAATCTCCAAGTGAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGA
 50 ATTACTAGCAGCTAAAGCAAAACAAGCACAACCTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCAT
 CGTTGAAAGCCGCACTGCACCAGACAGAAGCCTTAGCAGAGCAAGCCGCAGCCAGAGTGACAGCACTGGTG
 GCTAAAAAGCTCATTGCAATATCTAAGGGACTTTAAATTTGAATCCTAACCCGCTTCAAGTGATACGTGA
 GCGCATTGATAATACTAAGCAAGATTTGGCTAAAACTACCTCATCTTTGTTAAATGCACAAGAAAGCTTTAG
 55 CAGCCTTACAAGCTAAACAAAGCAGCTAGAAAGCTACTATTGCTACCACAGAACACCAGTTGACTTTGCTT
 AAAACCTTAGCTAACGAAAAGGAATATCGCCACTTAGACGAAGATATAGCTACTGTGCCTGATTTGCAAGT
 AGCTCCACCTCTTACGGGCGTAAACCGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAA
 TGGTTAAAGAAACGAACAACCTATTAGAAGCTTCAGCAAGATTAGCTGCTGAAAATACAGTCTTGTAGCA
 GAAGCGCTTGTGGCCAAACCTCTGAAATGGTAGCAAGTAATGCCATTGTGTCTAAATCACATCTTCGAT
 TACTCAGCCCTCATCTAAGACATCTTATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTG
 60 ATGAAAGTACTCAAGT

SEQ ID NO: 28 secuencia de aminoácidos que comprende HIS-40a NH

5 M G S H H H H H H A S S V G V S H Q V K A D D R A S G E T K A S N T H D
D S L P K P E T I Q E A K A T I D A V E K T L S Q Q K A E L T E L A T A
L T K T T A E I N H L K E Q Q D N E Q K A L T S A Q E I Y T N T L A S S
10 E E T L L A Q G A E H Q R E L T A T E T E L H N A Q A D Q H S K E T A L
S E Q K A S I S A E T T R A Q D L V E Q V K T S E Q N I A K L N A M I S
N P D A I T K A A Q T A N D N T K A L S S E L E K A K A D L E N Q K A K
V K K Q L T E E L A A Q K A A L A E K E A E L S R L K S S A P S T Q D S
I V G N N T M K A P Q G Y P L E E L K K L E A S G Y I G S A S Y N N Y Y
K E H A D Q I I A K A S P G N Q L N Q Y Q D I P A D R N R F V D P D N L
T P E V Q N E L A Q F A A H M I N S V R R Q L G L P P V T V T A G S Q E
15 F A R L L S T S Y K K T H G N T R P S F V Y G Q P G V S G H Y G V G P H
D K T I I E D S A G A S G L I R N D D N M Y E N I G A F N D V H T V N G
I K R G I Y D S I K Y M L F T D H L H G N T Y G H A I N F L R V D K H N
P N A P V Y L G F S T S N V G S L N E H F V M F P E S N I A N H Q R F N
K T P I K A V G S T K D Y A Q R V G T V S D T I A A I K G K V S S L E N
R L S A I H Q E A D I M A A Q A K V S Q L Q G K L A S T L K Q S D S L N
20 L Q V R Q L N D T K G S L R T E L L A A K A K Q A Q L E A T R D Q S L A
K L A S L K A A L H Q T E A L A E Q A A A R V T A L V A K K A H L Q Y L
R D F K L N P N R L Q V I R E R I D N T K Q D L A K T T S S L L N A Q E
A L A A L Q A K Q S S L E A T I A T T E H Q L T L L K T L A N E K E Y R
H L D E D I A T V P D L Q V A P P L T G V K P L S Y S K I D T T P L V Q
25 E M V K E T K Q L L E A S A R L A A E N T S L V A E A L V G Q T S E M V
A S N A I V S K I T S S I T Q P S S K T S Y G S G S S T T S N L I S D V
D E S T Q R

SEQ ID NO: 29 secuencia de polinucleótidos que comprende HIS-40a CH

30
35 ATGGCTAGTAGTGTAGGCGTATCTCACCAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAG
TAATACTCAGCAGATAGTTTACCAAAACCAGAAACAATTCAGAGGCAAAGGCAACTATTGATGCAGTTG
AAAAAACTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGTCTACCGCTCTGACGAAAACTACTGCTGAA
ATCAACCATTTAAAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAATTTACACTAA
TACTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCCAAGGAGCCGAACATCAAAGAGAGTTAACAGCTACTG
AAACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCATTGTCAGAACAAAAAGCTAGC
40 ATTTACAGCAGAACTACTCGAGCTCAAGATTTAGTGGAAACAAGTCAAAACGCTCTGAACAAAATATTGTAA
GCTCAATGCTATGATTAGCAATCCTGATGCTATCATTAAAGCAGCTCAAAACGGCTAATGATAATACAAAAG
CATTAAGCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAAATCAAAAAGCTAAAGTTAAAAAGCAATTG
ACTGAAGAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCAGAACTTAGTCGTCTTAAATCCTC
AGCTCCGTCTACTCAAGATAGCATTGTGGGTAATAATACCATGAAAGCACCGCAAGGCTATCCTCTTGAAG
AACTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAAGAGCATGCA
45 GATCAAATTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATATTCCAGCAGATCGTAA
TCGCTTTGTTGATCCCATAATTGACACCAGAAGTGCAAAATGAGCTAGCGCAGTTTGACAGCTCACATGA
TTAATAGTGTAAGAAGACAATTAGGTCTACCACCAGTTACTGTTACAGCAGGATCACAAGAATTTGCAAGA
TTACTTAGTACCAGCTATAAGAAAATCATGGTAAATACAAGACCATCATCTGTCTACGGACAGCCAGGGGT
ATCAGGGCATTATGGTGTGGGCCCTCATGTAAAACTATTTATGAAGACTCTGCCGGAGCGTCAGGGCTCA
50 TTCGAAATGATGATAACATGTACGAGAATATCGGTGCTTTTAAACGATGTGCATACTGTGAATGGTATTAAA
CGTGGTATTTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAAATACATACGGCCATGC
TATTAACCTTTTACGTGTAGATAAACATAACCCTAATGCGCCTGTTTACCTTGGATTTTCAACCAGCAATG
TAGGATCTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATTGCTAACCATCAACGCTTTAATAAG
55 ACCCTATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAAGAGTAGGCACTGTATCTGATACTATTGC
AGCGATCAAGGAAAAGTAAGTCAATTAGAAAATCGTTTGTGCGCTATTTCATCAAGAAGCTGATATTATGG
CAGCCCAAGCTAAAGTAAGTCAACTTCAAGGTAAATTAGCAAGCACACTTAAGCAGTCAGACAGCTTAAT
CTCCAAGTGAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTAAAGCAAAACA
AGCACAACTCGAAGCTACTCGTGATCAATCATTAGCTAGCATCGTTGAAAGCCGCATGACCCAGCA
60 CAGAAGCCTTAGCAGAGCAAGCCGCAGCCAGAGTGACAGCACTGGTGGCTAAAAAGCTCATTTGCAATAT
CTAAGGGACTTTAAATTGAATCCTAACCCGCTTCAAGTGATACGTGAGCGCATTGATAATACCTAAGCAAGA
TTTGGCTAAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCTAAACAAAGCA
GTCTAGAAGCTACTATTGCTACCACAGAACACCAGTTGACTTTGCTTAAACCTTAGCTAACGAAAAGGAA
65

TATCGCCACTTAGACGAAGATATAGCTACTGTGCCTGATTTGCAAGTAGCTCCACCTCTTACGGGCGTAAA
 ACCGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAAGAAATGGTTAAAGAAACGAAACAACCTAT
 TAGAAGCTTCAGCAAGATTAGCTGCTGAAAATACAAGTCTTGTAGCAGAAGCGCTTGTGGCCAAACCTCT
 GAAATGGTAGCAAGTAATGCCATTGTGTCTAAAATCACATCTTCGATTACTCAGCCCTCATCTAAGACATC
 TTATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAACGtCGCGCCG
 CACTCGAGCACCACCACCACCACCAC

SEQ ID NO: 30 secuencia de aminoácidos que comprende HIS-40a CH

M A S S V G V S H Q V K A D D R A S G E T K A S N T H D D S L P K P E T
 I Q E A K A T I D A V E K T L S Q Q K A E L T E L A T A L T K T T A E I
 N H L K E Q Q D N E Q K A L T S A Q E I Y T N T L A S S E E T L L A Q G
 A E H Q R E L T A T E T E L H N A Q A D Q H S K E T A L S E Q K A S I S
 A E T T R A Q D L V E Q V K T S E Q N I A K L N A M I S N P D A I T K A
 A Q T A N D N T K A L S S E L E K A K A D L E N Q K A K V K K Q L T B E
 L A A Q K A A L A E K E A E L S R L K S S A P S T Q D S I V G N N T M K
 A P Q G Y P L E E L K K L E A S G Y I G S A S Y N N Y Y K E H A D Q I I
 A K A S P G N Q L N Q Y Q D I P A D R N R F V D P D N L T P E V Q N E L
 A Q F A A H M I N S V R R Q L G L P V T V T A G S Q E F A R L L S T S
 Y K K T H G N T R P S S V Y G Q P G V S G H Y G V G P H D K T I I E D S
 A G A S G L I R N D D N M Y E N I G A F N D V H T V N G I K R G I Y D S
 I K Y M L F T D H L H G N T Y G H A I N F L R V D K H N P N A P V Y L G
 P S T S N V G S L N E H F V M F P E S N I A N H Q R F N K T P I K A V G
 S T K D Y A Q R V G T V S D T I A A I K G K V S S L E N R L S A I H Q E
 A D I M A A Q A K V S Q L Q G K L A S T L K Q S D S L N L Q V R Q L N D
 T K G S L R T E L L A A K A K Q A Q L E A T R D Q S L A K L A S L K A A
 L H Q T E A L A E Q A A A R V T A L V A K K A H L Q Y L R D F K L N P N
 R L Q V I R E R I D N T K Q D L A K T T S S L L N A Q E A L A A L Q A K
 Q S S L E A T I A T T E H Q L T L L K T L A N E K E Y R H L D E D I A T
 V P D L Q V A P P L T G V K P L S Y S K I D T T P L V Q E M V K E T K Q
 L L E A S A R L A A E N T S L V A E A L V G Q T S E M V A S N A I V S K
 I T S S I T Q P S S K T S Y G S G S S T T S N L I S D V D E S T Q R A A
 A L E H H H H H H

SEQ ID NO: 31 secuencia de polinucleótidos que comprende HIS-40a-RR NH

ATGGGATCGCATCACCATCACCATCAGCTAGTAGTGTAGGCGTATCTCACCAGTCAAAGCAGATGATAG
 AGCCTCAGGAGAAACGAAGGCGAGTAATACTCAGCAGATAGTTTACCAAAACAGAAACAAATCAAGAGG
 CAAAGGCAACTATTGATGTCAGTTGAAAAAATCTCAGTCAACAAAAGCAGAAGTACAGAGCTTGCTACC
 GCTCTGACAAAACCTACTGCTGAAATCAACCACCTTAAAAGAGCAGCAAGATAATGAACAAAAGCTTTAAC
 CTCGACACAAGAAATTTACACTAATACTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAAC
 ATCAAAGAGAGTTAACAGCTACTGAAACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACT
 GCATTGTCAGAACAAAAGCTAGCATTTAGCAGAACTACTCGAGCTCAAGATTTAGTGGAAACAAGTCAA
 AACGCTCTGAACAAAATATTGCTAAGCTCAATGCTATGATTAGCAATCTGATGCTATCACTAAAGCAGCTC
 AAACGGCTAATGATAATACAAAAGCATTAAAGCTCAGAATTGGAGAGGCTAAAGCTGACTTAGAAAAATCAA
 AAAGCTAAAGTTAAAAAGCAATTGACTGAAGAGTTGGCAGCTCAGAAAAGCTGCTCTAGCAGAAAAAGAGGC
 AGAAGCTTAGTCGCTTTAAATCTCAGCTCCGCTACTCAAGATAGCATTTGTTGGGTAATAATACCATGAAAG
 CACCGCAAGGCTATCCTCTTGAAGAACTTAAAAAATAGAAAGCTAGTGGTTATATTGGATCAGCTAGTTAC
 AATAATTATTACAAAGAGCATGCAGATCAAATTTATGCCAAAGCTAGTCCAGGTAATCAATTAATCAATA
 CCAAGATATTCAGCAGATCGTAATCGCTTTGTTGATCCCGATAATTTGACACCAGAAGTGCAAAATGAGC
 TAGCGCAGTTTGCAGCTCACATGATTAATAGTGTAcGtGtCAATTAGGTCTACCACCAGTTACTGTTACA
 GCAGGATCACAAGAAATTTGCAAGATTACTTAGTACCAGCTATAAGAAAACTCATGGTAATACAAGACCATC
 ATTTGTCTACGGACAGCCAGGGGTATCAGGGCATTATGGTGTGGGCCCTCATGATAAACTATTATTGAAG
 ACTCTGCGGAGCGTCAGGGCTCATTTCGAAATGATGATAACATGTACGAGAATATCGGTGCTTTTAAAGAT
 GTGCATACTGTGAATGGTATTAAACGTGGTATTATGACAGTATCAAGTATATGCTCTTTACAGATCATTT
 ACACGGAAATACATACGGCCATGCTATTAACTTTTACGTGTAGATAAACATAACCCTAATGCGCCTGTTT
 ACCTTGGATTTTCAACCAGCAATGTAGGATCTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATT
 GCTAACCATCAACGCTTTAATAAGACCCCTATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAAGAGT

AGGCCTGTATCTGATACTATTGTCAGCGATCAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTGCGGCTA
 TTCATCAAGAAGCTGATATTTATGGCAGCCCAAGCTAAAGTAAGTCAACTTCAAGGTAATTAGCAAGCACA
 CTTAAGCAGTCAGACAGCTTAAATCTCCAAGTGAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGA
 ATTACTAGCAGCTAAAGCAAAACAAGCACAACCTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCAT
 CGTTGAAAGCCGCACTGCACCAGACAGAAGCCTTAGCAGAGCAAGCCGAGCCAGAGTGACAGCACTGGTG
 GCTAAAAAAGCTCATTGCAATATCTAAGGGACTTTAAATTTGAATCCTAACCGCCTTCAAGTGATACGTGA
 GCGCATTGATAATACTAAGCAAGATTGGCTAAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAG
 CAGCCTTACAAGCTAAACAAAGCAGTCTAGAAGCTACTATTGCTACCACAGAACACCAGTTGACTTTGCTT
 AAAACCTTAGCTAACGAAAAGGAATATCGCCACTTAGACGAAGATATAGCTACTGTGCCTGATTTGCAAGT
 AGCTCCACCTCTTACGGGCGTAAACCGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAA
 TGGTTAAAGAAACGAAACAACCTATTAGAAGCTTCAGCAAGATTAGCTGCTGAAAATACAAGTCTTGTAGCA
 GAAGCGCTTGTGGCCAAACCTCTGAAATGGTAGCAAGTAATGCCATTGTGTCTAAAAATCACATCTTCGAT
 TACTCAGCCCTCATCTAAGACATCTTATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTG
 ATGAAAGTACTCAACGt

SEQ ID NO: 32 secuencia de aminoácidos que comprende HIS-40a-RR NH

M G S H H H H H H A S S V G V S H Q V K A D D R A S G E T K A S N T H D
 D S L P K P E T I Q E A K A T I D A V E K T L S Q Q K A E L T E L A T A
 L T K T T A E I N H L K E Q Q D N E Q K A L T S A Q E I Y T N T L A S S
 E E T L L A Q G A E H Q R E L T A T E T E L H N A Q A D Q H S K E T A L
 S E Q K A S I S A E T T R A Q D L V E Q V K T S E Q N I A K L N A M I S
 N P D A I T K A A Q T A N D N T K A L S S E L E K A K A D L E N Q K A K
 V K K Q L T E E L A A Q K A A L A E K E A E L S R L K S S A P S T Q D S
 I V G N N T M K A P Q G Y P L E E L K K L E A S G Y I G S A S Y N N Y Y
 K E H A D Q I I A K A S P G N Q L N Q Y Q D I P A D R N R F V D P D N L
 T P E V Q N E L A Q F A A H M I N S V R R Q L G L P V T V T A G S Q E
 F A R L L S T S Y K K T H G N T R P S F V Y G Q P G V S G H Y G V G P H
 D K T I I E D S A G A S G L I R N D D N M Y E N I G A F N D V H T V N G
 I K R G I Y D S I K Y M L F T D H L H G N T Y G H A I N F L R V D K H N
 P N A P V Y L G F S T S N V G S L N E H F V M F P E S N I A N H Q R F N
 K T P I K A V G S T K D Y A Q R V G T V S D T I A A I K G K V S S L E N
 R L S A I H Q E A D I M A A Q A K V S Q L Q G K L A S T L K Q S D S L N
 L Q V R Q L N D T K G S L R T E L L A A K A K Q A Q L E A T R D Q S L A
 K L A S L K A A L H Q T E A L A E Q A A A R V T A L V A K K A H L Q Y L
 R D F K L N P N R L Q V I R E R I D N T K Q D L A K T T S S L L N A Q E
 A L A A L Q A K Q S S L E A T I A T T E H Q L T L L K T L A N E K E Y R
 H L D E D I A T V P D L Q V A P P L T G V K P L S Y S K I D T T P L V Q
 E M V K E T K Q L L E A S A R L A A E N T S L V A E A L V G Q T S E M V
 A S N A I V S K I T S S I T Q P S S K T S Y G S G S S T T S N L I S D V
 D E S T Q R

SEQ ID NO: 33 secuencia de polinucleótidos que comprende 40N-HIS

ATGCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAATACTCACGACGATAGTTTACC
 AAAACCAGAAACAATTCAAGAGGCAAGGCAACTATTGATGCAGTTGAAAAAATCTCAGTCAACAAAAAG
 CAGAACTGACAGAGCTTGCTACCGCTCTGACAAAAAATCTGCTGAAATCAACCACTTAAAAGAGCAGCAA
 GATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAATTTACACTAATACTCTTGCAAGTAGTGAGGAGAG
 GCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTAACAGCTACTGAAACAGAGCTTCATAATGCTCAAG
 CAGATCAACATTCAAAGAGAGCTGCATTGTCAGAACAAAAAGCTAGCATTTTCAGCAGAACTACTCGAGCT
 CAAGATTTAGTGGAACAAGTCAAACGCTCTGAACAAAATATTGCTAAGCTCAATGCTATGATTAGCAATCC
 TGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAGCATTAAGCTCAGAAATTGGAGAAGG
 CTAAAGCTGACTTAGAAAAATCAAAGGCTAAAGTTAAAAAGCAATTGACTGAAGAGTTGGCAGCTCAGAAA
 GCTGCTCTAGCAGAAAAAGAGGCAGAACTTAGTCGTCTTAAATCCTCAGCTCCGCTACTCAAGATAGCAT
 TGTGGGTAATAATACCATGAAAGCACCAGGCTATCCTCTTGAAGAACTTAAAAAATTAGAAGCTAGTG
 GTTATATTGGATCAGCTAGTTACAATAATTATTACAAAGAGCATGCAGATCAAATATTGCCAAAGCTAGT
 CCAGGTAATCAATTAAATCAATACCAAGCGGCCGCACTCGAGCACCACCACCACCACCACCAC

SEQ ID NO: 34 secuencia de aminoácidos que comprende 40N-HIS

5 M Q V K A D D R A S G E T K A S N T H D D S L P K P E T I Q E A K A T I
D A V E K T L S Q Q K A E L T E L A T A L T K T T A E I N H L K E Q Q D
N E Q K A L T S A Q E I Y T N T L A S S E E T L L A Q G A E H Q R E L T
A T E T E L H N A Q A D Q H S K E T A L S E Q K A S I S A E T T R A Q D
10 L V E Q V K T S E Q N I A K L N A M I S N P D A I T K A A Q T A N D N T
K A L S S E L E K A K A D L E N Q K A K V K K Q L T E E L A A Q K A A L
A E K E A E L S R L K S S A P S T Q D S I V G N N T M K A P Q G Y P L E
E L K K L E A S G Y I G S A S Y N N Y Y K E H A D Q I I A K A S P G N Q
L N Q Y Q A A A L E H H H H H H H

15 **SEQ ID NO: 35** secuencia de aminoácidos que comprende GAS 117

20 MTLKKHYLLSLLALVTVGAAFNTSQSVSQAQVYSNEGYHQHLTDEKSHLQYSKDNAQLQLRNILDGYQNDL
GRHYSSYYYYNLRTVMGLSSEQDIEKHYEELKNKLHDMYNHY

SEQ ID NO: 36 secuencia de polinucleótidos que codifica GAS 117

25 ATGACACTAAAAAACACTATTATCTTCTCAGCCTGCTAGCTCTTGTAACGGTGGTGCTGCCTTTAACAC
AAGCCAGAGTGTCAGTGACACAAGTTTATAGCAATGAAGGGTATCACCAGCATTGACTGATGAAAAATCAC
ACCTGCAATATAGTAAAGACAACGCACAACCTCAATTGAGAAATATCCTTGACGGCTACCAAAATGACCTA
GGGAGACACTACTCTAGCTATTATTACTACAACCTAAGAACC GTTATGGGACTATCAAGTGAGCAAGACAT
30 TGAAAAACACTATGAAGAGCTTAAGAACAAGTTACATGATATGTACAATCATTATTAA

SEQ ID NO: 37 secuencia líder de GAS 117 secuencia de aminoácidos que comprende la TLKKHYLLSLLALVTGA

35 **SEQ ID NO: 38** secuencia de aminoácidos que comprende el fragmento de GAS 117 en el que la secuencia líder está eliminada

40 AFNTSQSVSQAQVYSNEGYHQHLTDEKSHLQYSKDNAQLQLRNILDGYQNDLGRHYSSYYYYNLRTVMGLSS
EQDIEKHYEELKNKLHDMYNHY

SEQ ID NO: 39 secuencia de aminoácidos que comprende GAS 130

45 MSHMKKRPEVLSFAGTLEKLKVAIDYGADAVFVGQAYGLRSRAGNFSMEELQEGIDYAHARGAKVYVAAN
MVTHEGNEIGAGEWFRQLRDMGLDAVIVSDPALIVICSTEAPGLEIHLSTQASSTNYETPEFWKAMGLTRV
VLAREVNMAELAEIRKRTDVEIEAFVHGAMCISYSGRCVLSNHMSHRDANRGCSQSCRWKYDLYDMPFGG
ERRSLKGEIPEDYSMSVDMCMIDHIPDLIENGVDLSKIEGRMKS IHVSTVTNICYKAAVGAYMESPEAFY
50 AIKEELIDELWKVAQRELATGFYYGIPTENEQLFGARRKIPQYKFVGEVAFDSASMTATIRQRNVIMEGD
RIECYGPGRFHFETVVKDLHDADGQKIDRAPNPMELLTISLPREVKPGDMI RACKGLVNLYQKDGTSKTV
RT

55

60

65

SEQ ID NO: 40 secuencia de polinucleótidos que codifica GAS 130

ATGTCACATATGAAAAACGTCCCGAGGTCCTATCACCTGCTGGAACACTTGAAAAATTTAAAGTTGCGAT
 TGACTATGGCGCAGATGCTGTTTTTGTGGAGGGCAGGCCATGGCCTAAGAAGCCGCGCTGGTAACTTCT
 CTATGGAAGAAATGCAAGAAGGCATTGATTATGCACATGCGCGTGGAGCTAAGGTCATGTTGCTGCTAAC
 ATGGTTACCCACGAAGGGAACGAAATGGGTGCGGGCGAGTGGTTTCGTCAACTGCGTGATATGGGGCTTGA
 TGGCGTCATTGTTTCAGATCCAGCCTTGATTGTTATTTGTTCAACAGAAGCCCCAGGTTTGGAAATTCATT
 TGTC AACGCAAGCTTCACTACCAATTACGAGACCTTTGAATTTTGGAAAGCCATGGGCTTGACCCGAGTT
 GTTTTAGCTCGCGAGGTTAATATGGCCGAGTTAGCAGAAATCCGCAAGCGGACAGATGTGGAAATGAAGC
 CTTTGTCCATGGAGCCATGTGTATCTCTTATTAGGCCGCTGTGTTTGTCAAACCATGAGTCACCGTG
 ATGCCAACAGGGGCGGCTGCTCACAGTCTTGCCGCTGGAAGTATGATTTGTATGACATGCCATTGGAGGA
 GAGCGCGCTCCTTAAAGGGGAAATCCAGAAGACTATTCTATGTCCTCTGTTGACATGTGTATGATTGA
 CCATATTCTGACCTGATTGAAAATGGGGTTGATAGCTTAAAAAATTGAAGGCCGAATGAAATCTATCCACT
 ACGTCTCAACCGTAACCAACTGTTACAAGGCGGCTGTAGGTGCTTACATGGAAAGCCAGAAGCTTTTAT
 GCTATCAAAGAGGAATTGATTGACGAGTTGTGGAAGGTTGCCAGCGCGAGTTGGCTACAGGTTTACTA
 TGGTATCCCAACTGAAAATGAACAATTTATTTGGTGCTCGCCGCAAAATTCACAATATAAATTTGTCGGAG
 AAGTAGTTGCCTTTGACTCAGCTAGCATGACAGCGACCATTCGTGAGCGTAATGTCATCATGGAAGGCGAT

 CGGATTGAATGTTATGGACCAGGTTTCCGTCATTTTGAACGTTGTTAAGGACTTACATGATGCGGATGG
 CCAAAAGATTGACCGTGCCCCAAATCCAATGGAACCTTAACCATCTCTTTACCGAGAGAAGTTAAGCCAG
 GGGATATGATTAGGCTTGCAAGGAAGGCTGGTTAACTCTATCAAAAAGATGGCACCAGTAAACTGTT
 AGAACATAG

SEQ ID NO: 41 secuencia de aminoácidos que comprende GAS 277

MTTMQKTIISLLSLALLIGLLGTSGKAISVYAQDQHTDNVIAESTISQVSVEASMRGTEPYIDATVTTDQPV
 RQPTQATITLKDASDNTINSWVYTMAAQORRFTAFWDLTGQKSGDYHVTVTVHTQEKAVTGQSGTVHFDQN
 KARKTPTNMQQKDTSKAMTNSVDVDTKAQTNQSANQEIDSTSNPFRSATNHRSTSLKRSTKNEKLPTASN
 SQKNGSNKTKMLVDKEEVKPTSKRGFPWVLLGLVVS LAAGLFIAIQVSRK

SEQ ID NO: 42 secuencia de polinucleótidos que codifica GAS 277

ATGACAACATATGCAAAAAACAATTAGCTTATTATCAGTACTTTACTTATTGGTTTGTGCGGACTTCTGG
 CAAAGCCATATCTGTGTATGCACAAGATCAGCACACTGATAATGTTATAGCTGAATCAACTATTAGTCAGG
 TCAGTGTTGAAGCCAGTATGCGTGGAAACAGAACCCTTATATTGATGCTACAGTCACCACAGATCAACCTGTC
 AGACAACCAACTCAGGCAACGATAACACTTAAAGACGCTAGTGATAATACTATTAATAGTTGGGTATATAC
 TATGGCAGCGCAACAGCGTCGTTTTACAGCTTGGTTTGATTAACTGGACAAAAGAGTGGTGACTATCATG
 TAACTGTCAACCGTTCATACTCAAGAAAAGGCAGTAACCTGGTCAATCAGGAACGTTCATTTTGATCAAAAC
 AAAGCTAGAAAAACACCAACTAATATGCAACAAAAGGATACTTCTAAAGCAATGACGAATTCAGTCGATGT
 AGACAACAAAAGCTCAAAACAAATCAATCAGCTAACCAAGAAATAGATTCTACTTCAAATCCTTTTCAGATCAG
 CTACTAATCATCGATCAACTTCCTTAAAGCGATCTACTAAAAATGAGAACTTACACCAACTGCTAGTAAT
 AGCCAAAAAACGGTAGCAACAAGACAAAAATGCTAGTGGACAAAGAGGAAGTAAACCTACTTCAAAAAAG
 AGGATTCCCTTGGGTCTTATTAGGTCTAGTAGTCAGTTAGCTGCAGGTTTATTTATAGCTATTCAAAAAAG
 TATCTAGACGAAAATAA

SEQ ID NO: 43 secuencia de aminoácidos que comprende la secuencia líder N-terminal de GAS 277
 TTMQKTIISLLSLALLIGLLGTSGKAISVYA

SEQ ID NO: 44 secuencia de aminoácidos que comprende el fragmento de GAS 277 en el que la secuencia líder
 N-terminal está eliminada

QDQHTDNVIAESTISQVSVEASMRGTEPYIDATVTTDQPVVRQPTQATITLKDASDNTINSWVYTMAAQORR
 FTAFWDLTGQKSGDYHVTVTVHTQEKAVTGQSGTVHFDQNKARKTPTNMQQKDTSKAMTNSVDVDTKAQTN
 QSANQEIDSTSNPFRSATNHRSTSLKRSTKNEKLPTASNSQKNGSNKTKMLVDKEEVKPTSKRGFPWVLL
 GLVVS LAAGLFIAIQVSRK

SEQ ID NO: 45 secuencia de aminoácidos que comprende GAS 236

MTQMNYTGKVKRVAIIANGKYQSKRVASKLFSVPKDDPDFYLSKKNPDIVISIGGDGMLLSAFHMYEKELD
 KVRFGIHTGHLGFYTDYRDFEVDKLDNLRKDKGEQISYPILKVAITLDDGRVVKARALNEATVKRIEKT
 MVADVIINHVFESFRGDGIVSTPTGSTAYNKS LGGAVLHPTIEALQLTEISSLNRRVFRITLGSIIIPK
 KDKIELVPRKLGIIYISIDNKTYQLKNVTKEVYFIDDEKIHVSSPSHTSFWERVKDAFIGEIDS

