

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 552 773**

51 Int. Cl.:

C07K 1/113 (2006.01)

A61K 38/00 (2006.01)

C12N 15/63 (2006.01)

G01N 33/68 (2006.01)

G01N 33/15 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.02.2010 E 10746935 (5)**

97 Fecha y número de publicación de la concesión europea: **12.08.2015 EP 2403864**

54 Título: **Motivos estructurales de polipéptidos asociados con la actividad de señalización celular**

30 Prioridad:

27.02.2009 US 156370 P

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

02.12.2015

73 Titular/es:

**ATYR PHARMA, INC. (100.0%)
3565 General Atomics Court, Suite 103
San Diego, CA 92121, US**

72 Inventor/es:

GREENE, LESLIE ANN

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 552 773 T3

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

DESCRIPCIÓN

Motivos estructurales de polipéptidos asociados con la actividad de señalización celular

Antecedentes de la invención

Campo de la invención

- 5 La presente descripción se refiere en general a polipéptidos que contienen motivos estructurales asociados con la señalización celular y otras actividades biológicamente relevantes, a las composiciones que comprenden tales polipéptidos, y a los métodos de identificación y uso de los mismos.

Descripción de la técnica relacionada

10 Las proteínas aminoacil-ARNt sintetasas (AARS), que catalizan la aminoacilación de moléculas de ARNt, son esenciales para descodificar la información genética durante el proceso de traducción. Cada una de las ARNt sintetasas eucarióticas consiste en una enzima núcleo, que está íntimamente relacionada con las ARNt sintetasas procarionóticas, así como con dominios adicionales que están agregados al extremo amino terminal, al extremo carboxilo terminal o insertados a una región interna de la enzima núcleo. La tirosil-ARNt sintetasa (TyrRS) humana, por ejemplo, tiene un dominio carboxilo terminal que no es parte de moléculas de TyrRS procarionóticas ni eucarióticas inferiores. Se ha demostrado que varias aminoacil-ARNt sintetasas tienen funciones no canónicas distintas de su implicación en la traducción. Por ejemplo, la mini-tirosil ARNt sintetasa (mini-TyrRS), el dominio N-terminal de TyrRS que corresponde a los residuos de aminoácido 1-364 y es escindido por la elastasa y la plasmina de las células polimorfonucleares, muestra actividades no canónicas no encontradas en la proteína completa. *In vitro*, se ha demostrado que la mini-TyrRS estimula la activación y la quimiotaxis de neutrófilos, la proliferación y migración de las células endoteliales, y es pro-angiogénica en la membrana corioalantoica de pollo (CAM) y en análisis de matrigel de ratón. La mini-TyrRS tiene un motivo ELR que, como las quimioquinas CXC tales como IL-8, está implicado en muchas de sus actividades quimioquina y angiogénicas. Como en otras citoquinas que contienen ELR, la mutación de este motivo inhibe la unión de mini-TyrRS y la estimulación de leucocitos y la angiogénesis.

25 Además, se ha demostrado que las formas truncadas de TrpRS tienen propiedades anti-angiogénicas. En células humanas normales, existen dos formas de TrpRS que pueden ser detectadas: una forma principal que consiste en la molécula completa (residuos de aminoácido 1-471) y una forma minoritaria truncada. La forma minoritaria se genera por delección de un dominio amino terminal a través de un corte y empalme alternativo del pre-ARNm y se denomina mini-TrpRS. Se ha determinado que el extremo amino de miniTrpRS es el residuo de metionina de la posición 48 de la molécula de TrpRS completa. Alternativamente, la TrpRS truncada puede ser generada por proteólisis. Por ejemplo, la TrpRS bovina es altamente expresada en el páncreas y es secretada en el jugo pancreático, dando como resultado de ese modo la producción de una molécula de TrpRS truncada. Otros estudios indican que la miniTrpRS inhibe la proliferación y la migración celulares inducidas por VEGF (Wakasugi et al., Proc. Natl. Acad. Sci. 99: 173-177 (2002)). En particular, un análisis de CAM de pollo demuestra que la miniTrpRS bloquea la actividad angiogénica de VEGF. En contraste, la TrpRS completa no inhibe la angiogénesis. De este modo, la eliminación de los primeros 48 residuos de aminoácido expone la actividad anti-angiogénica de TrpRS. Por lo tanto, como con TyrRS, ciertas formas de TrpRS poseen actividades distintas de la aminoacilación de ARNt.

40 También se ha demostrado que otras proteínas citoplásmicas tienen isoformas con actividades biológicas extracelulares. Por ejemplo, la tioredoxina (Hu-Trx) humana cumple un papel clave en las actividades redox en el citoplasma de las células, pero tiene una forma secretada truncada que es una citoquina mitogénica (Arner, 2000; Eur J biochem 267:6102-6109).

Desafortunadamente, no hay métodos fiables para pronosticar o identificar qué proteínas citoplásmicas o fragmentos polipeptídicos derivados de proteínas AARS y/u otras proteínas pueden tener actividades novedosas y previamente no apreciadas. La presente descripción aborda estas necesidades y ofrece otras ventajas relacionadas.

45 Park et al. 2005 (Proc. Natl. Acad. Sci. 102: 6356-6361) describen que el polipéptido lisil-ARNt humano induce la secreción de TNF-alfa.

El documento US 2006/078553 describe composiciones que comprenden fragmentos de ARNt sintetasas, tales como Trp-ARNt sintetasa para inhibir la angiogénesis.

Yang et al. 2002 (Proc. Natl. Acad. Sci. 99: 15369-15374) describen la estructura de un fragmento de tyr-ARNt sintetasa que tiene tres hebras beta antiparalelas y dos hélices alfa y que tiene actividad citoquina.

50 Xie et al. 2007 (Proc. Natl. Acad. Sci. 104: 9976-9981) describen la estructura de GlyRS, que comprende tres hebras beta antiparalelas flanqueadas por dos hélices alfa.

Yadong et al. 2004 (J. Biol. Chem. 279: 8378-8388), Guijarro et al. 2002 (Structure 10: 311-317) y Goldgur et al. 1997 (Structure 5: 59-68) describen polipéptidos que contienen varias láminas beta y hélices alfa que muestran una actividad de señalización celular.

Datson et al. 2007 (BMC Genomics vol. 8, núm. 1, pág. 190) describen un fragmento de 180 aminoácidos de Lys-ARNt sintetasa.

Compendio de la invención

5 La presente descripción se refiere a polipéptidos aislados que comprenden o consisten esencialmente en motivos estructurales concretos, como se describe en la presente memoria, que muestran al menos una actividad de señalización celular y/u otra actividad no canónica de relevancia biológica. La presente descripción se refiere adicionalmente a polinucleótidos que codifican tales polipéptidos, agentes de unión que se unen a tales polipéptidos, análogos, variantes y fragmentos de tales polipéptidos, etc., así como composiciones y métodos de identificación y uso de cualquiera de los anteriores.

10 Por lo tanto, de acuerdo con un aspecto de la presente descripción, se proporcionan polipéptidos aislados que comprenden o consisten esencialmente en tres láminas β y dos hélices α , en donde los polipéptidos muestran una actividad de señalización celular y/u otra actividad no canónica relativa a la proteína de la cual se obtuvieron. El orden y la orientación precisos de las láminas β y las hélices α requeridas dentro de un polipéptido de la presente descripción pueden variar al tiempo que todavía originan las actividades de señalización celular y/u otras actividades no canónicas deseadas. Sin embargo, en una realización particularmente ilustrativa, un polipéptido de la presente descripción consiste esencialmente en tres láminas β antiparalelas flanqueadas en cada extremo por una hélice α .

15 El tamaño de un polipéptido de la presente descripción puede variar al tiempo que todavía conserva los rasgos estructurales y funcionales deseados descritos en la presente memoria. No obstante, en ciertas realizaciones ilustrativas de la presente descripción, un polipéptido aislado descrito en la presente memoria tendrá un tamaño en el intervalo de aproximadamente 40-400, 50-300 o 60-200 residuos de aminoácido.

20 En muchas realizaciones descritas en la presente memoria, un polipéptido de la presente descripción será un fragmento contiguo de una proteína de mamífero (p. ej., una proteína humana), o una secuencia que comparte una identidad estructural sustancial (p. ej., al menos 70%, 80% o 90%) con dicho fragmento contiguo. No obstante, en otras realizaciones de la presente descripción, se entenderá que el polipéptido puede estar formado por dos o más fragmentos no contiguos de una proteína, o secuencias que comparten una identidad de secuencia sustancial con dichos fragmentos no contiguos.

25 En una realización específica de la presente descripción, un polipéptido aislado descrito en la presente memoria es un polipéptido GlyRS (GRS) que comprende o consiste esencialmente en los residuos 345-438 del SEQ ID NO: 1, cuyo fragmento ha sido identificado en la presente memoria por contener un motivo estructural de la presente descripción y por tener una actividad de señalización celular. En una realización relacionada, el polipéptido aislado es un fragmento activo o una variante de un polipéptido que comprende los residuos 345-438 del SEQ ID NO: 1, p. ej., un fragmento del polipéptido que conserva la misma actividad de señalización celular o una actividad similar o una variante que comparte una identidad de secuencia sustancial con éste (p. ej., al menos 80% o 90%) y que conserva la misma actividad de señalización celular o una actividad similar.

30 En otra realización específica de la presente descripción, el polipéptido es un polipéptido AspRS (DRS) que comprende o consiste esencialmente en los residuos 367-448 de SEQ ID NO: 2 o un fragmento activo o una variante del mismo.

35 En otra realización específica de la presente descripción, el polipéptido es un polipéptido HisRS (HRS) que comprende o consiste esencialmente en los residuos 294-372 del SEQ ID NO: 3 o un fragmento activo o una variante del mismo.

40 En otra realización específica más de la presente descripción, el polipéptido es un polipéptido ThrRS (TRS) que comprende o consiste esencialmente en los residuos 469-586 del SEQ ID NO: 4 o un fragmento activo o una variante del mismo.

45 En otra realización específica de la presente descripción, el polipéptido es un polipéptido GluProRS (EPRS) que comprende o consiste esencialmente en los residuos 1171-1253 del SEQ ID NO: 5 o un fragmento activo o una variante del mismo.

50 En otra realización específica más de la presente descripción, el polipéptido es un polipéptido SerRS (SRS) que comprende o consiste esencialmente en los residuos 325-410 del SEQ ID NO: 6 o un fragmento activo o una variante del mismo.

55 En otra realización específica más de la presente descripción, el polipéptido es un polipéptido PheRS (FRS) que comprende o consiste esencialmente en los residuos 380-449 del SEQ ID NO: 7 o un fragmento activo o una variante del mismo.

En otra realización específica más de la presente descripción, el polipéptido es un polipéptido LysRS (KRS) que comprende o consiste esencialmente en los residuos 425-523 del SEQ ID NO: 8 o un fragmento activo o una variante del mismo.

En otra realización específica más de la presente descripción, el polipéptido es un polipéptido AspRS (NRS) que comprende o consiste esencialmente en los residuos 416-494 del SEQ ID NO: 9 o un fragmento activo o una variante del mismo.

5 En otra realización específica más de la presente descripción, el polipéptido es un polipéptido AlaRS (ARS) que comprende o consiste esencialmente en los residuos 148-258 del SEQ ID NO: 10 o un fragmento activo o una variante del mismo.

10 En otra realización específica de la presente descripción, el polipéptido es un polipéptido de tiorredoxina que comprende o consiste esencialmente en los residuos 20-105 del SEQ ID NO: 11 o un fragmento activo o una variante del mismo. En otra realización, el polipéptido es un polipéptido Trx80 que comprende o consiste esencialmente en los residuos 20-84 del Trx80, que es una forma secretada de tiorredoxina que contiene los 84 residuos de aminoácido N-terminales de la tiorredoxina.

En otra realización específica de la presente descripción, el polipéptido es un polipéptido del factor inhibidor de macrófagos que comprende o consiste esencialmente en los residuos 1-90 del SEQ ID NO: 12 o un fragmento activo o una variante del mismo.

15 En otra realización específica más de la presente descripción, el polipéptido es un polipéptido de isoforma B de peroxirredoxina 5 humana que comprende o consiste esencialmente en los residuos 32-68 y 125-161 del SEQ ID NO: 13 o un fragmento activo o una variante del mismo.

20 En ciertas realizaciones descritas en la presente memoria, la actividad de señalización celular de un polipéptido de la presente descripción es una actividad quimioquina y/o citoquina y se puede determinar y/o confirmar utilizando cualquiera de una variedad de análisis ilustrativos conocidos y establecidos en la técnica. En realizaciones específicas descritas en la presente memoria, la actividad de señalización celular se puede determinar utilizando esencialmente cualquier análisis que mida la actividad citoquina, la actividad quimioquina, la quimiotaxis, la migración celular, la liberación de citoquina, la diferenciación celular y/o la toxicidad celular. En realizaciones más específicas descritas en la presente memoria, la actividad citoquina se puede determinar, por ejemplo, utilizando un análisis que mide la quimiotaxis de monocitos, la liberación de interleuquinas, y/o la apoptosis dependientes de GPCR.

30 Ciertas realizaciones descritas en la presente memoria incluyen métodos de modulación de la actividad quimioquina, tales como la secreción de TNF- α . Las realizaciones específicas descritas en la presente memoria incluyen métodos de inducción de la secreción de TNF- α . Ciertas realizaciones descritas en la presente memoria incluyen métodos de modulación de la quimiotaxis de las células inmunitarias. Las realizaciones específicas descritas en la presente memoria incluyen métodos de inducción de la quimiotaxis de monocitos. Ciertas realizaciones descritas en la presente memoria incluyen métodos de modulación de la señalización de receptores de tipo Toll. Las realizaciones específicas descritas en la presente memoria incluyen métodos de inducción de la señalización de receptores 2 de tipo Toll.

35 De acuerdo con otro aspecto de la presente descripción, se proporciona un polinucleótido aislado que codifica un polipéptido descrito en la presente memoria.

De acuerdo con otro aspecto más, la presente descripción proporciona una composición farmacéutica que comprende un polipéptido como se describe en la presente memoria, o un polinucleótido que codifica un polipéptido como se describe en la presente memoria.

40 De acuerdo con otro aspecto, la presente descripción proporciona a vector que comprende un polinucleótido aislado como se describe en la presente memoria, así como también una célula anfitriona que comprende semejante vector.

45 De acuerdo con otro aspecto más, la presente descripción proporciona métodos de escrutinio para identificar proteínas que contienen un motivo estructural como se describe en la presente memoria y de ese modo identificar fragmentos de polipéptidos novedosos dentro de dichas proteínas que poseen actividades de señalización celular y/u otras actividades no canónicas. Por ejemplo, en una realización, la presente descripción proporciona un método para identificar un fragmento de polipéptido que tiene una actividad citoquina identificando una secuencia de proteína que contiene un motivo estructural como se describe en la presente memoria, *p. ej.*, formado por dos hélices α y tres láminas β , determinando los límites de los residuos de aminoácido del motivo estructural dentro de dicha proteína, e identificando de ese modo un fragmento de polipéptido de la proteína que tiene una actividad citoquina.

50 En ciertas realizaciones de los métodos de escrutinio descritos, la etapa de identificación de una secuencia de proteína que contiene un motivo estructural formado por dos hélices α y tres láminas β anti-paralelas se lleva a cabo utilizando un método de predicción de la estructura secundaria conocido y disponible en la técnica, que puede incluir de manera ilustrativa, pero no limitado a, PHDsec, NSSP, SOPM, DSC, SSPRED, MultiPredict, PSA, NNPREDICT, APSSP, GOR, HNN, HTMSRAP, Jpred, JUFO, nnPredict, Porter, PredictProtein, Prof, PSIPred, SOPMA, SSpro y DLP-SVM.

En otros aspectos más de la presente descripción, los polipéptidos, anticuerpos y/u otras composiciones descritos en la presente memoria se pueden utilizar esencialmente en cualquier tipo de análisis de escrutinio conocido y disponible en la técnica. Por ejemplo, las composiciones de la presente descripción (*p. ej.*, polipéptidos, polinucleótidos y/o anticuerpos) se pueden utilizar junto con cualquier metodología de escrutinio esencialmente conocida con el fin de identificar tipos de células adecuados y/o estados de enfermedad susceptibles de tratamiento de acuerdo con la presente descripción. En otros ejemplos, las composiciones de la presente descripción (*p. ej.*, polipéptidos, polinucleótidos y/o anticuerpos) se pueden utilizar junto con metodologías de escrutinio bien conocidas con el fin de identificar compañeros de unión, inhibidores competitivos, efectores celulares, y similares, que median o modulan, bien directamente bien indirectamente, las actividades de señalización celular y/o actividades no canónicas de las composiciones de la presente memoria. Por ejemplo, en una realización concreta, se proporciona un método de escrutinio para identificar compuestos de ensayo como inhibidores, o alternativamente, potenciadores, de una interacción entre una composición de la presente descripción y uno o más de sus compañeros de unión, efectores celulares y/o tipos de células sujetos a modulación. Esto puede incluir, por ejemplo, etapas de formación de una mezcla de reacción que incluye: (i) una composición de la presente descripción, (ii) un compañero de unión, un efector celular y/o un tipo de célula que se sabe que están modulados por dicha composición, y (iii) un compuesto de ensayo; y detectar la interacción del compuesto de ensayo con el compañero de unión, el efector celular y/o el tipo de célula. Un cambio estadísticamente significativo (potenciación o inhibición) en la actividad o la modulación en presencia del compuesto de ensayo, con respecto a la interacción en ausencia del compuesto de ensayo, indica un agonista potencial (mimético o potenciador) o antagonista (inhibidor) de la actividad.