SEQ ID NO: 46 secuencia de polinucleótidos que codifica GAS 236

5
10
15
ATGACACAGATGAATTATACAGGTAAGGTAAAACGAGTTGCTATTATTGCAAATGGTAAGTACCAAAGTAA
ACGCGTCGCCCTCCAAACTTTCTCCGTATTTAAAGATGATCCTGATTCTATCTTTCAAAGAAAAATCCGG
ATATTGTGATTTCTATTGGCGGAGATGGGATGCTCTTATCTGCCTTTCACATGTATGAAAAAGAATTAGAT
AAGGTACGTTTGTAGGAATCCACACCGGTATCTTGGCTTTTATACCGATTATAGGGATTTTGAAGTTGA
TAAATTAATTGATAATTTAAGAAAAAGACAAGGGAGAACAAATCTCTTATCCGATTTTAAAAGTTGCTATTA
CTTTAGATGATGGTCGTGTGGTTAAAGCGCGTGCTTTGAATGAAGCGACGGTTAAGCGTATTGAAAAACG
ATGGTAGCAGATGTTATTATTAAACCATGTCAAATTTGAAAGCTTCCGAGGTGATGGGATTTTCAATATCGAC
CCCGACAGGGAGCACAGCCTACAATAAATCTTTAGGTGGTGCTGTCTTGATCCGACGATTGAAGCGCTGC
AATTGACGGAAATTTCCAGTCTTAATAACCGTGCTTTAGAACCTTGGGCTCATCAATCATTATTCCCAA
AAAGATAAGATTGAGTTAGTGCCAAAACGATTAGGAATTTATACCATTTCATTGATAATAAAACCTATCA
GTTAAAAAATGTGACGAAGGTGGAGTATTTATCGACGATGAGAAAATTCATTTTGTTCCTCTCCGAGTC
ATACGAGCTTTTGGGAAAGGGTCAAGGATGCCTTTATTGGAGAGATTGACTCATGA

SEQ ID NO: 47 secuencia de aminoácidos que comprende la secuencia líder N-terminal de GAS 236 MTQM

20 **SEQ ID NO: 48** secuencia de aminoácidos que comprende un fragmento de GAS 236 en el que la secuencia líder N-terminal está eliminada

25
NYTGKVKRVAIIANGKYQSKRVASKLFSVFKDDPDFYLSKKNPDIVISIGDGMLLSAFHMVEKELDKVRF
VGINTGHLGFTDYRDFEVDKLDNLRKDKGEQISYPILKVAITLDDGRVVKARALNEATVKRIEKTMTVAD
VIINHVKFESFRGDGISVSTPTGSTAYNKS LGGAVLHPTIEALQLTEISSLNRRVPRTLGSIIIPKKDKI
ELVPKRLGIYTTISIDNKTYQLKNVTKVEYFIDDEKIHVSSPSHTSFWERVKDAFIGEIDS

SEQ ID NO: 49 secuencia de aminoácidos que comprende GAS 389

30
35
40
45
50
55
60
65
MRNEMAKIMNVTGEEVIALAATYMTKADVAFVAKALAYATAAHFYQVRKS GEPYIVHPIQVAGILADLHLD
AVTVACGFLHDVVEDTDITLDEIADFGHDARDIVDGVTKLGEVEYKSHERQLAENHRKMLMAMSKDIRVI
LVKLADRLHNMRTLKHLRKDKQERISRETMEIYAPLAHRLGISRIKWELEDLAFRYLNETEFYKISHMMKE
KRREREALVEAIVSKVKTYTTQQLFGDVYGRPKHIYSIYRKMRDKKRFDQIFDLIAIRCVMETQSDVYA
MVGYIHELWRPMPGRFKDYIAAPKANGYQSIHTTVYGPKEPIEIQIRTKDMHQVAEYGVAAHWAYKGVGR
KVNQAEQAVGMNWIKELVELQDASNGDAVDFVDSVKEDIFSERIYVFTPTGAVQELPKESGPIDPAYAIHT
QIGEKATGAKVNGRMVPLTAKLKTGDVVEIITNANSFGPSRDWVKLVKTNKARNKIRQFFKNQDKELSVNK
GRDLLVSYFQEQGYVANKYLDKKRIEAILPKVSVKSEESLYAAVGFDISPISVFNKLTEKERREERAKA
KAEAEELVKGGEVKHENKDVVKVRSNGVVIQAGSGLLMRIAKCCNPVPGDPIDGYITKGRGIAIHRSDCH
NIKSQDGYQERLIEVEWDLNSSKDYQAEIDIYGLNRSGLLNDVLQILSNSTKSISTVNAQPTKDMKFANI
HVSFGIPNLTHLTTVVEKIKAVPDVYSVKRTNG

SEQ ID NO: 50 secuencia de polinucleótidos que codifica GAS 389

5 ATGAGGAACGAAATGGCAAAATAATGAACGTAACAGGAGAAGAAGTCATTGCCTTAGCGGCCACCTATAT
GACCAAGGCTGATGTGGCTTTTGTGGCAAAGGCCTTAGCATATGCAACAGCGGCCCATTTCTACCAAGTGA
GAAAGTCAGGCGAACCCCTATATCGTCCATCCGATTTCAGGTGGCGGGGATTCTGGCTGATTTGCATCTGGAT
10 GCTGTGACAGTTGCTTGTGGCTTTTACATGATGTCTGTAAGAATACGGATATTACCTTAGATGAGATCGA
AGCAGACTTTGGCCATGATGCTCGTGATATCGTTGATGGTGTCAACCAAGTTAGGTGAAGTTGAGTACAAAT
CTCATGAGGAGCAACTCGCCGAAAACCATCGCAAAATGCTGATGGCTATGTCCAAAGATATTTCGCGTGATT
TTGGTGAAATTTGGCTGACCGCTGCATAATATGCGCACCCCTCAAACATTTGCGCAAGGACAAACAAGAGCG
15 CATTTCGCGCGAAACCATGGAAATCTATGCCCCCTTGGCGCATCGTTTGGGGATTAGTCGCATCAAATGGG
AACTAGAAGATTTGGCTTTTCGTTACCTCAATGAAACCGAATTTTACAAAATTTCCCATATGATGAAAGAA
AAACGTCGCGAGCGTGAAAGCTTTGGTAGAGGCTATTGTTCAGTAAGGTCAAACCTATACGACACAACAAGG
GTTGTTTGGAGATGTGTATGGCCGACCAAAACACATTTATTTCGATTTATCGGAAAATGCGGGACAAAAAGA
AAGATTCGATCAGATTTTGTGATGTCATTCGCTGCTGTCATGGAACGCAAGCGTGTCTATGCTGCT
20 ATGGTTGGCTATATTCATGAGCTTTGGCGTCCCATGCCAGGCCGCTTCAAGGATTATATTGCAGCTCCTAA
AGCTAATGGCTACCACTCTATTATACCAACCGTGTATGGGCCAAAAGGACCTATTGAGATTCAAATCAGAA
CTAAGGACATGCATCAAGTGGCTGAGTACGGGGTGTCTGCTCACTGGGCTTATAAAAAAGGCGTGCCTGGT
AAGGTCAATCAAGCTGAGCAAGCCGTTGGCATGAACTGGATCAAAGAGCTGGTAGAATTGCAAGATGCCTC
AAATGGCGATGCAGTGGACTTTGTGGATTCCGTCAAAAGAAGACATTTTTTCTGAACGGATTTATGTCTTTA
25 CACCGACAGGGGCCGTTTCAAGGATTACCAAAAGAAATCAGGTCTTATTGATTTTGTCTATGCGATCCATACG
CAATTCGGTGAAAAAGCAACAGGTGCCAAAGTCAATGGACGTATGGTTCTCTCACTGCCAAGTTAAAAAC
AGGAGATGTGGTTGAAATCATCACCAATGCCAATTCCTTTGGCCCTAGTCGAGACTGGGTAAACTGGTCA
AAACCAATAAGGCTCGCAACAAAATTCGTGAGTTCTTTAAAAATCAAGACAAGGAATTGTCAAGTGAATAAA
GGCCGTGATTTGTTGGTGTCTTATTTTCAAGAGCAGGGCTACGTTGCCAATAAATACCTTGACAAAAAACG
30 CATTGAAGCCATCCTTCCAAAAGTCAGTGTGAAGAGCGAAGAATCACTCTATGCAGCCGTTGGGTTTGGTG
ACATTAGTCTTATCAGTGTCTTTAACAAGTTAACCGAAAAAGAGCGCCGTGAAGAAGAAAGGCCAAGGCT
AAAGCAGAAGCTGAAGAATTGGTTAAGGGCGGTGAGGTCAAACACGAAAAACAAAGATGTCTCAAGGTTTCG
CAGTGAAGATGGAGTCATTATCCAAGGAGCATCAGGCCTCTTGATGCGGATTGCCAAGTGTGTAATCCTG
TACCTGGTGTCTTATGACGGCTACATTACCAAGGGCGTGGCATTGCGATTACAGATCGGACTGTCTAT
35 AACATTAAGAGTCAAGATGGCTACCAAGAACGCTTGATTGAGGTGAGTGGGATTTGGACAATTTCGAGTAA
AGATTATCAGGCTGAAATGATATCTATGGGCTCAATCGTAGTGGTCTGCTTAATGATGTCTCCAAATTT
TATCAAATCAACCAAGAGCATATCGACAGTCAATGCTCAGCCGACCAAGGACATGAAGTTTGCTAATATT
CACGTGAGCTTTGGCATTCCAAATCTGACGCATCTGACCACTGTTGTGCAAAAAATCAAGGCAGTTCAGAA
TGTTTATAGCGTGAAGCGGACCAATGGCTAA

SEQ ID NO: 51 secuencia de aminoácidos que comprende GAS 504

40 MKTRITELLNIDYPIFQGGMAWVADGDLGAVSNAGGLGIIGGGNAPKEVVKANIDRVKAITDRPFGVNIM
LLSPFADDIVDLVIEEGVKVVTGAGNPGKYMERLHQAGIIVVPVPSVALAKRMEKLGVDVIAEGMEAG
GHIGKLTMSLVRQVVEAVSIPVIAAGGIADGHGAAAFMLGAEAVQIGTRFVVAKESNAHQNFKDKILAA
KDIDTVISAQVVGHVPRSIKNKLTSAKAKAFLIGQKTATDIEEMGAGSLRHAVIEGDDVNGSVMAGQI
45 AGLVRKEESCETILKDIYYGAARVIQNEAKRWQSVSIEK

SEQ ID NO: 52 secuencia de polinucleótidos que codifica GAS 504

50 ATGAAAAACAGTATTACAGAATTACTTAATATTGATTACCCCATTTTTCAAGGAGGAATGGCTTGGGTTGC
TGATGGTGATTTAGCAGGTGCAGTTTCTAATGCTGGTGGTTTAGGCATTATAGGTGGTGGCAATGCTCCCA
AAGAAGTCGTTAAAGCTAATATTGATCGTGTCAAAGCTATTACTGATAGACCTTTTGGGGTTAATATCATG
CTTTTATCTCCTTTTGTGCTGATGATATCGTTGATCTGGTCATTGAAGAAGGTGTTAAAGTAGTAACAACAGG
55 CGCAGGAAATCCAGGAAAGTATATGGAAAGACTGCACCAGGCGGGTATAATCGTTGTTCCTGTTGTGCCAA
GCGTTGCGCTAGCCAAACGTATGGAAAAGCTTGGGGTAGATGCTGTTATTGCTGAGGGTATGGAAGCTGGA
GGACATATTGGCAAGTTAACGACTATGTCTTTAGTAAGACAAGTTGTTGAAGCGGTTTCGATTCTGTCTCAT
TGCGGCAGGTGTTATAGCTGATGGTTCATGGTGCAGCAGCAGCATTATGTTAGGAGCAGAGGCTGTTCAA
TTGGAACCTCGCTTTGTTGTTGCTAAAGAATCCAATGCTCACCAAAATTTTAAAGATAAAATCTTAGCAGCA
60 AAAGATATTGATACGGTGATTCTGCGCAGGTTGTGGGCCACCCGTCCGTTCTATTAATAAATAAATTGAC
CTCAGCTTACGCTAAAGCAGAAAAAGCATTTTTAATTGGTCAAAAAACAGCTACTGATATTGAAGAAATGG
GAGCAGGATCGCTTCGACACGCTGTTATTGAAGGCGATGTAGTCAATGGATCTGTTATGGCTGGCCAAATT
GCAGGGCTTGTGAGAAAAGAGAAAGCTGTGAACGATTTTAAAGATATTATTATGGTGCAGCTCGTGT
TATTCAAAATGAAGCTAAGCGCTGGCAATCTGTTTCAATAGAAAAGTAG

SEQ ID NO: 53 secuencia de aminoácidos que comprende GAS 509

MTKIYKTITELVGQTPPIIKLNRLIPNEAADVYVKLEAFNPGSSVKDRIALSMIEAAEAEGLISPGDVIIIE
PTSGNTGIGLAWVGAAGYRVIIIVMPETMSLERRQIIQAYGAEVLVTPGAEGMKGAIAKAETLAIELGAN
MPMQFNPNPANPSIHEKTTAQEILEAFKEISLDAFVSGVGTGGTSLGVSHVLKKNANPETVIYAVEAEESAV
LSGQEPGPHKIQGISAGFIPNTLDTKAYDQIIIRVKSDALETARLTGAKEGFLVGISSGAALYAAIEVAK
QLGKGKHVLTILPDNGERYLSTELYDVPVIKTK

SEQ ID NO: 54 secuencia de polinucleótidos que codifica GAS 509

ATGACTAAAATTTACAAAACCTATAACAGATTAGTAGGTCAAACACCTATTATCAAACCTAACCCTTTAA
TTCCAAACGAAGCTGCTGACGTTTATGTAAATTTAGAAAGCTTTTAAACCAGGATCTTCTGTAAAGATCG
TATTGCTTTATCGATGATTGAAGCTGCTGAAGCTGAAGGTCTGATAAGTCCTGGTGACGTTATTATCGAA
CCAACAAGTGGTAATACAGGTATTGGTCTTGCAATGGGTAGGTGCTGCTAAAGGGTATCGAGTCATTATTG
TTATGCCCCGAAACTATGAGCTTGGAAGACGGCAATCATTACAGGCTTATGGTGCAGAGCTTGTCTTAAC
ACCTGGAGCAGAAGGTATGAAAGGGGCTATTGCAAAAGCTGAAACTTTAGCAATAGAACTAGGTGCTTGG
ATGCCATATGCAATTTAATAACCTGCCAATCCAAGCATCCATGAAAAACAACAGCTCAAGAAATTTTGG
AAGCTTTTAAGGAGATTTCCTTAGATGCATTCGTATCTGGTGTGGTACTGGAGGAACACTTCTGGTGT
TTCACATGCTCTTGAAGAAAGCTAACCCTGAACTGTTATCTATGCTGTGAAGCTGAAGAATCTGCTGTC
TTATCTGGTCAAGAGCCTGGACCACATAAAATTCAGGTATATCAGCTGGATTTATCCCAAACACGTTAG
ATACCAAAGCCTATGACCAAATTATCCGTGTTAAATCGAAAGATGCTTTAGAACTGCTCGACTAACAGG
AGCTAAGGAAGGCTTCCCTGGTTGGGATTTCTTCTGGAGCTGCTCTTTACGCCGCTATTGAAGTCGCTAAA
CAGTTAGGAAAAGGCAACATGTGTTAACTATTTACCAGATAATGGCGAACGCTATTTATCGACTGAAC
TCTATGATGTACCAGTAATTAAGACGAAATAA

SEQ ID NO: 55 secuencia de aminoácidos que comprende la región transmembrana C-terminal de GAS 509
FLVGISSGAALYAAIEVAKQLGKGKHVLTILPDNGERYLSTELYDVPVIKTK

SEQ ID NO: 56 secuencia de aminoácidos que comprende un fragmento de GAS 509 en el que la región
transmembrana C-terminal está eliminada

MTKIYKTITELVGQTPPIIKLNRLIPNEAADVYVKLEAFNPGSSVKDRIALSMIEAAEAEGLISPGDVIIIEP
TSGNTGIGLAWVGAAGYRVIIIVMPETMSLERRQIIQAYGAEVLVTPGAEGMKGAIAKAETLAIELGAWMP

SEQ ID NO: 57 secuencia de aminoácidos que comprende GAS 366

MKVISNFQNKILILGLAKSGEAAKLLTKLGLVTVNDSKPFQNPAAQALLEBGIKVICGSHPEVLLDE
NPEYVMKNPGIPYDNPMVKRALAKEIPILTEVELAYFVSEAPIIGITGSNGKTTTTMIADVLNAGGQSAL
LSGNIGYPASKVVQKAIAGDTLVMELSSFLVGVNAFRPHIAVITNLMPTHLDYHGSFEDYVAAKWMIAQ
MTESDYILNANQEISATLAKTTKATVIPFSTQKVVVDGAYLKDGIYFKEQAIIAATDLGVPGSHNIENAL
ATIAVAKLSGIADDIIAQCLSHFGGVKHLQVRVQIKDITFYNDKSTNIIATQKALSGFDNSRLILIAGG
LDRGNEFDDLVDPDLLGLKQMIILGESARMKRAANKAEVSYLEARNVAEATELAFKLAQTGDTILLSPAN
ASWDMYPNFEVRGDEFLATFDCLRGDA

SEQ ID NO: 58 secuencia de polinucleótidos que codifica GAS 366

5 ATGAAAGTGATAAGTAATTTTCAAAACAAAAAATATTAATATTGGGGTTAGCCAAATCGGGCGAAGCAGC
AGCAAAATTTATGACCAAACTTGGTGCTTTAGTGACTGTTAATGATAGTAAACCATTTGACCAAAATCCAG
CGGCACAAGCCTTGTTGGAAGAGGGGATTAAAGGTCATTTGTGGTAGCCACCCAGTAGAATTATTAGATGAG
AACTTTGAGTACATGGTTAAAAACCCCTGGGATTCTTATGATAATCCTATGGTTAAACCGCGCCCTTGCAAA
GGAAATTTCCATCTTGACTGAAGTAGAATTGGCTTATTTCTGATCTGAAGCGCCTATTATCGGGATTACAG
GATCAAACGGGAAGACAACACACAGCAATGATTGCCGATGTTTTGAATGCTGGCGGGCAATCTGCACATC
TTATCTGGAACATTGGTTATCCTGCTTCAAAAGTTGTTCAAAAAGCAATTGCTGGTGATACCTTTGGTGAT
GGAATTTGCTCTCTTTTCAATTAGTGGGAGTGAATGCTTTTCGCCCTCATATTGCTGTCATCACTAATTTAA
TGCCGACTCACCTGGACTATCATGGCAGTTTGTAGGATTATGTTGCTGCTAAATGGATGATTCAAGCTCAG
ATGACAGAATCAGACTACCTTATTTAAATGCTAATCAAGAGATTTCAGCAACTCTAGCTAAGACCACCAA
AGCAACAGTGATTCCTTTTCAACTCAAAAAGTGGTTGATGGAGCTTATCTGAAGGATGGAATACTCTATT
TAAAGAACAGGCGATTATAGCTGCACTGACTTAGGTGTCCAGGTAGCCACAACATTGAAAATGCCCTA
GCAACTATTGCAGTTGCCAAGTTATCTGGTATTGCTGATGATATTATTGCCAGTGCCCTTTCACATTTTGG
AGGCGTTAAACATCGTTTGCAACGGGTTGGTCAAATCAAAGATATTACCTTCTACAAATGACAGTAAGTCAA
CCATATTTTAGCCACTCAAAAAGCTTTATCAGGTTTGTATAACAGTCGCTTGATTTTGATTGCTGGCGGT
CTAGATCTGCGCAATGAATTTGACGATTGGTGCCAGACCTTTTAGGACTTAAGCAGATGATTATTTTGGG
AGAATCCCGCAGAGCGTATGAAGCGAGCTGCTAACAAAGCAGAGGTCCTTATCTTGAAGCTAGAAATGTGG
CAGAAGCAACAGAGCTTGCTTTTAAGCTGGCCCAACAGGCGATACTATCTTGCTTAGCCAGCCAATGCT
AGCTGGGATATGTATCTTAATTTTGAAGTTCGTGGGGATGAATTTTTTGCAACCTTTGATTGTTTAAGAGG
AGATGCCTAA

SEQ ID NO: 59 secuencia de aminoácidos que comprende la secuencia líder N-terminal de GAS 366
 MKVISNFQNKILILGLAKSGEAAA

25 **SEQ ID NO: 60** secuencia de aminoácidos que comprende un fragmento de GAS 366 en el que la secuencia
 líder N-terminal está eliminada

30 KLTLKLGLVTVNDSKPFQNPAAQALLEEGIKVICGSHPVVELLDENFEYMVKNPGIPYDNPVMKRALAKE
IPILTEVELAYFVSEAPIIGITGSNGKTTTMIADVLNAGGQSALLSGNIGYPASKVVQKAIAGDTLVME
LSSFQLVGVNAFRPHIAVITNLMPTHLDYHGSFEDYVAAKWMIAQMTESDYLIILNANQEISATLAKTTKA
TVIPFSTQKVVDGAYLKDGLYFKEQAIIAATDLGVPGSHNIENALATIAVAKLSGIADDIIAQCLSHFGG
VKHLRQVRVQIKDITFYNDKSTNILATQKALSGFDNSRLILLAGGLDRGNEFDDLVPDLLGLKQMIILGE
SAERMKRAANKAEVSYLEARNVAEATELAFKLAQTGDTILLSPANASWDMYPNFEVRGDEPLATFDCLRGD
A

SEQ ID NO: 61 secuencia de aminoácidos que comprende GAS 159

40 MRKLYSFLAGVLGVIVILTSLSFILQKKSQSGSGSDKLVIYNWGDYIDPALLKKFTKETGIEVQYETFDSN
EAMYTKIKQGGTTYDIAVPSDYTTIDKMIKENLLNKLDKSKLVGMDNIGKEFLGKSFDPQNDYSLPYFWGTV
GIVYNDQLVDKAPMHWEDLWRPEYKNSIMLIDGAREMLGVGLTTFGYSVNSKNLEQLQAAERKLQQLTPNV
KAIVADEMKGYMIQGDAAIGITFSGEASEMLDSNEHLHYIVPSEGSNLWFDNLVLPKTMKHEKEAYAFLENF
INRPENAAQNAAYIGYATPNKKAKALLPDEIKNDPAFYPTDDIIKKLEVYDNLGSRWLGIIYNDLYLQFKMY
RK

SEQ ID NO: 62 secuencia de polinucleótidos que codifica GAS 159

50 ATGCGTAAACTTTATTCTCTTCTAGCAGGAGTTTTGGGTGTTATTGTTATTTTAAACAAGTCTTTCTTTTCAT
CTTGACAGAAAAATCGGGTTCTGGTAGTCAATCGGATAAAATTAGTTATTTATAACTGGGGAGATTACATTG
ATCCAGCTTTGCTCAAAAAATTCACCAAAGAAACGGGCATTGAAGTGCAGTATGAAACTTTTCGATTCCAAT
GAAGCCATGTACACTAAAATCAAGCAGGGCGGAACCACTTACGACATTGCTGTTCTTAGTGATTACACCAT
TGATAAAATGATCAAAGAAAACCTACTCAATAAGCTTGATAAGTCAAAATTAGTTGGCATGGATAATATCG
GGAAAGAATTTTAGGGAAGGCTTTGACCCACAAAACGACTATTCTTTGCCTTATTTCTGGGGAACCGTT
GGGATTGTTTATAATGATCAATTAGTTGATAAGGCGCCATGCACTGGGAAGATCTGTGGCGTCCAGAATA
TAAAAATAGTATTATGCTGATTGATGGAGCGCGTGAAATGCTAGGGGTTGGTTTAACTTTTGGTTATA
GTGTGAATTCTAAAAATCTAGAGCAGTTGCAGGCAGCCGAGAGAAAACTGCAGCAGTTGACGCCGAATGTT
AAAGCCATTGTAGCAGATGAGATGAAAGGCTACATGATTCAGGTGACGCTGCTATTGGAATTACCTTTTC
TGGTGAAGCCAGTGAGATGTTAGATAGTAACGAACACCTTCACTACATCGTGCCTTCAGAAGGGTCTAACC
TTTGGTTTGATAATTTGGTACTACCAAAAACCATGAAACACGAAAAAGAAGCTTATGCTTTTGTGAACCTT
ATCAATCGTCTGAAAATGCTGCGCAAAATGCTGCATATATTGGTTATGCGACACCAAAATAAAAAGCCAA
GGCCTTACTTCCAGATGAGATAAAAAATGATCCTGCTTTTATCCAACAGATGACATTATCAAAAAATTGG
AAGTTTATGACAATTTAGGGTCAAGATGGTTGGGGATTATAATGATTTATACCTCCAATTTAAATGTAT
CGCAAAATAA

SEQ ID NO: 63 secuencia de aminoácidos que comprende la secuencia líder N-terminal de GAS 159
MRKLYSFLAGVLGVIVILTSLFSI

SEQ ID NO: 64 secuencia de aminoácidos que comprende un fragmento de GAS 159 en el que la secuencia líder N-terminal está eliminada

LQKKSGSGSQSDKLVIYNWGDYIDPALLKKFTKETGIEVQYETFDSDNEAMYTKIKQGGTTYDIAVPSDYTI
DKMIKENLLNKLDKSKLVGMDNIGKEFLGKSFDPQNDYSLPYFWGTGIVYNDQLVDKAPMHWEDLWRPEY
KNSIMLIDGAREMLGVGLTTFGYSVNSKNLEQLQAAERKLQQLTPNVKAIVADEMKGYMIQGDAAGITFS
GEASEMLDSNEHLHYIVPSEGSNLWFDNLVLPKTMKHEKEAYAFNFINRPNENAAQNAAYIGYATPNKKAK
ALLPDEIKNDPAFYPTDDI IKKLEVDNLGSRWLGIYNDLYLQFKMYRK

SEQ ID NO: 65 secuencia de aminoácidos que comprende la secuencia hidrófoba C-terminal de GAS 159
WLGIVNDLYLQFKMYRK

SEQ ID NO: 66 secuencia de aminoácidos que comprende un fragmento de GAS 159 en el que la región hidrófoba C-terminal está eliminada

MRKLYSFLAGVLGVIVILTSLFSFILQKKSGSGSQSDKLVIYNWGDYIDPALLKKFTKETGIEVQYETFDSDN
EAMYTKIKQGGTTYDIAVPSDYTI DKMIKENLLNKLDKSKLVGMDNIGKEFLGKSFDPQNDYSLPYFWGTG
GIVYNDQLVDKAPMHWEDLWRPEYKNSIMLIDGAREMLGVGLTTFGYSVNSKNLEQLQAAERKLQQLTPNV
KAIVADEMKGYMIQGDAAGITFSGEASEMLDSNEHLHYIVPSEGSNLWFDNLVLPKTMKHEKEAYAFNFINR
PNENAAQNAAYIGYATPNKKAKALLPDEIKNDPAFYPTDDI IKKLEVDNLGSR

SEQ ID NO: 67 secuencia de aminoácidos que comprende un fragmento de GAS 159 en el que la secuencia líder N-terminal y la región hidrófoba C-terminal están eliminadas

LQKKSGSGSQSDKLVIYNWGDYIDPALLKKFTKETGIEVQYETFDSDNEAMYTKIKQGGTTYDIAVPSDYTI
DKMIKENLLNKLDKSKLVGMDNIGKEFLGKSFDPQNDYSLPYFWGTGIVYNDQLVDKAPMHWEDLWRPEY
KNSIMLIDGAREMLGVGLTTFGYSVNSKNLEQLQAAERKLQQLTPNVKAIVADEMKGYMIQGDAAGITFS
GEASEMLDSNEHLHYIVPSEGSNLWFDNLVLPKTMKHEKEAYAFNFINRPNENAAQNAAYIGYATPNKKAK
ALLPDEIKNDPAFYPTDDI IKKLEVDNLGSR

SEQ ID NO: 68 secuencia de aminoácidos que comprende GAS 217

MAQRIVITGASGGLAQAIVKQLPKEDSLILLGRNKRLEHCYQHIDNKECLELDITNPVAIEKMVAQIYQ
RYGRIDVLINNAGYGAFKGFEEFSAQEIADMFQVNTLASIHFACLIGQKMAEQGGHILINIVSMAGLIASA
KSSIYSATKFALIGFSNALRLELADKGVYVTTVNPFGPIATKFFDQADPSGHYLESVGKFTLQPNQVAKRLV
SIIGKNKRELNLPPSLAVTHQFYTLFPKLSDYLARKVFNKY

SEQ ID NO: 69 secuencia de polinucleótidos que codifica GAS 217

ATGGCACAAAGAATCATTGTTATCACGGGAGCTTCTGGAGGACTGGCTCAGGCAATTGTTAAGCAGTTACC
CAAGGAAGACAGCTTGATTTTATTAGGACGTAACAAAGAACGCCTAGAACACTGTTATCAGCATATTGACA
ACAAAGAATGCCTCGAGTTGGATATTACCAATCCAGTAGCCATTGAGAAAATGGTCGCCAGATTACCAG
CGCTATGGCCGTATTGATGTCTTGATTAATAATGCTGGCTACGGAGCTTTCAAAGGCTTTGAAGAGTTTTT
TGCCCAAGAAATAGCTGATATGTTTACGGTTAACACCCTAGCGAGCATTCACTTTGCTTGCTTGATTGGTC
AGAAAATGGCAGAGCAGGGGCAAGGTCACCTTATTAATATTGTGTCCATGGCAGGCTTGATTGCGTCAGCC
AAATCGAGCATTATTACGCCACCAAGTTTGCCTTATCGGATTTTCCAATGCCCTTCGCTTAGAATTAGC
GGATAAAGGGGTTTACGTGACCACCGTGAATCCAGGTCCCATGGCCACCAAGTTTTTGGACCAAGCTGACC
CGTCTGGACATTATTTGGAAAGCGTTGGTAAATTTACTCTCCAACCAATCAAGTGGCTAAGCGTTTTGGTT
TCTATTATCGGGAAAAATAAACGAGAATTGAATTTGCCCTTTAGTTTAGCGGTGACCCATCAATTTTACAC
CCTTTTCCCTAAATTATCTGATTATCTTGCAAGAAAGGTATTTAATTATAAATGA

SEQ ID NO: 70 secuencia de aminoácidos que comprende GAS 309

5 MIEKYLESSIESKQLIVLFFPKTSYLPITEVAEKTGLTFLQLNHYCEELNAFFPGSLSMTIQKRMISCQFT
HPFKETYLYQLYASSNVLQLLAFLIKNGSHSRPLTDFARSHFLSNSSAYRMREALIPLLRNFELKLSKNKI
VGEYRIRYLIALLYSKFGIKVYDLTQQDKNTIHSFLSHSSTHLKTSFWLSESFYDILLALSWKRHQFS
VTIPQTRIFQQLKKLFVYDSLKSSHDIIETVCQLNFSAGDLDYLYLIYITANNSFASLQWTPHEHIRQYCQ
10 LFEENDTFRLLLNPIITLLPNLKEQKASLVKALMFFSKSFLFNLQHFIPETNLFVSPYYKGNQKLYTSLKL
IVEEWMAKLPGKRDNLNKHHLFCHYVEQSLRNIQPLVVVVFVASFNAHLLTDSFPRYFSDKSIDPHSY
YLLQDNVYQIPDLKPDVLVTHSQLIPFVHHELTGKIAVAEISFDESILSIQELMYQVKEBKFPQADLTQKLT

SEQ ID NO: 71 secuencia de polinucleótidos que codifica GAS 309

15 TTGATAGAAAAATACTTGAATCATCAATCGAATCAAAATGTCAGTTAATTGTCTTGTGTTTTTAAGACATC
TTATTTGCCAATAACTGAGGTAGCAGAAAAAACTGGCTTAACCTTTTTACAACTAAACCATTTATGTGAGG
AACTGAATGCCTTTTTCCCTGGTAGTCTGTCTATGACCATCCAAAAAGGATGATATCTTGCCAATTTACA
20 CATCCTTTTAAAGAACTTATCTTTACCAACTCTATGCATCATCTAATGTCTTACAATTACTAGCCTTTTT
AATAAAAAATGGTTCCCACTCTCGTCCCTTACGGATTTTGCAGAAAGTCATTTTTTATCAAACTCCTCAG
CTTATCGGATGCGCGAAGCATTTGATTCCTTTATTAAGAACTTTGAATTAAACTCTCTAAGAACAAGATT
GTCGGTGAGGAATATCGCATCCGTTACCTCATCGCTCTGCTATATAGTAAGTTTGGCATTAAAGTTTATGA
CTTGACGCAGCAAGACAAAAACACTATTATAGCTTTTATCCCATAGTTCCACCCACCTTAAACCTCTC
CTTGGTTATCGGAATCGTTTTCTTTCTATGACATTTTATTAGCTTTATCGTGAAGCGGCATCAATTTTCG
25 GTAATATTTCCCAAAACAGAAATTTTTCAACAATTAACCAACTTTTTGTCTACGATTCTTTGAAAAAAG
TAGCCATGATATTATCGAACTTACTGCCAATAAATTTTACAGCAGGAGATTGGACTACCTCTATTATA
TTTATATCACCGCTAATAATCTTTTGCAGCTTACAATGGACACCTGAGCATATCAGACAATATTGTCAA
CTTTTTGAAGAAATGATACTTTTCGCCTGCTTTTAAATCCTATCATCACTCTTTTACCTAACCTAAAAAGA
GCAAAAGGCTAGTTTAGTAAAGCTCTTATGTTTTTTTCAAAATCATTCTTGTTTAATCTGCAACATTTTA
TTCCTGAGACCAACTTATTCGTTTCTCCGTACTATAAAGGAAACCAAAACTCTATACGTCCTTAAAGTTA
30 ATTGTCGAAGAGTGGATGGCCAACTTCTCGTGAAGCGTGACTTGAACCATAAGCATTPTCATCTTTTTTG
CCATATGTCGAGCAAAGTCTAAGAAATATCCAACCTCCTTTAGTTGTTGTTTTTCGTAGCCAGTAATTTTA
TCAATGCTCATCTCCTAACGGATTCTTTTCCAAGGTATTTCTCGGATAAAAGCATTGATTTTCATTCCTAT
TATCTATTGCAAGATAATGTTTATCAAATTCCTGATTTAAAGCCAGATTGGTTCATCACTCACAGTCAACT
GATTCCTTTTGTTCACCATGAACCTACAAAAGGAATTGCTGTTGCTGAAATATCTTTTGATGAATCGATTC
35 TGTCTATCCAAGAATTGATGTATCAAGTTAAAGAGGAAAAATCCAAGCTGATTTAACCAAGCAATTAACA
TAA

SEQ ID NO: 72 secuencia de aminoácidos que comprende GAS 372

40 MIQIGKLFAGRYRILKSIGRGGMADVYLANDLILDNEDVAIKVLRITNYQTDQVAVARFQREARAMAELNHP
NIVAIRDIGEEDGQQLVMEYVDGADLKRYIQNHAPLSNNEVVRIMEEVL SAMTLAHQKGI VHRDLKPQNI
LLTKEGVVKVTDGFI AVAFAETSLTQTNSMLGSHVHLSPEQARGSKATI QSDIYAMGIMLFEMLTGHI PYD
GDSAVTIALQHFQKPLPSII EENHNVPQALENVVIRATAKKLSDRYGSTFEMSRLMTALSYNRSRERKII
45 FENVESTKPLPKVASGPTASVKLSPTPTVLTQESRLDQTNQTDALQPPTKKKSGRFLGTLFKILFSFFI
VGVALFTYLILTKPTSVKVPNVAGTSLKVAKQELYDVGLKVGKIRQIESD TVAEGNVVRTPDKAGTAKRQG
SSITLYVSI GNKGFDMENYKGLDYQEAMNSLIETYGVPSKIKIERIVTNEY PENTVISQSPSAGDKFNP
GKSKITLSVAVSDTITMPMVTEYSYADAVNTLTALGIDASRIKAYVPSSSSATGFPVPIHSPSSKAIVSGQS
PYYGTSLSLSDKGEISLYLYPEETHSSSSSSSSSTSSSSSSSINDSTAPGSNTELSPSETTSQTP

50

55

60

65

SEQ ID NO: 73 secuencia de polinucleótidos que codifica GAS 372

ATGATTCAGATTGGCAAATTATTTGCTGGTCGTTATCGCATTCTGAAATCTATTGGCCGCGGTGGTATGGC
GGATGTTTATTTAGCAAATGACTTGATCTTGATAATGAAGACGTTGCAATCAAGGTCTTGCCTACCAATT
ATCAAACAGATCAGGTAGCAGTTGCGCGTTTCCAACGAGAAGCGCGGGCCATGGCTGAATTGAACCATCCC
AATATTGTTGCCATCCGGGATATAGGTGAAGAAGACGGACAGCAATTTTTAGTAATGGAATATGTGGATGG
TGCTGACCTAAAGAGATACATTCAAATCATGCTCCATTATCTAATAATGAAGTGGTTAGAATTATGGAAG
AAGTCTTTCTGCTATGACTTTAGCCCAACAAAAAGGAATTGTACACAGAGATTTAAACCTCAAAATATC
CTACTAACTAAGGAGGGTGTGTCAAAGTAACTGATTTCCGGCATCGCAGTAGCCTTTGCAGAAAACAGCTT
GACACAACTAATTCGATGTTAGGCAGTGTTCATTACTTGTCTCCAGAACAGGCTCGCGGCTCCAAAGCGA
CGATTCAAAGTGATATTTATGCGATGGGGATTATGCTCTTTGAGATGTTGACAGGCCATATCCCTTATGAC
GGCGATAGTGCTGTTACGATTGCCTTGCAACATTTTCAAAAGCCTCTTCCATCTATTATCGAGGAGAACCA
CAATGTGCCACAAGCTTTGGAGAATGTTGTTATTTCGAGCAACAGCCAAGAAATTAAGTGATCGTTACGGGT
CAACCTTTGAAATGAGTCGTGACTTAATGACGGCGCTTAGTTATAATCGTAGTCGGGAGCGTAAGATTATC
TTTGAGAATGTTGAAAGTACCAAAACCCCTCCCAAAAGTGGCCTCAGGTCCACCGCTTCTGTAAAATTGTC
TCCCCCTACCCCAACAGTGTAAACACAGGAAAGTCGATTAGATCAAATAATCAAACAGATGCTTTACAGC
CCCCCAACAAAAAGAAAAAGTGGTCGTTTTTTAGGTACTTTATTCAAATTTCTTTTCTTTCTTTATT
GTAGGTGTAGCACTCTTACTTATCTTATACTAACTAAACCAACTTCTGTGAAAGTTCCTAATGTAGCAGG
CACTAGTCTTAAAGTTGCCAAACAAGAACTGTATGATGTTGGGCTAAAAGTGGGTAAAATCAGGCAAAATTG
AGAGTGATACGGTTGCTGAGGGAATGTAGTTAGAACAGATCTAAAGCAGGAACAGCTAAGAGGCAAGGC
TCAAGCATTACGCTTTATGTGTCAATTTGGAACAAAGGTTTTGACATGGAAGAACTACAAAGGACTAGATT
TCAAGAAGCTATGAATAGTTTGTATAGAACTTATGGTGTCCAAAATCAAAAATCAAAATTGAGCGCATTG
TAACTAATGAATATCCTGAAAATACAGTCATCAGTCAATCGCCAAGTGCGGGTGATAAATTTAATCCAAAC
GGAAAGTCTAATAATTACGCTCAGTGTGCTGTTAGTGATACGATCACTATGCCTATGGTAACAGAATATAG
TTATGCAGATGCAGTCAATACCTTAACAGCTTTAGGTATAGATGCATCTAGAATAAAAGCTTATGTGCCAA
GCTCTAGCTCAGCAACGGGCTTTGTGCAATTCATTCTCTAGTTCTAAAGCTATTGTGCTAGTGGTCAATCT
CCTTACTATGGAACGCTTTTGAGTCTGTCTGATAAAGGAGAGATTAGTCTTTACCTTTATCCAGAAGAAAC
ACACTCTTCTAGTAGCTCATCGAGTTCAACGTCAAGTTCAAACAGTTCCTCAATAAATGATAGTACTGCAC
CAGGTAGCAACACTGAATTAAGCCCATCAGAACTACTTCTCAAACACCTTAA

SEQ ID NO: 74 secuencia de aminoácidos que comprende GAS 39

MDLILFLVLVLLGLGAYLLFKVNLQHQQLAQTLEGNADNLSQMTYQLDTANKQQLLELTQLMNRQOAGL
YQQLTDIRDVLHRSLSRDRSDRLEKINQOVNQLKNMQESNEKRLKMRQIVEEKLEETLKNRLHASF
DSVSKQLESVNKGLGEMRSVAQDVGTLNKVLSTKTRGILGELQLGQI IEDIMTSSQYEREFVTVSGSSER
VEYAIKLPNGNQGGYIYLPIDSKFPLEDYRLEDAYEVDKLAIEASRKALLAAIKRFAKDIHKKYNLPPE
TTNFGVMFLPTEGLYSEVVRNASFFDSLREENIVVAGPSTLSALLNSLSVGFKTLNMQNADDISKILGN
VKLEFDKFGGLLAKAQKQMNNTANNTLDQLI STRTNAIVRALNTVETYQDQATKSLNMLPLEEENNEN

SEQ ID NO: 75 secuencia de polinucleótidos que codifica GAS 39

ATGGACCTTATCTTGTTCCTTTTGGTCTTGGTTCTCTTAGGTTTAGGGGCTTATCTGTTGTTCAAAGTCAA
CGGCTTCAACATCAGCTTGCCCAACCTTAGAAGGCAACGCGGATAATTGTCTGACCAATGACCTTACC
AGTTGGATACAGCTAACAAACAACAATTGTTAGAGCTAACACAGCTGATGAACCGACAACAAGCAGGCCCT
TACCAACAATTAACAGATATTCGTGACGTCTTGCACCGTAGTTTGTCTGATAGTAGGGACCGGTCTGACAA
ACGCTTAGAAAAAATTAACCAGCAGGTCAACCAATCGCTCAAAAATATGCAAGAATCTAACGAAAAACGTT
TGGAGAAAAATGCGCCAGATCGTTGAAGAAAAATTTGAAGAAACCTTAAAAAATCGTCTGCACGCCTCTTTC
GATCTGTATCCAAGCAACTAGAAAGTGTCAATAAAGGCTTGGGAGAAATGCGTAGCGTGGCTCAAGATGT
GGGTACTTTAAATAAGGTTTGTCCAAATACCAAAACACGAGGCATTTTAGGCGAACTTCAACTAGGCCAAA
TCATTGAGGATATCATGACATCAAGCCAGTACGAAAGAGAATTTGTAACGGTTAGTGGTTCTAGTGAACGC
GTAGAATATGCGATTAAAGCTCCAGGAAATGGTCAAGGCGGTTATATTTACCTACCGATTGACTCAAAAT
CCCTCTTGAAGATTATACCGATTAGAAGATGCTTACGAAGTTGGTGATAAACTGGCCATCGAGGCTAGCC
GAAAAGCACTTCTGGCAGCTATCAAACGCTTTGCCAAGACATTCAATAAAAGTACTTGAACCCCCCAGAG
ACGACCAATTTCCGAGTTATGTTCTTACCAACAGAAGGTCTTTATTCAGAAGTGGTCAGAAATGCGTCTTT
CTTTGATAGCTTTCGCGGAAGAAAAATTTGTTGGTTGACAGGCCCTTCGACCTGTCTGCTTTGCTGAATT
CCTTATCTGTTGGTTTCAAGACCTTAAATATCAAAAAAATGCTGTATGACATCAGTAAATTTTAGGCAAT
GTCAAGTTAGAATTCGATAAATTTGGCGGCCTGCTTGCCAAGGCTCAAAAACAAATGAATACAGCTAATAA
TACGCTGGATCAGCTCATTTCAACAAGGACAAATGCCATTGTTGAGGCTTGAATACCGTTGAAACTTATC
AAGACCAAGCAACAAATCTCTCTTGAACATGCCCTTATTAGAAGAGGAAAAATTAATGAAATTA

SEQ ID NO: 76 secuencia de aminoácidos que comprende GAS 42

MTKEKLVAFSQAHAEPAWLQERRLAALAEIPNLELPTIERVKFHRWNLGDGTLTENESLASVPDFIAIGDN
PKLVQVGTQTVLEQLPMALIDKGVVFSDFYTALEEIPEVIEAHFGQALAFDEDKLAAYHTAYFNSAAVLVY
PDHLEITTPIEAIFLQSDSDVPFNKHVLIAGKESKFTYLERFESIGNATQKISANISVEVIAQAGSQIK
FSAIDRLGPSVTTYISRRGRLEKDANIDWALAVMNEGNVIADFDSDLIGQGSQADLKVVAASSGRQVQGID
TRVTNYGQRTVGHIHQHVILERGTLTFTNGIGHILKDAKGADAQQESRVLMLSDQARADANPILLIDENEV
TAGHAASIGQVDPEDMYLMSRGLDQETAERLVIRGFLGAVIAEIPISVRQEIIKVLDEKLLNR

SEQ ID NO: 77 secuencia de polinucleótidos que codifica GAS 42

ATGACAAAAGAAAACTAGTGGCTTTTTCGCAAGCCCACGCTGAGCCTGCTTGGCTGCAAGAACGGCGTTT
AGCGGCATTAGAAGCCATTCCAAATTTGGAATTACCAACCATCGAAAGGGTTAAATTTACCGTTGGAATC
TAGGAGATGGTACCTTAACAGAAAATGAAAGTCTAGCTAGTGTTCAGATTTTATAGCTATTTGGAGATAAC
CCAAAGCTTGTTCAGGTAGGCACGCAACAGTCTTAGAAGCAGTTACCAATGGCGTTAATTGACAAGGGAGT
TGTTTTTCACTGATTTTATACGGCGCTTGAGGAAATCCCAGAAAGTAATTGAAGCTCATTTTGGTCAGGCAT
TAGCTTTTGTATGAAGACAACTAGCTGCCTACCACACTGCTTATTTTAAATAGCGCAGCGTGTCTACGTT
CCTGATCACTTGGAAATCACAACCTCCTATTGAAGCTATTTTCTTACAAGATAGTGACAGTGACGTTCCCTT
TAACAAGCATGTTCTAGTGATTGCAGGAAAAGAAAGTAAGTTCACCTATTTAGAGCGTTTGAATCTATTG
GCAATGCCACTCAAAAGATCAGCGCTAATATCAGTGATAGAAGTGATTGCTCAAGCAGGCAGCCAGATTAAA
TTCTCGGCTATCGACCGCTTAGGTCCTTCAGTGACAACCTATATTAGCCGTCGAGGACGTTTAGAGAAGGA
TGCCAAACATTGATTGGGCCTTAGCTGTGATGAATGAAGGCAATGTCTATTGCTGATTTTGACAGTGATTGA
TTGGTCAGGGCTCACAAGCTGATTGAAAGTTGTTGCAGCCTCAAGTGGTCGTCAGGTACAAGGTATTGAC
ACGCGCGTGACCAACTATGGTCAACGTACGGTCGGTCATATTTTACAGCATGGTGTGATTTTGGAACTGG
CACCTTAACGTTTAAACGGGATTGGTCATATTCTAAAGACGCTAAGGGAGCTGATGCTCAACAAGAAAGCC
GTGTTTTGATGCTTTCTGACCAAGCAAGAGCCGATGCCAATCCAATCCTCTTAATTGATGAAAATGAAGTA
ACAGCAGGTCATGCAGCTTCTATCGGTCAGGTTGACCCGTAAGATATCTATTACTTGATGAGTCGAGGACT
GGATCAAGAAACAGCAGAACGATTGGTTATTAGAGGATTCCTAGGAGCGGTTATCGCTGAAAATTCCTATTC
CATCAGTCCGCCAAGAGATTATTAAGGTTTTAGATGAGAAATTGCTTAATCGTTAA

SEQ ID NO: 78 secuencia de aminoácidos que comprende GAS 58

MKWSGFMKTKSKRFLNLATLCLALLGTTLLMAHPVQAEVISKRDYMTFRFGLGDLEDDSANYPNLEARYKG
YLEGYEKGLKGDDIPERPKIQVPEDVQPSDHGDYRDGYEEGFGEQHKRDPLETEAEDDSQGGREQRQGH
QEGADSSDLNVEESDGLSVIDEVGVYQAFSTIWTYLSGLF

SEQ ID NO: 79 secuencia de polinucleótidos que codifica GAS 58

ATGAAATGGAGTGGTTTATGAAAACAAAATCAAAACGCTTTTAAACCTAGCAACCTTTGCTTGGCCCT
ACTAGGAACAACCTTGTCTAATGGCACATCCCGTACAGCGGAGGTGATATCAAAAAGAGACTATATGACTC
GCTTCGGGTTAGGCGATTTAGAAGATGATTACAGCTAACTATCCTTCAAATTTAGAAGCTAGATATAAAGGA
TATTTAGAGGGATATGAAAAGGCTTAAAAGGAGATGATATACCGAACGGCCCAAGATTCAGGTTCCCTGA
GGATGTTTACGCCATCTGACCATGGCGACTATAGAGATGGTTATGAGGAAGGATTTGGAGAAAGGACAACATA
AACGTGATCCATTAGAAACAGAAGCAGAAGATGATTCTCAAGGAGGACGTCAGAAGGACGTCAGGACAT
CAAGAAGGAGCAGATTCTAGTGATTTGAACGTTGAAGAAAGCGACGTTTGTCTGTTATTGATGAAGTAGT
TGGAGTAATTTATCAAGCATTTAGTACTATTTGGACATACTTAAGCGGTTTGTCTAA

SEQ ID NO: 80 secuencia de aminoácidos que comprende la secuencia líder N-terminal de GAS 58
MKWSGFMKTKSKRFLNLATLCLALLGTTLLMA

SEQ ID NO: 81 secuencia de aminoácidos que comprende un fragmento de GAS 58 en el que la secuencia líder
N-terminal está eliminada

HPVQAEVISKRDYMTFRFGLGDLEDDSANYPNLEARYKGYLEGYEKGLKGDDIPERPKIQVPEDVQPSDHG
DYRDGYEEGFGEQHKRDPLETEAEDDSQGGREQRQGHQEGADSSDLNVEESDGLSVIDEVGVYQAFS
TIWTYLSGLF

SEQ ID NO: 82 secuencia de aminoácidos que comprende GAS 290

5

MKHILFIVGSLREGSFPNHQLAAQAQKALEHQAVVSYLNWKDVPLNQDIEANAPLPVVDARQAVQSADAIW
IFTPVYNFSIPGSVKNLLDWSRALDLSIPTGSAIGGKVTVSSVANGGHDQVFDQFKALLPFIRTSVAG
EFTKATVNPDAWGTGRLEISKETKANLLSQAEALLAAI

SEQ ID NO: 83 secuencia de polinucleótidos que codifica GAS 290

10

ATGAAACATATTTTATTTATTGTTGGCTCGCTTCGTGAAGGGTCTTTTAACCATCAATTAGCGGCTCAAGC
ACAAAAGCTCTGGAACATCAAGCAGTTGTATCTTACTTAAATTGGAAAGACGTTCCGTGTTTGAATCAAG
ATATCGAAGCTAATGCACCTTTACAGTTGTTGACGCTCGTCAAGCTGTTTCAGTCAGCGGATGCTATCTGG
ATTTTACACCAGTTTACAACTTCTCTATTCCAGGTTCTGTTAAAAACCTGCTAGACTGGTTGTCTCGTGC
TCTTGATTTGTCTGATCCGACGGGCCCATCTGCTATTGGCGGTAAGGTGGTTACGGTCTCTTCAGTTGCAA
ATGGCGGGCATGATCAAGTATTTGATCAGTTTAAAGCACTATTGCCGTTTATCCGAACCTCAGTAGCAGGA
GAGTTTACAAAAGCAACTGTGAATCCTGATGCCTGGGGAACAGGAAGGCTTGAGATTTCAAAGAGACAAA
AGCAAACTTGCTATCTCAGGCAGAGGCTCTTTTAGCGGCTATTTAG

15

20

SEQ ID NO: 84 secuencia de aminoácidos que comprende GAS 511

25

MTDVSRIKLEARDQGRLLTLDYANLIPDDFMELHGDHRHFSDDGAIVGGLAYLAGQPVTVIGIQKGNLQDN
LARNFGQPNPEGYRKALRLMKQAEKFGRPVVTFFINTAGAYPGVGAEEERGQGEAIKLNLMEMSDLKVPILAI
IIGEGSGGALALAVADQVWMLNENTMYAVLSPEGFASILWKDGSRAEAAELMKITAGELYKMGIVDRIIP
EHGYFSSEIVDIIKANLIEQITSLQAKPLDQLLDERYQRFKY

SEQ ID NO: 85 secuencia de polinucleótidos que codifica GAS 511

30

ATGACAGATGTATCAAGAATTTTAAAGAAGCGCGTGATCAAGGGCGTTTAAACAACCTTTGGATTACGCCAA
CCTTATTTTCGATGACTTTATGGAAGTGCATGGCGATCGCCATTTTCAGATGATGGTGCCATTGTAGGTG
GCCTAGCTTATTTGGCGGGACAACCTGTTACGGTCATTGGTATTCAAAAAGGTAAGAATTTACAGGATAAT
TTGGCAAGGAATTTGGCCAGCCCAATCCAGAAGGTTATCGTAAAGCTTTGCGCCTTATGAAACAGGCAGA
AAAATTTGGACGACCAAGTTGTTACGTTTATCAATACTGCAGGAGCCTATCCAGGTGTCGGTGCGGAAGAAC
GAGGACAGGGTGAGGCCATTGCTAAAAATTTGATGGAAATGAGTGATCTCAAGGTTCCCATTTATCGCCATC
ATTATTGGTGAAAGGAGGCTCTGGTGGTGCATTAGCCTTAGCGGTTGCCGATCAGGTCTGGATGCTTGAAA
TACTATGTATGCGGTTCTTAGCCCAAGAGGCTTTGCTTCTATTTTATGGAAGGATGGTTCAAGGGCGACCG
AGGCCGCTGAATFGATGAAAATCACAGCGGGTGAACCTACAAAATGGGAATAGTAGACCGTATTATTTCCA
GAACATGGTTATTTTCAAGTGAAATCGTTGACATCATCAAAGCTAACCTCATCGAACAAATAACAGTTT
GCAAGCTAAGCCATTAGACCAATTATTAGATGAGCGCTACCAACGCTTTCGTAAATATTAA

40

SEQ ID NO: 86 secuencia de aminoácidos que comprende GAS 533

45

MAITVADIRREVKEKNVTFRLMFTDIMGVMKNVEIPATKEQLDKVLSNKMVDFGSSIEGFVRINESDMYL
YPDLDTWIVFPWGDENGAVAGLICDIYTAEGKPFAGDPRGNLKRALKHMNEIGYKSFNLGPEPEFFFLFKMD
DKGNPTLEVNDNGGYFDLAPIDLADNTRREIVNITKMGFEVEASHHEVAVGQHEIDFKYADVLRKACDNIQ
IFKLVVKTIAREHGLYATFMAKPKFGIAGSGMHCMNSLFDNQGNNAFYDEADKRGMLSEDAIYFLGLMK
HAYNYTAITNPTVNSYKRLVPGYEAPVYVWAGSNRSLIRVPASRGMGTRLELRSDPTANPYLALAVLL
EAGLDGIINKIEAPEPVEANIYTMTEERNEAGIIDLPSTLHNALKALKQKDDVVQKALGYHIYTNFLEAKR
IEWSSYATFVSQWEIDHYIHNY

50

55

60

65

SEQ ID NO: 87 secuencia de polinucleótidos que codifica GAS 533

ATGGCAATAACAGTAGCTGACATTTCGTCGTGAAGTCAAAGAAAAAATGTAACGTTTCTTCGCTTGATGTT
 CACTGATATCATGGGCGTTATGAAAAATGTGGAGATTCTGCAACTAAAGAACAGTTAGACAAAGTATTGT
 5 CTAACAAGGTTATGTTTGTATGGTTCATCTATCGAAGGTTTGTACGGATCAATGAGTCAGATATGTACCTT
 TACCCCGATTAGACACTTGGATTGTTTTCCTGGGGAGATGAAAATGGAGCAGTTGCAGGTTTAATTTG
 TGATATTTATACAGCAGAAGGAAAGCCTTTTCAGGAGATCCTAGAGGAAATTTAAAAAGAGCCCTGAAAC
 ACATGAACGAGATCGGCTACAAATCATTTAATCTTGGACCAGAACCAGAATTTTTCCTTTTAAAGATGGAT
 10 GATAAAGGTAATCCGACACTTGAAGTTAACGATAATGGTGGTTATTTGATTTAGCGCCAATTGACTTAGC
 AGACAACACGCGCCGTGAAATTGTGAATATTTTAACGAAAATGGGTTTGAAGTGGAAGCTAGTCATCATG
 AAGTGGCTGTTGGTCAACATGAGATTGATTTTAAATATGCAGATGTTTGAAGCTTGTGATAATATTCAA
 ATTTTAAAGCTAGTTGTAAAAACGATTGCCCGTGAACATGGACTTTATGCTACTTTTCATGGCTAAACCAA
 ATTTGGAATAGCTGGATCAGGGATGCATGTAACATGTCTTTGTTTGATAACCAAGGTAATATGCTTTT
 15 ATGATGAAGCTGATAAGCGAGGGATGCAGTTATCAGAAGATGCTTATTATTTCTTGGGAGGACTAATGAAG
 CATGCTTATAACTACACTGCTATCACTAACCTACAGTGAATTCCTATAAACGATTAGTTCAGGTTATGA

GGCACCTGTTTATGTCGCTTGGGCTGGAAGTAATCGTTACCGCTTATCCGTGTTCCAGCATCACGTGGTA
 TGGGAACGCGTTTGGAGTTACGTTTCGTTGATCCGACAGCTAATCCTTATTTAGCCTTGGCTGTTCTCTTG
 20 GAAGCTGGATTAGATGGTATCATTAACAAAATGAAGCTCCAGAACCCGTTGAAGCTAACATTTATACCAT
 GACAATGGAAGAACGAAATGAAGCAGGCATTATTGATTTGCCATCAACGCTTCATAATGCCCTTAAAGCTC
 TTCAAAAAGATGATGTGGTACAAAAGGCACCTAGGTTACCATATCTACACTAATTTCTTAGAAGCAAAACGA
 ATTGAATGGTCTTCTATGCACTTTTGTTCCTCAATGGGAAATTGACCATTATATTCATAATTATTAG