20 Breve descripción de los identificadores de secuencia

SEQ ID NO: 1 es la secuencia de aminoácidos completa de la glicil-ARNt sintetasa (GlyRS) humana.
 SEQ ID NO: 2 es la secuencia de aminoácidos completa de aspartil-ARNt sintetasa (AspRS) humana.
 SEQ ID NO: 3 es la secuencia de aminoácidos completa de la histidil-ARNt sintetasa (HisRS) humana.
 SEQ ID NO: 4 es la secuencia de aminoácidos completa de la treonil-ARNt sintetasa (ThrRS) humana.
 25 SEQ ID NO: 5 es la secuencia de aminoácidos completa de la glutamil/prolil-ARNt sintetasa (GluProRS) humana.
 SEQ ID NO: 6 es la secuencia de aminoácidos completa de la seril-ARNt sintetasa (SerRS) humana.
 SEQ ID NO: 7 es la secuencia de aminoácidos completa de la fenilalanil-ARNt sintetasa (FRSa) humana.
 SEQ ID NO: 8 es la secuencia de aminoácidos completa de la lisil-ARNt sintetasa (KRS) humana.
 SEQ ID NO: 9 es la secuencia de aminoácidos completa de la asparraginil-ARNt sintetasa (NRS) humana.
 30 SEQ ID NO: 10 es la secuencia de aminoácidos completa de la alanil-ARNt sintetasa (ARS) humana.
 SEQ ID NO: 11 es la secuencia de aminoácidos completa de la tiorredoxina humana.
 SEQ ID NO: 12 es la secuencia de aminoácidos completa del factor inhibidor de macrófagos humano.
 SEQ ID NO: 13 es la secuencia de aminoácidos completa de la isoforma B de la peroxirredoxina 5 humana.

Breve descripción de los dibujos

35 La FIG. 1 muestra un motivo estructural ilustrativo de la presente descripción que comprende tres láminas β antiparalelas flanqueadas por una hélice α en cada extremo.

Las FIG. 2A-2C muestran la identificación de un fragmento de GlyRS humana que contiene un motivo estructural de la presente descripción y la demostración de que el fragmento tiene actividad quimioquina. La FIG. 2A muestra la localización del fragmento G6 y el motivo estructural dentro de GlyRS completa. La FIG. 2B muestra un modelo estructural de los residuos 344-420 de GlyRS. La FIG. 2C muestra que el fragmento G6 induce la migración de monocitos THP-1 dependiente de GPCR.

40 La FIG. 3 muestra que el motivo estructural identificado se encuentra en otras proteínas AARS de clase II, que incluye GlyRS (GRS), ThrRS (TRS), HisRS (HRS), AspRS (DRS), GluProRS (EPRS), SerRS (SRS), AsnRS (NRS), AlaRS (ARS), PheRS (FRS) y LysRS (KRS).

45 La FIG. 4 muestra la conservación estructural de un motivo de la presente descripción a través de múltiples proteínas AARS de clase II, a pesar de su relativamente baja conservación de secuencia.

La FIG. 5 muestra que el motivo estructural de la presente descripción también se encuentra en proteínas que no son AARS, incluyendo (A) tiorredoxina humana, (B) proteína inhibidora de macrófagos (residuos 1-90), y (C) isoforma B de peroxirredoxina 5 (residuos 32-68 y 125-161). Una característica de las proteínas que comparten este motivo estructural es que se secretan de una manera no convencional.

La FIG. 6 muestra que un motivo estructural de la presente descripción se encuentra dentro del fragmento de una seril ARNt sintetasa (SerRS) humana que muestra actividad de señalización celular. Esta figura muestra la localización del fragmento S3 (residuos 253-484 de la SerRS completa) y el motivo estructural (residuos 325-410) dentro de la secuencia de polipéptidos de SerRS completa.

5 Las FIG. 7A-7C muestran que el fragmento S3 (residuos 253-484 de SerRS completa) que contiene un motivo estructural de la presente descripción se une a monocitos y células B humanos y estimula la secreción de TNF- α procedente de una línea celular monocítica humana. La FIG. 7A muestra la unión de S3 a monocitos humanos, la FIG. 7B muestra la unión de S3 a células B humanas, y la FIG. 7C muestra la inducción de la secreción de TNF- α procedente de una línea celular monocítica (THP-1) humana en comparación con un control de LPS.

10 La Figura 8 muestra que el fragmento S3 señalizaba a través de un receptor 2 de tipo Toll para lograr una secreción fuerte de fosfatasa alcalina de una manera dependiente de la dosis (eje x = dosificación en nM; eje y = DO₆₀₀).

La Figura 9 muestra que la señalización mediada por S3 a través del receptor 2 de tipo Toll (TLR2) era inhibida por el pre-tratamiento con un anticuerpo monoclonal que bloquea la unión a mTLR2.

Descripción detallada de la invención

15 La presente invención es como se proporciona en las reivindicaciones.

La práctica de la presente descripción empleará, a no ser que se indique específicamente lo contrario, métodos convencionales de biología molecular y técnicas de ADN recombinante que están dentro del conocimiento práctico de la técnica, muchos de los cuales se describen más adelante con fines ilustrativos. Dichas técnicas se explican totalmente en la bibliografía. Véanse, por ejemplo, Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (2ª Edición, 1989); Maniatis et al., *Molecular Cloning: A Laboratory Manual* (1982); *ADN Cloning: A Practical Approach*, vol. I y II (D. Glover, ed.); *Oligonucleotide Synthesis* (N. Gait, ed., 1984); *Nucleic Acid Hybridization* (B. Hames y S. Higgins, eds., 1985); *Transcription and Translation* (B. Hames y S. Higgins, eds., 1984); *Animal Cell Culture* (R. Freshney, ed., 1986); *A Practical Guide to Molecular Cloning* (B. Perbal, ed., 1984).

20 Según se utiliza en esta memoria descriptiva y en las reivindicaciones adjuntas, las formas en singular "un," "una", "uno" y "el" y "la" incluyen las referencias en plural a menos que el contenido dicte claramente lo contrario.

A lo largo de esta memoria descriptiva, a no ser que el contexto requiera lo contrario, se entenderá que las palabras "comprenden", "comprende", y "que comprende" implican la inclusión de una etapa o elemento o grupo de etapas o elementos indicado pero no la exclusión de cualquier otra etapa o elemento o grupo de etapas o elementos.

30 Por "consiste en" se quiere significar que incluye, y se limita a, lo que sea que sigue la expresión "que consiste en". Así, la expresión "que consiste en" indica que los elementos listados se requieren o son obligatorios, y que no pueden estar presentes otros elementos. Por "consiste esencialmente en" se quiere significar que incluye cualesquiera elementos listados después de la expresión, y que está limitado a otros elementos que no interfieren con o contribuyen a la actividad o acción especificada en la descripción para los elementos listados. Así, la expresión "que consiste esencialmente en" indica que los elementos listados se requieren o son obligatorios, pero que otros elementos son opcionales y pueden o no estar presentes dependiendo de si afectan o no materialmente a la actividad o acción de los elementos listados.

35 Un "agonista" se refiere a una molécula que intensifica o imita la actividad biológica no canónica de una AARS. Los agonistas pueden incluir proteínas, ácidos nucleicos, carbohidratos, pequeñas moléculas, o cualquier otro compuesto o composición que module la actividad de una AARS o bien interaccionando directamente con la AARS o su compañero de unión, o bien actuando sobre componentes de la ruta biológica en la cual participa la AARS. Se incluyen los agonistas parciales y totales.

40 El término "antagonista" se refiere a una molécula que inhibe o atenúa la actividad biológica no canónica de una AARS. Los antagonistas pueden incluir proteínas tales como anticuerpos, ácidos nucleicos, carbohidratos, moléculas pequeñas, o cualquier otro compuesto o composición que module la actividad de una AARS o su compañero de unión, o bien interaccionando directamente con la AARS o su compañero de unión o bien actuando sobre los componentes de la ruta biológica en la cual participa la AARS. Se incluyen los antagonistas parciales y totales.

45 El término "modular" incluye "incrementar" o "estimular", así como también "disminuir" o "reducir", típicamente en una cantidad estadísticamente significativa o fisiológicamente significativa en comparación con un control. Una cantidad "incrementada" o "aumentada" es típicamente una cantidad "estadísticamente significativa", y puede incluir un incremento que es 1,1, 1,2, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30 o más veces (*p. ej.*, 500, 1000 veces) (incluyendo todos los números enteros y puntos decimales intermedios y superiores a 1, *p. ej.*, 1,5, 1,6, 1,7, 1,8, etc.) la cantidad producida sin composición (en ausencia de un agente o compuesto) o una composición de control. Una cantidad "disminuida" o reducida es típicamente una cantidad "estadísticamente significativa", y puede incluir un descenso de 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, o 100% en la cantidad producida sin

composición (en ausencia de un agente o compuesto) o una composición de control, incluyendo todos los números enteros intermedios.

5 Un "sujeto", según se utiliza en la presente memoria, incluye cualquier animal que muestra un síntoma, o se encuentra en riesgo de mostrar un síntoma, que puede ser tratado o diagnosticado con una composición de la presente descripción. Los sujetos adecuados (pacientes) incluyen animales de laboratorio (tales como ratón, rata, conejo, o cobaya), animales de granja, y animales domésticos o mascotas (tales como un gato o perro). Están incluidos los primates no humanos y, preferiblemente, los pacientes humanos.

10 "Tratamiento" o "tratar", según se utiliza en la presente memoria, incluye cualquier efecto deseable sobre los síntomas o la patología de una enfermedad o afección que se puede ver afectada por las actividades moduladoras de la célula de una composición como se describe en la presente memoria, y puede incluir incluso cambios mínimos o mejoras en uno o más marcadores medibles de la enfermedad o afección que están siendo tratadas. "Tratamiento" o "tratar" no indican necesariamente la erradicación completa o cura de la enfermedad o afección, o síntomas asociados de ésta. El sujeto que recibe este tratamiento es cualquier sujeto que lo necesite. Los marcadores ilustrativos de mejora clínica serán evidentes para los expertos en la técnica.

15 Según se utilizan en la presente memoria, los términos "polipéptido" y "proteína" se utilizan de acuerdo con el significado convencional, es decir, como una secuencia de aminoácidos. Los polipéptidos no se limitan a una longitud específica, pero, en el contexto de la presente descripción, representan típicamente un fragmento de una proteína completa, y pueden incluir modificaciones post-traduccionales, por ejemplo, glicosilaciones, acetilaciones, fosforilaciones y similares, así como también otras modificaciones conocidas en la técnica, tanto de origen natural
20 como de origen no natural. Los polipéptidos y proteínas de la presente descripción se pueden preparar utilizando cualquiera de una variedad de mecanismos recombinantes y/o sintéticos bien conocidos, cuyos ejemplos ilustrativos se comentan adicionalmente más abajo.

25 La presente descripción se refiere en general a la identificación de motivos estructurales de polipéptidos asociados con actividades biológicas inesperadas de relevancia terapéutica. Por ejemplo, como se describe adicionalmente en la presente memoria, se identificó un motivo estructural de proteína dentro de un fragmento aislado de la proteína glicil-ARNt sintetasa (GRS) humana, y se cree que este motivo estructural es responsable de la actividad de señalización celular inesperada que se ha encontrado que es inducida por el fragmento GlyRS. Otro análisis del motivo estructural identificado reveló su presencia no solamente en otras proteínas AARS humanas, sino también en
30 otras proteínas no relacionadas con la síntesis de ARNt. De este modo, basándose en este descubrimiento, la presente descripción proporciona un medio para identificar fragmentos de polipéptidos que tenían actividades de señalización celular y/u otras actividades biológicas no reconocidas previamente, y proporciona adicionalmente los fragmentos de polipéptidos identificados de ese modo. Por ejemplo, en una realización ilustrativa, al identificar la presencia de un motivo estructural como se describe en la presente memoria dentro de proteínas citoplásmicas que están siendo escrutadas, es posible delinear fragmentos de polipéptidos secretados novedosos de esas proteínas
35 citoplásmicas que tienen actividades no apreciadas previamente no relacionadas con sus actividades citoplásmicas canónicas.

40 Por consiguiente, la presente descripción proporciona en sentido amplio polipéptidos aislados que contienen un motivo estructural como se describe en la presente memoria, en donde los polipéptidos muestran al menos una actividad de señalización celular u otra actividad no canónica (*p. ej.*, una actividad que no se observa para la proteína completa a partir de la cual se obtuvo el polipéptido o una actividad que es significativamente diferente (*p. ej.*, incrementada o disminuida) con respecto a la proteína completa de la cual se obtuvo).

45 En una realización concreta de la presente descripción, se proporcionan polipéptidos aislados que contienen un motivo estructural formado por dos hélices α y tres láminas β , en donde los polipéptidos muestran al menos una actividad de señalización u otra actividad no canónica. En una realización específica de la presente descripción, se proporcionan polipéptidos aislados que contienen un motivo estructural formado por dos hélices α y tres láminas β antiparalelas, en donde los polipéptidos muestran al menos una actividad de señalización celular u otra actividad no canónica. En una realización incluso más específica de la presente descripción, se proporcionan polipéptidos aislados que contienen un motivo estructural formado por tres láminas β antiparalelas flanqueadas por una hélice α en cada extremo (*p. ej.*, como se representa ilustrativamente en la FIG.1), en donde los polipéptidos muestran al
50 menos una actividad de señalización celular u otra actividad no canónica.

55 Un motivo estructural de la presente descripción, *p. ej.*, que tiene tres láminas β y dos hélices α , se puede identificar dentro de una proteína utilizando esencialmente cualquiera de una variedad de metodologías establecidas y bien conocidas disponibles en la técnica para la predicción de la estructura secundaria de una proteína. La predicción de la estructura secundaria consiste en un conjunto de técnicas que pueden predecir con una elevada exactitud las estructuras secundarias locales de las proteínas basándose en el conocimiento de su secuencia de aminoácidos primaria y/u otros parámetros. Para las proteínas, una predicción consiste generalmente en la asignación de regiones de la secuencia de aminoácidos que es probable que adopten una estructura secundaria concreta, *p. ej.*, hélices α , hebras β , láminas β , etc. El éxito de la predicción se puede evaluar, por ejemplo, comparándola con los resultados del algoritmo DSSP aplicado a la estructura cristalina de la proteína (si estuviera disponible). Se han desarrollado algoritmos especializados para la detección de estos y otros patrones bien definidos en las proteínas.
60

Por consiguiente, utilizando tales metodologías, se puede predecir con gran exactitud la presencia de motivos estructurales de la presente descripción dentro de las proteínas que están siendo escrutadas utilizando cualquiera de los numerosos programas disponibles públicamente y bien conocidos y/o herramientas de modelado para la predicción de la estructura secundaria, de los cuales se exponen varios ejemplos ilustrativos más abajo:

- 5 • ESyPred3D (<http://www.fundp.ac.be/sciences/biologie/urbm/bioinfo/esypred/>), Lambert C, Leonard N, De Bolle X, Depiereux E. ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics*. 2002 Sep;18(9):1250-1256, que es un programa de modelado por homología automatizado;
- I_Tasser (<http://zhang.bioinformatics.ku.edu/I-TASSER/>), que construye modelos basados en alineamientos múltiples por diseño de homología remota;
- 10 • Bhageerath (<http://www.scfbio-iitd.res.in/bhageerath/index.jsp>) un servidor de predicción de la estructura de proteínas basado en la energía;
- PHDsec (<http://www.embl-heidelberg.de/predictprotein/>), que es un sistema de red neuronal basado en el alineamiento múltiple;
- 15 • NSSP (<http://dot.imgen.bcm.tmc.edu:9331/pssp/pssp.html>), que es un método del vecino más próximo basado en el alineamiento múltiple;
- SOPM (<http://www.ibcp.fr/predict.html>), que es un método basado en el alineamiento múltiple que combina diversos programas de predicción;
- DSC (http://bonsai.lif.icnet.uk/bmm/dsc/dsc-read_align.html), que es un programa basado en el alineamiento múltiple que utiliza la estadística;
- 20 • SSPRED: (http://www.embl-heidelberg.de/ssp/ssp_mul.html), que es un programa basado en el alineamiento múltiple que utiliza la estadística;
- MultiPredict (<http://kestrel.ludwig.ucl.ac.uk/zpred.html>), que es un método basado en el alineamiento múltiple que utiliza la información fisicoquímica de un conjunto de secuencias alineadas y constantes de decisión estadísticas de la estructura secundaria;
- 25 • PSA (<http://bmerc-www.bu.edu/psa/>), que analiza las secuencias de aminoácidos para predecir las estructuras secundarias y las clases de plegamientos;
- NNPREDICT (<http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html>), que es una predicción de red neuronal basada en la secuencia individual;
- 30 • APSSP (<http://imtech.res.in/raghava/apssp/>), que es conocido como Advanced Protein Secondary Structure Prediction Server;
- GOR (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html), que cuenta con Teoría de Información/Inferencia Bayesiana (Garnier et al, 1996);
- HNN (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_nn.html) que cuenta con un método Hierarchical Neural Network (Guermeur, 1997);
- 35 • HTMSRAP (<http://biotechnology.tbzmed.ac.ir/htmsrap/index.htm>), que emplea Helical TransMembrane Segment Rotational Angle Prediction;
- Jpred (<http://www.compbio.dundee.ac.uk/~www-jpred/>), que implica la predicción de la estructura secundaria basada en una red neuronal;
- 40 • JUFO (<http://www.meilerlab.org/view.php?section=0&page=6>), que implica la predicción de la estructura secundaria de proteínas basada en una red neuronal;
- nnPredict (<http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html>);
- Porter (<http://distill.ucd.ie/porter/>);
- PredictProtein (<http://www.predictprotein.org/>), que incluye PHDsec, PHDacc, PHDhtm, PHDtopology, PHDthreader, MaxHom y EvalSec;
- 45 • Prof (<http://www.aber.ac.uk/~phiwww/prof/>), que utiliza clasificadores múltiples en cascada para la predicción de la estructura secundaria;

- *PSIpred* (<http://bioinf.cs.ucl.ac.uk/psipred/>), que incluye diversos métodos de predicción de la estructura de proteínas;
- *SOPMA* (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) (Geourjon y Deléage, 1995)
- 5 • *SSpro* (<http://www.igb.uci.edu/?page=tools&subPage=psss>), que incluye una predicción de la estructura secundaria utilizando redes neuronales recurrentes bidireccionales; y,
- *DLP-SVM* (<http://www.tuat.ac.jp/~domserv/cgi-bin/DLP-SVM.cgi>), que implica una predicción de conectores de dominios utilizando SVM.

Se puede encontrar información adicional referente a algunos de estos y otras metodologías de predicción de la estructura secundaria para secuencias individuales, por ejemplo, en Chou et al. (1974) *Biochemistry*, 13, 211-222; Lim (1974) *Journal of Molecular Biology*, 88, 857-872; Garnier et al. (1978) *Journal of Molecular Biology*, 120, 97-120; Kabsch et al. (1983) *FEBS Letters*, 155, 179-182; Deleage et al. (1987) *Protein Engineering*, 1, 289-294 (DPM; http://www.ibcp.fr/serv_pred.html); Presnell et al. (1992) *Biochemistry*, 31, 983-993; Holley et al. (1989) *Proceedings of the National Academy of Science*, 86, 152-156; King et al. (1990). *Journal of Molecular Biology*, 216, 441-457; y Kneller et al. (1990) *Journal of Molecular Biology*, 214, 171-182 (NNPRED; <http://www.cmpfarm.ucsf.edu/~nomi/nnpredict.html>).

Además, se puede encontrar información adicional referente a algunos de estos y otros métodos automatizados para predecir la estructura secundaria de secuencias de proteínas alineadas de manera múltiple, por ejemplo, en Zvelebil et al. (1987) *Journal of Molecular Biology*, 195, 957-961 (ZPRED); Rost et al. (1993) *Journal of Molecular Biology*, 232, 584-599 (PHD; <http://www.embl-heidelberg.de/predictprotein/predictprotein.html>); Salamov et al. (1995) *Journal of Molecular Biology*, 247, 1 (NNSSP; <http://dot.imgen.bcm.tmc.edu:9331/seq-search/struc-predict.html>); Geourjon et al. (1994), *Protein Engineering*, 7, 157-16 (SOPMA; http://www.ibcp.fr/serv_pred.html); Solov'yev et al. (1994) *Computer Applications in the Biosciences*, 10, 661-669. (SSP; <http://dot.imgen.bcm.tmc.edu:9331/seq-search/struc-predict.html>); Wako et al. (1994) *Journal of Molecular Biology*, 238, 693-708; Mehta et al. (1995) *Protein Science* 4, 2517-25 (SSPRED; http://www.embl-heidelberg.de/sspred/sspred_info.html); y King et al. (1996) *Protein Sci* 5, 2298-2310. (DSC; http://www.bmm.icnet.uk/dsc/dsc_form_align.html).

También se puede utilizar un soporte lógico de visualización de la estructura, si se desea, junto con estructuras cristalinas disponibles públicamente para proteínas AARS o distintas de AARS (p. ej., a través de la base de datos Swiss Prot en el servidor NIH Entrez) para visualizar una proteína de interés e identificar motivos estructurales (p. ej., Accelrys Software Inc., Discovery Studio Modeling Environment, Release 2.0, San Diego: Accelrys Software Inc., 2007). Otras herramientas de soporte lógico ilustrativas disponibles para identificar motivos estructurales secundarios incluyen ESyPred3d (Lambert 2002; ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* 18:1250-1256). Las predicciones de la estructura se pueden inspeccionar después visualmente (p. ej., en Accelrys Discovery Visualizer) para determinar los dominios peptídicos que tienen el motivo estructural deseado.

Después de identificar una proteína o polipéptido que contiene un motivo estructural deseado como se describe en la presente memoria (p. ej., que contiene tres láminas β flanqueadas en cada extremo por hélices α), utilizando una o más técnicas adecuadas tales como las comentadas más arriba, se determinan los límites de los aminoácidos del motivo y se pueden producir fragmentos de polipéptidos aislados que contienen el motivo (p. ej., sintéticamente o recombinantemente) y se someten a ensayo para la confirmación de la actividad de señalización celular.

Se dice que un polipéptido de la presente descripción tiene una "actividad de señalización celular" cuando el polipéptido muestra una o más actividades comúnmente asociadas con y/o inducidas por proteínas de señalización celular, ya sea directamente o indirectamente. En ciertas realizaciones de la presente descripción, la actividad de señalización celular es una actividad citoquina y/o quimioquina. Las citoquinas son una categoría de moléculas de señalización celular y las quimioquinas son una familia de citoquinas pequeñas, o proteínas secretadas por las células.

Las proteínas se clasifican generalmente como quimioquinas de acuerdo con características estructurales compartidas tales como el tamaño pequeño (todas tienen un tamaño de aproximadamente 8-10 kilodalton), con la presencia de cuatro residuos de cisteínas en localizaciones conservadas que son clave para formar su conformación tridimensional. Una actividad distintiva de las quimioquinas es su capacidad para actuar como quimioatrayentes para inducir y/o controlar la migración de las células. Las células que son atraídas por las quimioquinas generalmente siguen una señal de concentración que quimioquina creciente hacia la fuente de la quimioquina.

Las quimioquinas sirven para varios papeles biológicos. Algunas quimioquinas controlan las células del sistema inmunitario durante los procesos de vigilancia inmunitaria, tales como el direccionamiento de linfocitos a los ganglios linfáticos de manera que se puedan escrutar para determinar la invasión de patógenos interactuando con células presentadoras de antígenos que residen en estos tejidos. Estas son conocidas como quimioquinas homeostáticas y son producidas y secretadas sin ninguna necesidad de estimular sus células de origen. Algunas quimioquinas tienen papeles en el desarrollo, tales como la promoción de la angiogénesis (el crecimiento de nuevos vasos sanguíneos),

o la orientación de células a tejidos que proporcionan señales específicas críticas para la maduración celular. Otras quimioquinas son inflamatorias y son liberadas de una amplia variedad de células en respuesta a la infección bacteriana, a virus y a agentes que causan un daño físico. Su liberación es estimulada a menudo por citoquinas pro-inflamatorias tales como la interleuquina 1. Las quimioquinas inflamatorias funcionan principalmente como quimioatrayentes para los leucocitos, reclutando monocitos, neutrófilos y otras células efectoras desde la sangre a sitios de infección o lesión tisular. Ciertas quimioquinas inflamatorias activan las células para iniciar una respuesta inmunitaria o promover la curación de heridas. Éstas son liberadas de muchos tipos de células diferentes y sirven para guiar a las células tanto del sistema inmunitario innato como del sistema inmunitario adaptativo. Las quimioquinas ejercen muchos de los efectos biológicos interactuando con receptores transmembrana unidos a proteína G denominados receptores de quimioquinas, que se encuentran selectivamente en las superficies de sus células diana.

Dados estos papeles y actividades biológicas, las actividades quimioquina, citoquina o de señalización celular de un polipéptido de la presente descripción se pueden determinar utilizando cualquiera de los numerosos análisis de rutina y reconocidos en la técnica. En cierta realización, la actividad de señalización celular se determina utilizando esencialmente cualquier análisis que mida la quimiotaxis, la migración celular, la liberación de citoquina, la diferenciación celular y/o la viabilidad celular. En realizaciones más específicas de la presente descripción, la actividad quimioquina se determina, por ejemplo, utilizando un análisis que mide la quimiotaxis de monocitos dependiente de GPCR, la liberación de interleuquinas y/o la apoptosis.

Los polipéptidos aislados de la presente descripción pueden tener esencialmente cualquier longitud y pueden ser esencialmente de cualquier origen con tal que proporcionen los elementos necesarios para constituir un motivo estructural que posea la actividad de señalización celular u otra actividad no canónica deseadas.

En ciertas realizaciones ilustrativas descritas en la presente memoria, un polipéptido de la presente descripción tendrá un tamaño que oscilará de aproximadamente 30-100, 30-200, 30-300, 30-400, 30-500, 40-100, 40-200, 40-300, 40-400, 40-500, 50-100, 50-200, 50-300, 50-400, 50-500, 60-100, 60-200, 60-300, 60-400 o 60-500 aminoácidos.

En ciertas realizaciones descritas en la presente memoria, un polipéptido de la presente descripción es una proteína de mamífero truncada (*p. ej.*, humana) o una variante activa de la misma. Una proteína truncada hace referencia a un polipéptido que es más corto que su correspondiente proteína completa, por ejemplo, debido a la eliminación de aminoácidos de sus extremos N y/o C terminales. El grado de truncamiento, esto es, el número residuos de aminoácido N y/o C terminales eliminados de la proteína completa puede variar considerablemente siempre que todavía proporcione los efectos celulares deseados cuando se administren a una célula, tejido o sujeto, como se describe en la presente memoria. En ciertas realizaciones de la presente descripción, se truncan al menos aproximadamente 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 300, 350 aminoácidos, o más, incluyendo todas las longitudes intermedias, del extremo N y/o C de una proteína completa. Se pretende que las longitudes intermedias incluyan todos los números enteros intermedios, por ejemplo, 6, 7, 8, *etc.*, 51, 52, 53, *etc.*, 201, 202, 203, *etc.*

En otras realizaciones descritas en la presente memoria, un polipéptido de la presente descripción está formado por uno o más fragmentos de una proteína de mamífero, o variantes activas de los mismos. Por ejemplo, en una realización ilustrativa, un polipéptido de la presente descripción está formado por un tramo lineal de aminoácidos contiguos (*p. ej.*, con una longitud dentro de los intervalos indicados anteriormente) derivado de una proteína de mamífero, tal como una proteína humana. Alternativamente, un polipéptido de la presente descripción puede estar formado por fragmentos no contiguos de una proteína de mamífero, en donde los fragmentos no contiguos son suficientes para constituir un motivo estructural como se describe en la presente memoria.

En una realización específica de la presente descripción, el polipéptido es un fragmento de una proteína GlyRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica descrita en la presente memoria, el polipéptido es una forma truncada de una proteína GlyRS que comprende los residuos de aminoácido 345-420 de la proteína GlyRS humana mostrada en el SEQ ID NO: 1, o un fragmento activo o una variante del mismo.

En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína AspRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es una forma truncada de una proteína AspRS que comprende los residuos de aminoácido 367-448 de la proteína AspRS humana mostrada en el SEQ ID NO: 2, o un fragmento activo o una variante del mismo.

En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína HisRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es una forma truncada de una proteína HisRS que comprende los residuos de aminoácido 294-372 de la proteína HisRS humana mostrada en el SEQ ID NO: 3, o un fragmento activo o una variante del mismo.

- 5 En otra realización más de la presente descripción, el polipéptido es un fragmento de una proteína ThrRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es una forma truncada de una proteína ThrRS que comprende los residuos de aminoácido 469-586 de la proteína ThrRS humana mostrada en el SEQ ID NO: 4, o un fragmento activo o una variante del mismo.
- 10 En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína GluProRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es una forma truncada de una proteína GluProRS que comprende los residuos de aminoácido 1171-1253 de la proteína GluProRS humana mostrada en el SEQ ID NO: 5, o un fragmento activo o una variante del mismo.
- 15 En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína SerRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es una forma truncada de una proteína SerRS que comprende los residuos de aminoácido 325-410 de la proteína SerRS humana mostrada en el SEQ ID NO: 6, o un fragmento activo o una variante del mismo.
- 20 En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína PheRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es un polipéptido PheRS (FRS) que comprende los residuos 380-449 del SEQ ID NO: 7 o un fragmento activo o una variante del mismo.
- 25 En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína LysRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es un polipéptido LysRS (KRS) que comprende los residuos 425-523 del SEQ ID NO: 8 o un fragmento activo o una variante del mismo.
- 30 En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína AsnRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es un polipéptido AsnRS (NRS) que comprende los residuos 416-494 del SEQ ID NO: 9 o un fragmento activo o una variante del mismo.
- 35 En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína AlaRS, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es un polipéptido AlaRS (ARS) que comprende los residuos 148-258 del SEQ ID NO: 10 o un fragmento activo o una variante del mismo.
- 40 En otra realización más de la presente descripción, el polipéptido es un fragmento de una proteína tiorredoxina, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es una forma truncada de una proteína tiorredoxina que comprende los residuos de aminoácido 20-105 de la proteína tiorredoxina humana mostrada en el SEQ ID NO: 11, o un fragmento activo o una variante del mismo.
- 45 En otra realización de la presente descripción, el polipéptido es un fragmento de una proteína de factor inhibidor de macrófagos, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es una forma truncada de una proteína de factor inhibidor de macrófagos que comprende los residuos de aminoácido 1-90 de la proteína de factor inhibidor de macrófagos humana mostrada en el SEQ ID NO: 12, o un fragmento activo o una variante del mismo.
- 50 En otra realización más de la presente descripción, el polipéptido es un fragmento de una proteína de isoforma B de peroxirredoxina 5 humana, en donde el fragmento tiene una actividad de señalización celular u otra actividad no canónica. En una realización más específica de la presente descripción, el polipéptido es una forma truncada de una isoforma B de peroxirredoxina 5 humana que comprende los residuos de aminoácido 32-68 y 125-161 de la secuencia mostrada en el SEQ ID NO: 13, o un fragmento activo o una variante del mismo.
- 55 En otra realización específica de la presente descripción, el polipéptido no es un fragmento de una proteína TrpRS o TyrRS.
- En otra realización de la presente descripción, el polipéptido no es un fragmento de HisRS que consiste en los primeros 48 aminoácidos de la proteína HisRS.
- En otra realización de la presente descripción, el polipéptido no es un fragmento de tiorredoxina que consiste en los primeros 80 u 84 aminoácidos N-terminales de la proteína tiorredoxina.
- La presente descripción proporciona adicionalmente variantes de los polipéptidos descritos en la presente memoria. Las variantes de polipéptidos incluidas en la presente descripción mostrarán típicamente una identidad de al menos aproximadamente 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o 99% (determinada,

por ejemplo, como se describe más abajo), en toda su longitud, con la región correspondiente de una secuencia de tipo salvaje o de referencia a partir de la cual derivan.

5 Una variante de polipéptido puede diferir de un polipéptido natural de la presente descripción en una o más sustituciones, deleciones, adiciones y/o inserciones. Tales variantes pueden ser de origen natural o pueden ser generadas sintéticamente, por ejemplo, modificando una o más de las secuencias polipeptídicas anteriores de la descripción y evaluando su actividad biológica como se describe en la presente memoria utilizando cualquiera de los numerosos mecanismos bien conocidos en la técnica.

10 En otras realizaciones ilustrativas descritas en la presente memoria, la variante puede ser una variante de empalme, ya sea de origen natural o no natural, en donde la variante de empalme posee al menos una actividad no canónica, p. ej., como se describe en la presente memoria.

En otras realizaciones ilustrativas descritas en la presente memoria, la variante contiene una o más mutaciones puntuales con respecto a una secuencia de polipéptido de tipo salvaje o de referencia, ya sea de origen natural o no natural, en donde el polipéptido variante posee al menos una actividad no canónica, p. ej., como se describe en la presente memoria.

15 En ciertas realizaciones descritas en la presente memoria, una variante contendrá sustituciones conservativas. Una "sustitución conservativa" es aquella en la un aminoácido es sustituido por otro aminoácido que tiene propiedades similares, de manera que un experto en la técnica esperaría que la estructura secundaria y la naturaleza hidropática del polipéptido permanecieran sustancialmente inalteradas. Se pueden realizar modificaciones en la estructura de los polinucleótidos y polipéptidos de la presente descripción y obtener todavía una molécula funcional que codifique un polipéptido variante o derivado con características deseables. Cuando se desea alterar la secuencia de aminoácidos de un polipéptido para crear una variante de un polipéptido equivalente, o incluso mejorada, de la presente descripción, un experto en la técnica, por ejemplo, puede cambiar uno o más codones de la secuencia de ADN codificante de acuerdo con la Tabla 1.

25 Por ejemplo, ciertos aminoácidos pueden ser sustituidos por otros aminoácidos en una estructura de proteína sin una pérdida apreciable de capacidad de unión interactiva con estructuras tales como, por ejemplo, receptores, regiones de unión a antígeno de anticuerpos o sitios de unión sobre una molécula sustrato. Puesto que son la capacidad interactiva y la naturaleza de una proteína las que generalmente definen la actividad funcional biológica de la proteína, se pueden realizar ciertas sustituciones en la secuencia de aminoácidos de una secuencia de proteína, y, por supuesto, su secuencia codificante de ADN subyacente, y sin embargo obtener una proteína con propiedades similares. De este modo se contempla que se puedan realizar varios cambios en las secuencias de polipéptidos de las composiciones descritas, o secuencias de ADN correspondientes que codifican dichos polipéptidos sin una pérdida apreciable de su utilidad o actividad deseadas.

Tabla 1.

Aminoácidos			Codones					
Alanina	Ala	A	GCA	GCC	GCG	GCU		
Cisteína	Cys	C	UGC	UGU				
Ácido aspártico	Asp	D	GAC	GAU				
Ácido glutámico	Glu	E	GAA	GAG				
Fenilalanina	Phe	F	UUC	UUU				
Glicina	gly	G	GGA	GGC	GGG	GGU		
Histidina	His	H	CAC	CAU				
Isoleucina	Ile	I	AUA	AUC	AUU			
Lisina	Lys	K	AAA	AAG				
Leucina	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU

Aminoácidos			Codones					
Metionina	Met	M	AUG					
Asparragina	Asn	N	AAC	AAU				
Prolina	Pro	P	CCA	CCC	CCG	CCU		
Glutamina	Gln	Q	CAA	CAG				
Arginina	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serina	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Treonina	Thr	T	ACA	ACC	ACG	ACU		
Valina	Val	V	GUA	GUC	GUG	GUU		
Triptófano	Trp	W	UGG					
Tirosina	Tyr	Y	UAC	UAU				

Al realizar tales cambios, también se puede considerar el índice hidropático de los aminoácidos. La importancia del índice hidropático de los aminoácidos al conferir la función biológica interactiva a una proteína es generalmente comprendida en la técnica (Kyte y Doolittle, 1982). Por ejemplo, se sabe que el carácter hidropático relativo del aminoácido contribuye a la estructura secundaria de la proteína resultante, que a su vez define la interacción de la proteína con otras moléculas, por ejemplo, enzimas, sustratos, receptores, ADN, anticuerpos, antígenos, y similares. Cada aminoácido tiene un índice hidropático asignado basándose en sus características de carácter hidrófobo y carga (Kyte y Doolittle, 1982). Estos valores son: isoleucina (+4,5); valina (+4,2); leucina (+3,8); fenilalanina (+2,8); cisteína/cistina (+2,5); metionina (+1,9); alanina (+1,8); glicina (-0,4); treonina (-0,7); serina (-0,8); triptófano (-0,9); tirosina (-1,3); prolina (-1,6); histidina (-3,2); glutamato (-3,5); glutamina (-3,5); aspartato (-3,5); asparragina (-3,5); lisina (-3,9); y arginina (-4,5).