SEQ ID NO: 88 secuencia de aminoácidos que comprende GAS 527

MTEISILNDVQKIIIVLDYGSQYNQLIARRIREFGVFSSELKSHKITAQELREINPIGIVLSGGPNSVYADNA
 FGIDPEIFELGIPILGICYGMLITHKLGGKVVPAGQAGNREYQSTLHLRETSKLSFGTFQEQLVLSHG
 30 DAVTEIPEGFHLVGDSNDPCPYAAIENTEKNLYGIQPHPEVRHSVYVNDILKNFAISICGARGDWSMDNFID
 MEIAKIRETVGDRKVLGLSGGVDSVVGVLQKAITGDQLTCIFVDHGLLRKDEGDQVMGLGGKFLNLI
 RVDASKRFLDLLADVEDPEKKRKIIIGNEFVYVFDDEASKLKGVDFLAQGTLTYTDIIIESGTETAQTIKSHHN
 VGGLPEDMQFELIEPLNTLFKDEVRLGIALGMPEBIVWRQFPFPGPLAIRVMGAITEEKLETVRESDAIL
 35 REEIAKAGLDRDVRVWQYFTVNTGVRVSVGMGDGRTYDYTTIAIRAITSIDGMTADFAQLPNDVLKKISTRIVN
 EVDHVNRIYDITSKPPATVEWE

SEQ ID NO: 89 secuencia de polinucleótidos que codifica GAS 527

ATGACTGAAATTTCAATTTTGAATGATGTTCAAAAAATTATCGTTCTTGATTATGGTAGCCAGTACAATCA
 GCTTATTGCTAGACGTATTCGAGAGTTTGGTGTTCCTCCGAACTAAAAAGCCATAAAATCACCCTCAAG
 40 AACTTCGTAGATCAATCCCATAGGTATCGTTTATCAGGAGGGCCTAACTCTGTTTACGCTGATAACGCC
 TTTGGCATTGACCTGAAATCTTTGAAC TAGGGATTCCGATTCTTGGTATCTGTTACGGTATGCAATTAAT
 CACCCATAAATTAGGTGTTAAAGTTGTTCTGCTGGACAAGCTGGTAATCGTGAATACGGTCAGTCAACCC
 45 TTCATCTTCGTGAAACGTCAAAATATTTCAGGCACACCTCAAGAACAACCTCGTTTGTATGAGCCATGGT
 GATGCTGTTACTGAAATCCAGAAGTTTCCACCTTGTGGAGACTCAAATGACTGTCCCTATGCAGCTAT
 TGAANAATACTGAGAAAAACCTTTACGGTATTCAGTTCCACCCAGAAGTGAGACACTCTGTTTATGGAAATG
 ACATTCTTAAAAACTTTGCTATATCAATTTGTGGCGCGCTGGTGATTGGTCAATGGATAATTTTATTGAC
 ATGGAAATTGCTAAAAATTCGTGAAACGTAGGCGATCGTAAAGTTCTTCTAGGTCTTTCTGGTGGAGTTGA
 50 TTCTTCAGTTGTTGGTGTCTACTTCAAAAAGCTATCGGTGACCAATTAACCTGTATTTTCGTTGATCAG
 GTCTTCTTCGTAAAGACGAGGGCGATCAAGTTATGGGAATGCTTGGGGGCAAAATTTGGCCATAATATTATC
 CGTGTGGATGCTTCAAAACGTTTCTTAGACCTTCTTGCAGACGTTGAAGATCCTGAGAAAAACGTAAAT
 TATTGGTAATGAATTTGTCTATGTTTGTGATGATGAAGCCAGCAAAATTAAGGTGTTGACTTCCTTGCCC
 55 AAGGAACACTTTATACTGATATCATTGAGTCAGGAACAGAACTGCTCAAACCATCAAATCACATCACAT
 GTGGGTGGTCTCCCCGAAGACATGCAGTTTGAATTGATTGAGCCCTTAAACACTCTTTTCAAAGATGAAGT
 TCGAGCGCTTGAATCGCTCTTGAATGCCTGAAGAAATTTGTTGGCGCAACCATTTCCAGGTCTGGAC
 TTGCTATCCGTGTCATGGGAGCAATTACTGAAGAAAACTTGAACCGTTTCGCGAATCAGACGCTATCCTT
 60 CGTGAAGAAATTTGCTAAGGCTGGACTTGATCGTGACGTGTGGCAATACTTTACAGTTAACACAGGTGTCCG
 TTCTGTAGCGCTCATGGAGATGGTCTGACTTATGATTATACCATCGCCATTCGTGCTATTACGCTATTG
 ATGGTATGACAGCTGACTTTGCTCAACTTCCTTGGGATGCTTGAAAAAAATCTCAACACGTATCGTAAAT
 65 GAAGTTGACCACGTTAACCGTATCGTCTACGACATCACAAGTAAACCACCCGCAACAGTTGAATGGGAATA
 A

SEQ ID NO: 90 secuencia de aminoácidos que comprende GAS 294

5 MSQSTATYINIVIGAGLAGSEAAAYQIAKRGIPVKLYEMRGVKATPQHKTTNFALVCSNSFRGDSLTVNAVGL
LKEEMRRLDSIIMRNGEANRVPAGGAMAVDREGYAESVTAELNHPLIEVIRGEITEIPDDAIVTIATGPL
TSDALAEKIHALNGGDGFYFYDAAPIIDKSTIDMSKVYLKSRYDKGEAAAYLNCMPMTKEEFMAFHEALTTA
EEAPLNAFEKEKYFEGCMPIEVMAKRGIKTMLYGPMPKPVGLEYPDDYTGPDRDGEFKTPYAVVQLRQDNAAG
10 SLYNIVGFQTHLKWGEQKRVFQMIPLGLENABFVRYGVMHRSYMDSPNLLTETFPQSRSNPNLFFAGQMTGV
EGYVESAAAGLVAGINAARLFKREEALIFPQTTAIGSLPHYVTHADSKHFQPMNVNFGIIEKELEGPRIRDK
KERYEAIASRALADLDTCLASL

SEQ ID NO: 91 secuencia de polinucleótidos que codifica GAS 294

15 TTGCTCTCAATCAACTGCAACTTATATTAATGTTATTGGAGCTGGGCTAGCTGGTTCTGAAGCTGCCTATCA
GATTGCTAAGCGCGGTATCCCCGTTAAATTGTATGAAATGCGTGGTGTCAAAGCAACACCGCAACATAAAA
CCACTAATTTTCCGAATTGGTCTGTTCCTCAACTCATTTCGTGGTGATAGCTTAACCAATGCAGTCGGTCTT
CTCAAAGAAGAAATGCGGCGATTAGACTCCATTATTATGCGTAATGGTGAAGCTAACCGCGTACCTGCTGG
20 GGGAGCAATGGCTGTTGACCGTGAGGGGTATGCAGAGAGTGTCACTGCAGAGTTGGAATAATCATCCTCTCA
TTGAGGTCAATCGTGGTGAAATTACAGAAATCCCTGACGATGCTATCACGGTTATCGCGACGGGACCGCTG
ACTTCGGATGCCCTGGCAGAAAAAATTCACGCGCTAAATGGTGGCGACGGATTCTATTTTTACGATGCAGC
AGCGCCTATCATTTGATAAATCTACCATTGATATGAGCAAGGTTTACCTTAAATCTCGCTACGATAAAGGCG
25 AAGCTGCTTACCTCAACTGCCCTATGACCAAGAAGAATTATGGCTTTCCATGAAGCTCTGACAACCGCA
GAAGAAGCCCCGCTGAATGCTTTGAAAAAGAAAAGTATTTTGAAGGCTGTATGCCGATTGAAGTTATGGC
TAAACGTGGCATTAAAAACCATGCTTTATGGACCTATGAAACCCGTTGGATTGGAATATCCAGATGCTATA
CAGGTCTCGCGATGGAGAATTTAAAACGCCATATGCCGTCGTGCAATGCGTCAAGATAATGCAGCTGGA
AGCCTTTTATAATATCGTTGGTTTCCAAACCCATCTCAAATGGGGTGAGCAAAAACGCGTTTTCCAAATGAT
30 TCCAGGGCTTGAAATGCTGAGTTTGTCCGCTACGGCGTCATGCATCGCAATTCCTATATGGATTACACAA
ATCTTTTAAACCGAAACCTTCCAATCTCGGAGCAATCCAAACCTTTTCTTTGCAGGTGAGTACTGGAGTT
GAAGGTTATGTCGAATCAGCTGCTTCAGGTTTATAGCAGGAATCAATGCTGCTCGTTTGTTCAAAAGAGA
AGAAGCACTTATTTTTCTCAGACAACAGCTATTGGGAGTTTGCCTCATTTATGTGATGCTGCCGACAGTA
AGCATTTCCAAACCAATGAACGTCAACTTTGGCATCATCAAAGAGTTAGAAGGCCACGCATTCGTGACAAA
35 AAAGAACGTTATGAAGCTATGCTAGTCTGCTTTGGCAGATTTAGACACCTGCTTAGCGTCGCTTTAA

SEQ ID NO: 92 secuencia de aminoácidos que comprende GAS 253

40 MPKILFTGGGTVGHVTLNLILIPKFIKDGWEVHYIGDKNGIEHTEIEKSGLDVTFHAIATGKLRRYFSWQ
NLADVFKVALGLLQSLFIVAKLRPQALFSKGGFVSVPPVAAKLLGKPVFIHESDRSMGLANKIAYKFATT
MYTTFEQEDQLSKVKHLGAVTKVFKDANQMPESTQLEAVKEYFSRDLKTLTFIGGSAGAHVFNQFISDHPE
LKQRYNIIINITGDPHLNELSSHLRYVDYVTDLYQPLMAMADLVVTRGGSNTLFELLAMAKHLIIVPLGKEA
SRGDQLENATYFEKRGYAKQLQEPDLTLHNFDQAMADLFEHQADYEATMLATKEIQSPDFYDILLRADISS
45 AIKEK

SEQ ID NO: 93 secuencia de polinucleótidos que codifica GAS 253

50 ATGCCAAGAAGATTTTATTTACAGGTGGTGGAACTGAGGTGATGTACCTTGAACCTCATTCTCATACC
AAAATTTATCAAGGACGGTTGGGAAGTACATTATATTTGGTGATAAAAATGGCATTGAACATACAGAAATTG
AAAAGTCAGGCCCTTGACGTGACCTTTTCATGCTATCGCGACAGGCAAGCTTAGACGCTATTTTTCATGGCAA
AATCTAGCTGATGTTTTTAAGGTTGCACTTGGCCTCCTACAGTCTCTCTTTATTGTTGCCAAGCTTCGCCC
TCAAGCCCTTTTTTCCAAAGGTGGTTTTGTCTCAGTACCGCCAGTTGTGGCTGCTAAATTGCTTGGTAAAC
CAGTCTTTATTCATGAATCAGATCGGTCAATGGGACTAGCAACAAGATTGCCTACAAATTTGCAACTACC
55 ATGTATACCACTTTTGAGCAGGAAGACAGTTGTCTAAAGTTAAACACCTTGGAGCGGTGACAAAAGGTTTT
CAAAGATGCCAACCAATGCCTGAATCAACTCAGTTAGAGGCGGTGAAAGAGTATTTTAGTAGAGACCTAA
AAACCTCTTGTTTTATGGTGGTTTCGGCAGGGGCGCATGTGTTTTAATCAGTTTATTAGTGATCATCCAGAA
TTGAAGCAACGTTATAATATCATCAATATTACAGGAGACCTCACCTTAATGAATTGAGTTCTCATCTGTA
TCGAGTAGATTATGTTACCGATCTCTACCAACCTTTGATGGCGATGGCTGACCTTGTAGTGACAAGAGGGG
60 GCTCTAATACACTTTTGGAGCTACTGGCAATGGCTAAGCTACACCTCATCTCCTTGGTAAAGAAGCT
AGCCGTGGCGATCAGTTAGAAAATGCCACTTATTTGAGAAGAGGGGCTACGCTAAACAATTAACAGGAACC
TGATTTAACTTTGCATAATTTTGATCAGGCAATGGCTGATTTGTTTGAACATCAGGCTGATTATGAGGCTA
CTATGTTGGCAACTAAGGAGATTCACTACCGGACTTCTTTTATGACCTTTTGAGAGCTGATATTAGCTCC
CGGATTAAGGAGAAGTAA

65

SEQ ID NO: 94 secuencia de aminoácidos que comprende GAS 529

MCGIVGVVGNRNATDILMQGLEKLEYRGVDSAGIFVANANQTNLIKSVGRIADLRKIGIDVAGSTGIGHT
 RWATHGQSTEDNAHPHTSQTGRFVLVHNGVIENYLHIKTEFLAGHDFKQTDTEIAVHLIGKFVEEDKLSV
 LEAFKKSLSIIIEGSYAFALMDSQATDTIYVAKNKSPLLIIGLEGYNMVCSDAMAMIRETSEFMEIHDKELV
 ILTKDKVTVTDYDGKELIRDSYTAELDLSDIGKGYPFYMLKEIDEQPTVMRQLISTYADETGNVQVDPAI
 ITSIQEADRLYLILAAGTSYHAGFATKNMLEQLTDTPELVGVASENGYHMPLLSKKPMFILLSQSGETADSR
 QVLVKANAMGIPSLVTNVPGSTLSREATYTMLIHAGPEIAVASTKAYTAQIAALAFKAVGEANGKQEA
 LDFNLVHELSSLVAQSI EATLSEKDLVAEKVQALLATTRNAFYIGRGNDYVAMEAALKKEISYIQCEGFA
 AGELKHGTISLIEEDTPVIALISSSQLVASHTRGNIQEVAARGAHVLTVVEEGLDREGDDIIVNKVHPFLA
 PIAMVIPTQLIAYYASLQRLDVKPRNLAKAVTVE

SEQ ID NO: 95 secuencia de polinucleótidos que codifica GAS 529

ATGTGTGGAATTGTTGGAGTTGTTGGAAATCGCAATGCAACGGATATTTTAATGCAAGGCCCTTGAAAAGCT
 TGAATACCGGGTTATGATTGAGCAGGAATTTTGTGGCTAATGCCAATCAAACAACTTGATTAAATCAG

TGGGGCGGATTGCTGATTTGCGTGCCAAAGATTGGCATTGATGTTGCTGGTTCAACAGGGATTGGTCACACC
 CGTTGGGCAACGCATGGCCAATCAACAGAGGATAATGCCCATCCTCACACGTCACAACTGGACGTTTTGT
 ACTTGTTCATAATGGTGTGATTGAAAATTACCTTCACATTAAACAGAGTTCTTAGCTGGACATGATTTTA
 AGGGGCAGACAGATACTGAGATTGCAGTACACTTGATTGGAAAATTTGTGGAAGAAGACAAGTTGTCAGTA
 CTGGAAGCTTTTAAAAAATCTTTAAGCATTATTGAAGGTTCTTACGCTTTGTCATTAAATGGATAGCCAAGC
 AACTGATACTATTTATGTGGCTAAAAACAAGTCTCCATTGTTGATTGGACTTGGTGAAGGTTACAACATGG
 TTTGTTTCAGATGCCATGGCCATGATTCGTGAAACCAAGTGAATTTATGGAAATTCATGATAAGGAGCTAGTT
 ATTTTAAACCAAGATAAGGTAACGTGTTACAGACTACGATGGTAAAGAGCTGATACGAGATTCTTACACTGC
 TGAATTAGACTTATCTGATATTGGCAAGGGACTTATCCTTTCTATATGCTGAAAGAAATTGATGAGCAAC
 CAACCGTAATGCGTCAATTAATTTCAACTTATGCAGATGAAACTGGTAACGTACAGGTTGATCCGGCTATC
 ATTACCTCTATCCAAGAGGCTGACCGCTCTTATATTTTAGCGGCAGGGACTTCTTACCATGCTGGTTTTC
 AACAAAAATATGCTTGAGCAATTGACAGATACACCAGTTGAGTTGGGCGTGGCTTCTGAGTGGGGTTACC
 ACATGCCTCTGCTTAGCAAGAAACCAATGTTTATCTACTAAGCCAATCAGGAGAAACCGCAGATAGTCGT
 CAAGTTTGTAGTAAAGGCAAAATGCTATGGGCATTCCGAGTTTGACAGTAACCTAACGTTCCAGGATCAACCTT
 ATCACGTGAAGCAACATACACCATGTTGATTGATGCTGGACCTGAAATGCTGTTGCGTCTACAAAAGCTT
 ACACCTGCACAAATGCTGCCCTTGGCTTTTGGCTAAGGCAGTTGGTGAGGCAATGGTAAGCAAGAAGCT
 CTTGACTTTAACTTGGTACATGAGTTGTCATTGGTTGCCCAATCTATGAGGCGACTTTGTCTGAAAAAGA
 TCTCGTGGCAGAAAGGTTCAAGCTTTGCTAGCTACTACTCGTAATGCTTTTACATCGGGCGTGGCAATG
 ATTATTACGTTGCGATGGAAGCTGCTTTGAAATTAAGAGAGATTCTTATATTCAATGCGAAGGCTTTGCG
 GCTGGTGAATTGAAACATGGAACCATTTTCATTAATTGAGGAGGACACGCCAGTAATCGCTTTAATATCGTC
 TAGTCAGTTGGTTGCCCTCTCATACGCGTGGTAATATTCAAGAAGTTGCTGCCCGTGGGGCTCATGTTTTAA
 CAGTTGTGGAAGAAGGCTTGACCGTGAGGGAGATGACATTATTGTCAATAAGGTTTCATCTTTCTTAGCC
 CCGATTGCTATGGTCATTCCAACCTCAACTGATTGCTTACTACGCTTCATTACAACGTGGACTTGATGTTGA
 TAAGCCACGTAATTTGGCTAAAGCTGTAACAGTAGAATAA

SEQ ID NO: 96 secuencia de aminoácidos que comprende la GAS 45

VTFMKKSKWLAAVSVAILSVSALAACGNKNASGGSEATKTYKYVFVNDPKSLDYILTNGGGTTDVIITQMVD
 GLENDHEYGNLVP SLAKDWKVS KDGLTYTYTLRDGVS WYTADGEEYAPVTAEDFVTGLKHA VDDKSDALYV
 VEDSIK NLKAYQNGEVDFKEVG VKALDDKT VQYTLNKPESYWN SKTYSVLFPVNAKFLKSKGKDFGTTDP
 SSILVNGAYFLSAFTSKSSMEFHKNENYWDANKV GIESVKLTYS DGSDPGSFYKNPDKGEFSVARLYPNDP
 TYKSAKKNYADNITYGMLTGDIRHLTWNLNRTSFKNTKKDPAQQDAGKKALNNKDFRQAIQFAFDRASFQA
 QTAGQDAKTALRNMLVPPTFVTIGESDFGSEVEKEMAKLGDEWKDVNLADAQDGFYNPEKAKAEFAKAKE
 ALTAEGVTFFPVQLDYPVDQANAATVQEAQSFKQSVEASLGKENVIVNULETETSTHEAQGFYAETPEQQDY
 DIISWNGPDYQDPRTYLDIMSPVGGG SVIQKLGIKAGQNKDVVAAAGLDYQTLLEDEAAAITDDNDARYK
 AYAKAQAYLTDNAVDIPVVALGGTPRVTKAVPFSGGFSWAGSKGPLAYKGMKLQDKPVTVKQYEKAKEKWM
 KAKAKSNAKYAEKLADIIVEK

SEQ ID NO: 97 secuencia de polinucleótidos que codifica GAS 45

GTGACTTTTATGAAGAAAAGTAAATGGTTGGCAGCTGTAAGTGTTCGATCTTGTTCAGTATCCGCTTTGGC
AGCTTGTGGTAATAAAATGCTTCAGGTGGCTCAGAAGCTACAAAAACCTACAAGTACGTTTTTGTAAACG
ATCCAAAATCATTTGGATTATATTTTGTACTAATGGCGGTGGAACGACTGATGTGATAACACAAAATGGTTGAT
GGTCTTTTGGAAAACGATGAGTATGGTAATTTAGTACCATCACTTGTCTAAAGATTGGAAGGTTTCAAAAGA
CGGTCTGACTTATACTTATCTCTTCGCGATGGTGTCTCTTGGTATACGGCTGATGGTGAAGAATATGCCC
CAGTAACAGCAGAAGATTTTGTGACTGGTTTGAAGCAGCGGTTGACGATAAATCAGATGCTCTTTACGTT
GTTGAAGATTCAATAAAAACTTAAAGGCTTACCAAAATGGTGAAGTAGATTTTAAAGAAGTTGGTGTCAA
AGCCCTTGACGATAAACTGTTTCAGTATACTTTGAACAAGCCTGAAAGCTACTGGAATTCAAAAACAATT
ATAGTGTGCTTTTCCAGTTAATGCGAAATTTTGAAGTCAAAAGGTAAAGATTTTGGTACAACCGATCCA
TCATCAATCCTTGTAAATGGTGTCTTCTTTCGAGCGCTTCACCTCAAAATCATCTATGGAATTCATATA
AATGAAAACTACTGGGATGCTAAGAATGTTGGGATAGAATCTGTTAAATTGACTTAACTCAGATGGTTTCAG
ACCCAGGTTCTTCTACAAGAACTTTGACAAGGGTGAGTTTCAGCGTTGCACGACTTTACCCAAATGACCCCT
ACCTACAAATCAGCTAAGAAAACTATGCTGATAACATTTACTTACGGAATGTTGACTGGAGATATCCGTCA
TTTAACATGGAAATTTGAACCGTACTTCTTTCAAAAACACTAAGAAAGACCCCTGCACAACAAGATGCCGGTA
AGAAAGCTCTTAACAACAAGGATTTTCGTCAGCTATTTCAGTTTGTCTTTGACCGAGCGCTATTCCAAGCA
CAAACCTGCAGGTCAAGATGCCAAAAACAAAGCCTTACGTAACATGCTTGTCCCACCAACATTTGTGACCAT
TGGAGAAAGTGATTTTGGTTTCAGAAGTTGAAAAGGAAATGGCAAACTTGGTGTGAATGGAAAGACGTTA
ACTTAGCTGATGCTCAAGATGGTTTCTATAATCTGAAAAAGCAAAAGCTGAGTTTGCAAAAGCCAAAGAA

SECuencia de Datos

GCTTTAACAGCTGAAGGTGTAACCTTCCAGTTCAATTAGATTACCCTGTTGACCAAGCAAAAGCAGCAAC
TGTTTCAGGAAGCCAGTCTTTCAAACAATCTGTTGAAGCATCTCTTGGTAAAGAGAATGTCATTGTCAATG
TTCTTGAAACAGAAACATCAACTCACGAAGCCCAAGGCTTCTATGCTGAGACCCAGAAACAAGACTAC
GATATCATTTTCATCATGGTGGGACCAGACTATCAAGATCCACGGACCTACCTTGACATCATGAGTCCAGT
AGGTGGTGGATCTGTTATCCAAAACTTGAATCAAAGCAGGTCAAATAAGGATGTTTGGCAGCTGCAG
GCCTTGATACCTTACCAAACTCTTCTTGATGAAGCAGCAGCAATTACAGACGACAACGATGCCGCTATAAA
GCTTACGCAAAAGCACAAGCCTACCTTACAGATAATGCCGCTAGATATTCAGATTGTGGCATTGGGTGGCAC
TCCACGAGTTACTAAAGCCGTTCCATTAGCGGGGGCTTCTCTTGGGCAGGCTCTAAAGGTCCTCTAGCAT
ATAAAGGAATGAACTTCAAGACAAACCTGTACAGTAAACAATACGAAAAAGCAAAAGAAAAATGGATG
AAAGCAAAAGGCTAAGTCAAATGCAAAATATGCTGAGAAGTTAGCTGATCACGTTGAAAAA

SEQ ID NO: 98 secuencia de aminoácidos que comprende una secuencia líder N-terminal de GAS 45
VTFMKKSKWLAAVSVAILSVSALAA

SEQ ID NO: 99 secuencia de aminoácidos que comprende un fragmento de GAS 45 en el que la secuencia líder
N-terminal está eliminada

CGNKNASGSGSEATKTYKYVFNPKSLDYILTNNGGTTDVIQMDGLLENDEYGNLVP SLAKDWKVS KDG
LTYTYTLRDGVSWYTADGEEYAPVTAEDFVTGLKHAVDDKSDALYVVEDSIKNLKAYQNGEVDPFKEVGVA
LDDKTQYQTLNKPESYWSKTTYSVLFVNAKFLKSKGKDFGTTDPSSILVNGAYFLSAFTSKSSMEFHKN
ENYWDARNVIGIESVKLTYS DSGDPGSFYKNFDKGEFSVARLYPNDPTYKSAKKNYADNITYGMLTGDIRHL
TWNLNRTSFKNTKKDPAQQDAGKALNNKDFRQAIQFAFDRASFQAQTAGQDAKTKALRNMLVPPTFTVIG
ESDFGSEVEKEMAKLGDEWKDVLNADAQDGFYNPEKAKAEPAKAKEALTAEGVTPFPVQLDYPVDQANAATV
QEAQSFQKQSVESLGKENVIVNLETETSTHEAQGFYAETPEQQDYDISSWNGPDYQDPRTYLDIMSPVG
GGSVIQKLGIKAGQNKDVVAAAGLDYQTLLEAAAITDDNDARYKAYAKAQAYLTDNAVDIPVVALGGTF
RVTKAVPFGSGFSWAGSKGPLAYKGMKLQDKPVTVKQYKAKKWKAKAKSNAKYAEKLADHVEK

SEQ ID NO: 100 secuencia de aminoácidos que comprende GAS 95

MKIGKKIVLMFTAIVLTTVLALGVYLT SAYTFSTGELSKTFKDFSTSSNKSDAIKQTRAPSIILMGVDTGS
SERASKWEGNSDSMILVTVPKTKKTTMTSLERDTLTTLSPKNNEMNGVEAKLNAAYAAGGAQMAIMTVQ
DLLNITIDNYVQINMQGLIDLNVAVGGITVTNEFDFFISIAENEPEYQATVAPGTHKINGEQALVYARMRY
DDPEGDYGRQKRQREVIQKVLKILALDISSYRKILSAVSSNMQTNIEISSRTIPSLGLYRDALRTIKTY
QLKGEDATLSDGGSYQIVTSNHLLEIQNRIRTELGLHKVNQLKTNAVTYENLYGSTKQTVNMNYDSSGQA
PSYSDSHSSYANYSSGVDTGQSASTDQDSTASSHRPATPSSSSDALADESSSSGSGSLVPPANINPQT

SEQ ID NO: 101 secuencia de polinucleótidos que codifica GAS 95

5 ATGAAAATFGAAAAAAATAGTTTAAATGTTTACAGCTATTGTGTTAACTGTCTTGGCATTAGGTGT
CTATCTAACTAGTGTCTTATACCTTCTCAACAGGAGAATTATCAAAGACCTTTAAAGATTTCGACATCTT
CAAACAAAAGTGATGCCATTAAACAAACAAGAGCTTTTCTATCTTGTGATGGGTGTTGATACAGGCTCT
TCAGAGCGTGCCTCCAAGTGGGAAGGAAACAGTGATTCGATGATTTTGGTTACGGTTAATCCAAAGACCAA
GAAAACAACTATGACTAGTTTAGAACGAGATACCTTAACCACGTTATCTGGACCCAAAAATAATGAAATGA
ATGGTGTGGAAGCTAAGCTTAACGCTGCTTATGCAGCAGGTGGCGCTCAGATGGCTATTATGACCGTGCAA
10 GATCTTTTGAATATCACCATTGATAACTATGTTCAAATTAATATGCAAGGCCTTATTGATCTTGTGAATGC
AGTTGGAGGGATTACAGTTACAAATGAGTTTGATTTTCTATCTCGATTGCTGAAAACGAACCTGAATATC
AAGCTACTGTTGCGCCTGGAACACACAAAAATTAACGGTGAACAAGCTTTGGTTTATGCTCGTATGCGTTAT
GATGATCCTGAGGGAGATTATGGTTCGACAAAAGCGTCAACGTGAAGTCATTCAAAGGTATTGAAAAAAT
CCTTGCTCTTGATAGCATTAGCTCTTATCGGAAGATTTTATCTGCTGTAAGTAGTAATATGCAAACGAATA
15 TCGAAATCTCTTCTCGCACTATCCCTAGTCTATTAGGTTATCGTGACGCACTTAGAACTATTAAGACTTAT
CAACTAAAAGGAGAAGATGCCACTTTATCAGATGGTGGATCATAACCAATTGTTACCTCTAATCATTTGTT
AGAAATCCAAAATCGTATCCGAACAGAATTAGGACTTCATAAGGTTAATCAATTAACCAAAATGCTACTG
TTTATGAAAATTTGTATGGGTCAACTAAGTCTCAGACAGTAAACAACAATATGACTCTTCAGGCCAGGCT
CCATCTTATCTGATAGTCATAGCTCTTACGCTAATTATTCAAGTGGAGTAGATACCGCCAGAGTGCTAG
20 TACAGACCAGGACTCTACTGCTTCAAGCCATAGGCCAGCTACGCCGTCTCTTCATCAGATGCTTTAGCAG
CTGATGAGTCTAGCTCATCAGGGTCTGGATCATTAGTTCCCTCTGCTAATATCAACCCTCAGACCTAA

SEQ ID NO: 102 secuencia de aminoácidos que comprende la secuencia líder N-terminal de GAS 95
MKIGKKIVLMFTAIVLTTLALGVYLTSAITFS

25 **SEQ ID NO: 103** secuencia de aminoácidos que comprende un fragmento de GAS 95 en el que la secuencia
líder N-terminal está eliminada.

30 TGELSKTFKDFSTSSNKSDAIKQTRAFSILLMGVDTGSSERASKWEGNSDSMILVTNPKTKTMTSLER
DTLTTLSGPKNNEMNGVEAKLNAAYAAGGAQMAIMTVQDLLNITIDNYVQINMQGLIDLNAVGGITVTNE
FDFPISIAENEPEYQATVAPGTHKINGEQALVYARMRYDDPEGDYGRQKRQREVIQKVLKKILALDSISSY
RKILSAVSSNMQTNIEISSRTIPSLGTYDALRTIKTYQLKGEDATLSDGGSYQIVTSNHLLEIQNRIRTE
LGLHKVNQLKTNATVYENLYGSTKSQTVNMNYDSSGQAPSYSDSHSSYANYSSGVDTGQSASTDQDSTASS
HRPATPSSSSDALAADESSSSSGSLVPPANINPQT

35 **SEQ ID NO: 104** secuencia de aminoácidos que comprende GAS 193

40 MKKRKLLAVTLLSTILLNSAVPLVVDTSLRNSTSSSTDQPTTADTDDESETPKKDKKSKETASQHDTQK
DHKPSHTHTPTPSNDTKQTDQASSEATDKPNKDKNDTKQPDSSDQSTPSPKDQSSQKESQNKDGRPTPSPD
QQKDQTPDKTPEKSADKTPEKGPEKATDKTPEPNRDAPKPIQPPLAAAPVFIWRESKDLKLPSSRSS
AAVVRHWTGDSAYTHNLLSRRYGITAEQLDGFNLGLIHYDKERLNGKRLLLEWEKLTGLDVRAIVAIAMAE
SSLGTQGVAKKEGANMFGYGAFDFNPNNAKKYSDEVAIRHVMEDTIIANKNQTFERQDLKAKKWSLGQLDT
LIDGGVYFTDTSGSGQRRADIMTKLDQWIDDHGSTPEIPEHLKITSGTQFSEVPVGYKRSQFQNVLTYSKSE
45 TYSFGQCTWYAYNRVKELGYQVDYMGNGGDWQRKPGFVTHKPKVGYVVSFAPGQAGADATYGHVAVVEQ
IKEDGSILISESNVMGLGTISYRTFTAEQASLLTYVVGDKLPRP

50

55

60

65

SEQ ID NO: 105 secuencia de polinucleótidos que codifica GAS 193

5 ATGAAGAAAAGGAAATTGTTAGCAGTAACACTATTAAGTACCATACTCTTAAACAGTGCAGTGCCATTAGT
TGTTGCTGATACCTCCTTGGCGTAATAGCACATCATCCACTGATCAGCCTACTACAGCAGATACTGATACGG
10 ATGACGAGAGTGAAACACCAAAAAAGACAAAAAAGCAAGGAAACAGCGTCGCAGCAGACACCAAAAA
GACCATAAGCCATCACACACTCACCCAACCCCCCTTCAAATGATACTAAGCAGACCGATCAGGCATCATC
TGAAGCTACTGACAAACCAAATAAAGACAAAAACGACACCAAGCAACCAGACAGCAGTGATCAATCCACCC
CATCTCCCAAAGACCAGTCTGCTCAAAAAAGAGTCACAAAAACAAAGACGGCCGACCTACCCCATCACCTGAT
15 CAGCAAAAAAGATCAGACACCTGATAAAAAACCCAGAAAAATCAGCTGATAAAACCCCTGAAAAAGGACCAGA
AAAAGCAACTGATAAAACACCAGAGCCAAATCGTGACGCTCCAAAACCCATCCAACCTCCTTTAGCAGCTG
CTCCTGTCTTTATACCTTTGGAGAGAAAGTGACAAAGACCTGAGCAAGCTAAAACCAAGCAGTCGCTCATCA
CTGGCTTACGTGAGACACTGGACAGGTGACTCTGCCTACACTCAACCTGTTGTACGCCGCTTATGGGAT
TACTGCTGAACAGCTAGATGGTTTTTTGAACAGTCTAGGTATTCACTATGATAAAGAAGCCTTAAACGGAA
20 AGCGTTTATTAGAATGGGAAAACTAACAGGACTAGACGTTTCGAGCTATCGTAGCTATTGCAATGGCAGAA
AGCTCACTAGGTACTCAGGGAGTTGCTAAAGAAAAAGGAGCCAAATATGTTTGGTTATGGCGCCTTTGACTT
CAACCCAAACAATGCCAAAAATACAGCGATGAGGTTGCTATTCTGTCACATGGTAGAAGACACCATCATTG
CCAACAAAAACCAACCTTTGAAAGACAAGACCTCAAAGCAAAAAAATGGTCACTAGGCCAGTTGGATACC
TTGATTGATGGTGGGGTTTACTTTACAGATACAAGTGGCAGTGGGCAAGACGAGCAGATATCATGACCAA
25 ACTGATGACCAATGGATAGATGATGATGGAAGCACACCTGAGATTCCAGAACATCTCAAGATAAATTCCGGGA
CACAAATTTAGCGAAGTGCCCGTAGGTTATAAAAGAAGTCAGCCACAAAACGTTTTGACCTACAAGTCAGAG
ACCTACAGCTTTGGCCAAATGCACTTGGTACGCCTATAATCGTGTCAAAGAGCTAGGTTATCAAGTCGACAG
GTACATGGGTAAACGGTGGCGACTGGCAGCGCAAGCCAGGTTTTGTGACCACCCATAAACCTAAAGTGGGCT
ATGTCGTCTCATTTGCACCAGGCCAAGCAGGAGCAGATGCAACCTATGGTCACGTTGCTGTTGTAGAGCAA
30 ATCAAAGAAGATGGTTCATCTTAATTTAGAGTCAAATGTTATGGGACTAGGCACCATTTCCTATCGGAC
GTTACAGCTGAGCAGGCTAGTTTGTGACCTATGTCGTAGGGGACAAACTCCCAAGACCATAA

SEQ ID NO: 106 secuencia de aminoácidos que comprende GAS 137

30 MSDKHINLVIVTGMMSGAGKTVAIQSFEDLGYFTIDNMPPALVVPKFLELIEQTNNRRLVALVVDMRSLFFK
EINSTLDSIESNPSIDFRILFLDATDGLVSRKETRSHPLAADGRVLDGIRLERELLSPLKSMSQHVVD
TTKLTPRQLRKTISDQFSEGSNQASFRIEVMSFGFKYGLPLDADLVDPVRLPNPYQVELREKTGLDEDV
35 FNYVMSPPESEVFYKHLNLIVPILPAYQKEGKSVLTVAIGCTGGQHRVFAHCLAESLATDWSVNESH
DQNRKETVNR

SEQ ID NO: 107 secuencia de polinucleótidos que codifica GAS 137

40 ATGTCAGACAAACACATTAATTTAGTTATTGTGACAGGAATGAGCGGCGCTGGAAAAACAGTTGCCATTCA
GTCTTTTGAGGATCTAGGCTACTTTACCATTGATAATATGCCCCAGCCTTGGTTCCAAAATTTTGAAT
TAATTTGAACAAACCAATGAAAATCGTAGGGTGGCTTTGGTTGTCGATATGAGAAGTCGTTTGTTCGAA
45 GAAATTAATTCACCTTAGATAGTATTGAAAGCAATCCTAGCATTGATTTTCGGATTCTTTTTTTGGATGC
AACGGATGGAGAATTGGTGTACGCTATAAAGAAACCAGACGGAGCCACCCCTTTGGCTGCGGACGGTCGTG
TGCTTGATGGTATTTCGATTGGAAGAGAACTCCTATCTCCTTTGAAAAGCATGAGCCAACATGTGGTGGAT
ACAACAAAATTGACCCCTAGACAATTGCGTAAACCAATTTAGACCAGTPTTCTGAAGGGTCTAATCAAGC
50 CTCTTTCCGTATTGAAGTGATGAGCTTTGGGTTCAAATATGGTCTTCCCTTTGGATGCGGATTGGTTTTTG
ATGTGCGTTTTCTACCCAATCCTTATTATCAGGTAGAGCTTCGTGAAAAACAGGACTAGATGAGGACGTT
TTTAATTATGTGATGTCTCACCCAGAATCAGAGGTGTTTTACAAGCATTTGTTAAACCTTATTGTCCCTAT
CTTACCGGCTTACCAAAAAGAAGGGAAGTCTGTCTTGACGGTGGCTATTGGCTGCACAGGAGGCCAACACC
55 GCAGCGTTGCCCTTGCCCATTTGCTTGGCAGAAAGTCTGGCAACAGATTGGTCGGTTAATGAAAGCCATCGT
GATCAAAATCGTCGTAAGGAAACGGTGAATCGTTCATGA

SEQ ID NO: 108 secuencia de aminoácidos que comprende GAS 84

60 MIIKKRTVAILAIASSFFLVACQATKSLKSGDANGVYQKQKSITVGFNDTFVPMGYKDESGRCKGFDIDLA
KEVPHQYGLKVNFPQAINWDMKEAELNNGKIDVIWNGYSITKERQDKVAFDTSYMRNEQIIIVVKRSDIKTI
SDMKHKVLGAQSASSGYDSLRLTPKLLKDFIKNKDANQYETFTQAFIDLKSDRIDGILIDKVYANYLLAKE
65 GQLENYRMIPTTFENEAFSVGLRKEDKTLQAKINRAFRVLYQNGKFQAISEKWFDDVATANIKS

SEQ ID NO: 109 secuencia de polinucleótidos que codifica GAS 84

5 ATGATTATAAAAAAAGAACCCTAGCAATTTTAGCCATAGCTAGTAGCTTTTCTTGGTAGCTTGTCAGC
TACTAAAAGTCTTAAATCAGGAGATGCTTGGGGAGTTTACCAAAAGCAAAAAAGTATTACAGTTGGTTTTG
ACAATACGTTTGTTCCTATGGGCTATAAGGATGAAAGCGGCAGATGCAAGGTTTGTATATTGATTGGCT
10 AAAGAAGTTTTTACCAATATGGACTCAAGGTTAACTTTCAAGCTATTAATFGGGACATGAAAGAAGCAGA
ACTAAACAATGGTAAATTTGATGTAATCTGGAATGGTTATTCAATAACTAAGGAGCGTCAGGATAAGGTTG
CCTTTACTGATTCTTACATGAGAAATGAACAAATTATGTTGTCAAAAAAGATCTGATATTAACAATA
TCAGATATGAAACATAAAGTGTAGGAGCACAAATCAGCTTCATCAGGTTATGACTCCTTGTTAAGAACTCC
TAAACTGCTGAAAGATTTTATTAAAAATAAAGACGCTAATCAATATGAAACCTTTACACAAGCTTTTATTG
ATTTAAATCAGATCGTATCGATGGAATATTGATTGACAAAGTATATGCCAATTACTATTTAGCAAAAGAA
GGGCAATTAGAGAATTATCGGATGATCCCAACGACCTTTGAAATGAAGCATTTTCGTTGGACTTAGAAA
15 AGAAGACAAAACGTTGCAAGCAAAAATTAATCGTGCTTTTCAGGGTGCTTTATCAAAATGGCAAAATTTCAAG
CTATTTCTGAGAAATGGTTTGGAGATGATGTTGCCACTGCCAATATTAAATCTTAA

SEQ ID NO: 110 secuencia de aminoácidos que comprende la secuencia líder N-terminal de GAS 84
MIKKRTVAILAIASSFLVA

SEQ ID NO: 111 secuencia de aminoácidos que comprende un fragmento de GAS 84 en el que la secuencia
líder N-terminal está eliminada

25 CQATKSLKSGDAWGVYQKQKSITVGFNTFVPMGYKDESGRCKGFDIDLAKVFPHQYGLKVNQAINWDMK
EAEIWNNGKIDVIWNGYSITKERQDKVAFTDSYMRNEQIIIVVKRSDIKTISDMKHKVLGAQSASSGYDSLL
RTPKLLKDFIKNKDANQYETPTQAFIDLKSDRIDGILIDKVYANYLLAKEGQLENYRMIPTTFENEAFSVG
LRKEDKTLQAKINRAFRVLYQNGKFQAISEKWFQDDVATANIKS

SEQ ID NO: 112 secuencia de aminoácidos que comprende GAS 384

35 MKTLAFDTSNKTLSLAILDDETLADMTLNIQKKHSVSLMPAIDFLMTCTDLKPQDLERIVVAKGPGSYTG
LRVAVATAKTALAYSLNIALVGISSLYALAASTCKQYPNTLVVPLIDARRQNAVYVGYRQKSVMPQAHASL
EVIIEQLVEEGQLIFVGETAPFAEKIQKLPQAILLPTLPSAYECGLLGQSLAPENVDAFVFPQYLKRVEAE
ENWLKDNEIKDDSHYVKRI

SEQ ID NO: 113 secuencia de polinucleótidos que codifica GAS 384

40 ATGAAGACACTTGCATTTGATACCTCAAATAAAACCTTGTCCTTGCTATACCTTGATGATGAGACACTTCT
AGCAGATATGACCTTAAACATTCAGAAAAACATAGTGTTAGCCTTATGCCTGCTATTGATTTTTTGATGA
CTTGACTGATCTTAAACCTCAAGATTTAGAAAGAATAGTGGTTGCAAAAGGCCCTGGATCTTACACAGGT
45 TTACGAGTGGCAGTTGCTACTGCAAAACGTTAGCGTACAGTTTAAATATGCAATGGTTCGGGATTTTCGAG
TCTATATGCTTTGGCTGCGTCTACTTGTAAACAGTATCCAAATACCTTGGTGGTGCCATTGATTGATGCTA
GAAGGCAAAATGCGTATGTAGGTTATTATCGGCAAGGAAATCAGTGTGCCACAAGCCCATGCTTCACTA
GAAGTTATTATAGAACAATTAGTAGAAGAAGGACAGCTGATTTTTGTTGGGGAGACTGCTCCTTTTGCTGA
50 GAAATTCAAAAGAACTACCTCAGGCGATACTACTTCCAACCTTCTCTGCTTACGAATGTGGTCTTT
TGGGGCAAAGTTTGGCACCAGAAATGTAGACGCCTTTGTCCCTCAATATCTCAAGAGAGTGGAAGCTGAA
GAAACTGGCTCAAAGATAATGAGATAAAAGATGATAGTCACTACGTTAAGCGAATCTAA

SEQ ID NO: 114 secuencia de aminoácidos que comprende GAS 202

55 MLKRLWLILGPLLIAFVLVVITIFSFPTQLDHSIAQEKANAVAITDSSFKNGLIKRQALSDETCRFVPPFG
SSEWSRMDSMHPSVLAERYKRSYRPFLIGKRGASLSHYYGIIQOITNEMQKKKALFVVSQWFTAQGINPS
AVQMYLSNTQVIEFLKARTDKESQFAAKRLELNPGVSKSNLLKKVSKGKSLSRDLRAILKQHQVALRE
60 ESLFSFLGKSTNYEKRIIPRVKGLPKVFSYKQLNALATKRGQLATTNNRFGIKNTFYRKRIAPKYNLYKNF
QVNSYLASPEYNDFQLLLSEFAKRKTDVLFVITPVNKAWADYTGILNQDKYQAAVRKIKFQLSQGFHRIA
DFSKDGGESYFMQDTIHLGWNGWLAFDKKVQPFLETKQVPVNYKMPYFYSKIWANRKLQ

SEQ ID NO: 115 secuencia de polinucleótidos que codifica GAS 202

5 ATGCTTAAGAGACTCTGGTTAATTCTAGGTCCCTCTTCTTATTGCCTTTGTTTTAGTAGTGATTACTATTTT
TAGTTTTCTACACAACTTGATCATTCCATAGCTCAGGAAAAAGCAAATGCCGTTGCGATCACAGATAGTT
CTTTTAAAAATGGTTTGATTAAAAGACAAGCTTTATCAGATGAGACTTGTCGTTTTGTGCCTTTTTTTGGT
TCTAGCGAATGGAGTCGAATGGATAGTATGCACCCCTTCGGTGCTTGCGAGCGCTACAAGCGGAGCTATAG
10 ACCATTTTTAAATGGTAAGAGAGGATCAGCATCTTTGTCGCATTATTATGGTATACAACAAATTACCAATG
AAATGCAAAAGAAAAAGCCATCTTTGTAGTATCTCCTCAATGGTTTACTGCTCAAGGGATTAATCCTAGT
GCGGTTGAGATGACTTGTCTAACACTCAAGTGATTGAATTTTTACTAAAAGCTAGAACTGATAAAGAATC
ACAGTTTGCAGCAAAGCGTTTGCTTGAGCTTAACCCCTGGTGTGTCTAAATCAAACCTTATTGAAAAAAGTAA
GTAAGGGTAAGTCTCTTAGTCGGTTAGACAGAGCTATTTTGAAATGTCAACATCAAGTAGCATTGAGAGAA
GAGTCCCTTTTTAGTTTTTTAGGCAAACTTACTAATATGAAAAAGAATTTTGCCCTCGCGTTAAGGGATT
15 ACCTAAAGTATTTTCGTATAAACAATTGAATGCATTAGCAACTAAGAGAGGCCAATTAGCAACAACCAACA
ACGTTTTTGGGATTAATAAATACATTTTATCGTAAACGAATAGCACCTAAATACAATCTTTATAAGAAATTC
CAAGTTAATTATAGTTTACCTGGCGTACCAGAAATACAATGATTTTTCAGCTTTTATATATCAGAATTTGCTAA
ACGAAAAACAGATGTACTCTTTGTTATAACTCCTGTTAATAAAGCTTGGGCGGATTATACCGGCTTAAATC
AAGATAAGTATCAAGCGGCAGTTTCGTAAATAAAATTCAGTTAAAGTCACAAGGATTTTCATCGCATTGCT
20 GACTTCTCAAAGATGGTGGTGAGTCCTACTTTATGCAAGATACCATCCATCTCGGTGGGAATGGCTGGTT
AGCTTTTGATAAGAAAGTGCAACCATTCTAGAAAAGCAAGCAGCCAGTGCCCACTATAAAATGAACCCCTT
ATTTTTATAGTAAAATTTGGGCAAAATAGGAAAGACTTGCAATAG

SEQ ID NO: 116 secuencia de aminoácidos que comprende GAS 057

25 MEKKQRFSLRKYKSGTFSVLIGSVFLVMTTVAADELSTMSEPTITNHAQQQAQHL/TNTELSAESKSQDT
SQITLKTNREKEQSQDLVSEPTTTELDADTAASMAN TGSDATQKSASLPPVNTDVHDWVKTKGAWDKGYKG
30 QGKVVAVIDTGIDPAHQSMRISDVSTAKVSKEDMLARQKAAGINYGSWINDKVVFAHNYVENSNDNIKENQ
FEDFDEDWENFEFDAAEPKAIKKHKIYRPQSTQAPKETVIKTEETDGSHDIDWTQTDDDTKYESHGMHVT
GIVAGNSKEAAATGERFLGIAPEAQVMFMRVVFANDIMGSAESLFKAIEDIAVALGADVINSLSLTANGAQL
SGSKPLMEAIEKAKKAGVSVVVAAGNERVYGS DHDDPLATNP DYGLVGS PSTGRPTPTSVAAINSKWVIQRL
MTVKELNLRADLNHGKAIYSESVD FKDIKDSLGYDKSHQFAYVKESTDAGYNAQDVKGKIALIERDPNKTY
35 DEMIALAKKHGALGVLI FNNKPGQSNRSMRLTANGMGIPSAFISHEFGKAMSQNLNGNGTGSLEPDSVVS KA
PSQKGNEMNHFSNWGLTSDGYLKPDIAPGGDIYSTYNDNHYGSQTGTSMAS PQIAGASLVVKQYLEKTQP
NLPKEKIADIVKNLLMSNAQIHVNPETKTTTSPRQQAGLLNIDGAVTSGLYVTGKDNYSISLGNITDTM
TFDVTVHNLSNKDKTLRYDTELLTDHVD PQKGRFTLTSHSLKTYQGGEVTVPANGKVTVRVTMDVSQPTKE
LTKQMPNGYYLEGFVRFRDSQDDQLNRVNI PFVGFPGKQFENLAVAEE SIYRLKSQKGTGFYFDES GPKDDI
40 YVGKHFTGLVTLGSETNVSTKTI SDNGLHTLGT FKNADGKFILEKNAQGNPVLAISPNGDNNQDFAAPKGV
FLRKYQGLKASVYHASDKEHKNPLWVSPESFKGDKNFNSDIRFAKSTTLLGTAFSGKSLTGAELPDGHYHY
VVSYPDVVGAQRQEMTFDMILDRQKPVLSQATFDPETNRFKPEPLKDRGLAGVRKDSVFYLERKDNKPYT
VTINDSYKYVSVEDNKTFVERQADGSFILPLDKALGDFYMYVEDFAGNVAIAKLGDHLPQTLGKTPIKLK
LTDGNYQTKETLKDNL EMTQSDTGLVTNQAQLAVVHRNQPSQLTKMNQDFFISPNE DGNKDFVAFKGLKN
NVYNDLTVNVYAKDDHQKTPIWSSQAGASVSAIESTAWYGITARGSKVMPGDYQYVVYTYRDEHGKEHQKQ
45 YTISVNDKKPMITQGRFDTINGVDHFTPDKTKALDSSGIVREEVFYLA KKNRKFVDVTEGKDGITVSDNKV
YIPKNPDGSYITISKRDGVTLSDYYYLVEDRAGNVSFATLRDLKAVGKDKAVVNFGLDLPVPEDKQIVNFTY
LVRDADGKPIENLEYNNNSGNSLILPYGKYTVELLTYDTNAAKLES DKIVSFTLSADNNFQQVTFKITMLA
TSQITAHFDHLLPEGSRVSLKTAQDQLIPLEQSLVVPKAYGKTVQEGTYEVVVS LPKYRIEGNTKVN TLP
50 NEVHEL SLRLVKVG DASDSTGDHKVMSKNNSQALTASATPTKSTTSATAKAL PSTGERKMLKLRI VGLVLL
GLTCVFSRKKSTKD

SEQ ID NO: 117 secuencia de polinucleótidos que codifica GAS 057

GTGGAGAAAAAGCAACGTTTTTCCCTTAGAAAAATACAAATCAGGAACGTTTTCGGTCTTAATAGGAAGCGT
 TTTCTTGGTGATGACAACAACAGTAGCAGCAGATGAGCTAAGCACAATGAGCGAACCAACAATCAGGAATC
 ACGCTCAACAACAAGCGCAACATCTCACCATACAGAGTTGAGCTCAGCTGAATCAAAATCTCAAGACACA
 TCACAAATCACTCTCAAGACAAATCGTGAAAAAGAGCAATCACAAGATCTAGTCTCTGAGCCAACCACAAC
 TGAGCTAGCTGACACAGATGCAGCATCAATGGCTAATACAGGTTCTGATGCGACTCAAAAAAGCGCTTCTT
 TACCGCCAGTCAATACAGATGTTTACGATTTGGGTAAAAACCAAAGGAGCTTGGGACAAGGGATACAAAGGA
 CAAGGCAAGGTTGTGCGAGTTATTGACACAGGGATCGATCCGGCCCATCAAAGCATGCGCATCAGTGATGT
 ATCAACTGCTAAAGTAAAAATCAAAAGAGACATGCTAGCACGCCAAAAAGCCCGGTATTAAATTATGGGA
 GTTGGATAAAATGATAAAGTTGTTTTTGCACATAATTATGTGGAAAAATAGCGATAATATCAAAGAAAAATCAA
 TTCGAGGATTTTGATGAGGACTGGGAAAACTTTGAGTTTGATGCAGAGGCAGAGCCAAAAGCCATCAAAAA
 ACACAAGATCTATCGTCCCAATCAACCCAGGCACCGAAAGAACTGTTATCAAAACAGAAGAAACAGATG
 GTTCACATGATATTGACTGGACACAAACAGACGATGCACCAAAATACGAGTCAACACGGTATGCATGTGACA
 GGTAATTGTAGCCGGTAATAGCAAAGAACCCGCTGCTACTGGAGAACGCTTTTATAGGAATTGCACCAGAGGC
 CCAAGTCATGTTTCATGCGTGTTTTTGCCAACGACATCATGGGATCAGCTGAATCACTCTTTATCAAAGCTA
 TCGAAGATGCCGTGGCTTTAGGAGCAGATGTGATCAACCTGAGTCTTGGAAACCGCTAATGGGGCAGAGCTT
 AGTGGCAGCAAGCCTCTAATGGAAGCAATTGAAAAAGCTAAAAAAGCCGGTGTATCAGTTGTTGTAGCAGC
 AGGAAATGAGCGCGCTCTATGGATCTGACCATGATGATCCATTGGCGACAAATCCAGACTATGGTTTGGTCG
 GTTCTCCCTCAACAGGTCGAACACCAACATCAGTGGCAGCTATAAACAGTAAGTGGGTGATTCAACGTCTA
 ATGACGGTCAAAGAATTAGAAAACCGTGCCGATTTAACCATGGTAAAGCCATCTATTTCAGAGTCTGTGCA
 CTTTAAAGACATAAAAGATAGCCTAGGTTATGATAAATCGCATCAATTTGCTTATGTCAAAGAGTCAACTG
 ATGCGGGTTATAACGCACAAGACGTTAAAGGTAAATTTGCTTTAATTGAACGTGATCCCAATAAACCTAT
 GACGAAATGATTGCTTTGGCTAAGAAACATGGAGCTCTGGGAGTACTTATTTTAATAACAAGCCTGGTCA
 ATCAAAACCGCTCAATGCGTCTAACAGCTAATGGGATGGGGATACCATCTGCTTTTCATATCGCAGCAATTTG
 GTAAGGCCATGTCCCAATTAAATGGCAATGGTACAGGAAGTTTAGAGTTTGACAGTGTGGTCTCAAAAGCA
 CCGAGTCAAAGAGCAATGAATGAATCATTTTTCAAAATGGGGCCTAACCTCTGATGGCTATTTAAACC
 TGACATTACTGCACCAGGTGGCGATATCTATTCTACCTATAACGATAACCACTATGGTAGCCAAACAGGAA
 CAAGTATGGCCTCTCTCAGATTGCTGGCGCCAGCCTTTTGGTCAAACAATACCTAGAAAAGACTCAGCCA
 AACTTGCCAAAAGAAAAAATTGCTGATATCGTTAAGAACCTATTGATGAGCAATGCTCAAATTCATGTTAA
 TCCAGAGACAAAAACGACCCTCACCGCTCAGCAAGGGGCGAGGATTACTTAATATTGACGGAGCTGTCA
 CTAGCGGCCTTTTATGTGACAGGAAAAGACAACATAGGCAGTATATCATTAGGCAACATCACAGATACGATG
 ACGTTTGATGTGACTGTTTACAACCTAAGCAATAAAGACAAAAATACGTTATGACACAGAATTGCTAAC
 AGATCATGTAGACCCACAAAAGGGCCGCTTCACCTTTGACTTCTCACTCCTTAAAAACGTACCAAGGAGGAG
 AAGTTACAGTCCAGCCAATGGAAGAGTGAAGGGTTACCATGGATGTCTCAGAGTTCACAAAAGAG
 CTAACAAAAACAGATGCCAAATGGTTACTATCTAGAAGGTTTTGTCCGCTTTAGAGATAGTCAAGATGACCA
 ACTAAATAGAGTAAACATTCCTTTTGTGGTTTTAAAGGGCAATTTGAAAACCTTAGCAGTTGCAGAAGAGT
 CCATTTACAGATTAAAAATCTCAAGGCAAACTGGTTTTTACTTTGATGAATCAGGTCCAAAAGACGATATC
 TATGTGCGGTAAACACTTTACAGGACTTGTCACTCTTGGTTTCAAGAGCAATGTGTCAACCAAAACGATTTT
 TGACAATGGTCTACACACACTTGGCACCTTTAAAAATGCAGATGGCAAATTTATCTTAGAAAAAATGCCC
 AAGGAAACCTGTCTTAGCCATTTCTCCAAATGGTGACAACAACCAAGATTTTGCAGCCTTCAAAGGTGTT
 TTCTTGAGAAAAATATCAAGGCTTAAAGCAAGTGTCTACCATGCTAGTGACAAGGAACACAAAAATCCACT
 GTGGGTGAGCCAGAAAGCTTTAAAGGAGATAAAAACTTTAATAGTGACATTAGATTGCAAAATCAACGA
 CCCTGTTAGGCACAGCATTTTCTGGAAAATCGTTAACAGGAGCTGAATTACAGATGGGCATTATCATTAT
 GTGGTGTCTTATTACCCAGATGTGGTGGTCCAAACGTCAAGAAATGACATTTGACATGATTTTAGACCG
 ACAAAAACCGGTACTATACAAGCAACATTTGATCCTGAAACAAACCGATTCAAACAGAACCCCTAAAAG
 ACCGTGGATTAGCTGGTGTTCGCAAGACAGTGTCTTTTATCTAGAAAGAAAAGACAACAAGCCTTATACA
 GTTACGATAAACGATAGCTACAAATATGTCTCAGTAGAAGACAATAAAACATTTGTGGAGCGACAAGCTGA
 TGGCAGCTTTATCTTGCCGCTTGATAAAGCAAAATTAGGGGATTTCTATTACATGGTCGAGGATTTTGCAG
 GGAACGTGGCCATCGCTAAGTTAGGAGATCACTTACCACAAACATTAGGTAAACACCAATTTAAACTTAAG
 CTTACAGACGGTAATTATCAGACCAAAAGAAACGCTTAAAGATAATCTTGAAATGACACAGTCTGACACAGG
 TCTAGTCACAAATCAAGCCAGCTAGCAGTGGTGACCGCAATCAGCCGCAAAGCCAGCTAACAAAGATGA
 ATCAGGATTTCTTTATCTCACCACCAAGATGGGAATAAAGACTTTGTGGCCTTTAAAGGCTTGAAAAAT
 AACGTGTATAATGACTTAACGGTTAACGTATACGCTAAAGATGACCACCAAAAACAAACCCCTATCTGGTC