Se sabe en la técnica que ciertos aminoácidos pueden ser sustituidos por otros aminoácidos que tienen un índice o puntuación hidropáticos similares y todavía dan como resultado una proteína con una actividad biológica similar, esto es, todavía se obtiene una proteína biológicamente funcional equivalente. Al realizar tales cambios, se prefiere la sustitución de aminoácidos cuyos índices hidropáticos están en ± 2 , son particularmente preferidos aquellos en ± 1 , y son incluso más concretamente preferidos aquellos en $\pm 0,5$.

También se entiende en la técnica que la sustitución de aminoácidos similares se puede llevar a cabo eficazmente basándose en el carácter hidrófilo. Como se detalla en la Patente de los Estados Unidos 4.554.101, se han asignado los siguientes valores de carácter hidrófilo a los residuos de aminoácido: arginina (+3,0); lisina (+3,0); aspartato (+3,0 ± 1); glutamato (+3,0 ± 1); serina (+0,3); asparragina (+0,2); glutamina (+0,2); glicina (0); treonina (-0,4); prolina (-0,5 ± 1); alanina (-0,5); histidina (-0,5); cisteína (-1,0); metionina (-1,3); valina (-1,5); leucina (-1,8); isoleucina (-1,8); tirosina (-2,3); fenilalanina (-2,5); triptófano (-3,4). Se entiende que se puede sustituir un aminoácido por otro que tenga un valor de carácter hidrófilo similar y todavía obtener una proteína biológicamente equivalente. En tales cambios, se prefiere la sustitución de aminoácidos cuyos valores de carácter hidrófilo están en ± 2 , son particularmente preferidos aquellos en ± 1 , y son incluso aún más particularmente preferidos aquellos en $\pm 0,5$.

Como se ha esbozado anteriormente, las sustituciones de aminoácidos se pueden basar en la similitud relativa de los sustituyentes de las cadenas laterales de los aminoácidos, por ejemplo, su carácter hidrófobo, su carácter hidrófilo, su carga, su tamaño, y similares. Las sustituciones ilustrativas que tienen en consideración varias de las características anteriores son bien conocidas por los expertos en la técnica e incluyen: arginina y lisina; glutamato y aspartato; serina y treonina; glutamina y asparragina; y valina, leucina e isoleucina.

Las sustituciones de aminoácidos se pueden realizar adicionalmente basándose en la similitud en la polaridad, la carga, la solubilidad, el carácter hidrófobo, el carácter hidrófilo y/o la naturaleza anfipática de los residuos. Por ejemplo, los aminoácidos cargados negativamente incluyen ácido aspártico y ácido glutámico; los aminoácidos cargados positivamente incluyen lisina y arginina; y los aminoácidos con grupos de cabeza polar no cargados que tienen valores de carácter hidrófilo similares incluyen leucina, isoleucina y valina; glicina y alanina; asparragina y

glutamina; y serina, treonina, fenilalanina y tirosina. Otros grupos de aminoácidos que pueden representar cambios conservativos incluyen: (1) ala, pro, gly, glu, asp, gin, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; y (5) phe, tyr, trp, his. Una variante también puede contener, o alternativamente contiene, cambios no conservativos. En una realización preferida, los polipéptidos variantes difieren de una secuencia nativa por la sustitución, delección o adición de cinco aminoácidos o menos. Las variantes también se pueden modificar (o alternativamente se modifican), por ejemplo, por medio de la delección o adición de aminoácidos que tienen una influencia mínima sobre la estructura secundaria y la naturaleza hidropática del polipéptido.

Los polipéptidos pueden comprender una secuencia señal (o líder) en el extremo N-terminal de la proteína, que dirige co-traduccionalmente o post-traduccionalmente la transferencia de la proteína. El polipéptido también se puede conjugar con un conector o con otra secuencia para facilitar la síntesis, purificación o identificación del polipéptido (*p. ej.*, poli-His), o para intensificar la unión del polipéptido a un soporte sólido. Por ejemplo, se puede conjugar un polipéptido a una región Fc de inmunoglobulina.

Cuando se comparan secuencias de polipéptidos, se dice que dos secuencias son "idénticas" si las secuencias de aminoácidos de las dos secuencias son iguales cuando se alinean para una correspondencia máxima, como se describe más abajo. Las comparaciones entre dos secuencias se llevan a cabo típicamente comparando las secuencias a lo largo de una ventana de comparación para identificar y comparar las regiones locales de similitud de secuencia. Una "ventana de comparación" según se utiliza en la presente memoria, se refiere a un segmento de al menos aproximadamente 20 posiciones contiguas, normalmente de 30 a aproximadamente 75, de 40 a aproximadamente 50, en el que una secuencia se puede comparar con una secuencia de referencia del mismo número de posiciones contiguas después de alinear óptimamente las dos secuencias.

El alineamiento óptimo de secuencias para la comparación se puede llevar a cabo, por ejemplo, utilizando el programa Megalign del conjunto de soporte lógico bioinformático Lasergene (DNASTAR, Inc., Madison, WI), utilizando los parámetros por defecto. Este programa incorpora varios esquemas de alineamiento descritos en las siguientes referencias: Dayhoff, M.O. (1978) A model of evolutionary change in proteins - Matrices for detecting distant relationships. En Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Supl. 3, págs. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenesis págs. 626-645 Methods in Enzymology vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. y Sharp, P.M. (1989) CABIOS 5:151-153; Myers, E.W. y Muller W. (1988) CABIOS 4:11-17; Robinson, E.D. (1971) Comb. Theor 11:105; Santou, N. Nes, M. (1987) Mol. Biol. Evol. 4:406-425; Sneath, P.H.A. y Sokal, R.R. (1973) Numerical Taxonomy - the Principles and Practice of Numerical Taxonomy, Freeman Press, San Francisco, CA; Wilbur, W. J. y Lipman, D. J. (1983) Proc. Nat'l Acad., Sci. USA 80:726-730.

Alternativamente, el alineamiento óptimo de secuencias para su comparación se puede llevar a cabo mediante el algoritmo de identidad local de Smith y Waterman (1981) Add. APL. Math 2:482, mediante el algoritmo de alineamiento de identidad de Needleman y Wunsch (1970) J. Mol. Biol. 48:443, mediante la búsqueda de métodos de similitud de Pearson y Lipman (1988) Proc. Nat'l Acad. Sci. USA 85: 2444, mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, BLAST, FASTA, y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), o mediante inspección.

Los ejemplos de los algoritmos que son adecuados para determinar el porcentaje de identidad de secuencia y la similitud de secuencia son los algoritmos BLAST y BLAST 2.0, que han sido descritos por Altschul et al. (1977) Nucl. Acids Res. 25:3389-3402 y Altschul et al. (1990) J. Mol. Biol. 215:403-410, respectivamente. Se pueden utilizar BLAST y BLAST 2.0, por ejemplo con los parámetros descritos en la presente memoria, para determinar el porcentaje de identidad de secuencia para los polinucleótidos y polipéptidos de la presente descripción. El soporte lógico para llevar a cabo los análisis BLAST está disponible públicamente a través del National Center for Biotechnology Information. Para las secuencias de aminoácidos, se puede utilizar una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabra en cada dirección se para cuando: la puntuación acumulativa de alineamiento cae por la cantidad X desde su máximo valor conseguido; la puntuación acumulativa va a cero o menos, debido a la acumulación de uno o más alineamientos de residuos con puntuación negativa; o se alcance el final de cualquiera de las secuencias. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y la velocidad del alineamiento.

En un enfoque ilustrativo, el "porcentaje de identidad de secuencia" se determina comparando dos secuencias óptimamente alineadas a lo largo de una ventana de comparación de al menos 20 posiciones, en donde la porción de la secuencia de polipéptidos en la ventana de comparación puede comprender adiciones o delecciones (esto es, huecos) de 20 por ciento o menos, normalmente 5 a 15 por ciento, o 10 a 12 por ciento, en comparación con las secuencias de referencia (que no comprenden adiciones ni delecciones) para el alineamiento óptimo de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las cuales existen residuos de aminoácidos idénticos en ambas secuencias para proporcionar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la secuencia de referencia (esto es, el tamaño de la ventana) y multiplicando los resultados por 100 para obtener el porcentaje de identidad de las secuencias.

En ciertas realizaciones de la presente descripción, se proporcionan polipéptidos de fusión, y polinucleótidos que codifican los polipéptidos de fusión. Los polipéptidos de fusión hacen referencia a polipéptidos de la presente descripción que se han conectado covalentemente, o bien directamente o bien indirectamente a través de un conector de aminoácidos, a una o más secuencias de polipéptidos heterólogas (compañeros de fusión). Los polipéptidos que forman la proteína de fusión están unidos típicamente del extremo C al extremo N, aunque también pueden unirse del extremo C al extremo C, extremo N a extremo N, o extremo N a extremo C. Los polipéptidos de la proteína de fusión pueden estar en cualquier orden.

El compañero de fusión puede ser diseñado e incluido esencialmente para cualquier propósito deseado siempre que no afecte adversamente a la actividad deseada del polipéptido. Por ejemplo, en una realización, un compañero de fusión comprende una secuencia que ayuda a expresar la proteína (un intensificador de la expresión) a rendimientos superiores a los de la proteína recombinante nativa. Otros compañeros de fusión pueden seleccionarse para incrementar la solubilidad de la proteína o para permitir que la proteína sea dirigida a compartimentos intracelulares deseados. Otros compañeros de fusión adicionales incluyen etiquetas de afinidad, que facilitan la purificación de la proteína.

Las proteínas de fusión pueden prepararse generalmente usando técnicas convencionales. Por ejemplo, las secuencias de ADN que codifican los componentes del polipéptido de una fusión deseada pueden ensamblarse separadamente, y ligarse en un vector de expresión apropiado. El extremo 3' de la secuencia de ADN que codifica un componente del polipéptido se liga, con o sin un conector peptídico, al extremo 5' de una secuencia de ADN que codifica el segundo componente del polipéptido de manera que los marcos de lectura de las secuencias están en fase. Esto permite la traducción en una única proteína de fusión que retiene la actividad biológica de ambos polipéptidos componentes.

Puede emplearse una secuencia de conector peptídico para separar el primer y segundo componentes del polipéptido por una distancia suficiente para asegurar que cada polipéptido se pliega en sus estructuras secundaria y terciaria, si se desea. Dicha secuencia de conector peptídico se incorpora a la proteína de fusión usando mecanismos convencionales muy conocidos en la técnica. Determinadas secuencias de conector peptídico pueden elegirse tomando como base los factores siguientes: (1) su capacidad de adoptar una conformación extendida flexible; (2) su incapacidad de adoptar una estructura secundaria que podría interaccionar con epítomos funcionales en el primer y segundo polipéptido; y (3) la ausencia de residuos hidrófobos o cargados que podrían reaccionar con los epítomos funcionales del polipéptido. Las secuencias de conector peptídico preferidas contienen residuos Gly, Asn y Ser. También pueden usarse otros aminoácidos casi neutros, tales como Thr y Ala en la secuencia conectora. Las secuencias de aminoácidos que pueden emplearse de manera útil como conectores incluyen las descritas en Maratea et al., Gene 40:39 46 (1985); Murphy et al., Proc. Natl. Acad. Sci. USA 83:8258 8262 (1986); Patente de los Estados Unidos Núm. 4.935.233 y Patente de los Estados Unidos Núm. 4.751.180. La secuencia conectora puede tener una longitud generalmente de 1 a aproximadamente 50 aminoácidos. Las secuencias conectoras no se requieren cuando el primer y segundo polipéptido tienen regiones de aminoácidos no esenciales N-terminales que pueden usarse para separar los dominios funcionales y evitar la interferencia estérica.

Las secuencias de ADN ligadas están conectadas operablemente a elementos reguladores de la transcripción o la traducción adecuados. Los elementos reguladores responsables de la expresión del ADN están localizados 5' con respecto a la secuencia de ADN que codifica los primeros polipéptidos. De manera similar, los codones de parada requeridos para terminar la traducción y las señales de terminación de la transcripción están presentes 3' con respecto a la secuencia de ADN que codifica el segundo polipéptido.

En general, los polipéptidos y polipéptidos de fusión (así como sus polinucleótidos codificantes) están aislados. Un polipéptido o polinucleótido "aislado" es uno que se retira de su entorno original. Por ejemplo, una proteína natural está aislada si se separa de parte o todos los materiales que coexisten en el sistema natural. Preferiblemente, dichos polipéptidos son al menos aproximadamente 90% puros, más preferiblemente al menos aproximadamente 95% y lo más preferiblemente al menos aproximadamente 99% puros. Un polinucleótido se considera que está aislado si, por ejemplo, se clona en un vector que no es parte del entorno natural.

En otras realizaciones más descritas en la presente memoria, un polipéptido de la presente descripción puede ser parte de un dímero. Los dímeros pueden incluir, por ejemplo, homodímeros entre dos polipéptidos AARS idénticos, heterodímeros entre dos polipéptidos AARS diferentes (*p. ej.*, un polipéptido GlyRS completo y un polipéptido GlyRS truncado, o dos polipéptidos AARS truncados diferentes), y/o heterodímeros entre un polipéptido AARS y un polipéptido heterólogo. Los monómeros y/o dímeros pueden ser solubles y pueden ser aislados o purificados hasta la homogeneidad. Ciertos heterodímeros, tales como aquellos entre un polipéptido AARS y un polipéptido heterólogo, pueden ser bifuncionales.

Asimismo se describen en la presente memoria proteínas monoméricas o sustancialmente monoméricas. En ciertas realizaciones descritas en la presente memoria, una proteína monomérica tiene una capacidad reducida para dimerizarse consigo misma (es decir, homodimerizarse) y/o para dimerizar con otro polipéptido AARS (esto es, heterodimerizar).

En otras realizaciones descritas en la presente memoria, un polipéptido de la presente descripción puede ser parte de un complejo multi-unitario. Un complejo multi-unitario de la presente descripción puede incluir, por ejemplo, al menos 2, 3, 4, o 5 o más monómeros. Los monómeros y/o los complejos multi-unitarios pueden ser solubles y se pueden aislar o purificar hasta su homogeneidad. Las unidades monoméricas de un complejo multiunitario pueden ser diferentes, homólogas, sustancialmente homólogas, o idénticas entre sí. No obstante, un complejo multi-unitario de la presente descripción incluye al menos un monómero que comprende un polipéptido como se describe en la presente memoria o, en otras realizaciones descritas en la presente memoria, al menos dos o más polipéptidos, como se describe en la presente memoria.

Los monómeros conectados covalentemente se pueden conectar directamente (por medio de enlaces) o indirectamente (*p. ej.*, a través de un conector). Para conectar directamente los monómeros de polipéptidos en la presente memoria, puede resultar beneficioso modificar los polipéptidos de la presente memoria para intensificar la dimerización o la multimerización. Por ejemplo, se pueden modificar uno o más residuos de aminoácido de un polipéptido AARS mediante la adición o sustitución por una o más cisteínas. Los métodos para crear sustituciones de aminoácidos, tales como sustituciones de cisteína, u otras modificaciones para facilitar la conexión, son bien conocidos por los expertos en la técnica.

Ciertas realizaciones de la presente descripción también contemplan el uso de AARS modificados u otros polipéptidos, incluyendo modificaciones que mejoran las características deseadas de un AARS u otro polipéptido, como se describe en la presente memoria. Las modificaciones ilustrativas de los polipéptidos AARS de la presente descripción incluyen, pero no se limitan a, derivatizaciones químicas y/o enzimáticas en uno o más aminoácidos constitutivos, incluyendo modificaciones de la cadena lateral, modificaciones de la cadena principal, y modificaciones N y C terminales incluyendo, hidroxilación, metilación, amidación, y el anclaje de radicales carbohidratados o lipídicos, cofactores, y similares. Las modificaciones ilustrativas también incluyen la pegilación de polipéptidos (véase, *p. ej.*, Veronese y Harris, *Advanced Drug Delivery Reviews* 54: 453-456, 2002).

En ciertos aspectos de la presente descripción, se puede utilizar la tecnología de ligación quimioselectiva para modificar polipéptidos truncados de la presente descripción, por ejemplo atrayendo polímeros de una manera específica del sitio y controlada. Dicha tecnología se basa típicamente en la incorporación de anclajes quimioselectivos en el núcleo de la proteína bien por medios químicos o recombinantes, y la modificación posterior con un polímero que porta un conector complementario. Como resultado, el proceso de ensamblaje y la estructura covalente del conjugado proteína-polímero resultante puede controlarse, permitiendo la optimización racional de las propiedades del fármaco, tal como eficacia y propiedades farmacocinéticas (véase, por ejemplo, Kochendoerfer, *Current Opinion in Chemical Biology* 9:555-560, 2005).

Los polipéptidos AARS descritos en la presente memoria se pueden preparar mediante cualquier procedimiento adecuado conocido por los expertos en la técnica, tal como mediante mecanismos recombinantes. Por ejemplo, se pueden preparar polipéptidos mediante un procedimiento que incluye las etapas de: (a) preparar un constructo que comprende una secuencia de polinucleótidos que codifica un polipéptido AARS de la presente descripción y que está conectado operablemente a un elemento regulador; (b) introducir el constructo en una célula anfitriona; (c) cultivar la célula anfitriona para expresar el polipéptido; y (d) aislar el polipéptido de la célula anfitriona. Los polipéptidos AARS recombinantes se pueden preparar convenientemente utilizando protocolos convencionales como describen por ejemplo en Sambrook, *et al.*, (1989, *supra*), en particular en las Secciones 16 y 17; Ausubel *et al.*, (1994, *supra*), en particular los Capítulos 10 y 16; y Coligan *et al.*, *Current Protocols in Protein Science* (John Wiley & Sons, Inc. 1995-1997), en particular los Capítulos 1, 5 y 6.

Además de los métodos de producción recombinante, los polipéptidos de la presente descripción pueden ser producidos mediante síntesis peptídica directa utilizando técnicas en fase sólida (Merrifield, *J. Am. Chem. Soc.* 85:2149-2154 (1963)). La síntesis de proteínas puede llevarse a cabo usando técnicas manuales o automatizadas. La síntesis automatizada puede conseguirse, por ejemplo, usando el Sintetizador de Péptidos Applied Biosystems 431A (Perkin Elmer). Alternativamente, varios fragmentos pueden sintetizarse químicamente separadamente y combinarse usando métodos químicos para producir la molécula deseada.