5 TAGTCAAGCAGGCGCTAGTGTATCCGCTATTGAAAGTACAGCCTGGTATGGCATAACAGCCCGAGGAAGCA
 AGGTGATGCCAGGTGATTATCAGTATGTTGTGACCTATCGTGACGAACATGGTAAAGAACATCAAAAGCAG
 TACACCATATCTGTGAATGACAAAAACCAATGATCACTCAGGGACGTTTTGATACCATTAATGGCGTTGA
 CCACTTTACTCCTGACAAGACAAAAGCCCTTGACTCATCAGGCATTGTCCGCGAAGAAGTCTTTTACTTTGG
 CCAAGAAAAATGGCCGTAAATTTGATGTGACAGAAGGTAAAGATGGTATCACAGTTAGTGACAATAAGGTG
 10 TATATCCCTAAAAATCCAGATGGTCTTACACCATTTCAAAAAGAGATGGTGTACACTGTTCAGATTATTA
 CTACCTTGTCGAAGATAGAGCTGGTAATGTGTCTTTTGCTACCTTGCGTGACCTAAAAGCGGTGCGAAAAG
 ACAAAGCAGTAGTCAATTTTGGATTAGACTTACCGGTCCCTGAAGACAAACAANTAGTGAAGTTTACCTTAC
 CTTGTGCGGGATGCAGATGGTAAACCGATTGAAAACCTAGAGTATTATAATAACTCAGGTAACAGTCTTAT
 CTTGCCATACGGCAAATACACGGTCGAATTGTTGACCTATGACACCAATGCAGCCAACTAGAGTCAGATA
 15 AAATCGTTTCCTTTACCTTGTTCAGCTGATAACAACCTCCAACAAGTTACCTTTAAGATAACGATGTTAGCA
 ACTTCTCAAATAACTGCCCACTTTTGATCATCTTTTGCCAGAAGGCAGTCGCGTTAGCCTTAAAACAGCTCA
 AGATCAGCTAATCCCGCTTGAACAGTCTTGTATGTGCTTAAAGCTTATGGCAAAACCGTTCAAGAAGGCA
 CTTACGAAGTTGTTGTTCAGCTGCCTAAAGGCTACCGTATCGAAGGCAACACAAAGGTGAATACCTACCA
 AATGAAGTGCACGAAGTATCATTACGCTTGTCAAAGTAGGAGATGCCTCAGATTCAACTGGTGATCATAA
 GGTATGTCAAAAAATTAATTCACAGGCTTTGACAGCCTCTGCCACACCAACCAAGTCAACGACCTCAGCAA
 20 CAGCAAAAGCCCTACCATCAACGGGTGAAAAAATGGGTCTCAAGTTGCGCATAGTAGGTCTTGTGTTACTC
 GGACTTACTTGCGTCTTTAGCCGAAAAAATCAACCAAGATTGA

SEQ ID NO: 118 secuencia de aminoácidos que comprende la secuencia líder N-terminal de GAS 57
 MEKKQRFSLRKYKSGTFSVLIGSVFLVMTTVA

25 **SEQ ID NO: 119** secuencia de aminoácidos que comprende un fragmento de GAS 57 en el que la secuencia
 líder N-terminal está eliminada

30 ADELSTMSEPTITNHAQQQAQHLTNTELSSAESKSQDTSQITLKTNREKEQSQDLVSEPTTTELADTDAAS
 MANTGSDATQKSASLPVNTDVHVDWVKTKGAWDKYKGQGVVAVIDTGIDPAHQSMRISDVSTAKVKSKE
 DMLARQKAAGINYGSWINDKVVFVFAHNYVENSNDNIENQFEDFEDWENFEFDAEAPKAIKKHKIYRPQST
 QAPKETVIKTEETDGSHDIDWTQTDDEDTKYESHGMHVTGIVAGNSKEAAATGERFLGIAPEAQVMFMRVFA
 NDIMGSAESLFIKAIEDAVALGADVNLNLTGTANGAQLSGSKPLMEALIEKAKKAGVSVVVAAGNERVYGS
 35 HDDPLATNPDYGLVGSPTGRTPTSVAAINSKWVQRLMTVKELENRADLNHGKAIYSESVDKDIKDSL
 YDKSHQFAYVKESTDAGYNAQDVKGKIALIERDPNKTYDEMIALAKKHGALGVLPNNKPGQSNRSMRLTA
 NGMGIPSAFISHEPGKAMSQNLNGTGSLEFDSVVSQKAPSKQGNEMNHFSNNGLTSDGYLKPDIITAPGGDI
 YSTYNDNHYSQTGTSMASTPQIAGASLLVKQYLEKTQPNLPKEKIADIVKNLLMSNAQIHVNPETKTTTSP
 RQQGAGLLNIDGAVTSGLYVTGKDNYSISLGNITDTMTFDVTVHNLNNDKTLRYDTELLTDHVDPOKGR
 40 FTLTSHSLKTYQGGEVTVFANGKVTVRVTMDVVSQFTKELTKQMPNGYYLEGFVRFRDSQDDQLNRVNI
 PFV GFKGQFENLAVAEESIYRLKSQKGTGFYFDESQPKDDIYVGKHFTGLVTLGSETNVSTKTI
 SDNGLHTLGT FKNADGKFILEKNAQGNPVLAI SPNGDNNQDFAAPKGVFLRKYQGLKASVYHASDKEHKNPLWV
 SPESFKG DKNFNSDIRFAKSTTLLGTAFIGKSLTGAELPDGHHYVVSYYPDVVGAKRQEMTFDMILDRQK
 PVLQAT FDPETNRFKPEPLKDRGLAGVRKDSVFYLERKDNKPYTVTINDSYKYVSVEDNKT
 FVERQADGSFILPLDK AKLGDFYYMVEDFAGNVAIAKLGDHLPQTLGKTPIKLKLTDGNYQTKETL
 45 KDNLEMTQSDTGLVTNQAQLA VVHRNQPSQLTKMNQDFFISPNEDGNKDFVAFKGLKNNVYNDLTVN
 VYAKDDHQKQTFIWSQAGASVSA IESTAWYGITARGSKVMPGDYQYVVVYRDEHGEH
 QKQYTI SVNDKKPMITQGRFDTINGVDHFTPDKTKA LDSSGIVREEV
 FYLAKKNRKFVTEGKDGITVSDNKVYI PKNPDGSYTI SKRDGVTLS
 DYYYLVEDRAGN VSFATRLDLKAVGKDKAVVNFGLDLVPEDKQIVNFTYLV
 RADGKPIENLEYNNNSGNLILPYGKYTVE LLTYDTNAAKLES
 DKIVSFTLSADNNFQVTFKITMLATSQITAHFDHLLPEGSRVSLKTAQDQLI
 50 PLEQS LYVPKAYGKTVEGTYEVVVS LPKGYRIEGNTKVNTLPNEVHEL
 SLRLVKVGDA SDSTGDHKVMSKNNSQA LTASATPTKSTTSATAKALPSTGEK
 MGLKLRLVGLVLLGLTCVFSRKKSTKD

SEQ ID NO: 120 secuencia de aminoácidos que comprende la región hidrófoba C-terminal
 LPSTGEKMGMLKLRLVGLVLLGLTCVFSRKKSTKD

SEQ ID NO: 121 secuencia de ácido aminoácidos que comprende un fragmento de GAS 57 en el que la región hidrófoba C-terminal está eliminada

5 MEKKQRFSLRKYKSGTFSVLIGSVFLVMTTVADELSTMSEPTITNHAQQQAQHLTNTELSSAESKSQDT
SQITLKTNRKEQSQDLVSEPTTTELADTDAASMANTGSDATQKSASLPPVNTDVHDWVKTGAWDKGYKG
QGKVVAVIDTGIDPAHQSMRISDVSTAKVKSKEMLARQKAAGINYGSWINDKVVFAHNVENSNDNIKENQ

10 FEDFDEDWENFEFDAAEPKAIKKHKIYRPQSTQAPKETVIKTEETDGSHDIDWTQTD DDTKYESHGMHVT
GIVAGNSKEAAATGERFLGIAPEAQVMFMRVVFANDIMGSAESLFIKAIEDAVALGADV INLSLGTANGAQL
SGSKPLMEAIEKAKKAGVSVVVAAGNERVYGSDDHDDPLATNPDYGLVGSPTGRTPPTSVAAINSKWVIQRL
MTVKELNLRADLNHGKAIYSESVDKDIKDSLGYDKSHQFAYVKESTDAGYNAQDVKGKIALIERDPNKTY
DEMIALAKKHGALGVLIFFNNKPGQSNRSMRLTANGMGIPSAFISHEFGKAMSQNLNGNGTGSLEFDSVVSKA
15 PSQKGNEMNHFNSWGLTSDGYLKPDIAPGGDIYSTYNDNHYGSQTGTSMASQIAGA SLLVKQYLEKTQP
NLPKEKIADIVKNLLMSNAQIHVNPEKTTTSPRQQGAGLLNIDGAVTSGLYVTGKDNYGSI SLGNITDTM
TFDVTVHNLSNKKDLRYDTTELLTDHVDPPQGRFTLTSHSLKTYQGGEVTPANGKVTVRVTMDVSQFTKE
LTKQMPNGYYLEGFVRFDSQDDQLNRVNI PFVGFKGQFENLAVAEESIYRLKSQKGTGFYFDESQPKDDI
YVGKHFTGLVTLGSETNVSTKTI SDNGLHTLGTFFKNADGKFILEKNAQGNPVLAI SPNGDNNQDFAAFKGV
20 FLRKYQGLKASVYHASDKEHKNPLVSPESFKGDKNFNSDIRFAKSTTLLGTAFSGKSLTGAE LDPDGHYHY
VVSYPDVVGAKRQEMTFDMLDRQKPVLSQATFDPETNRFPKPEPLKDRGLAGVRKDS VFYLERKDNKPYT
VTINDSYKYVSVEDNKTPTVERQADGSPFILPLDKAKLGDFFYMVEDFAGNVIAIAGLGDHL PQTLGKTP IKLK
LTDGNYQTKETLKDNLMTQSDTGLVTNQAQLAVVHRNQPSQLTKMNQDFFISPNE DGNKDFVAPKGLKN
NVYNDLTVNVYAKDDHQKQTP IWSSQAGASVSAIESTAWYGITARGSKVMPGDYQYVV TYRDEHGKEHQKQ
YTI SVNDKKPMITQGRFDTINGVDHFTPDKTKALDSSGIVREEVFYLA KNGRKFVDVTEGKDGITVSDNKV
25 YIPKNPDGYSYTI SKRDGVTLSDYIYLVEDRAGNVSFATLRDLKAVGKDKAVVNFGLDLPVPEDKQIVNFYI
LVRDADGKPIENLEYNNNSGNLSILPYGKYTVELLTYDTNAAKLES DKIVSFTLSADNNFQQVTFKITMLA
TSQITAHFDHLLPEGSRVSLKTAQDQLI PLEQSLYVVPKAYGKTVQEGTYEVVVS LPKG YRIEGNTKVNTLP
NEVHEL SLRLVKVGDA SDSTGDHKVMSKNNSQAL TASATPTKSTTSATAKA

SEQ ID NO: 122 secuencia de aminoácidos que comprende un fragmento de GAS 57 en el que la secuencia líder N-terminal y la región hidrófoba C-terminal están eliminadas

35 ADELSTMSEPTITNHAQQQAQHLTNTELSSAESKSQDTSQITLKTNRKEQSQDLVSEPTTTELADTDAAS
MANTGSDATQKSASLPPVNTDVHDWVKTGAWDKGYKGQGKVVAVIDTGIDPAHQSMRI SDVSTAKVKSKE
DMLARQKAAGINYGSWINDKVVFAHNVENSNDNIKENQFEDFDEDWENFEFDAAEPKAIKKHKIYRPQST
QAPKETVIKTEETDGSHDIDWTQTD DDTKYESHGMHVTGIVAGNSKEAAATGERFLGIAPEAQVMFMRVFA
NDIMGSAESLFIKAIEDAVALGADV INLSLGTANGAQLSGSKPLMEAIEKAKKAGVSVVVAAGNERVYGS
YDKSHQFAYVKESTDAGYNAQDVKGKIALIERDPNKTYDEMIALAKKHGALGVLIFFNNKPGQSNRSMRLTA
40 NGMGIPSAFISHEFGKAMSQNLNGNGTGSLEFDSVVS KAPSQKGNEMNHFNSWGLTSDGYLKPDIAPGGDI
YSTYNDNHYGSQTGTSMASQIAGASLLVKQYLEKTQPNLPKEKIADIVKNLLMSNAQIHVNPEKTTTSP
RQQGAGLLNIDGAVTSGLYVTGKDNYGSI SLGNITDTMTFDVTVHNLSNKKDLRYDTTELLTDHVDPPQGR
FTLTSHSLKTYQGGEVTPANGKVTVRVTMDVSQFTKELTKQMPNGYYLEGFVRFDSQDDQLNRVNI PFV
GFKGQFENLAVAEESIYRLKSQKGTGFYFDESQPKDDIYVGKHFTGLVTLGSETNVSTKTI SDNGLHTLGT
FKNADGKFILEKNAQGNPVLAI SPNGDNNQDFAAFKGVFLRKYQGLKASVYHASDKEHKNPLVSPESFKG
45 DKNFNSDIRFAKSTTLLGTAFSGKSLTGAE LDPDGHYHYVVSYPDVVGAKRQEMTFDMLDRQKPVLSQAT
FDPETNRFPKPEPLKDRGLAGVRKDSVFYLERKDNKPYT VTINDSYKYVSVEDNKTPTVERQADGSPFILPLDK
AKLGDFFYMVEDFAGNVIAIAGLGDHL PQTLGKTP IKLKLTLDGNYQTKETLKDNLMTQSDTGLVTNQAQLA
VVHRNQPSQLTKMNQDFFISPNE DGNKDFVAPKGLKNVYNDLTVNVYAKDDHQKQTP IWSSQAGASVSA
IESTAWYGITARGSKVMPGDYQYVV TYRDEHGKEHQKQYTI SVNDKKPMITQGRFDTINGVDHFTPDKTKA
50 LDSSGIVREEVFYLA KNGRKFVDVTEGKDGITVSDNKVYI PKNPDGYSYTI SKRDGVTLSDYIYLVEDRAGN
VSFATLRDLKAVGKDKAVVNFGLDLPVPEDKQIVNFYI LVRDADGKPIENLEYNNNSGNLSILPYGKYTVE
LLTYDTNAAKLES DKIVSFTLSADNNFQQVTFKITMLATSQITAHFDHLLPEGSRVSLKTAQDQLI PLEQS
LYVVPKAYGKTVQEGTYEVVVS LPKG YRIEGNTKVNTLPNEVHEL SLRLVKVGDA SDSTGDHKVMSKNNSQA
LTASATPTKSTTSATAKA

SEQ ID NO: 123 secuencia de aminoácidos de una proteína M de GAS

60 MAKNNNTNRHYSRLKLTGTASVAVALTVLGAGFANQTEVKANGDGNPREVIEDLAANNPAIQNIRLRYENK
DLKARLENAMVAGRDFKRAEELEKAKQALEQQRKDLTKLKELOQDYDLAKESTSWDRQRLEKELEEKKE
ALELAIDQASRDYHRATALEKELEEKKALELAIDQASQDYNRANVLEKELETTREBQINRNLGNAKLE
LDQLSSEKEQLTIEKAKLEEEKQISDASRQSLRRDLDA SREAKKQVEKDLANLTAE LDKVKEEKQISDASRQGLRRDLDA SREAKKQVEKALBEANSKLA
65 QGLRRDLDA SREAKKQVEKDLANLTAE LDKVKEEKQISDASRQGLRRDLDA SREAKKQVEKALBEANSKLA
ALEKLNKELEESKKLTEKEKAE LQAKLEAEAKALKEQLAKQAEELAKLRAGKASDSQTPIDTKPGNKAVPGK
GQAPQAGTKPNQNKAPMKETKRQLPSTGETANPFFTAALTVMATAGVAAVVKRKEEN

SEQ ID NO: 124 secuencia de aminoácidos de SfbI de GAS

5 MSFDGFFLHHLTNELKENLLYGRIQKVNQPFERELVLTIRNHRKNYKLLLSAHPVFGRVQITQADFQNPQV
 PNTFTMIMRKYLQGAVIEQLEQIDNDRIIEIKVSNKNEIGDAIQATLIIEIMGKHSNIILVDRAENKIIES
 IKHVGFSSQNSYRTILPGSTYIEPPKTAAVNPFTITDVPLFEILQTQELTVKSLQQHFQGLGRDTAKELAE
 LTDDKLKRFREFFARPTQANLTTASFAPVLFSDSHATFETLSDMLDHFYQDKAERDRINQQASDLIHRVQT
 ELDKNRNKLKQEAELLATENAELEPRQKGELLTTYLSLVFNNQDSVILDNYYTGKIEIALDKALTPNQNA
 10 QRYFKKYQKLKEAVKHLISGLIADTKQSITYFESVDYNLSQASIDDIETREELYQAGFLKSRQDKRHKRK
 KPEQYLASDGTITILMVGRNNLQNEELTFKMAKKGELWFHAKDIPGSHV I IKDNLDPSEVKTDAAELAAYY
 SKARLSNLVQVDMIEAKKLHKPSGAKPGFVTTYTGQKTLRVTPDQAKILSMKLS

SEQ ID NO: 125 secuencia de aminoácidos de una proteína Shp de GAS

15 MTKVVIKQLLQVIVVFMISLSTMTNLVYADKQGIYGCIIQRNYRHPISGQIEDSGGEHSFDIGQGMVEGT
 YSDAMLEVS DAGKIVLTFRMSLADYSGNYQFWIQPGGTGSFQAVDYNITQKGTDTNGTTLDIAISLPTVNS
 IIRGSMFVEPMGREVVFYLSASELIQKYSNMLAQLVTETDNSQNQEVKDSQKPVDTKLGESQDESHTGAM
 20 ITQNKPKANSNNKSLSDKKILPSKMGLTTSLELKKEDKFRSKKDLSIMIYYFPTFFLMLGGFAVWVKKR
 KKNDKTM

SEQ ID NO: 126 aminoácidos 10 a 30 de la proteína SagA de GAS
 FSIATGSGNSQGGSGSYTPGKC

25

30

35

40

45

50

55

60

65

SEQ ID NO: 127 secuencia de polinucleótidos que comprende el constructo de fusión 117-40a-RR

5 ATGGCCCTTTAACACAAGCCAGAGTGTCTAGTGCACAAGTTTATAGCAATGAAGGGTATCACCAGCATTGAC
TGATGAAAAATCACACCTGCAATATAGTAAAGACAACGCACAACCTCAATTGAGAAATATCCTTGACGGCT
ACCAAAATGACCTAGGGAGACACTACTCTAGCTATTATTACTACAACCTAAGAACCCTTATGGGACTATCA
AGTGAGCAAGACATTGAAAAACACTATGAAGAGCTTAAGAACAGTTACATGATATGTACAATCATTATgc
tagcgggtggcggatccATGAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAA
CGAAGGCGAGTAATACTCACGACGATAGTTTACCAAAACAGAAACAATTCAAGAGGCAAAGGCAACTATT
10 GATGCAGTTGAAAAAACTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAAAC
TACTGCTGAAATCAACCACCTTAAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAA
TTTACACTAATACTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTA
ACAGCTACTGAAACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAGAGAGCTGCATTGTCAGAACAA
AAAAGCTAGCATTTACAGCAGAACTACTCGAGCTCAAGATTTAGTGGAAACAAGTCAAAACGCTCTGAACAAA
ATATTGCTAAGCTCAATGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGAT
15 AATACAAAAGCATTAAAGCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAATCAAAAAGCTAAAGTTAA
AAAGCAATTGACTGAAGAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAGAGGCAGAACTTAGTCGTC
TTAAATCCTCAGCTCCGCTCTACTCAAGATAGCATTGTGGGTAATAATACCATGAAAGCACCGCAAGGCTAT
CCTCTTGAAGAACTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAA
AGAGCATGCAGATCAAATTTATGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATATTCCAG
20 CAGATCGTAATCGCTTTGTTGATCCCGATAATTTGACACCAGAAGTGCAAAATGAGCTAGCGCAGTTTGCA
GCTCACATGATTAATAGTGTacGtcGtCAATTAGGTCTACCACCAGTTACTGTTACAGCAGGATCACAAGA
ATTTGCAAGATTACTTAGTACCAGCTATAAGAAAACCTCATGGTAATACAAGACCATCATTGTCTACGGAC
AGCCAGGGGTATCAGGGCATTATGGTGTGGGCCCTCATGATAAACTATTTATGAAGACTCTGCCGGAGCG
TCAGGGCTCATTCGAAATGATGATAACATGTACGAGAATATCGGTGCTTTTAAACGATGTGCATACTGTGAA
25 TGGTATTAAACGTGGTATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAAATACAT
ACGGCCATGCTATTAACTTTTCAGTGTAGATAAAACATAACCCTAATGCGCCTGTTTACCTTGGATTTTCA
ACCAGCAATGTAGGATCTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATTGCTAACCATCAACG
CTTTAATAAGACCCCTATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAAGAGTAGGCACTGTATCTG
ATACTATTGCAGCGATCAAAGGAAAAGTAAGCTCATTAAGAAAATCGTTTGTGCGCTATTATCAAGAAGCT
30 GATATTATGCGAGCCCAAGCTAAAGTAAGTCAACTTCAAGGTAAATTAGCAAGCACACTTAAGCAGTCAGA
CAGCTTAAATCTCCAAGTGAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTA
AAGCAAAACAAGCACAACTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCA
CTGCACCAGACAGAAGCCTTAGCAGAGCAAGCCGACAGAGTGACAGCACTGGTGGCTAAAAAGCTCA
TTTGCAATATCTAAGGGACTTTAAATTGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATTGATAATA
CTAAGCAAGATTTGGCTAAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCT
35 AAACAAAGCAGTCTAGAAGCTACTATTGCTACCAAGAACACCAGTTGACTTTGCTTAAACCTTAGCTAA
CGAAAAGGAATATCGCCACTTAGACGAAGATATAGCTACTGTGCCTGATTTGCAAGTAGCTCCACCTCTTA

CGGGCGTAAAACCGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAATGGTTAAAGAAACG
40 AAACAACCTATTAGAAGCTTCAGCAAGATTAGCTGCTGAAATAACAAGTCTTGTAGCAGAAGCGCTTGTGG
CCAAACCTCTGAAATGGTAGCAAGTAATGCCATTGTGTCTAAAAATCACATCTTCGATTACTCAGCCCTCAT
CTAAGACATCTTATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAA
cGtgcgccgcactcgagCACACCACCACCACCAC

45

50

55

60

65

SEQ ID NO: 128 secuencia de aminoácidos que comprende el constructo de fusión 117-40a-RR

5
10
15
20
25

```

MAFN T S Q S V S A Q V Y S N E G Y H Q H L T D E K S H L Q Y S K D N
A Q L Q L R N I L D G Y Q N D L G R H Y S S Y Y Y Y N L R T V M G L S S
E Q D I E K H Y E E L K N K L H D M Y N H Y A S G G G S M S V G V S H Q
V K A D D R A S G E T K A S N T H D D S L P K P E T I Q E A K A T I D A
V E K T L S Q Q K A E L T E L A T A L T K T T A E I N H L K E Q Q D N E
Q K A L T S A Q E I Y T N T L A S S E E T L L A Q G A E H Q R E L T A T
E T E L H N A Q A D Q H S K E T A L S E Q K A S I S A E T T R A Q D L V
E Q V K T S E Q N I A K L N A M I S N P D A I T K A A Q T A N D N T K A
L S S E L E K A K A D L E N Q K A K V K K Q L T E E L A A Q K A A L A E
K E A E L S R L K S S A P S T Q D S I V G N N T M K A P Q G Y P L E E L
K K L E A S G Y I G S A S Y N N Y Y K E H A D Q I I A K A S P G N Q L N
Q Y Q D I P A D R N R F V D P D N L T P E V Q N E L A Q F A A H M I N S
V R R Q L G L P P V T V T A G S Q E F A R L L S T S Y K K T H G N T R P
S F V Y G Q P G V S G H Y G V G P H D K T I I E D S A G A S G L I R N D
D N M Y E N I G A F N D V H T V N G I K R G I Y D S I K Y M L F T D H L
H G N T Y G H A I N F L R V D K H N P N A P V Y L G F S T S N V G S L N
E H F V M F P E S N I A N H Q R F N K T P I K A V G S T K D Y A Q R V G
T V S D T I A A I K G K V S S L E N R L S A I H Q E A D I M A A Q A K V
S Q L Q G K L A S T L K Q S D S L N L Q V R Q L N D T K G S L R T E L L
A A K A K Q A Q L E A T R D Q S L A K L A S L K A A L H Q T E A L A E Q
A A A R V T A L V A K K A H L Q Y L R D F K L N P N R L Q V I R E R I D
N T K Q D L A K T T S S L L N A Q E A L A A L Q A K Q S S L E A T I A T
T E H Q L T L L K T L A N E K E Y R H L D E D I A T V P D L Q V A P P L
T G V K P L S Y S K I D T T P L V Q E M V K E T K Q L L E A S A R L A A
E N T S L V A E A L V G Q T S E M V A S N A I V S K I T S S I T Q P S S
K T S Y G S G S S T T S N L I S D V D E S T Q R A A A L E H H H H H

```

30 **SEQ ID NO: 129** secuencia de aminoácidos que comprende un conector en el constructo 117-40a-RR
YASGGGS

SEQ ID NO: 130 secuencia de polinucleótidos que comprende el constructo de fusión 40a-RR-117

35
40
45
50

```

ATGAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAATAC
TCACGACGATAGTTTACCAAAACCAGAAACAATTCAAGAGGCAAAGGCAACTATTGATGCAGTTGAAAAAA
CTCTCAGTCAACAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAACTACTGCTGAAATCAAC
CACTTAAAAGAGCAGCAAGATAATGAACAAAAGCTTTAACTCTGCACAAGAAATTTACACTAATACTCT
TGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTAACAGCTACTGAAACAG
AGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCAATGTCAGAACAAAAGCTAGCATTTC
GCAGAACTACTCGAGCTCAAGATTTAGTGGAACAAGTCAAACGCTCTGAACAAAATATTGCTAAGCTCAA
TGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGATAATACAAAAGCATTAA
GCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAATCAAAAAGCTAAAGTTAAAAAGCAATTGACTGAA
GAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCAGAACTTAGTCGTCTTAAATCCTCAGCTCC
GTCTACTCAAGATAGCATTGTGGGTAATAATACCATGAAAGCACCGCAAGGCTATCCTCTTGAAGAACTTA
AAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAAGAGCATGCAGATCAA
ATTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATATTCCAGCAGATCGTAATCGCTT
TGTTGATCCCGATAATTTGACACCAGAAGTGCAAAATGAGCTAGCGCAGTTGTCAGCTCATATGATTAATA
GTGTAcGtcGtCAATTAGGTCTACCACCAAGTCTACTGTTACAGCAGGATCACAAGAATTGCAAGATTACTT
AGTACCAGCTATAAGAAAACCTCATGGTAATACAAGACCATCATTGTCTACGGACAGCCAGGGGTATCAGG
GCATTATGGTGTGGGCCCTCATGATAAACTATTATTGAAGACTCTGCCGGAGCGTCAGGGCTCATTCGAA
ATGATGATAACATGTACGAGAATATCGGTGCTTTTAACGATGTCATACCTGTGAATGGTATTAAACGTGGT

```

55

60

65

ATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAAATACATACGGCCATGCTATTAA
 CTTTTTACGTGTAGATAAACATAAACCTTAATGCGCCTGTTTACCTTGGATTTC AACCCAGCAATGTAGGAT
 CTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATTGCTAACCATCAACGCTTTAATAAGACCCCT
 ATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAAGAGTAGGCACTGTATCTGATACTATTGCAGCGAT
 CAAAGGAAAAAGTAAGCTCATAGAAAATCGTTTGTGCGCTATTTCATCAAGAAGCTGATATTATGGCAGCCC
 AAGCTAAAGTAAGTCAACTTCAAGGTAATTAGCAAGCACACTTAAGCAGTCAGACAGCTTAAATCTCCAA
 GTGAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTAAAGCAAAAACAAGCACA
 ACTCGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACTGCACCAGACAGAAG
 CTTAGCAGAGCAAGCCGAGCCAGAGTGACAGCACTGGTGGCTA AAAAAGCTCATTTGCAATATCTAAGG
 GACTTTAAATTGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATTGATAATACTAAGCAAGATTGGC
 TAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCTAAACAAAGCAGTCTAG
 AAGCTACTATGTCTACCACAGAACACCAAGTTGACTTTGCTTAAACCTTAGCTAACGAAAAGGAATATCGC
 CACTTAGACGAAGATATAGTACTGTGCTGATTTGCAAGTAGCTCCACCTCTTACGGGCGTAAACCGCT
 ATCATATAGTAAGATAGATACTACTCGCTTGTTCAGAAATGGTTAAAGAAACGAAACAACTATTAGAAG
 CTTAGCAAGATTAGCTGCTGAAAATACAGTCTTGTAGCAGAAGCGCTTGTGGCCAAACCTCTGAAATG
 GTAGCAAGTAATGCCATTGTGTCTAAATCACATCTTCGATTACTCAGCCCTCATCTAAGACATCTTATGG
 CTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAACGtgctagcgggtagc
 gatccatggcccttTAACACAAGCCAGAGTGTCAGTGACACAAGTTTATAGCAATGAAGGGTATCACCAGCAT
 TTGACTGATGAAAAATCACACCTGCAATATAGTAAAGACAACGCACAACCTTCAATTGAGAAATATCTTGA
 CGGCTACGAAAATGACCTAGGGAGACACTACTCTAGCTATTATTACTACAACCTAAGAACCTTATGGGAG
 TATCAAGTGAGCAAGACATTGAAAAACACTATGAAGAGCTTAAAGACAAGTTACATGATATGTACAATCAT
 TATgcgccgcgcaactcgagCACCACCACCACCACCAC