Composiciones de polinucleótido

La presente descripción también proporciona polinucleótidos aislados que codifican los polipéptidos de la presente descripción, así como también composiciones que comprenden tales polinucleótidos. Por ejemplo, las secuencias de polinucleótidos que codifican proteínas ARNt sintetasa, y otras proteínas descritas en la presente memoria, son fácilmente asequibles a través de cualquiera de las numerosas bases de datos de secuencias públicas (*p. ej.*, <http://www.ncbi.nlm.nih.gov>) y pueden ser identificadas, elaboradas y utilizadas en el contexto de la presente descripción utilizando mecanismos y metodologías descritos en la presente memoria y/o bien establecidos en la técnica.

Según se utilizan en la presente memoria, los términos "ADN" y "polinucleótido" y "ácido nucleico" se refieren a una molécula de ADN que se ha aislado y que carece de ADN genómico total de un especie concreta. Por lo tanto, un segmento de ADN que codifica un polipéptido se refiere a un segmento de ADN que contiene una o más secuencias codificantes y que aún así está sustancialmente aislado de, o purificado careciendo de, ADN genómico total de la

especie del que se obtiene el segmento de ADN. Incluidos en los términos "segmento de ADN" y "polinucleótido" están segmentos de ADN y fragmentos menores de dichos segmentos, y también vectores recombinantes, incluyendo, por ejemplo, plásmidos, cósmidos, fagémidos, fagos, virus, y similares.

5 Como comprenderán los expertos en la técnica, las secuencias de polinucleótidos de esta descripción pueden incluir secuencias genómicas, secuencias extragenómicas y codificadas por plásmidos y segmentos génicos modificados más pequeños que expresan, o se pueden adaptar para que expresen, proteínas, polipéptidos, péptidos y similares. Dichos segmentos pueden aislarse naturalmente, o modificarse sintéticamente por la mano del hombre.

10 Como reconocerá el experto en la técnica, los polinucleótidos pueden ser monocatenarios (codificantes o antisentido) o bicatenarios, y pueden ser moléculas de ADN (genómico, ADNc o sintético) o ARN. Pueden estar presentes, pero no es necesario, otras secuencias codificantes o no codificantes, en un polinucleótido de la presente descripción, y un polinucleótido puede, pero no es necesario, estar conectado, a otras moléculas y/o materiales de soporte.

15 Los polinucleótidos pueden comprender una secuencia nativa (esto es, una secuencia endógena que codifica un polipéptido de la presente descripción o una porción del mismo) o pueden comprender una variante, o un equivalente funcional biológico de semejante secuencia. Las variantes de polinucleótidos pueden contener una o más sustituciones, adiciones, deleciones y/o inserciones, como se describe adicionalmente más abajo, preferiblemente de manera que la actividad deseada del polipéptido codificado no disminuya sustancialmente con respecto al polipéptido no modificado. El efecto sobre la actividad del polipéptido codificado se puede evaluar generalmente como se describe en la presente memoria.

20 En realizaciones adicionales descritas en la presente memoria, la presente descripción proporciona polinucleótidos aislados que comprenden varias longitudes de tramos contiguos de secuencia idéntica o complementaria a un polinucleótido que codifica un polipéptido como se describe en la presente memoria.

25 Por ejemplo, esta descripción proporciona polinucleótidos que codifican al menos aproximadamente 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450 o 500, o más, residuos de aminoácido contiguos de un polipéptido de la presente descripción, así como también todas las longitudes intermedias. Se comprenderá fácilmente que "longitudes intermedias", en este contexto, puede ser cualquier longitud entre los valores citados, tales como 101, 102, 103, etc.; 151, 152, 153, etc.; 201, 202, 203, etc.

30 Los polinucleótidos de la presente descripción, con independencia de la longitud de la propia secuencia codificante, se pueden combinar con otras secuencias de ADN, tales como promotores, señales de poliadenilación, sitios para enzimas de restricción adicionales, sitios de clonación múltiples, otros segmentos codificantes, y similares, de manera que su longitud global puede cambiar considerablemente. Se contempla, por lo tanto, que puede emplearse un fragmento de polinucleótido de casi cualquier longitud, estando limitada la longitud total preferiblemente para facilitar la preparación y uso en el protocolo de ADN recombinante pretendido.

35 Por otra parte, los expertos en la técnica apreciarán que, como resultado de la degeneración del código genético, existen muchas secuencias de nucleótidos que codifican un polipéptido como se describe en la presente memoria. Algunos de estos polinucleótidos guardan una homología mínima con la secuencia de nucleótidos de cualquier gen nativo. A pesar de ello, los polinucleótidos que varían debido a diferencias en el uso de codones son específicamente contemplados por la presente descripción, por ejemplo los polinucleótidos que están optimizados para la selección de codones en seres humanos y/o primates. Adicionalmente, los alelos de los genes que comprenden las secuencias de polinucleótidos proporcionadas en la presente memoria están dentro del alcance de la presente descripción. Los alelos son genes endógenos que están alterados como resultado de una o más mutaciones, tales como deleciones, adiciones y/o sustituciones de nucleótidos. El ARNm y la proteína resultantes pueden, pero no necesitan, tener una estructura o función alterada. Los alelos pueden ser identificados utilizando técnicas convencionales (tales como hibridación, amplificación y/o comparación de secuencias de bases de datos).

45 Los polinucleótidos y fusiones de éstos pueden prepararse, manipularse y/o expresarse usando cualquiera de una variedad de mecanismos bien establecidos conocidos y disponibles en la técnica. Por ejemplo, se pueden utilizar secuencias de polinucleótidos que codifican los polipéptidos de la presente descripción, o proteínas de fusión o equivalentes funcionales de las mismas, en moléculas de ADN recombinantes para dirigir la expresión de un polipéptido en células anfitrionas apropiadas. Debido a la degeneración inherente del código genético, pueden producirse otras secuencias de ADN que codifican sustancialmente la misma secuencia de aminoácidos o una funcionalmente equivalente y estas secuencias pueden usarse para clonar y expresar un polipéptido dado.

50 Como entenderán los expertos en la técnica, puede ser ventajoso en algunos casos producir secuencias de nucleótidos que codifican polipéptidos que poseen codones no naturales. Por ejemplo, los codones preferidos para un anfitrión procarionta o eucariota particular pueden seleccionarse para incrementar la proporción de la expresión de la proteína o para producir un transcrito de ARN recombinante que tiene propiedades deseables, tales como una vida media que es mayor que la de un transcrito generado a partir de una secuencia natural.

55 Por otra parte, las secuencias de polinucleótidos de la presente descripción se pueden modificar utilizando métodos generalmente conocidos en la técnica con el fin de alterar las secuencias que codifican los polipéptidos por una

variedad de razones, incluyendo pero no limitadas a, alteraciones que modifican la clonación, el procesamiento, la expresión y/o la actividad del producto génico.

Con el fin de expresar un polipéptido deseado, se pueden insertar una secuencia de nucleótidos que codifica el polipéptido, o un equivalente funcional, en un vector de expresión apropiado, es decir, un vector que contiene los elementos necesarios para la transcripción y traducción de la secuencia codificante insertada. Los métodos que son muy conocidos para los expertos en la técnica pueden usarse para construir vectores de expresión que contienen secuencias que codifican un polipéptido de interés y elementos de control apropiados de la transcripción y traducción. Estos métodos incluyen técnicas de ADN recombinante *in vitro*, técnicas sintéticas, y recombinación genética *in vivo*. Dichas técnicas se describen en Sambrook et al., *Molecular Cloning, A Laboratory Manual* (1989), y Ausubel et al., *Current Protocols in Molecular Biology* (1989).

Se conoce una variedad de sistemas de vectores de expresión/anfitrión y pueden utilizarse para contener y expresar secuencias de polinucleótido. Éstos incluyen, pero no están limitados a, microorganismos tales como bacterias transformadas con bacteriófago recombinante, vectores de expresión de ADN en plásmido, o cósmido; levaduras transformadas con vectores de expresión de levadura; sistemas de células de insecto infectadas con vectores de expresión de virus (p. ej., baculovirus); sistemas de células de plantas transformadas con vectores de expresión de virus (p. ej., virus del mosaico de la coliflor, CaMV; virus del mosaico del tabaco, TMV) o con vectores de expresión bacterianos (p. ej., plásmidos Ti o pBR322); o sistemas de células animales.

Los "elementos de control" o "secuencias reguladoras" presentes en un vector de expresión son aquellas regiones no traducidas del vector—intensificadores, promotores, regiones no traducidas 5' y 3'—que interaccionan con proteínas celulares del anfitrión para llevar a cabo la transcripción y la traducción. Dichos elementos pueden variar en su fuerza y especificidad. Dependiendo del sistema vector y anfitrión utilizado, puede usarse cualquier número de elementos de la transcripción y traducción adecuados, incluyendo promotores constitutivos e inducibles. Por ejemplo, cuando se clonan en sistemas bacterianos, se pueden utilizar promotores inducibles tales como el promotor lacZ híbrido del fagémido PBLUESCRIPT (Stratagene, La Jolla, Calif.) o el plásmido PSPORT1 (Gibco BRL, Gaithersburg, Md.) y similares. En sistemas de células de mamíferos, se prefieren generalmente los promotores de genes de mamíferos o de virus de mamíferos. Si es necesario generar una línea celular que contiene múltiples copias de la secuencia que codifica un polipéptido, pueden usarse ventajosamente los vectores basados en SV40 o EBV con un marcador seleccionable apropiado.

En sistemas bacterianos, pueden seleccionarse varios vectores de expresión dependiendo del uso pretendido del polipéptido expresado. Por ejemplo, cuando se necesitan grandes cantidades, pueden usarse vectores que dirigen la expresión a alto nivel de proteínas de fusión que se purifican fácilmente. Dichos vectores incluyen, pero no están limitados a, la clonación multifuncional en *E. coli* y vectores de expresión tales como BLUESCRIPT (Stratagene), en el que la secuencia que codifica el polipéptido de interés puede ligarse en el vector en marco con secuencias para la Met amino-terminal y los 7 residuos posteriores de β -galactosidasa de manera que se produce una proteína híbrida; los vectores pIN (Van Heeke y Schuster, *J. Biol. Chem.* 264:5503-5509 (1989)); y similares, vectores pGEX (Promega, Madison, Wis.) también pueden usarse para expresar polipéptidos foráneos como proteínas de fusión con glutatión S-transferasa (GST). En general, dichas proteínas de fusión son solubles y pueden purificarse fácilmente de las células lisadas por adsorción a lechos de glutatión-agarosa seguido de elución en presencia de glutatión libre. Las proteínas preparadas en dichos sistemas pueden diseñarse para incluir sitios de escisión de heparina, trombina, o factor XA proteasa de manera que el polipéptido clonado de interés puede liberarse del radical GST cuando se desee.

En la levadura *Saccharomyces cerevisiae*, pueden usarse varios vectores que contienen promotores constitutivos o inducibles tales como factor alfa, alcohol oxidasa, y PGH. Para revisiones, véase Ausubel *et al.* (supra) y Grant et al., *Methods Enzymol.* 153:516-544 (1987).

En los casos en los que se usan vectores de expresión de plantas, la expresión de secuencias que codifican polipéptidos puede estar dirigida por cualquiera de varios promotores. Por ejemplo, pueden usarse promotores virales tales como promotores 35S y 19S de CaMV solos o combinados con la secuencia líder omega de TMV (Takamatsu, *EMBO J.* 6:307-311 (1987)). Alternativamente, pueden usarse promotores de plantas tales como la subunidad pequeña de RUBISCO o promotores de choque térmico (Coruzzi et al., *EMBO J.* 3:1671-1680 (1984); Broglie et al., *Science* 224:838-843 (1984); y Winter et al., *Results Probl. Cell Differ.* 17:85-105 (1991)). Estos constructos pueden introducirse en células de planta por transformación de ADN directa o transfección mediada por patógeno. Dichas técnicas se describen en varias revisiones disponibles generalmente (véase, p. ej., Hobbs en McGraw Hill, *Yearbook of Science and Technology*, págs. 191-196 (1992)).

También puede usarse un sistema de insecto para expresar un polipéptido de interés. Por ejemplo, en uno de dichos sistemas, se usa el virus de la polihedrosis nuclear de *Autographa californica* (AcNPV) como un vector para expresar genes foráneos en células de *Spodoptera frugiperda* o en larvas de *Trichoplusia*. Las secuencias que codifican el polipéptido pueden clonarse en una región no esencial del virus, tal como el gen de polihedrina, y ponerse bajo el control del promotor de polihedrina. La inserción exitosa de la secuencia que codifica el polipéptido dará lugar al gen de polihedrina inactivo y producirá virus recombinantes que carecen de proteína de la cubierta. Los virus

recombinantes pueden usarse para infectar, por ejemplo, células de *S. frugiperda* o larvas de *Trichoplusia* en las que el polipéptido de interés puede expresarse (Engelhard et al., Proc. Natl. Acad. Sci. U.S.A. 91:3224-3227 (1994)).

En células anfitrionas de mamífero, están disponibles generalmente varios sistemas de expresión basados en virus. Por ejemplo, en los casos en los que se usa un adenovirus como un vector de expresión, las secuencias que codifican un polipéptido de interés pueden ligarse en un complejo de transcripción/traducción de adenovirus que consiste en el promotor tardío y la secuencia líder tripartita. La inserción en una región no esencial E1 o E3 del genoma viral puede usarse para obtener un virus viable que es capaz de expresar el polipéptido en células anfitrionas infectadas (Logan y Shenk, Proc. Natl. Acad. Sci. U.S.A. 81:3655-3659 (1984)). Además, se pueden utilizar intensificadores de la transcripción, tales como el intensificador del virus del sarcoma de Rous (RSV), para incrementar la expresión en células anfitrionas de mamífero.

También pueden usarse señales de inicio específicas para conseguir una traducción más eficiente de secuencias que codifican un polipéptido de interés. Dichas señales incluyen el codón de inicio ATG y secuencias adyacentes. En los casos en los que se insertan secuencias que codifican el polipéptido, su codón de inicio, y secuencias aguas arriba en el vector de expresión adecuado, pueden no necesitarse señales de control de la transcripción o traducción adicionales. Sin embargo, en los casos en los que se inserta sólo la secuencia codificante, o una parte de ésta, deben proporcionarse señales de control de la traducción exógenas incluyendo el codón de inicio ATG. Además, el codón de inicio debe estar en el marco de lectura correcto para asegurar la traducción del inserto completo. Los elementos de la traducción exógenos y codones de inicio pueden tener varios orígenes, tanto natural como sintético. La eficacia de la expresión se puede intensificar por medio de la inclusión de intensificadores que son apropiados para el sistema de células concreto que se utilice, tales como los descritos en la bibliografía (Scharf et al., Results Probl. Cell Differ. 20:125-162 (1994)).

Además, una cepa de célula anfitriona puede elegirse por su capacidad de modular la expresión de las secuencias insertadas o para procesar la proteína expresada en la forma deseada. Dichas modificaciones del polipéptido incluyen, pero no están limitadas a, acetilación, carboxilación, glicosilación, fosforilación, lipidación, y acilación. El procesamiento posterior a la traducción que escinde una forma "prepro" de la proteína también puede usarse para facilitar la inserción, plegamiento y/o función correctos. Pueden elegirse diferentes células anfitrionas tales como CHO, HeLa, MDCK, HEK293, y W138, que tienen maquinaria celular específica y mecanismos característicos para dichas actividades posteriores a la traducción, para asegurar la modificación y procesamiento correctos de la proteína foránea.

Para una producción a largo plazo, con alto rendimiento, de las proteínas recombinantes, se prefiere generalmente la expresión estable. Por ejemplo, pueden transformarse líneas celulares que expresan de manera estable un polinucleótido de interés usando vectores de expresión que pueden contener orígenes virales de replicación y/o elementos de expresión endógenos y un gen marcador seleccionable en el mismo vector o en un vector separado. Después de la introducción del vector, puede dejarse que las células crezcan durante 1-2 días en un medio enriquecido antes de cambiarlas a medio selectivo. El propósito del marcador seleccionable es conferir resistencia a la selección, y su presencia permite el crecimiento y recuperación de las células que expresan con éxito las secuencias introducidas. Los clones resistentes de células transformadas de manera estable pueden proliferar usando técnicas de cultivo tisular apropiadas para el tipo celular.

Puede usarse cualquier número de sistemas de selección para recuperar las líneas celulares transformadas. Éstos incluyen, pero no están limitados a, los genes de timidina quinasa del virus del herpes simple (Wigler et al., Cell 11:223-232 (1977)) y adenina fosforibosiltransferasa (Lowy et al., Cell 22:817-823 (1990)) que pueden emplearse en células tk- o aprt-, respectivamente. También puede usarse resistencia a antimetabolito, antibiótico o herbicida como la base de la selección; por ejemplo, dhfr que confiere resistencia a metotrexato (Wigler et al., Proc. Natl. Acad. Sci. U.S.A. 77:3567-70 (1980)); npt, que confiere resistencia a los aminoglicósidos, neomicina y G-418 (Colbere-Garapin et al., J. Mol. Biol. 150:1-14 (1981)); y als o pat, que confieren resistencia a clorsulfurón y fosfinotricina acetiltransferasa, respectivamente (Murry, supra). Se han descrito genes seleccionables adicionales, por ejemplo, trpB, que permite a las células utilizar indol en lugar de triptófano, o hisD, que permite a las células utilizar histinol en lugar de histidina (Hartman & Mulligan, Proc. Natl. Acad. Sci. U.S.A. 85:8047-51 (1988)). El uso de marcadores visibles ha ganado popularidad con marcadores tales como antocianinas, β -glucuronidasa y su sustrato GUS, y luciferasa y su sustrato luciferina, siendo usados ampliamente no sólo para identificar transformantes, sino también para cuantificar la cantidad de expresión de proteína temporal o estable atribuible a un sistema de vector específico (Rhodes et al., Methods Mol. Biol. 55:121-131 (1995)).

En la técnica se conoce una variedad de protocolos para detectar y medir la expresión de productos codificados por polinucleótidos, usando bien anticuerpos policlonales o monoclonales específicos para el producto. Los ejemplos incluyen el ensayo inmunoabsorbente ligado a enzima (ELISA), radioinmunoensayo (RIA), y separación de células activada por fluorescencia (FACS). Éstos y otros ensayos se describen, entre otros lugares, en Hampton et al., Serological Methods, a Laboratory Manual (1990) y Maddox et al., J. Exp. Med. 158:1211-1216 (1983).

Los expertos en la técnica conocen una amplia variedad de marcadores y técnicas de conjugación y pueden usarse en varios ensayos de ácidos nucleicos y aminoácidos. Los medios para producir sondas marcadas de hibridación o PCR para detectar secuencias relacionadas con los polinucleótidos incluyen oligomarcaje, traslación de mella, y

marcaje terminal o amplificación por PCR usando un nucleótido marcado. Alternativamente, las secuencias, o cualquier parte de éstas pueden clonarse en un vector para la producción de una sonda de ARNm. Dichos vectores son conocidos en la técnica, están disponibles comercialmente, y pueden usarse para sintetizar sondas de ARN *in vitro* por la adición de una ARN polimerasa apropiada tal como T7, T3, o SP6 y nucleótidos marcados. Estos procedimientos pueden llevarse a cabo usando una variedad de kits disponibles comercialmente. Las moléculas informadoras o marcas adecuadas, que pueden usarse incluyen radionúclidos, enzimas, agentes fluorescentes, quimioluminiscentes, o cromogénicos así como sustratos, cofactores, inhibidores, partículas magnéticas, y similares.

Las células anfitrionas transformadas con una secuencia de polinucleótido de interés pueden cultivarse en condiciones adecuadas para la expresión y recuperación de la proteína del cultivo celular. La proteína producida por una célula recombinante puede secretarse o estar contenida intracelularmente dependiendo de la secuencia y/o del vector usado. Como comprenderán los expertos en la técnica, los vectores de expresión que contienen los polinucleótidos de la presente descripción se pueden diseñar para que contengan secuencias señal que dirijan la secreción del polipéptido codificado a través de una membrana celular procariótica o eucariótica. Pueden usarse otras construcciones recombinantes para unir secuencias que codifican un polipéptido de interés a la secuencia de nucleótidos que codifica un dominio de polipéptido que facilitará la purificación de proteínas solubles.

Además de los métodos de producción recombinante, los polipéptidos de la presente descripción, y los fragmentos de los mismos, pueden ser producidos mediante síntesis peptídica directa utilizando técnicas en fase sólida (Merrifield, J. Am. Chem. Soc. 85:2149-2154 (1963)). La síntesis de proteínas puede llevarse a cabo usando técnicas manuales o automatizadas. La síntesis automatizada puede conseguirse, por ejemplo, usando el Sintetizador de Péptidos Applied Biosystems 431A (Perkin Elmer). Alternativamente, se pueden sintetizar químicamente por separado diversos fragmentos y combinarlos utilizando métodos químicos para producir la molécula completa.

De acuerdo con otro aspecto de la presente descripción, los polinucleótidos que codifican polipéptidos de la presente descripción se pueden liberar en un sujeto *in vivo*, *p. ej.*, utilizando técnicas de terapia génica. La terapia génica se refiere en general a la transferencia de ácidos nucleicos heterólogos a ciertas células, células diana, de un mamífero, concretamente un ser humano, con un trastorno o afecciones para los cuales se procura semejante terapia. El ácido nucleico es introducido en las células diana seleccionadas de manera que se exprese el ADN heterólogo y se produzca un producto terapéutico codificado de ese modo.

Los diversos vectores virales que se pueden utilizar para la terapia génica como se ilustra en la presente memoria incluyen adenovirus, virus herpes, vaccinia, virus adeno-asociados (AAV), o, preferiblemente, un virus con ARN tal como un retrovirus. Preferiblemente, el vector retroviral es un derivado de un vector retrovirus murino o aviar, o es un vector lentiviral. El vector retroviral preferido es un vector lentiviral. Los ejemplos de los vectores retrovirales en los que se puede insertar el único gen foráneo incluyen, pero no se limitan a: virus de la leucemia murina de Moloney (MoMuLV), virus del sarcoma murino de Harvey (HaMuSV), virus del tumor mamario murino (MuMTV), SIV, BIV, VIH y virus del Sarcoma de Rous (RSV). Numerosos vectores retrovirales adicionales pueden incorporar múltiples genes. Todos estos vectores pueden transferir o incorporar un gen para un marcador seleccionable de manera que las células transducidas puedan ser identificadas y generadas. Insertando una secuencia de polipéptido de unión a ADN derivada de un dedo de cinc de interés en el vector viral, junto con otro gen que codifica el ligando para un receptor sobre una célula diana específica, por ejemplo, el vector se puede volver específico de la diana. Los vectores retrovirales se pueden volver específicos de la diana insertando, por ejemplo, un polinucleótido que codifica una proteína (dímero). Se puede lograr un direccionamiento ilustrativo utilizando un anticuerpo para dirigir el vector retroviral. Los expertos en la técnica sabrán, o podrán averiguar fácilmente sin experimentación indebida, secuencias de polinucleótidos específicas que se pueden insertar en el genoma retroviral para permitir la liberación específica de la diana del vector retroviral que contiene el polinucleótido de la proteína de unión al dedo de cinc.

Puesto que los retrovirus recombinantes son defectuosos, requieren ayuda para producir partículas de vector infecciosas. Esta ayuda puede ser proporcionada, por ejemplo, utilizando líneas celulares coadyuvantes que contienen plásmidos que codifican todos los genes estructurales del retrovirus bajo el control de secuencias reguladoras dentro de la LTR. En estos plásmidos falta una secuencia de nucleótidos que permita que el mecanismo de empaquetamiento reconozca un transcrito de ARN para la encapsulación. Las líneas celulares coadyuvantes que tienen deleciones de la señal de empaquetamiento incluyen, pero no se limitan a PSI.2, PA317 y PA12, por ejemplo. Estas líneas celulares producen viriones vacíos, ya que no hay genoma empaquetado. Si se introduce un vector retroviral en tales células en las que la señal de empaquetamiento está intacta, pero se reemplazan los genes estructurales por otros genes de interés, el vector puede ser empaquetado y el virión de vector producido. Los viriones de vector producidos mediante este método se pueden utilizar a continuación para infectar una línea celular de tejido, tal como las células NIH 3T3, para producir grandes cantidades de viriones retrovirales quiméricos.

También se pueden utilizar técnicas de liberación "no virales" para la terapia génica incluyendo, por ejemplo, complejos de ADN-ligando, complejos de adenovirus-ligando-ADN, inyección directa de ADN, precipitación de CaPO₄, técnicas de pistola génica, electroporación, liposomas, lipofección, y similares. Cualquiera de estos métodos está ampliamente disponible para un experto en la técnica y sería adecuado para su uso en la presente descripción. Otros métodos adecuados se encuentran disponibles para un experto en la técnica, y se debe entender que la presente descripción puede completarse utilizando cualquiera de los métodos de transfección disponibles. La

lipofección se puede completar encapsulando una molécula de ADN aislada dentro de una partícula liposomal y poniendo en contacto la partícula liposomal con la membrana celular de la célula diana. Los liposomas son partículas coloidales, auto-ensamblantes, en las que una bicapa lipídica, compuesta por moléculas anfífilas tales como fosfatidilserina o fosfatidilcolina, encapsula una porción del medio circundante de manera que la bicapa lipídica rodea un interior hidrófilo. Se pueden construir liposomas unilamelares o multilamelares de manera que el interior contenga un agente químico, un fármaco, o, como en la presente descripción, una molécula de ADN aislada, deseados.

Las realizaciones de la presente descripción también incluyen oligonucleótidos (*p. ej.*, oligómeros antisentido, sondas, cebadores), ya sea para la detección, para la amplificación, para terapias antisentido, o para otro fin. Los oligonucleótidos comprenden típicamente o son complementarios a al menos una porción de una secuencia de polinucleótido de AARS. Para estos fines y fines relacionados, se pretende que el término "oligonucleótido" u "oligo" u "oligómero" incluya un "oligonucleótido" singular, así como también "oligonucleótidos" plurales, y se refiera a cualquier polímeros de dos o más nucleótidos, nucleósidos, nucleobases o compuestos relacionados utilizados como reactivo en los métodos de amplificación de la presente descripción, así como también en los métodos de detección subsiguientes. El oligonucleótido puede ser ADN y/o ARN y/o análogos de los mismos.

En ciertas realizaciones descritas en la presente memoria, pueden ser adecuados oligómeros tales como oligómeros antisentido tan largos como de 40 bases, en los que al menos un número mínimo de bases, *p. ej.*, 10-12 bases, son complementarias a la secuencia diana (*p. ej.*, una secuencia de polinucleótido de AARS). En general, sin embargo, la absorción favorecida o activa en las células se optimiza a longitudes de oligómeros menores de aproximadamente 30. Para ciertos oligómeros, descritos adicionalmente más abajo, se produce un equilibrio óptimo de estabilidad de unión y absorción a longitudes de 18-25 bases. Se describen en la presente memoria oligómeros antisentido (*p. ej.*, PNA, LNA, 2'-OMe, MOE, morfolino) que consisten en aproximadamente 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, o 40 bases, en los que al menos aproximadamente 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, o 40 bases contiguas o no contiguas son complementarias a su secuencia diana de AARS, o variantes de la misma.

Ciertas realizaciones descritas en la presente memoria se refieren a agentes de interferencia de ARN (ARNi), tales como ARN de interferencia pequeño (ARNip) u otros agentes ARNip, que dirigen uno o más transcritos de ARNm de un polinucleótido de AARS. Para ciertas realizaciones relacionadas con ARNip descritas en la presente memoria, cada hebra de un agente ARNip puede tener una longitud igual o menor de 35, 30, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, o 15 nucleótidos. La hebra tiene preferiblemente al menos 19 nucleótidos de longitud. Por ejemplo, cada hebra puede tener entre 21 y 25 nucleótidos de longitud. Los agentes ARNip preferidos tienen una región dúplex de 17, 18, 19, 20, 21, 22, 23, 24, o 25 pares de nucleótidos, y uno o más salientes, preferiblemente uno o dos salientes 3', de 2-3 nucleótidos. También se describen en la presente memoria los métodos de uso de los mismos para modular los niveles de un transcrito de AARS seleccionado, tal como un polinucleótido de AARS que codifica un motivo como se describe en la presente memoria.

Como se ha observado anteriormente, los polinucleótidos de AARS de la presente descripción se pueden utilizar en cualquiera de los métodos de diagnóstico, descubrimiento de fármacos, o terapéuticos descritos en la presente memoria.

40 Composiciones de anticuerpo, fragmentos de los mismos y otros agentes de unión

De acuerdo con otro aspecto descrito en la presente memoria, la presente descripción proporciona adicionalmente agentes de unión, tales como anticuerpos y fragmentos de unión a antígeno de los mismos, receptores solubles, péptidos, peptidomiméticos, aptámeros, *etc.*, que muestran especificidad de unión a un polipéptido descrito en la presente memoria, o a una porción, variante o derivado de los mismos, y métodos de utilización de los mismos. También se describen en la presente memoria agentes de unión que muestran especificidad de unión para un compañero de unión celular de un polipéptido descrito en la presente memoria. Preferiblemente, tales agentes de unión son eficaces para modular una o más de las actividades no canónicas mediadas por un polipéptido de la presente descripción.

En ciertas realizaciones descritas en la presente memoria, por ejemplo, el agente de unión es uno que se une a un polipéptido de la presente descripción e inhibe su capacidad para unirse a uno o más de sus compañeros de unión celulares. Por consiguiente, tales agentes de unión se pueden utilizar para tratar o prevenir enfermedades, trastornos u otras afecciones que están mediadas por un polipéptido de la presente descripción suscitando antagonismo sobre su actividad.

Se dice que un anticuerpo, o fragmento de unión a antígeno del mismo, se "une específicamente", "se une inmunológicamente", y/o es "reactivo inmunológicamente" con un polipéptido de la presente descripción si reacciona a un nivel detectable (dentro, por ejemplo, de un análisis ELISA) con el polipéptido, y no reacciona detectablemente con polipéptidos no relacionados en condiciones similares.

La unión inmunológica, tal y como se usa en este contexto, se refiere generalmente a las interacciones no covalentes del tipo que ocurren entre una molécula de inmunoglobulina y un antígeno para el que la inmunoglobulina es específica. La fuerza, o afinidad de las interacciones de unión inmunológica pueden expresarse en términos de la constante de disociación (K_d) de la interacción, en donde una K_d menor representa una mayor afinidad. Las propiedades de la unión inmunológica de polipéptidos seleccionados pueden cuantificarse usando métodos muy conocidos en la técnica. Uno de dichos métodos conlleva medir las velocidades de la formación del complejo de sitio unión a antígeno/antígeno y disociación, en donde esas velocidades dependen de las concentraciones de las parejas del complejo, la afinidad de la interacción, y de parámetros geométricos que influyen igualmente en la velocidad en ambas direcciones. Así, tanto la "constante de asociación" (k_{on}) como la "constante de disociación" (k_{off}) pueden determinarse por el cálculo de las concentraciones y las velocidades reales de asociación y disociación. La proporción de k_{off}/k_{on} permite la cancelación de todos los parámetros no relacionados con la afinidad, y es así igual a la constante de disociación K_d . Véase, generalmente, Davies et al. (1990) Annual Rev. Biochem. 59:439-473.

Un "sitio de unión a antígeno," o una "porción de unión" de un anticuerpo se refiere a la parte de la molécula de inmunoglobulina que participa en la unión al antígeno. El sitio de unión al antígeno está formado por residuos de aminoácidos de las regiones variables N-terminales ("V") de las cadenas pesada ("H") y ligera ("L"). Tres tramos altamente divergentes dentro de las regiones V de las cadenas pesada y ligera son referidos como "regiones hipervariables" que están interpuestos entre los tramos flanqueantes más conservados conocidos como "regiones marco", o "FR". De este modo el término "FR" se refiere a secuencias de aminoácidos que se encuentran naturalmente entre y adyacentes a regiones hipervariables en las inmunoglobulinas. En una molécula de anticuerpo, las tres regiones hipervariables de una cadena ligera y las tres regiones hipervariables de una cadena pesada están dispuestas una respecto a otra en espacio tridimensional para formar una superficie de unión a antígeno. La superficie de unión a antígeno es complementaria a la superficie tridimensional de un antígeno unido, y las tres regiones hipervariables de cada una de las cadenas pesada y ligera se refieren como "regiones determinantes de la complementariedad," o "CDR."

Un agente de unión puede ser, por ejemplo, un ribosoma, con o sin un componente peptídico, una molécula de ARN o un polipéptido. En una realización preferida, un agente de unión es un anticuerpo o un fragmento de unión a antígeno del mismo. Los anticuerpos pueden prepararse por cualquiera de una variedad de mecanismos conocidos para los expertos en la técnica. Véase, *por ejemplo*, Harlow y Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. En general, los anticuerpos pueden producirse por técnicas de cultivo celular, incluyendo la generación de anticuerpos monoclonales como se describe en la presente memoria, o mediante transfección de genes de anticuerpo en anfitriones celulares bacterianos o de mamífero adecuados, con el fin de permitir la producción de anticuerpos recombinantes. En una técnica, un inmunógeno que comprende el polipéptido se inyecta inicialmente en cualquiera de una amplia variedad de mamíferos (por ejemplo, ratones, ratas, conejos, ovejas o cabras). En esta etapa, los polipéptidos de esta descripción pueden servir como inmunógeno sin modificación. Alternativamente, particularmente para polipéptidos relativamente cortos, puede incitarse una respuesta inmune superior si el polipéptido se une a una proteína vehicular, tal como albúmina de suero bovino o hemocianina de lapa californiana. El inmunógeno se inyecta en el anfitrión animal, preferiblemente según un esquema predeterminado incorporando una o más inmunizaciones de refuerzo, y se toman muestras de sangre de los animales periódicamente. Los anticuerpos policlonales específicos para el polipéptido pueden purificarse de dichos antisueros, por ejemplo, por cromatografía de afinidad usando el polipéptido acoplado a un soporte sólido adecuado.

Los anticuerpos monoclonales específicos para un polipéptido de interés se pueden preparar, por ejemplo, utilizando la técnica de Kohler y Milstein, Eur. J. Immunol. 6:511-519, 1976, y las mejoras del mismo. Brevemente, estos métodos implican la preparación de líneas celulares inmortales capaces de producir anticuerpos que tienen la especificidad deseada (es decir, reactividad con el polipéptido de interés). Dichas líneas celulares pueden producirse, por ejemplo, a partir de células del bazo obtenidas de un animal inmunizado como se ha descrito anteriormente. Las células del bazo se immortalizan, por ejemplo, por fusión con un compañero de fusión de células de mieloma, preferiblemente una que es singénica con el animal inmunizado. Puede emplearse una variedad de técnicas de fusión. Por ejemplo, las células de bazo y las células de mieloma pueden combinarse con un detergente no iónico durante unos pocos minutos y a continuación cultivarse en placa a baja densidad en un medio selectivo que soporta el crecimiento de las células híbridas, pero no de las células de mieloma. Una técnica de selección preferida usa selección HAT (hipoxantina, aminopterina, timidina). Después de un tiempo suficiente, habitualmente aproximadamente 1 a 2 semanas, se observan colonias de híbridos. Se seleccionan colonias individuales y sus sobrenadantes de cultivo se someten a ensayo para determinar la actividad de unión frente al polipéptido. Se prefieren los hibridomas que tienen alta reactividad y especificidad.

Los anticuerpos monoclonales pueden aislarse de los sobrenadantes de colonias de hibridoma en crecimiento. Además, pueden emplearse varias técnicas para aumentar el rendimiento, tal como inyección de la línea celular de hibridoma en la cavidad peritoneal de un anfitrión vertebrado adecuado, tal como un ratón. Los anticuerpos monoclonales pueden recogerse del fluido de ascites o la sangre. Los contaminantes pueden eliminarse de los anticuerpos por técnicas convencionales, tales como cromatografía, filtración en gel, precipitación, y extracción. Los polipéptidos de esta descripción se pueden utilizar en el proceso de purificación, por ejemplo, en una etapa de una cromatografía de afinidad.