SEQ ID NO: 131 secuencia de aminoácidos que comprende el constructo de fusión 40a-RR-117

M S V G V S H Q V K A D D R A S G E T K A S N T H D D S L P K P E T I Q
 E A K A T I D A V E K T L S Q Q K A E L T E L A T A L T K T T A E I N H
 L K E Q Q D N E Q K A L T S A Q E I Y T N T L A S S E E T L L A Q G A E
 H Q R E L T A T E T E L H N A Q A D Q H S K E T A L S E Q K A S I S A E
 T T R A Q D L V E Q V K T S E Q N I A K L N A M I S N P D A I T K A A Q
 T A N D N T K A L S S E L E K A K A D L E N Q K A K V K Q L T B E L A
 A Q K A A L A E K E A E L S R L K S S A P S T Q D S I V G N N T M K A P
 Q G Y P L E E L K K L E A S G Y I G S A S Y N N Y Y K E H A D Q I I A K
 A S P G N Q L N Q Y Q D I P A D R N R F V D P D N L T P E V Q N E L A Q
 F A A H M I N S V R R Q L G L P P V T V T A G S Q E F A R L L S T S Y K
 K T H G N T R P S F V Y G Q P G V S G H Y G V G P H D K T I I E D S A G
 A S G L I R N D D N M Y E N I G A F N D V H T V N G I K R G I Y D S I K
 Y M L F T D H L H G N T Y G H A I N F L R V D K H N P N A P V Y L G F S
 T S N V G S L N E H F V M F P E S N I A N H Q R F N K T P I K A V G S T
 K D Y A Q R V G T V S D T I A A I K G K V S S L E N R L S A I H Q E A D
 I M A A Q A K V S Q L Q G K L A S T L K Q S D S L N L Q V R Q L N D T K
 G S L R T E L L A A K A K Q A Q L E A T R D Q S L A K L A S L K A A L H
 Q T E A L A E Q A A A R V T A L V A K K A H L Q Y L R D F K L N P N R L
 Q V I R E R I D N T K Q D L A K T T S S L L N A Q E A L A A L Q A K Q S
 S L E A T I A T T E H Q L T L L K T L A N E K E Y R H L D E D I A T V P
 D L Q V A P P L T G V K P L S Y S K I D T T P L V Q E M V K E T K Q L L
 E A S A R L A A E N T S L V A E A L V G Q T S E M V A S N A I V S K I T
 S S I T Q P S S K T S Y G S G S S T E S N L I S D V D E S T Q R A S G G
 G S M A F N T S Q S V S A Q V Y S N E G Y H Q H L T D E S H L Q Y S K
 D N A Q L Q L R N I L D G Y Q N D L G R H Y S S Y Y Y N L R T V M G L
 S S E Q D I E K H Y E E L K N K L H D M Y N H Y A A A L E H H H H H

SEQ ID NO: 132 secuencia de polinucleótidos que comprende el constructo de fusión GAS 117 - 40a

ATGGCCTTTAACAACAAGCCAGAGTGTCAGTGCACAAAGTTTATAGCAATGAAGGGTATCACCAGCATTGAC
 TGATGAAAAATCACACCTGCAATATAGTAAAGACAAAGCACAACCTTCAATTGAGAAATATCCTTGACGGCT
 ACCAAAATGACCTAGGGAGACACTACTCTAGCTATTATTACTACAACCTAAGAAAGGTTATGGGACTATCA
 AGTGAGCAAGACATTGAAAAACACTATGAAGAGCTTAAAGAACAAGTTACATGATATGTACAATCATTATgc

TTTGCGGTGGCGGATCCATGAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAA
 CGAAGGCGAGTAATACTCACGACGATAGTTTACCAAAACCAGAAACAATTCAAGAGGCAAGGCAACTATT
 5 GATGCAGTTGAAAAAAGCTCTCAGTCAACAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAAAC
 TACTGCTGAAATCAACCACCTTAAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAA
 TTTACACTAATACTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTA
 ACAGCTACTGAAACAGAGCTTTCATAATGCTCAAGCAGATCAACATTCAAAGAGAGCTGCATTGTCAGAACA
 10 AAAAGCTAGCATTTCAGCAGAACTACTCGAGCTCAAGATTTAGTGGAACAAGTCAAACGCTCTGAACAAA
 ATATTGCTAAGCTCAATGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGAT
 AATACAAAAGCATTAAAGCTCAGAATTGGAGAAGGCTAAAGCTGACTTAGAAAATCAAAGCTAAAGTTAA
 AAAGCAATTGACTGAAGAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCAGAACTTAGTCGTC
 TTAATCCTCAGCTCCGCTACTCAAGATAGCATTGTGGGTAATAATACCATGAAAGCACCGCAAGGCTAT
 15 CCTCTTGAAGAACTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAA
 AGAGCATGCAGATCAAATTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATATTCCAG
 CAGATCGTAATCGCTTTGTTGATCCCGATAATTTGACACCAGAAGTGCAAAATGCTAGCGCAGTTTGCA
 GCTCACATGATTAATAGTGTAGAAGACAATTAGGCTTACCACCAGTTACTGTTACAGCAGGATCACAAGA
 ATTTGCAAGATTACTTAGTACAGCTATAAGAAAACATCATGGTAATACAAGACCATCATTTGTCTACGGAC
 AGCCAGGGGTATCAGGGCATTATGGTGTGGGCTCATGATAAACTATTATTGAAGACTCTGCCGGAGCG
 20 TCAGGGCTCATTCGAAATGATGATAACATGTACGAGAATATCGGTGCTTTAACGATGTGCATACTGTGAA
 TGGTATTAAACGTGGTATTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAATACAT
 ACGGCCATGCTATTAACTTTTACGTGTAGATAAACATAACCCTAATGCGCCTGTTTACCTTGGATTTTCA
 ACCAGCAATGTAGGATCTTTGAATGAACACTTTGTAATGTTTCCAGAGTCTAACATTGCTAACCATCAACG
 CTTTAATAAGACCCCTATAAAAGCCGTTGGAAGTACAAAAGATTATGCCCAAAGAGTAGGCACTGTATCTG
 25 ATACTATTGCAGCGATCAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTCCGCTATTTCATCAAGAAGCT
 GATATTATGGCAGCCCAAGCTAAAGTAAAGTCAACTTCAGGTAATTAGCAAGCACACTTAAGCAGTCAGA
 CAGCTTAAATCTCAAGTGAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGAACTACTAGCAGCTA
 AAGCAAAACAAGCACAACCTGAAGCTACTCGTGATCAATCATTAGCTAAGCTAGCATGTTGAAAGCCGCA
 CTGCACCAGACAGAAGCCTTAGCAGAGCAAGCCGAGCCAGAGTGACAGCACTGGTGCTAAAAAGCTCA
 TTTGCAATATCTAAGGGACTTTAAATGAATCCTAACCGCCTTCAAGTGATACGTGAGCGCATTGATAATA
 30 CTAAGCAAGATTGGGCTAAAACCTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAGCCTTACAAGCT
 AAACAAAGCAGTCTAGAAGCTACTATTGCTACCACAGAACACCAGTTGACTTTGCTTAAACCTTAGCTAA
 CGAAAAGGAATATCGCCACTTAGACGAAGATATAGCTACTGTGCTGATTTGCAAGTAGCTCCACCTCTTA
 CGGGCGTAAAACCGCTATCATATAGTAAGATAGATACTACTCCGCTTGTTCAGAAATGGTTAAAGAAACG
 35 AAACAACCTATTAGAAGCTTCAGCAAGATTAGCTGCTGAAAATACAAGTCTTGTAGCAGAAGCGCTTGTGG
 CCAAACCTCTGAAATGGTAGCAAGTAATGCCATTGTGTCTAAAATCACATCTTCGATTACTCAGCCCTCAT
 CTAAGACATCTTATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATGAAAGTACTCAA
 cGtgcgcccgagcactcgagcaccaccaccaccacc

SEQ ID NO: 133 secuencia de aminoácidos que comprende el constructo de fusión GAS 117-40a

M A F N T S Q S V S A Q V Y S N E G Y H Q H L T D E K S H L Q Y S K D N
 A Q L Q L R N I L D G Y Q N D L G R H Y S S Y Y Y Y N L R T V M G L S S
 E Q D I E K H Y E E L K N K L H D M Y N H Y A S G G S M S V G V S H Q
 45 V K A D D R A S G E T K A S N T H D D S L P K P E T I Q E A K A T I D A
 V E K T L S Q Q K A E L T E L A T A L T K T T A E I N H L K E Q Q D N E
 Q K A L T S A Q E I Y T N T L A S S E E T L L A Q G A E H Q R E L T A T
 E T E L H N A Q A D Q H S K E T A L S E Q K A S I S A E T T R A Q D L V
 E Q V K T S E Q N I A K L N A M I S N P D A I T K A A Q T A N D N T K A
 50 L S S E L E K A K A D L E N Q K A K V K Q L T E E L A A Q K A A L A E
 K E A E L S R L K S S A P S T Q D S I V G N N T M K A P Q G Y P L E E L
 K K L E A S G Y I G S A S Y N N Y Y K E H A D Q I I A K A S P G N Q L N
 Q Y Q D I P A D R N R F V D P D N L T P E V Q N G L A Q F A A H M I N S
 V R R Q L G L P P V T V T A G S Q E F A R L L S T S Y K K T H G N T R P
 S F V Y G Q P G V S G H Y G V G P H D K T I I E D S A G A S G L I R N D
 55 D N M Y E N I G A F N D V H T V N G I K R G I Y D S I K Y M L F T D H L
 H G N T Y G H A I N F L R V D K H N P N A P V Y L G F S T S N V G S L N
 E H F V M F P E S N I A N H Q R F N K T P I K A V G S T K D Y A Q R V G
 T V S D T I A A I K G K V S S L E N R L S A I H Q E A D I M A A Q A K V

5 S Q L Q G K L A S T L K Q S D S L N L Q V R Q L N D T K G S L R T E L L
A A K A K Q A Q L E A T R D Q S L A K L A S L K A A L H Q T E A L A E Q
A A A R V T A L V A K K A H L Q Y L R D F K L N P N R L Q V I R E R I D
N T K Q D L A K T T S S L L N A Q E A L A A L Q A K Q S S L E A T I A T
T E H Q L T L L K T L A N E K E Y R H L D E D I A T V P D L Q V A P P L
T G V K P L S Y S K I D T T P L V Q E M V K E T K Q L L E A S A R L A A
10 E N T S L V A E A L V G Q T S E M V A S N A I V S K I T S S I T Q P S S
K T S Y G S G S S T T S N L I S D V D E S T Q R A A A L E H H H H H

SEQ ID NO: 134 secuencia de polinucleótidos que comprende el constructo de fusión GAS 117-40N

15 ATGGCCTTTAACACAAGCCAGAGTGTCTAGTGCACAAGTTTATAGCAATGAAGGGTATCACCAGCATTGAC
TGATGAAAAATCACACCTGCAATATAGTAAAGACAACGCACAACTTCAATTGAGAAATATCCTTGACGGCT
ACCAAAATGACCTAGGGAGACACTACTCTAGCTATTATTACTACAACCTAAGAACCCTTATGGGACTATCA
20 AGTGAGCAAGACATTGAAAAACACTATGAAGAGCTTAAGAACAAGTTACATGATATGTACAATCATTATG
GagcggctggcggatcCATGAGTGTAGGCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAA
CGAAGGCGAGTAATACTCACGACGATAGTTTACCAAAACAGAAACAATTCAAGAGGCCAAAGGCAACTATT
GATGCAAGTTGAAAAACTCTCAGTCAACAAAAAGCAGAACTGACAGAGCTTGCTACCGCTCTGACAAAAAC
TACTGCTGAAATCAACCACTTAAAGAGCAGCAAGATAATGAACAAAAAGCTTTAACCTCTGCACAAGAAA
25 TTTACACTAATACTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAGTTA
ACAGCTACTGAAACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAGAGACTGCATTGTCTGAGAACA
AAAAGCTAGCATTTTACGAGAACTACTCGAGCTCAAGATTTAGTGGAACAAGTCAAAACGCTCTGAACAAA
ATATTGCTAAGCTCAATGCTATGATTAGCAATCCTGATGCTATCACTAAAGCAGCTCAAACGGCTAATGAT
AATACAAAAGCATTAAGCTCAGAAATTGAGAGAAGGCTAAAGCTGACTTAGAAAATCAAAGCTAAAGTTAA
AAAGCAATTGACTGAAGAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAAGAGGCAGAACTTAGTCGTC
30 TTAATCCTCAGCTCCGCTCTACTCAAGATAGCATTGTTGGGTAATAATACCATGAAAGCACCGCAAGGCTAT
CCTCTTGAAGAACTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAA
AGAGCATGCAGATCAAAATTATTGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGcggcgcgcac
tcgagCACCACCACCACCACCAC

SEQ ID NO: 135

35 M A F N T S Q S V S A Q V Y S N E G Y H Q H L T D E K S H L Q Y S K D N
A Q L Q L R N I L D G Y Q N D L G R H Y S S Y Y Y Y N L R T V M G L S S
40 E Q D I E K H Y E E L K N K L H D M Y N H Y A S G G G S M S V G V S H Q
V K A D D R A S G E T K A S N T H D D S L P K P E T I Q E A K A T I D A
V E K T L S Q Q K A E L T E L A T A L T K T T A E I N H L K E Q Q D N E
Q K A L T S A Q E I Y T N T L A S S E E T L L A Q G A E H Q R E L T A T
E T E L H N A Q A D Q H S K E T A L S E Q K A S I S A E T T R A Q D L V
45 E Q V K T S E Q N I A K L N A M I S N P D A I T K A A Q T A N D N T K A
L S S E L E K A K A D L E N Q K A K V K K Q L T E E L A A Q K A A L A E
K E A E L S R L K S S A P S T Q D S I V G N N T M K A P Q G Y P L E E L
K K L E A S G Y I G S A S Y N N Y Y K E H A D Q I I A K A S P G N Q L N
Q Y Q A A A L E H H H H H H

SEQ ID NO: 136
AGTTGGTA

REIVINDICACIONES

1. Composición inmunógena que comprende una combinación de antígenos de GAS, consistiendo dicha combinación en dos a diez antígenos de GAS, en la que dicha combinación incluye GAS40, siendo dicho "GAS40" una proteína que comprende una secuencia de aminoácidos (a) que tiene un 95% o más identidad con la SEQ ID NO: 1 y/o (b) que es un fragmento de al menos 200 aminoácidos consecutivos de la SEQ ID NO: 1 y en la que dicha combinación incluye GAS57, siendo dicho "GAS57" una proteína que comprende una secuencia de aminoácidos (a) que tiene un 95% o más identidad con la SEQ ID NO: 116 y/o (b) que es un fragmento de al menos 200 aminoácidos consecutivos de la SEQ ID NO: 116.
2. Composición según la reivindicación 1, en la que dicho antígeno GAS40 comprende una secuencia de aminoácidos que comprende una primera región con superenrollamiento y una segunda región con superenrollamiento.
3. Composición según la reivindicación 1, en la que el antígeno GAS40 comprende una secuencia de aminoácidos que comprende una primera región con superenrollamiento.
4. Composición según la reivindicación 3, en la que dicha región con superenrollamiento comprende una secuencia de aminoácidos que comprende la SEQ ID NO: 12.
5. Composición según la reivindicación 2, en la que dicha segunda región con superenrollamiento incluye una región con cremallera de leucina.
6. Composición según la reivindicación 5, en la que la segunda región con superenrollamiento comprende una secuencia de aminoácidos que comprende la SEQ ID NO: 13.
7. Composición según cualquiera de las reivindicaciones anteriores para uso terapéutico.
8. Método de fabricación de cualquiera de las composiciones inmunógenas según las reivindicaciones 1 a 6.
9. Kit que comprende un primer componente que comprende cualquiera de las composiciones inmunógenas según las reivindicaciones 1 a 6.

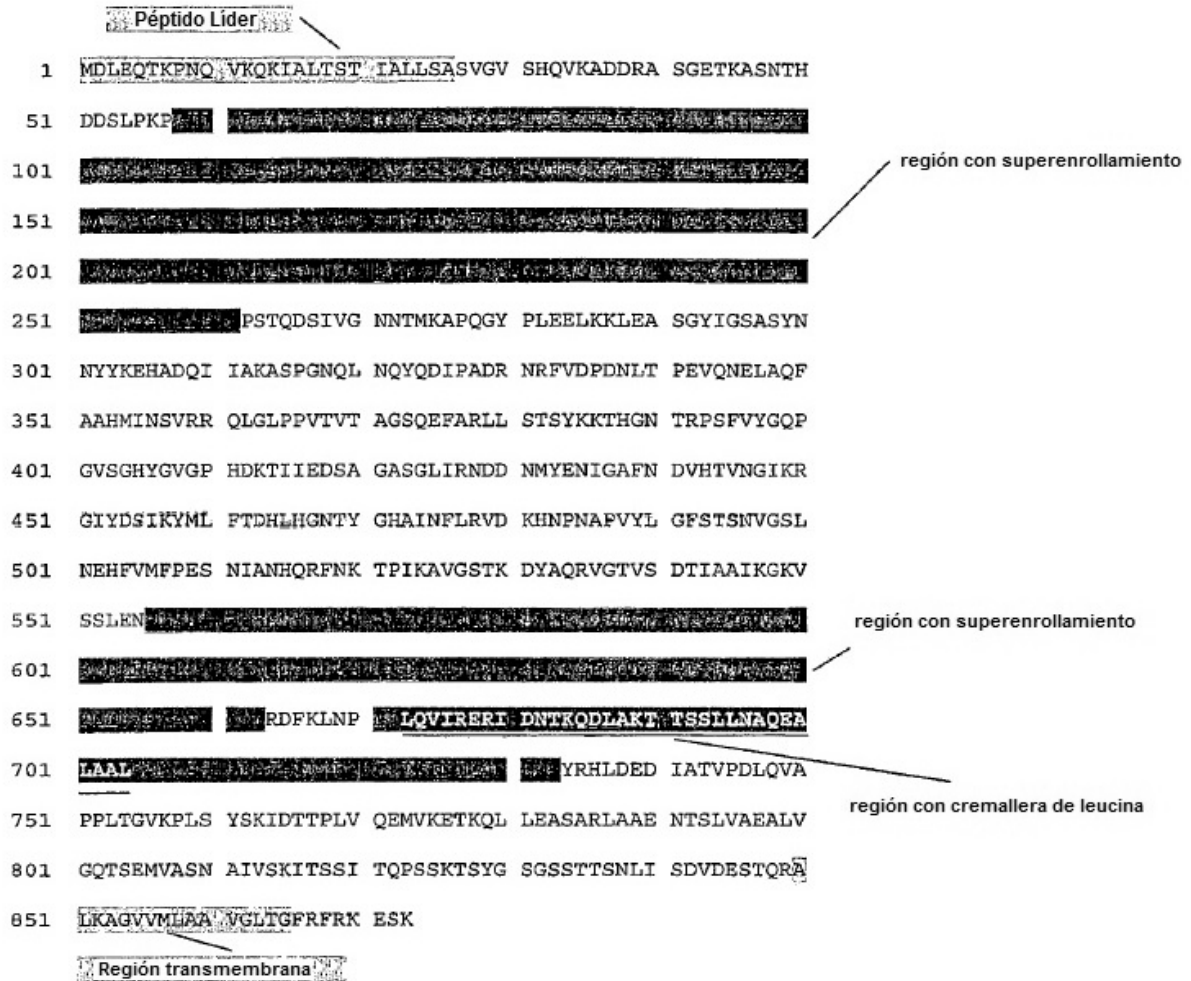
FIGURA 1: Anotación de GAS 40

FIGURA 2: Esquema de GAS40: supuesta proteína de exclusión superficial prgA (873 aa)

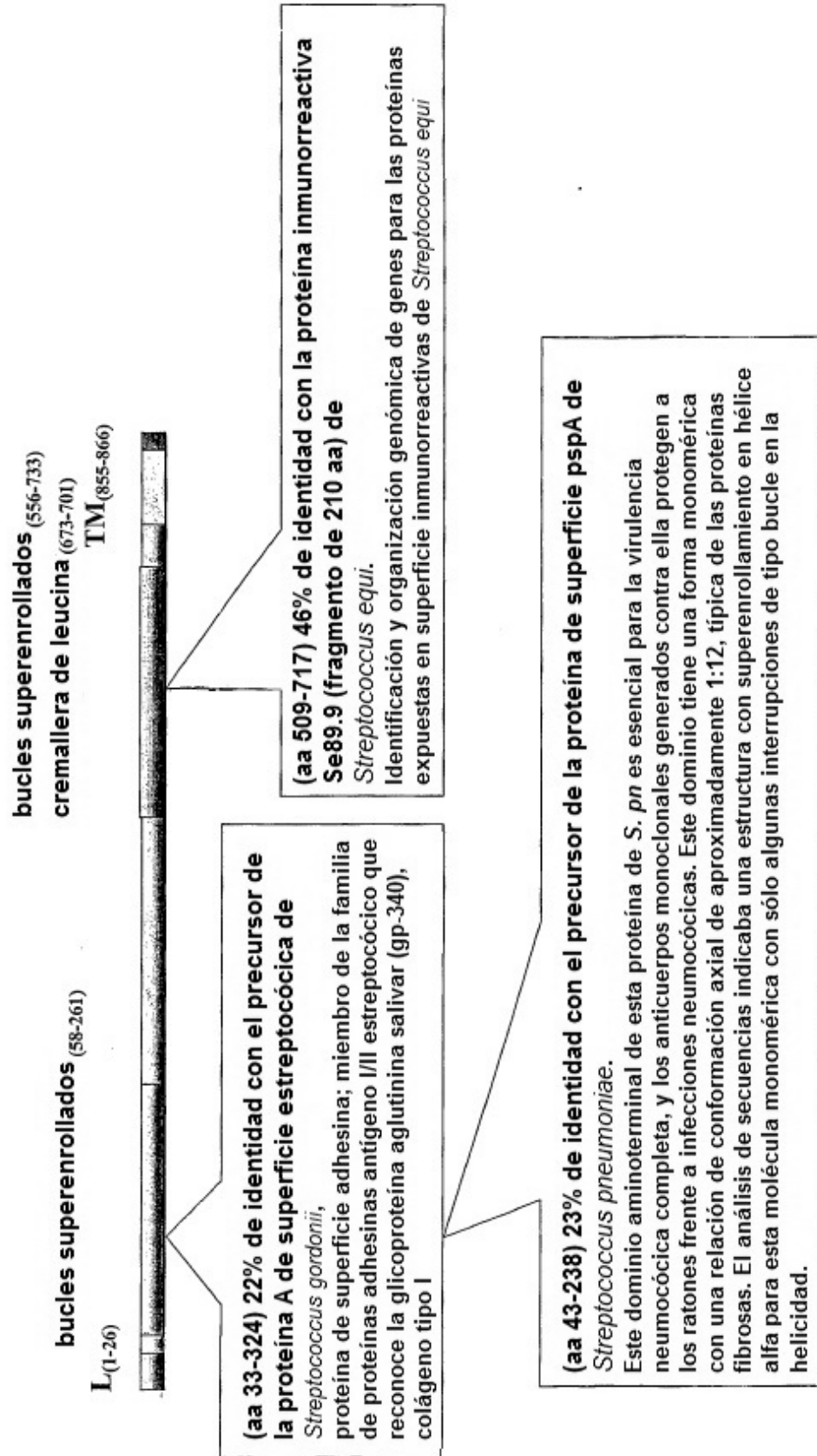


FIGURA 3: resultados BLAST de las regiones con superenrollamiento de GAS 40 con otras bacterias *Streptococcus*

3(a) alineamiento BLAST de la secuencia de aminoácidos de GAS 40 que incluye la primera región con superenrollamiento con el precursor SpA de *Streptococcus gordonii*

```
>gi|25990270|gb|AAC44101.3| precursor de proteína A de superficie estreptocócica
[Streptococcus gordonii]
    Longitud = 1575
>ref|NP_268623.1| supuesta proteína de exclusión superficial [Streptococcus pyogenes]
    Longitud = 873

Puntuación = 63.2 bits (152),    Prev = 5e-11
    Identidades = 65/293 (22%),    Positivos = 124/293 (42%), Huecos = 13/293
    (4%)

Consulta : 112 QDQTSKGTATTAAENAQKQAEIKSDYAKQA---EEIKKTTEAYKKEVEAHQAETDKIN
167
      Q + D+ + T A N + K + ++A + ++KT K E+ K
Objeto : 33 QVKADDRASGETKASNTHDSSLPKPETIQEAKATIDAVEKTLSSQKAELELATALTKTT
92

Consulta : 168 AENKAAEDKYQEDLKAHQAEVEKINTANATAKAEYAKLAQYQKDLAAVQKANEDSQLDY
227
      AE +++ + KA + E A+++ A+ A++Q++L A + ++Q D
Objeto : 93 ABINHLKBQDNEQKALTSAQEIYTNLASSSEETLLAQGAHQRELATETELHNAQADQ
152

Consulta : 228 QNKLSAYQAEELARVQKANAEAKEAYE--KAVKENTAKNAALQAENEAIKQRNETAKANYD
285
      +K +A + A + A++ E K ++N AK A+ + +AI + +TA N
Objeto : 153 HSKETALSEQKASISAETTRAQDLVEQVKTSEQNIAKLNAMISNPDAITKAAQTANDNTK
212

Consulta : 286 AAMKQYEADLAAIKKAKEDNDADYQAKLAAYQAEELARVQKANADAKAAYEKAVEENTAKN
345
      A + E A ++ K +LAA +A LA + + K++ + N
Objeto : 213 ALSSELEKAKADLENQKAKVKKQLTEELAAQKAALAEKEAELSRKSSAPSTQDSIVGNN
272

Consulta : 346 TAIQAEN---EAIKQRNAA---AKATYEAAALKQYEADLAAAKKANEDSDADYQ 392
T + E +K+ A+ A+Y K++ AD AK + + YQ
Objeto : 273 TMKAPQGYPLEELKKLEASGYIGSASYNYYKEH-ADQIIAKASPGNQLNQYQ 324
```

FIGURA 3: resultados BLAST de las regiones con superenrollamiento de GAS 40 con otras bacterias *Streptococcus*

3(b) alineamiento BLAST de la secuencia de aminoácidos de GAS 40 que incluye la primera región con superenrollamiento con el precursor SpB de *Streptococcus gordonii*

```
>gi|25055226|gb|AAC44102.3| precursor de proteína B de superficie estreptocócica
[Streptococcus gordonii]
    Longitud = 1499
>ref|NP_268623.1| supuesta proteína de exclusión superficial [Streptococcus
pyogenes]
    Longitud = 873

Puntuación = 54.3 bits (129), Prev = 2e-08
    Identidades = 53/226 (23%), Positivos = 98/226 (43%), Huecos = 13/226
(5%)

Consulta : 111 QDQTSKGTATTAAENAQKQAEIKSDYAKQA---EEIKKTTEAYKKEVEAHQAETDKIN
166
      Q + D+ + T A N + K + ++A + ++KT K E+ K
Objeto : 33 QVKADDRASGETKASNTTHDDSLPKPETIQEAKATIDAVEKTLSSQKAELETALATLTKTT
92

Consulta : 167 AENKAAEDKYQEDLKAHQAEVEKINTANATAKAEEYAKLAQYQKDLAAVQKANEDSQLDY
226
      AE +++ + KA + E A+++ A+ A++Q++L A + ++Q D
Objeto : 93 AEINHLKEQQDNEQKALTSQAEIYTNTLASSEETLLAQGAEHQRELTATETELHNAQADQ
152

Consulta : 227 QNKLSAYQAEELARV--QXXXXXXXXXXXXXXXXXNTAKNAALQAENEAIKQRNETAKANYD
284
      +K +A + A + + N AK A+ + +AI + +TA N
Objeto : 153 HSKETALSEQKASISAETTRAQDLVEQVKTSEQNIAKLNAMISNPDAITKAAQTANDNTK
212

Consulta : 285 AAMKQYE---ADL---AAIKKAKEDNDADYQAKLAAYQAEELARVQ 323
      A + E ADL A +KK + A +A LA +AEL+R++
Objeto : 213 ALSSELEKAKADLENQKAKVKQLTEELAAQKAALAEKEAELSLRK 258
```

FIGURA 3: resultados BLAST de las regiones con superenrollamiento de GAS 40 con otras bacterias *Streptococcus*

3(c) alineamiento BLAST de la secuencia de aminoácidos de GAS 40 que incluye la primera región con superenrollamiento con el precursor de la proteína de superficie PspA de *Streptococcus pneumoniae*

```
>gi| 282335 |pir|A41971 precursor de la proteína de superficie pspA - Streptococcus pneumoniae
>ref|NP_268623.1|          supuesta proteína de exclusión superficial [Streptococcus pyogenes]
Longitud = 873
```