En la técnica se conocen varias moléculas terapéuticamente útiles que comprenden sitios de unión a antígeno que son capaces de presentar propiedades de unión inmunológicas de una molécula de anticuerpo. La enzima proteolítica papaína escinde preferiblemente moléculas de IgG para rendir varios fragmentos, dos de los cuales (los fragmentos "F(ab)") comprenden cada uno un heterodímero covalente que incluye un sitio de unión a antígeno
 5 intacto. La enzima pepsina es capaz de escindir moléculas de IgG para proporcionar varios fragmentos, incluyendo el fragmento "F(ab)₂" que comprende ambos sitios de unión al antígeno. Un fragmento "Fv" puede producirse por escisión proteolítica preferencial de una IgM, y en raras ocasiones molécula de inmunoglobulina IgG o IgA. Los fragmentos Fv, sin embargo, se obtienen más comúnmente usando mecanismos recombinantes conocidos en la técnica. El fragmento Fv incluye un heterodímero V_H:V_L no covalente incluyendo un sitio de unión a antígeno que
 10 retiene la mayor parte de las capacidades de reconocimiento y unión a antígeno de la molécula de anticuerpo nativa. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710; y Ehrlich et al. (1980) Biochem 19:4091-4096.

Un polipéptido de cadena sencilla Fv ("sFv") es un heterodímero V_H:V_L unido covalentemente que se expresa a partir de una fusión génica incluyendo genes que codifican V_H y V_L unidos por un conector que codifica un péptido.
 15 Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. Se han descrito varios métodos para discernir estructuras químicas para convertir las cadenas de polipéptido de cadena ligera y pesada agregadas naturalmente pero químicamente separadas de una región V de anticuerpo en una molécula sFv que se plegará en un estructura tridimensional sustancialmente similar a la estructura de un sitio de unión a antígeno. Véanse, por ejemplo, las Patentes de los Estados Unidos Núms. 5.091.513 y 5.132.405, de Huston et al.; y Patente de los Estados Unidos
 20 Núm. 4.946.778, de Ladner et al.

Cada una de las moléculas anteriormente descritas incluye un conjunto de CDR de cadena pesada y cadena ligera, interpuestas respectivamente entre un conjunto de FR de una cadena pesada y una cadena ligera que proporciona soporte a las CDR y definen la relación espacial de las CDR entre sí. Tal y como se usa en la presente memoria, el término "conjunto de CDR" se refiere a las tres regiones hipervariables de una región V de cadena pesada o ligera.
 25 Desde el extremo N de una cadena pesada o ligera, estas regiones se indican como "CDR1," "CDR2," y "CDR3" respectivamente. Un sitio de unión a antígeno, por lo tanto, incluye seis CDR, que comprenden el conjunto de CDR de cada una de una región V de cadena pesada y cadena ligera. Un polipéptido que comprende una única CDR, (por ejemplo, una CDR1, CDR2 o CDR3) se refiere en la presente memoria como una "unidad de reconocimiento molecular." Los análisis cristalográficos de varios complejos antígeno-anticuerpo han demostrado que los residuos de aminoácidos de las CDR forman un contacto extenso con el antígeno unido, en el que el contacto más extenso con el antígeno es con CDR3 de cadena pesada. Así, las unidades de reconocimiento molecular son responsables principalmente de la especificidad de un sitio de unión a antígeno.

Tal y como se usa en la presente memoria, el término "conjunto de FR" se refiere a las cuatro secuencias de aminoácidos flanqueantes que enmarcan las CDR de un conjunto de CDR de una región V de cadena pesada o
 35 ligera. Algunos residuos de FR pueden contactar el antígeno unido; sin embargo, las FR son principalmente responsables del plegamiento de la región V en el sitio de unión al antígeno, concretamente los residuos de FR directamente adyacentes a las CDR. En las FR, determinados residuos de aminoácidos y determinadas características estructurales están altamente conservados. A este respecto, todas las secuencias de región V contienen un bucle disulfuro interno de aproximadamente 90 residuos de aminoácidos. Cuando las regiones V se pliegan en un sitio de unión, las CDR se presentan como restos de bucle proyectados que forman una superficie de
 40 unión a antígeno. Generalmente se reconoce que existen regiones estructurales conservadas de las FR que influyen en la conformación plegada de los bucles de CDR en ciertas estructuras "canónicas" con independencia de la secuencia de aminoácidos de la CDR precisa. Además, se sabe que determinados residuos de FR participan en contactos interdominio no covalentes que estabilizan la interacción de las cadenas pesada y ligera del anticuerpo.

Se han descrito varias moléculas de anticuerpo "humanizadas" que comprenden un sitio de unión a antígeno derivado de una inmunoglobulina no humana, incluyendo anticuerpos quiméricos que tienen regiones V de roedor y sus CDR asociadas fusionadas con dominios constantes humanos (Winter et al. (1991) Nature 349:293-299; Llobuglio et al. (1989) Proc. Nat. Acad. Sci. USA 86:4220-4224; Shaw et al. (1987) J Immunol. 138:4534-4538; y Brown et al. (1987) Cancer Res. 47:3577-3583), CDR de roedor injertadas en un FR de soporte humano antes de la
 50 fusión con un dominio constante de anticuerpo humano apropiado (Riechmann et al. (1988) Nature 332:323-327; Verhoeven et al. (1988) Science 239:1534-1536; y Jones et al. (1986) Nature 321:522-525), y CDR de roedor soportadas por FR de roedor recubiertas recombinantemente (Publicación de Patente Europea No. 519.596, publicada el 23 de dic., 1992). Estas moléculas "humanizadas" se diseñan para minimizar la respuesta inmunológica no deseada hacia moléculas de anticuerpos de roedores que limita la duración y la eficacia de las aplicaciones terapéuticas de esos radicales en los receptores humanos.
 55

Tal y como se usa en la presente memoria, los términos "FR recubiertas" y "FR recubiertas recombinantemente" se refieren al reemplazo selectivo de residuos de FR de, por ejemplo, una región V de cadena pesada o ligera de roedor, con residuos de FR humano con el fin de proporcionar una molécula xenogénica que comprende un sitio de
 60 unión a antígeno que retiene sustancialmente toda la estructura de plegamiento del polipéptido FR nativo. Las técnicas de recubrimiento se basan en la comprensión de que las características de unión a ligando de un sitio de unión a antígeno están determinadas principalmente por la estructura y disposición relativa de los conjuntos de CDR de cadena pesada y ligera en la superficie de unión a antígeno. Davies et al. (1990) Ann. Rev. Biochem. 59:439-473.

Así, la especificidad de la unión a antígeno puede conservarse en un anticuerpo humanizado sólo cuando se mantienen cuidadosamente las estructuras CDR, su interacción entre sí, y su interacción con el resto de los dominios de la región V. Mediante el uso de técnicas de recubrimiento, los residuos de FR exteriores (por ejemplo, accesibles a disolvente) que son encontrados fácilmente por el sistema inmune se reemplazan selectivamente por residuos humanos para proporcionar una molécula híbrida que comprende bien una superficie débilmente inmunogénica, o bien sustancialmente no inmunogénica.

En otra realización de la presente descripción, los anticuerpos monoclonales u otros agentes de unión de la presente descripción se pueden acoplar a uno o más agentes de interés. Por ejemplo, un agente terapéutico puede acoplarse (por ejemplo, unirse covalentemente) a un anticuerpo monoclonal adecuado bien directamente o bien indirectamente (por ejemplo, mediante un grupo conector). Una reacción directa entre un agente y un anticuerpo es posible cuando cada uno posee un sustituyente capaz de reaccionar con el otro. Por ejemplo, un grupo nucleofílico, tal como un grupo amino o sulfidrilo, en uno puede ser capaz de reaccionar con un grupo que contiene carbonilo, tal como un anhídrido o haluro de ácido, o con un grupo alquilo que contiene un buen grupo saliente (por ejemplo, un haluro) en el otro.

Alternativamente, puede ser deseable acoplar un agente terapéutico y un anticuerpo a través de un grupo conector. Un grupo conector puede funcionar como un espaciador para distanciar un anticuerpo de un agente con el fin de evitar interferencia con las capacidades de unión. Un grupo conector también puede servir para incrementar la reactividad química de un sustituyente en un agente o un anticuerpo, e incrementar así la eficiencia del acoplamiento. Un incremento en la reactividad química también puede facilitar el uso de agentes, o grupos funcionales en agentes, que de otra manera no sería posible.

Será evidente para los expertos en la técnica que pueden emplearse como grupo conector una variedad de reactivos bifuncionales o polifuncionales, tanto homo- como hetero-funcionales (tales como los descritos en el catálogo de Pierce Chemical Co., Rockford, IL). El acoplamiento puede efectuarse, por ejemplo, a través de grupos amino, grupos carboxilo, grupos sulfidrilo o residuos de carbohidrato oxidados. Existen numerosas referencias que describen dicha metodología, por ejemplo, la Patente de los Estados Unidos Núm. 4.671.958, de Rodwell et al.

Cuando un agente terapéutico es más potente cuando está libre de la porción de anticuerpo de los productos inmunoconjugados de la presente descripción, puede ser deseable utilizar un grupo conector que sea escindible durante o después de la internalización en una célula. Se han descrito varios grupos conectores escindibles diferentes. Los mecanismos para la liberación intracelular de un agente de estos grupos conectores incluyen escisión por reducción de un enlace disulfuro (por ejemplo, Patente de los Estados Unidos Núm. 4.489.710, de Spittler), por irradiación de un enlace fotolábil (por ejemplo, Patente de los Estados Unidos Núm.4.625.014, de Senter et al.), por hidrólisis de cadenas laterales de aminoácidos derivatizadas (por ejemplo, Patente de los Estados Unidos Núm. 4.638.045, de Kohn et al.), por hidrólisis mediada por complemento del suero (por ejemplo, Patente de los Estados Unidos Núm. 4.671.958, de Rodwell et al.), e hidrólisis catalizada por ácido (por ejemplo, Patente de los Estados Unidos Núm. 4.569.789, de Blattler et al.).

Puede ser deseable acoplar más de un agente a un anticuerpo. En una realización, se acoplan múltiples moléculas de un agente a una molécula de anticuerpo. En otra realización, puede acoplarse más de un tipo de agente a un anticuerpo. Independientemente de la realización particular, los productos inmunoconjugados con más de un agente pueden prepararse de varias maneras. Por ejemplo, puede acoplarse más de un agente directamente a una molécula de anticuerpo, o pueden usarse conectores que proporcionan múltiples sitios para la unión.

Como se ha observado anteriormente, los "péptidos" se incluyen como agentes de unión. El término péptido se refiere típicamente a un polímero de residuos de aminoácido y a variantes y análogos sintéticos de los mismos. En ciertas realizaciones descritas en la presente memoria, el término "péptido" se refiere a polipéptidos relativamente cortos, incluyendo péptidos que consisten en aproximadamente 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, o 50 aminoácidos, incluyendo todos los números enteros e intervalos (p. ej., 5-10, 8-12, 10-15) intermedios, e interaccionan con un polipéptido de AARS, su compañero de unión celular, o ambos. Los péptidos pueden estar compuestos por aminoácidos de origen natural y/o aminoácidos de origen no natural, como se describe en la presente memoria.

Un agente de unión puede incluir un peptidomimético u otra molécula pequeña. Una "molécula pequeña" hace referencia a un compuesto orgánico que es de origen sintético o biológico (biomolécula), pero no es típicamente un polímero. Compuestos orgánicos hace referencia a una gran clase de compuestos químicos cuyas moléculas contienen carbono, típicamente excluyendo aquellas que contienen solamente carbonatos, óxidos de carbono simples, o cianuros. Una "biomolécula" hace referencia generalmente a una molécula orgánica que es producida por un organismo vivo, incluyendo moléculas poliméricas grandes (biopolímeros) tales como péptidos, polisacáridos, y también ácidos nucleicos, y moléculas pequeñas tales como metabolitos primarios secundarios, lípidos, fosfolípidos, glicolípidos, esteroides, glicerolípidos, vitaminas, y hormonas. Un "polímero" hace referencia generalmente a una molécula grande o macromolécula compuesta por unidades estructurales repetitivas, que están conectadas típicamente por enlaces químicos covalentes.

En ciertas realizaciones descritas en la presente memoria, una molécula pequeña tiene un peso molecular de menos de 1000 Dalton, típicamente entre aproximadamente 300 y 700 Daltons, e incluyendo aproximadamente 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, o 1000 Dalton.

5 Los aptámeros también se incluyen como agentes de unión. Los ejemplos de los aptámeros incluyen aptámeros de ácidos nucleicos (p. ej., aptámeros de ADN, aptámeros de ARN) y aptámeros de péptidos. Los aptámeros de ácidos nucleicos se refieren en general a especies de ácidos nucleicos que han sido modificadas por medio de rondas repetidas de selección *in vitro* o un método equivalente, tal como SELEX (evolución sistemática de ligandos por enriquecimiento exponencial), para la unión a diversas dianas moleculares tales como moléculas pequeñas, proteínas, ácidos nucleicos, e incluso células, tejidos y organismos. Por consiguiente, se incluyen aptámeros de 10 ácidos nucleicos que se unen a polipéptidos de AARS descritos en la presente memoria y/o sus compañeros de unión celulares.

15 Los aptámeros de péptidos incluyen típicamente un bucle peptídico variable anclado en ambos extremos a un armazón de proteína, una doble restricción estructural que incrementa típicamente la afinidad de unión del aptámero de péptido a niveles comparables con los de un anticuerpo (p. ej., en el intervalo nanomolar). En ciertas realizaciones descritas en la presente memoria, la longitud del bucle variable puede estar compuesta por aproximadamente 10-20 aminoácidos (incluyendo todos los números enteros intermedios), y el armazón puede incluir cualquier proteína que tenga una buena solubilidad y propiedades de compacidad. Ciertas realizaciones 20 ilustrativas descritas en la presente memoria pueden utilizar la proteína bacteriana Tiorredoxina A como una proteína armazón, insertándose el bucle variable dentro del sitio activo reductor (bucle -Cys-Gly-Pro-Cys- en la proteína salvaje), siendo las dos cadenas laterales de las cisteínas capaces de formar un puente disulfuro. Por consiguiente, se encuentran incluidos los aptámeros de péptidos que se unen a los polipéptidos de AARS descritos en la presente memoria y/o sus compañeros de unión celulares. La selección de los aptámeros de péptidos se puede realizar utilizando diferentes sistemas conocidos en la técnica, incluyendo el sistema de dos híbridos de levadura.

25 Como se ha indicado anteriormente, los polipéptidos de AARS y los agentes de unión de la presente descripción se pueden utilizar en cualquiera de los métodos de diagnóstico, descubrimiento de fármacos, o terapéuticos descritos en la presente memoria.

Formulación y administración

30 Las composiciones de la presente descripción (p. ej., polipéptidos, polinucleótidos, anticuerpos, etc.) se formulan generalmente en soluciones farmacéuticamente aceptables o fisiológicamente aceptables para la administración a una célula, tejido o animal, ya sea sola, o combinada con una o más modalidades de terapia distintas. Asimismo se entenderá que, si se desea, las composiciones se pueden administrar también en combinación con otros agentes, tales como, p. ej., otras proteínas o polipéptidos o varios agentes farmacéuticamente activos. No existe virtualmente ningún límite para otros componentes que puedan ser incluidos en las composiciones descritas en la presente memoria, siempre que los agentes adicionales no afecten adversamente a las propiedades de un polipéptido de la 35 presente descripción.

40 En las composiciones farmacéuticas de la presente invención, la formulación de los excipientes y las soluciones de vehículo farmacéuticamente aceptables es muy conocida para los expertos en la técnica, como lo es el desarrollo de regímenes adecuados de dosificación y tratamiento para utilizar las composiciones particulares descritas en la presente memoria en una variedad de regímenes de tratamiento, incluyendo, p. ej., la administración y formulación oral, parenteral, intravenosa, intranasal, intracraneal e intramuscular.

45 En determinadas aplicaciones de la presente descripción, las composiciones farmacéuticas descritas en la presente memoria se pueden entregar mediante administración oral a un sujeto. Como tales, estas composiciones pueden formularse con un diluyente inerte o con un vehículo comestible asimilable, o pueden incluirse en cápsulas de gelatina de cubierta dura o blanda, o pueden comprimirse en comprimidos, o pueden incorporarse directamente con el alimento de la dieta.

50 En determinadas circunstancias, será deseable administrar las composiciones farmacéuticas descritas en la presente memoria parenteralmente, intravenosamente, intramuscularmente, o incluso intraperitonealmente como se describe, por ejemplo en Patente de los Estados Unidos Núm. 5.543.158; Patente de los Estados Unidos Núm. 5.641.515 y Patente de los Estados Unidos Núm. 5.399.363. Las disoluciones de compuestos activos como base libre o sales farmacológicamente aceptables pueden prepararse en agua mezclada adecuadamente con un tensoactivo, tal como hidroxipropilcelulosa. También pueden prepararse dispersiones en glicerol, polietilenglicoles líquidos, y mezclas de éstos y en aceites. En condiciones ordinarias de almacenamiento y uso, estas preparaciones contienen un conservante para prevenir el crecimiento de microorganismos.

55 Las formas farmacéuticas adecuadas para uso inyectable incluyen disoluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de disoluciones o dispersiones inyectables estériles (Patente de los Estados Unidos Núm. 5.466.468). En todos los casos la forma debe ser estéril y debe ser fluida hasta el punto que se pueda manejar fácilmente con una jeringa. Debe ser estable en las condiciones de fabricación y almacenamiento y debe preservarse frente a la acción contaminante de microorganismos, tales como bacterias y

hongos. El portador puede ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, etanol, poliol (p. ej., glicerol, propilenglicol, y polietilenglicol líquido, y similares), mezclas adecuadas de los mismos, y/o aceites vegetales. La fluidez apropiada se puede mantener, por ejemplo, mediante el uso de un recubrimiento tal como lecitina, manteniendo el tamaño de partícula requerido en caso de dispersión y mediante el uso de tensioactivos. La prevención de la acción de microorganismos se puede facilitar con diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico, timesoral, y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares o cloruro sódico. La absorción prolongada de las composiciones inyectables se puede lograr mediante el uso en las composiciones de agentes que retrasen la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Para administración parenteral en una disolución acuosa, por ejemplo, la disolución debe estar adecuadamente tamponada si es necesario y el diluyente líquido se debe hacer isotónico en primer lugar con disolución salina o glucosa suficiente. Estas disoluciones acuosas particulares son especialmente adecuadas para administración intravenosa, intramuscular, subcutánea e intraperitoneal. A este respecto, un medio acuoso estéril que puede emplearse será muy conocido para los expertos en la técnica a la vista de la presente descripción. Por ejemplo, una dosificación se puede disolver en 1 ml de solución isotónica de NaCl y se puede añadir a 1000 ml de fluido de hipodermoclasia o se puede inyectar en el sitio de infusión propuesto (véase, p. ej., Remington's Pharmaceutical Sciences, 15a Edición, pág. 1035-1038 y 1570-1580). Necesariamente se producirá alguna variación en la dosificación dependiendo de la afección del sujeto que se está tratando. La persona responsable de administración, determinará, en cualquier evento, la dosis apropiada para el sujeto individual. Además, para administración a seres humanos, las preparaciones deben cumplir los patrones generales de esterilidad, pirogenicidad, y de seguridad y pureza según requieren los patrones biológicos de la de FDA Office.