Puntuación = 48.1 bits (113), Prev = 6e-07
 Identidades = 46/200 (23%), Positivos = 89/200 (44%), Huecos = 23/200 (11%)

```
Consulta: 139 KTKFNTVRAMVVPEPEQLAETK-----KKSEEAKQKAPELTKKLEEAkakLEE-AEKK
190
      +TK +      +P+PE + E K      K + K + EL L + A++ E++
Objeto : 43 ETKASNTHTDLSLPKPETIQEAKATIDAVEKTLSSQQAELTELATALTKTTAEINHLKEQQ
102
```

```
Consulta: 191 ATEAKQKVDAAEVAPQAKIAELENQVHRLEQELKEIDSESESEDYAKEGFRAPLQSKLDAK
250
      E K A+E+      + E + + + +E+ +E+E + + + ++ L +
Objeto : 103 DNEQKALTSAQEIYTNLTASSEETLLAQGAHQRELATETELHNAQADQHSKETALSEQ
162
```

```
Consulta: 251 KAKLS----KLEELSDKIDELDAEIAKLEDQL-----KAAEENNNVEDYFKEGLEKTI
299
      KA +S      + ++L +++ + IAKL +      KAA+ N+      LEK
Objeto : 163 KASISAETTRAQDLVEQVKTSEQNIAKLNAMISNPDAITKAAQTANDNTKALSSELEKA-
221
```

```
Consulta: 300 AAKKAELEKTEADLKKAVNE 319
      KA+LE +A +KK + E
Objeto : 222 ---KADLENQKAKVKKQLTE 238
```


FIGURA 3: resultados BLAST de las regiones con superenrollamiento de GAS 40 con otras bacterias *Streptococcus*

3(d) alineamiento BLAST de la secuencia de aminoácidos de GAS 40 que incluye la segunda región con superenrollamiento con el precursor SpB de *Streptococcus gordonii*

```
>gi|23380384|gb|AAN18299.1|    proteína inmunorreactiva Se89.9 (fragmento)
[Streptococcus equi]
      Length = 210
>ref|NP_268623.1|              supuesta proteína de exclusión superficial  [Streptococcus
pyogenes]
      Longitud = 873

Puntuación = 173 bits (438),      Prev = 4e-45
Identidades = 98/209 (46%),      Positivos = 144/209 (68%)

Consulta : 1   ESDIVDATRFSTTEIPKSGQVIDRSASIQALTNDIASIKGKIASLESRLADPSSEAEVTA
60
      ES+I + RF+ T I   G   D + + +++ IA+IKGK++SLE+RL+   EA++ A
Objeto : 509  ESNIANHQRFNKTPIKAVGSTKDYAQRVGTVSDTIAAIKGVSSLENRLSAIHQEADIMA
568

Consulta : 61  AQAKISQLQHLEAAQAKSHKLDQQVEQLANTKDSLRTQLLAAKEEQAQLKANLKDALAL
120
      AQAK+SQLQ +L +   +S   L+ QV QL +TK SLRT+LLAAK +QAQL+A   D++LA
Objeto : 569  AQAKVSQQLQGKLASTLKQSDSLNLQVRQLNDTKGSLRTELLAAKAKQAQLEATRDQSLAK
628

Consulta : 121 LASSKATLHKLEAAMEEAKARVAGLASQKAQLEDLLAFEKNPNRIELAQEKVAAAKKALA
180
      LAS KA LH+ EA   E+A ARV   L ++KA L+ L   F+ NPNR+++ +E++   K+ LA
Objeto : 629  LASLKAALHQTEALAEQAAARVTALVAKKAHLQYLRDFKLNPNRLQVIRERIDNTKQDLA
688

Consulta : 181 DTEDKLLAAQASLSDLQAQRARLQLSIAT 209
      T   LL AQ +L+ LQA+++ L+ +IAT
Objeto : 689  KTTSSLLNAQEALAAALQAKQSSLEATIAT 717
```

Figura 4: predicción de estructura secundaria de GAS 40**Figura 4(a) Alineación de predicción de estructura secundaria con secuencia de aminoácidos de GAS 40**

```

      10      20      30      40      50      60      70
MDLEQTKPNQVKQKIALTSTIALLSASVGVSHQVKADDRASGETKASNTHDDSLPKPETIQEAKATIDAV
CCCCCCCCcchhhHHhhhhHHHHhhccccEEEEcCCcCCCCcCCCCCCCCCCCCcchHHHHHHHHHH
EKTLSQQAELTELATALTKTTAEINHLKEQQDNEQKALTSAQEIYTNLTASSEETLLAQGAHQRELTA
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHcchHHHHHHHHHHHHHHHH
TETELHNAQADQHSKETALSEQKASISAETTRAQDLVEQVICTSEQNIAKLNAMISNPDAITKAAQTANDN
HHHHHHHHHHcchHHHHHHhhccchhhHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHcCchHHHHHHHHhhc
TKALSSELEKAKADLENQAKAVKKQLTEELAAQKAALAEKEAELSRKSSAPSTQDSIVGNNTMKAPQGY
cHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHcCCCCCcCceCCCCCCCCCCC
PLEELKKLEASGYIGSASYNNYYKEHADQIIAKASPGNQLNQYQDIPADRNRFPVDPDNLTPVQNELAQF
CHHHHHHHHHcCccccccchHHHHHHHHHHHHHHcCchhhhhhhcCccccCCCCCCCCcchHHHHHHHH
AAHMINSVRRQLGLPPVTVTAGSQEFARLLSTSYKKTGNTTRPSFVYGQPGVSGHYGVGPHDKTIIEDSA
HHHHHHHHHHcCCCCcCchHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
GASGLIRNDDNMYENIGAFNDVHTVNGIKRGIYDSIKYMLFTDHLHGNTYGHAINFLRVDKHNPNAPVYL
CCCCcCchHHHHhhccccccccCccccHHHHHHHHhheccccCccchHHhheeeccCCCCCcEEE
GFSTSNVGSLSNEHFVMPFESNIANHQRFNKTPIKAVGSTKDYAQRVGTVSDTIAAIKGVSSLENRLSAI
EEEEcCccCccccccccccccchHHhhCCCCccccCCCCHHHHHHhhHHHHHHHHcCccHHHHHHHH
HQEADIMAAQAKVSQLQGLASTLKQSDSLNLQVRQLNDTKGSLRTELLAAKAKQAQLEATRDQSLAKLA
HHHHHHHHHHHHHHHHHHHHHHHHHHcCChhhHHhhhhcCcCHHHHHHHHHHHHHHHHHHHHHHHHH
SLKALHQTEALAEQAAARVTVLAKKAHLQYLRDFKLNPNRLQVIRERIDNTKQDLAKTTSSLLNAQEA
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHcCchHHHHHHHHHHHHHHHHHHHH
LAALQAKQSSLEATIATTEHQLTLKTLANEKEYRHLDEDIATVPDLQVAPPLTGVKPLSYSKIDTTPLV
HHHHHHhhcCccccchHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
QEMVKETKQLLEASARLAAENTSLVAELVGQTSEMVASNAIVSKITSSITQPSSKTSYSGSGSSTSNLI
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHcchHHHHhhchhhhhcEEcCCCCccccCccccCce
SDVDESTQRLKAGVVMALAAVGLTGFRFRKESK
cCChHHHHHHhhccccEEcccccccccccc

```

Longitud de secuencia : 873

PHD :

Hélice alfa	(Hh) :	525 es	60,14%
hélice 3 ₁₀	(Gg) :	0 es	0,00%
Hélice Pi	(Ii) :	0 es	0,00%
Puente Beta	(Bb) :	0 es	0,00%
Hebra prolongada	(Ee) :	63 es	7,22%
Giro Beta	(Tt) :	0 es	0,00%
Región de flexión	(Ss) :	0 es	0,00%
Enrollamiento al azar	(Cc) :	285 es	32,65%
Estados ambiguos	(?) :	0 es	0,00%

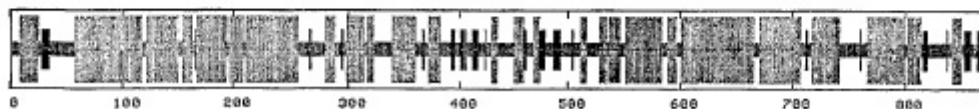
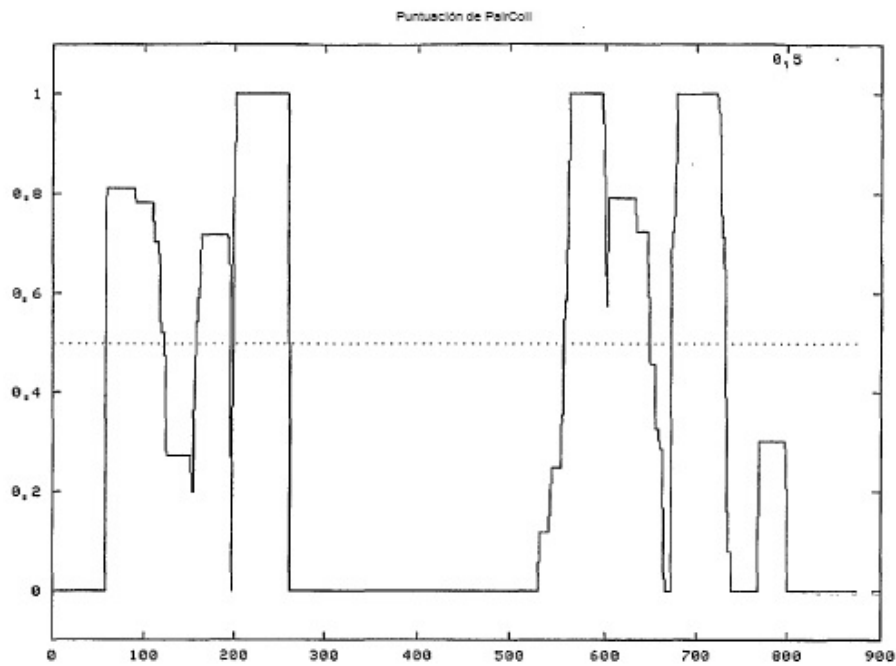


Figura 4(b): Predicción de estructura secundaria en base a puntuación de PairCoil



Superenrollamientos encontrados:

posiciones	58- 121	registro inicial	'f'	probabilidad	0,811
posiciones	156- 196	registro inicial	'e'	probabilidad	0,720
posiciones	198- 246	registro inicial	'e'	probabilidad	1,000
posiciones	247- 261	registro inicial	'a'	probabilidad	1,000
posiciones	556- 646	registro inicial	'e'	probabilidad	1,000
posiciones	671- 733	registro inicial	'f'	probabilidad	1,000

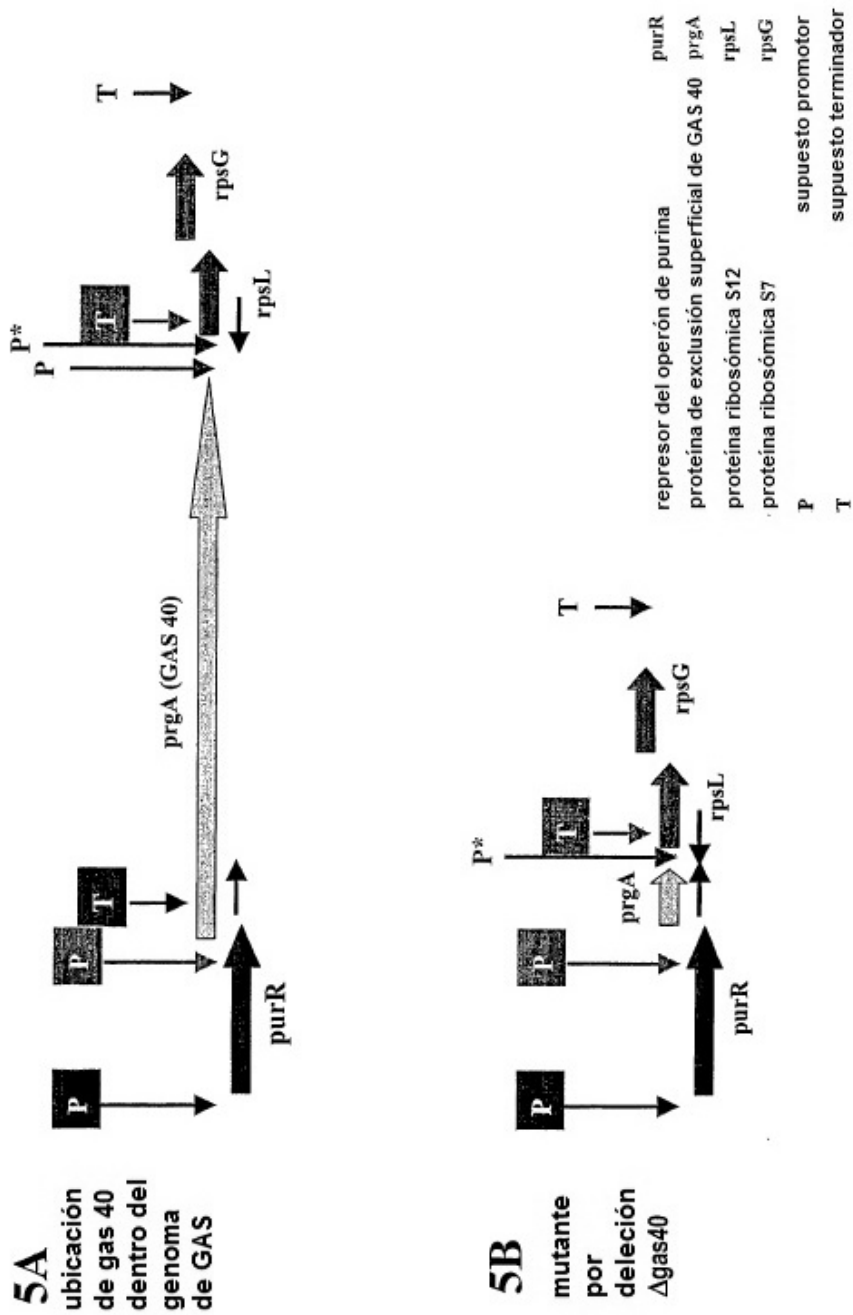
Figura 4(c): Predicción de estructura secundaria de cremallera de leucina dentro del superenrollamiento

```

673                               701
QYLRDFKLNPNRLQVIRERIDNTKQDLAKTTSSLNAQEALAAALQAKQSSLEATIATTEH
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
L-----I-----L-----L-----L-----L-----L-----L-----
OOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOO

```

Figura 5



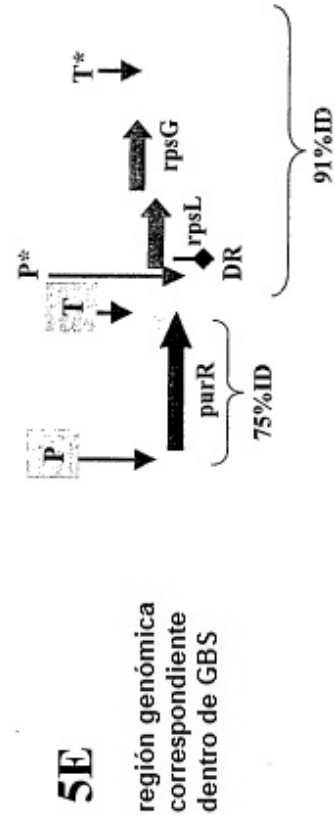
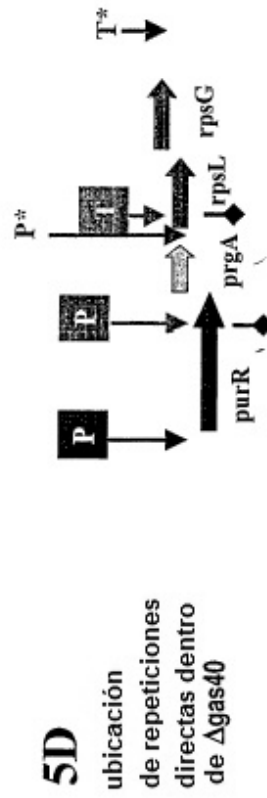
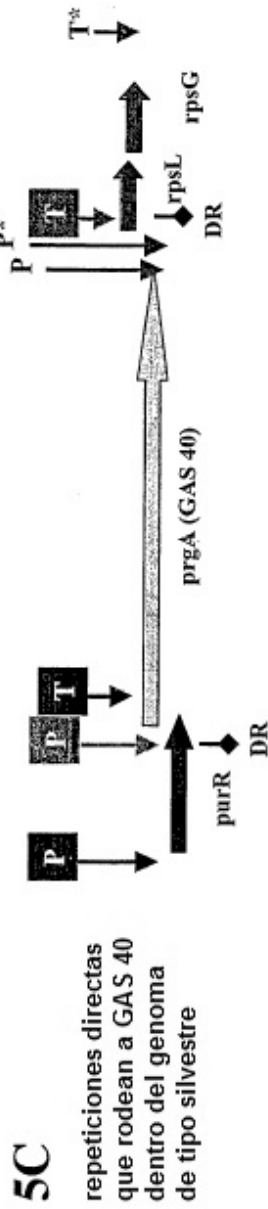


FIGURA 6: comparación FACS de GAS 40 en GAS de tipo silvestre y mutante por delección de GAS 40

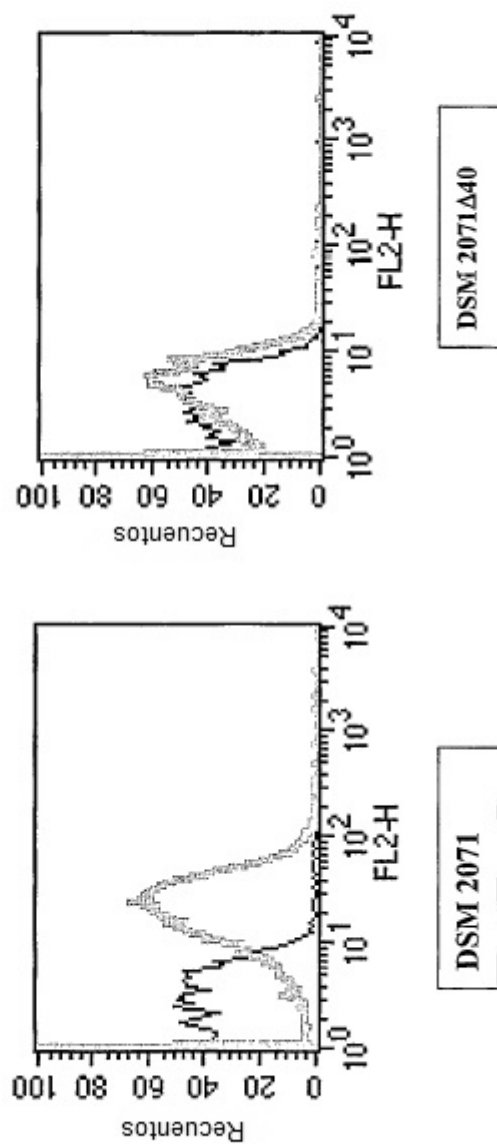
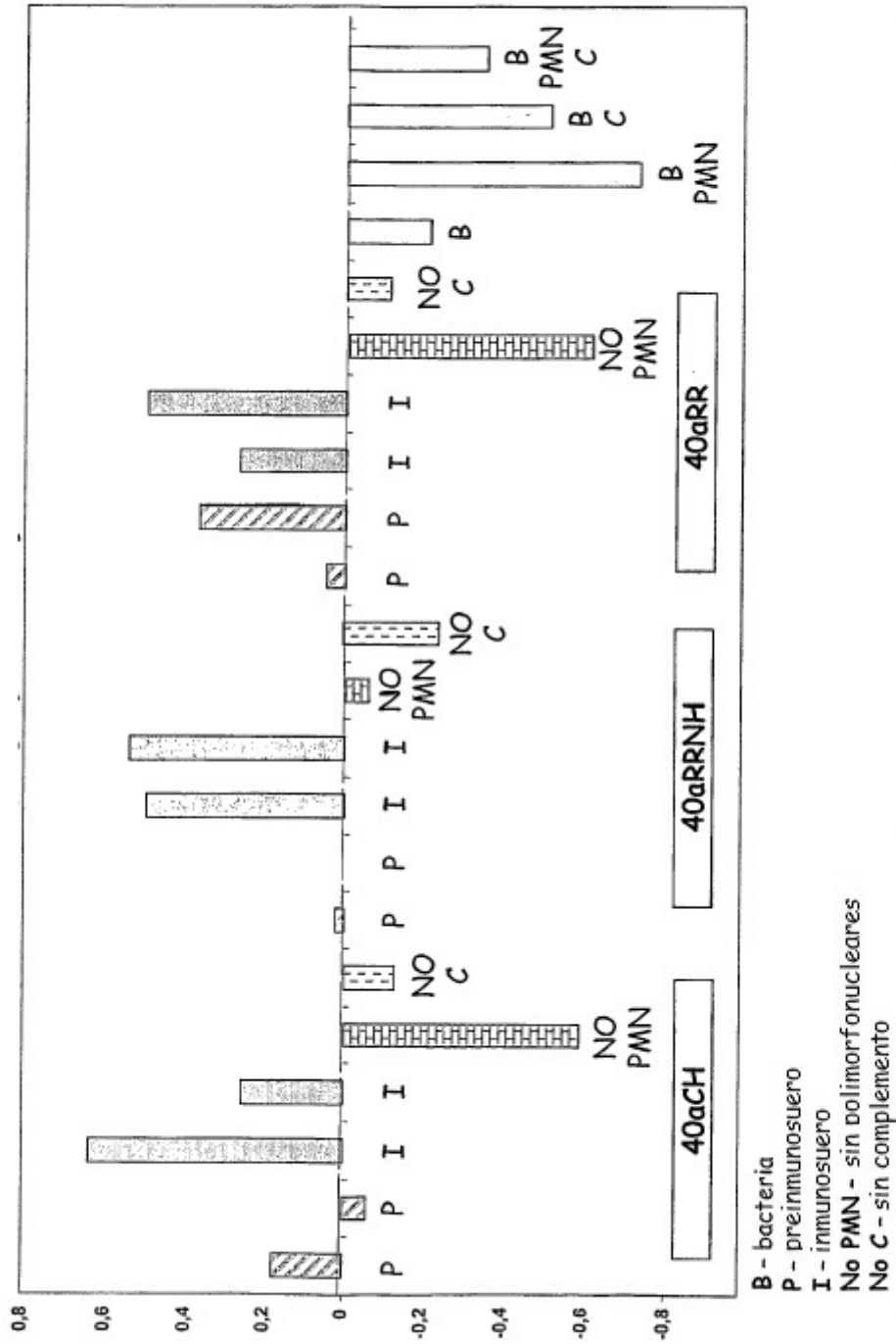
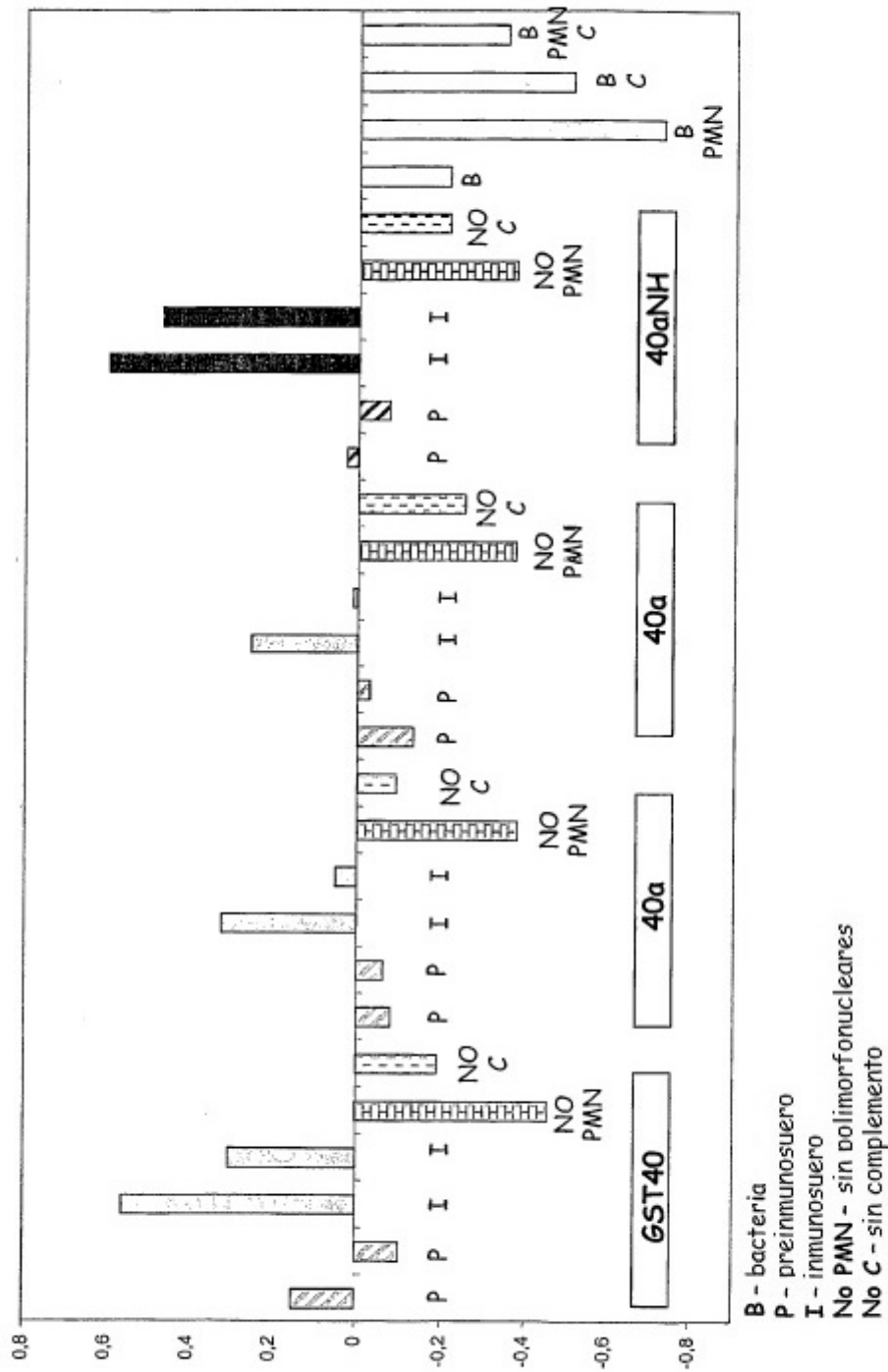


Figura 7(a): Ensayo de opsonofagocitosis bacteriana de constructos de GAS 40



Las barras del histograma representan la diferencia entre el logaritmo a T 0 (instante inicial) y T60 (UFC bacterianas contadas después de 60 minutos de incubación)

Figura 7(b): Ensayo de opsonofagocitosis bacteriana de constructos de GAS 40



Las barras del histograma representan la diferencia entre el logaritmo a T 0 (instante inicial) y T60 (UFC bacterianas contadas después de 60 minutos de incubación)

Figura 8: Inmunización en modelo murino (de ratón)

antígeno GAS	Supervivencia ratones ensayados			Protección	Valor p	Proteína pureza
	vivos	mueritos	ensayados			
				%	Chi-cuadrado	%
gst 40	67	63	130	51	0.000012	
253	14	36	50	28	0.006	15
253-urea	2	8	10	20		25
253-gst	2	8	10	20		30
39	9	31	40	22.5	0.09	20
39a	13	37	50	26	0.016	10
39a	10	30	40	25	0.039	
39a	12	28	40	30	0.0046	
urea 366	21	78	99	21.2	0.046	65
117	19	51	70	27	0.0036	15
117-urea	1	9	10	10		80
117-urea-2M	7	23	30	23.3	0.1	80
117-urea-2M (prep 117)	8	32	40	20	0.2	
urea 504	9	31	40	22.5	0.09	50
504	14	26	40	35	0.0003	40
504	7	33	40	17.5	0.4	80
urea 389	7	23	30	23	0.1	30
533	14	56	70	20	0.12	50
new 533	4	16	20	20	0.34	30
gst 57	12	48	60	20	0.14	60
57a	0	20	20	0		50
294	17	73	90	18.8	0.14	80
130	15	65	80	18.7	0.17	40
130	7	23	30	23.3	0.1	40
84	8	32	40	20	0.2	70
urea 159	7	33	40	17.5	0.4	5
159a	2	8	10	20		65
527	10	40	50	20	0.17	50
527	3	17	20	15		80
217	7	33	40	17.5	0.4	50
511	13	67	80	16.2	0.41	80
277	8	42	50	16	0.52	5
277a	2	28	30	6.6		50
gst 202	3	17	20	10	0.75	5
202a	5	25	30	16.6	0.53	5
45	5	25	30	16.6	0.53	80
urea 309	5	25	30	20	0.53	8
290	6	34	40	15	0.67	50
529	6	34	40	15	0.67	5
gst 58	10	60	70	14.2	0.71	30
384	7	43	50	14	0.78	80
384RR	1	19	20	5		80
urea 509	7	53	60	11.6	0.84	50
509-NH	2	8	10			75
509-CH	0	10	10			75
193	7	53	60	11.6	0.84	65
urea 372	4	25	29	13.7	0.85	20
gst 42	4	26	30	13.3	0.9	50
95	5	35	40	12.5	1	55
urea 236	5	35	40	12.5	1	80
new 236	2	8	10	20		70
137	5	35	40	12.5	1	75
Hls-Stop	29	201	230	12.96		