Las disoluciones inyectables estériles pueden prepararse incorporando los compuestos activos en la cantidad requerida en el disolvente apropiado con los diferentes otros ingredientes enumerados anteriormente, según se requiera, seguido de esterilización por filtración. Generalmente, las dispersiones se preparan incorporando los diferentes ingredientes activos esterilizados en un vehículo estéril que contiene el medio básico de dispersión y los demás ingredientes requeridos de los enumerados anteriormente. En el caso de polvos estériles para la preparación de disoluciones inyectables estériles, los métodos de preparación preferidos son técnicas de secado en vacío y de liofilización que rinden un polvo del ingrediente activo más cualquier ingrediente adicional deseado de una disolución de éstos previamente esterilizada por filtración.

Las composiciones descritas en la presente memoria pueden formularse en una forma neutra o de sal. Las sales farmacéuticamente aceptables, incluyen las sales de adición de ácido (formadas con los grupos amino libres de la proteína) y que se forman con ácidos inorgánicos tales como, por ejemplo, ácidos clorhídrico o fosfórico, o ácidos orgánicos tales como acético, oxálico, tartárico, mandélico, y similares. Las sales formadas con los grupos carboxilo libres también pueden derivarse de bases inorgánicas tales como, por ejemplo, hidróxidos de sodio, potasio, amonio, calcio, o férrico, y bases orgánicas tales como isopropilamina, trimetilamina, histidina, procaína y similares. Después de la formulación, las disoluciones se administrarán de una manera compatible con la formulación de dosificación y en una cantidad tal que sea terapéuticamente efectiva. Las formulaciones se administran fácilmente en una variedad de formas de dosificación tales como disoluciones inyectables, cápsulas de liberación de fármaco, y similares.

Tal y como se usa en la presente memoria, "portador" incluye cualquiera y todos los disolventes, medios de dispersión, vehículos, recubrimientos, diluyentes, agentes antibacterianos y antifúngicos, agentes isotónicos y que retrasan la absorción, tampones, disoluciones portadoras, suspensiones, coloides, y similares. El uso de dichos medios y agentes para sustancias farmacéuticas activas es muy conocido en la técnica. Excepto si cualquier medio o agente convencional es incompatible con el ingrediente activo, se contempla su uso en las composiciones terapéuticas. También pueden incorporarse ingredientes activos complementarios a las composiciones.

La expresión "farmacéuticamente aceptable" se refiere a entidades moleculares y composiciones que no producen una reacción alérgica o similar inapropiada cuando se administran a un ser humano. La preparación de una composición acuosa que contiene una proteína como un ingrediente activo es muy conocida en la técnica. Típicamente, dichas composiciones se preparan como inyectables, bien como disoluciones o suspensiones líquidas; también pueden prepararse formas sólidas adecuadas para disolución en, o suspensión en, líquido antes de la inyección. La preparación también puede emulsionarse.

En ciertas realizaciones descritas en la presente memoria, las composiciones farmacéuticas se pueden liberar mediante pulverizaciones intranasales, inhalación, y/u otros vehículos de liberación en aerosol. Los métodos para la liberación de genes, polinucleótidos, y composiciones de péptidos directamente a los pulmones a través de pulverizadores de aerosol nasal se han descrito, p. ej., en la Patente de los Estados Unidos Núm. 5.756.353 y la Patente de los Estados Unidos Núm. 5.804.212. Asimismo, la administración de fármacos utilizando resinas de micropartículas intranasales (Takenaga et al., 1998) y compuestos de lisofosfatidil-glicerol (Patente de los Estados Unidos Núm. 5.725.871) es muy conocida en las técnicas farmacéuticas. Asimismo, la administración transmucosal de fármacos en la forma de una matriz de soporte de politetrafluoroetileno se describe en Patente de los Estados Unidos Núm. 5.780.045.

En ciertas realizaciones descritas en la presente memoria, la liberación se puede producir mediante la utilización de liposomas, nanocápsulas, micropartículas, microsferas, partículas de lípidos, vesículas, y similares, para la introducción de las composiciones de la presente descripción en células anfitrionas adecuadas. En particular, las composiciones de la presente descripción se pueden formular para administración bien encapsuladas en una partícula lipídica, un liposoma, una vesícula, una nanoesfera, una nanopartícula o similares. La formulación y uso de dichos vehículos de administración pueden llevarse a cabo usando técnicas conocidas y convencionales.

Kits que comprenden las composiciones

La presente descripción, en otros aspectos, proporciona kits que comprenden uno o más recipientes cargados con uno o más de los polipéptidos, polinucleótidos, anticuerpos, complejos multiunitarios, composiciones de los mismos, etc., de la presente descripción, como se describe en la presente memoria. Los kits pueden incluir instrucciones por escrito sobre cómo utilizar tales composiciones (p. ej., para modular la señalización celular, la angiogénesis, el cáncer, afecciones inflamatorias, etc.).

Los kits en la presente memoria también pueden incluir uno o más agentes terapéuticos adicionales u otros componentes adecuados o deseados para la indicación que esté siendo tratada. Un agente terapéutico adicional puede estar contenido en un segundo recipiente, si se desea. Los ejemplos de los agentes terapéuticos adicionales incluyen, pero no se limitan a agentes antineoplásicos, agentes anti-inflamatorios, agentes antibacterianos, agentes antivirales, agentes angiogénicos, etc.

Los kits en la presente memoria también pueden incluir una o más jeringas u otros componentes necesarios o deseados para facilitar un modo de liberación pretendido (p. ej., dispositivos intraluminales, depósitos implantables, etc.).

Métodos de uso

Las realizaciones de la presente descripción también incluyen métodos de utilización de los "agentes" de aminoacil-ARNt (AARS) descritos en la presente memoria para fines de diagnóstico, descubrimiento de fármacos, y/o terapéuticos. El término "agentes" de AARS hace referencia en general a polinucleótidos de AARS, polipéptidos de AARS, agentes de unión tales como peptidomiméticos, y otros compuestos descritos en la presente memoria. Para fines de diagnóstico, se pueden utilizar los agentes de AARS en una variedad de formas no limitantes, por ejemplo para distinguir entre diferentes tipos de células o estados celulares diferentes, o para identificar sujetos que tienen una enfermedad o afección relevante. Con el fin de descubrir fármacos, se pueden utilizar los agentes de AARS para identificar uno o más "compañeros de unión" celulares de un polipéptido de AARS, caracterizar una o más actividades "no canónicas" de un polipéptido de AARS, identificar agentes que tienen un efecto agonístico o antagonístico selectivo o no selectivo sobre la interacción de un polipéptido de AARS con sus compañeros de unión, y/o identificar los agentes que tienen un efecto agonístico o antagonístico selectivo o no selectivo sobre una o más actividades "no canónicas" de un polipéptido de AARS. Para fines terapéuticos, se pueden utilizar los agentes o composiciones de AARS proporcionados en la presente memoria para tratar una variedad de enfermedades o afecciones, detalladas más abajo.

Descubrimiento de fármacos

Ciertas realizaciones de la presente descripción se refieren al uso de las secuencias de polipéptidos de aminoacil-ARNt sintetasa (AARS) descritas en la presente memoria en el descubrimiento de fármacos, típicamente para identificar agentes que modulan una o más de las actividades no canónicas del polipéptido. Por ejemplo, ciertas realizaciones de la presente descripción incluyen métodos de identificación de uno o más "compañeros de unión" de un polipéptido de AARS de la presente descripción, tal como una proteína celular u otra molécula del anfitrión que está asociada con el polipéptido y participa en su actividad o en sus actividades no canónicas. Asimismo se describen en la presente memoria los métodos para identificar un compuesto (p. ej., polipéptido) u otro agente que suscita agonismo o antagonismo de la actividad no canónica de un polipéptido de referencia o una variante activa del mismo, por ejemplo mediante interacción con el polipéptido y/o uno o más de sus compañeros de unión celulares.

Ciertas realizaciones de la presente descripción incluyen por lo tanto métodos de identificación de un compañero de unión de un polipéptido de AARS, que comprenden a) combinar el polipéptido de AARS con una muestra biológica en condiciones adecuadas, y b) detectar la unión específica del polipéptido de AARS a un compañero de unión, identificando de ese modo un compañero de unión que se une específicamente al polipéptido de AARS. Asimismo se describen en la presente memoria los métodos de escrutinio de un compuesto que se une específicamente a un polipéptido de AARS o un compañero de unión del polipéptido, que comprenden a) combinar el polipéptido o el compañero de unión con al menos un compuesto de ensayo en condiciones adecuadas, y b) detectar la unión del polipéptido o el compañero de unión al compuesto de ensayo, identificando de ese modo un compuesto que se une específicamente al polipéptido o su compañero de unión. En ciertas realizaciones descritas en la presente memoria, el compuesto es un polipéptido o un péptido. En ciertas realizaciones descritas en la presente memoria, el compuesto es una molécula pequeña u otro compuesto químico (p. ej., no biológicos). En ciertas realizaciones descritas en la presente memoria, el compuesto es un peptidomimético.

Se puede emplear cualquier método adecuado para detectar interacciones proteína-proteína para identificar proteínas celulares que interaccionan con un polipéptido de AARS, interaccionan con uno o más de sus compañeros de unión celulares, o ambos. Los ejemplos de los métodos tradicionales que se pueden emplear incluyen co-inmunoprecipitación, entrecruzamiento, y co-purificación a través de columnas con gradientes o cromatográficas de productos lisados celulares o proteínas obtenidas de productos lisados celulares, principalmente para identificar proteínas en el producto lisado que interaccionan con el polipéptido de AARS.

En estas realizaciones y en realizaciones relacionadas de la presente descripción, al menos una porción de la secuencia de aminoácidos de una proteína que interacciona con un polipéptido de AARS o su compañero de unión puede ser averiguada utilizando mecanismos bien conocidos por los expertos en la técnica, por ejemplo a través del mecanismo de degradación de Edman. Véase, p. ej., Creighton Proteins: Structures and Molecular Principles, W. H. Freeman & Co., N.Y., págs. 34-49, 1983. La secuencia de aminoácidos obtenida se puede utilizar como una guía para la generación de mezclas de oligonucleótidos que se pueden utilizar para escrutar secuencias génicas que codifican tales proteínas. El escrutinio se puede completar, por ejemplo, por medio de hibridación convencional o técnicas de PCR, como se describe en la presente memoria y se conoce en la técnica. Las técnicas para la generación de mezclas de oligonucleótidos y el escrutinio son bien conocidas. Véase, p. ej., Ausubel et al. Current Protocols in Molecular Biology Green Publishing Associates and Wiley Interscience, N.Y., 1989; e Innis et al., eds. PCR Protocols: A Guide to Methods and Applications Academic Press, Inc., Nueva York, 1990.

Adicionalmente, los métodos se pueden emplear en la identificación simultánea de genes que codifican el compañero de unión u otro polipéptido. Estos métodos incluyen, por ejemplo, el sondeo de bibliotecas de expresión, de una manera similar al mecanismo bien conocido de sondeo con anticuerpos de bibliotecas de lambda-gt11, utilizando una proteína AARS marcada, u otro polipéptido, péptido o proteína de fusión, p. ej., un polipéptido de AARS variante o un dominio de AARS fusionado a un marcador (p. ej., una enzima, flúor, una proteína luminiscente, o un colorante), o un dominio Fc de Ig.

Un método que detecta las interacciones de proteínas *in vivo*, el sistema de dos híbridos, se describe con detalle solamente como ilustración y no a modo de limitación. Un ejemplo de este sistema ha sido descrito (Chien et al., PNAS USA 88:9578-9582, 1991) y se encuentra disponible en el mercado en Clontech (Palo Alto, Calif.).

En resumen, utilizando semejante sistema, se pueden construir plásmidos que codifican dos proteínas híbridas: un plásmido consiste en nucleótidos que codifican el dominio de unión a ADN de una proteína activadora de la transcripción fusionada a una secuencia de nucleótidos de AARS de referencia (o, en ciertas realizaciones, su compañero de unión), o una variante de la misma, y el otro plásmido consiste en nucleótidos que codifican el dominio de activación de la proteína activadora de la transcripción fusionado a un ADNc (o una colección de ADNc) que codifica una o varias proteínas desconocidas que han sido recombinadas en el plásmido como parte de una biblioteca de ADNc. El plásmido de fusión del ADN-dominio de unión y la biblioteca de ADNc activador se pueden transformar en una cepa de la levadura *Saccharomyces cerevisiae* que contiene un gen informador (p. ej., HBS o lacZ) cuya región reguladora contiene el sitio de unión del activador de la transcripción. Cualquier proteína híbrida no puede activar la transcripción del gen informador: el híbrido de ADN-dominio de unión no puede porque no proporciona función de activación y el híbrido del dominio de activación no puede porque no puede localizar los sitios de unión del activador. La interacción de las dos proteínas híbridas reconstituye la proteína activadora funcional y da como resultado la expresión del gen informador, que es detectada por medio de un análisis para el producto génico del informador.

El sistema de dos híbridos u otra metodología semejante se pueden utilizar para escrutar las bibliotecas de dominios de activación para determinar las proteínas que interaccionan con el producto génico "cebo". A modo de ejemplo, y no de limitación, se puede utilizar un polipéptido de referencia de AARS o una variante como producto génico "cebo". También se puede utilizar un compañero de unión de AARS como producto génico "cebo". Las secuencias genómicas o de ADNc totales se fusionan al ADN que codifica un dominio de activación. Esta biblioteca y un plásmido que codifica un híbrido de un producto génico de AARS "cebo" fusionado al dominio de unión a ADN se transforman simultáneamente en una cepa informadora de levadura, y los transformantes resultantes se escrutan para determinar aquellos que expresan el gen informador.

Se puede elaborar una biblioteca de ADNc de la línea celular de la cual se van a detectar las proteínas que interaccionan con los productos génicos de AARS "cebo" utilizando métodos rutinarios puestos en práctica en la técnica. Por ejemplo, los fragmentos de ADNc se pueden insertar en un vector de manera que se fusionan tradicionalmente al dominio de activación transcripcional de GAL4. Esta biblioteca se puede transformar simultáneamente con el plásmido de fusión del gen "cebo"-GAL4 en una cepa de levadura, que contiene un gen lacZ dirigido por un promotor que contiene la secuencia de activación GAL4. Una proteína codificada por ADNc, fusionada al dominio de activación transcripcional GAL4, que interacciona con el producto génico "cebo" reconstituirá una proteína GAL4 activa y de ese modo dirigirá la expresión del gen HIS3. Las colonias, que expresan HIS3, se pueden detectar por medio de su crecimiento sobre placas Petri que contienen medios basados en agar semisólido carentes de histidina. El ADNc se puede purificar después a partir de estas cepas, y se puede utilizar para producir y aislar la proteína de interacción con el gen AARS "cebo" utilizando mecanismos rutinarios puestos en práctica en la técnica.

Asimismo se describen en la presente memoria sistemas de tres híbridos, que permiten la detección de interacciones ARN-proteína en levaduras. Véase, *p. ej.*, Hook et al., ARN. 11:227-233, 2005. Por consiguiente, se pueden utilizar estos métodos y métodos relacionados para identificar un compañero de unión celular de un polipéptido de AARS, y para identificar otras proteínas o ácidos nucleicos que interactúan con el polipéptido de AARS, el compañero de unión celular, o ambos.

Como se ha indicado anteriormente, una vez aislados, los compañeros de unión pueden ser identificados y, a su vez, pueden ser utilizados junto con mecanismos convencionales para identificar proteínas u otros compuestos con los cuales interactúan. Ciertas realizaciones de la presente descripción se refieren por lo tanto a métodos de escrutinio de un compuesto que se une específicamente a un compañero de unión de un polipéptido de referencia de AARS, que comprenden a) combinar el compañero de unión con al menos un compuesto de ensayo en condiciones adecuadas, y b) detectar la unión del compañero de unión al compuesto de ensayo, identificando de ese modo un compuesto que se une específicamente al compañero de unión. En ciertas realizaciones de la presente descripción, el compuesto de ensayo es un polipéptido. En ciertas realizaciones descritas en la presente memoria, el compuesto de ensayo es un compuesto químico, tal como un compuesto de molécula pequeña o un peptidomimético.

Ciertas realizaciones descritas en la presente memoria incluyen métodos de escrutinio de un compuesto que modula la actividad de un polipéptido de AARS, que comprenden a) combinar el polipéptido con al menos un compuesto de ensayo en condiciones permisivas para la actividad del polipéptido, b) evaluar la actividad del polipéptido en presencia del compuesto de ensayo, y c) comparar la actividad del polipéptido en presencia del compuesto de ensayo con la actividad del polipéptido en ausencia del compuesto de ensayo, en donde un cambio en la actividad del polipéptido en presencia del compuesto de ensayo es indicativa de un compuesto que modula la actividad del polipéptido. Ciertas realizaciones descritas en la presente memoria incluyen métodos de escrutinio de un compuesto que modula la actividad de un compañero de unión de un polipéptido de AARS, que comprenden a) combinar el polipéptido con al menos un compuesto de ensayo en condiciones permisivas para la actividad del compañero de unión, b) evaluar la actividad del compañero de unión en presencia del compuesto de ensayo, y c) comparar la actividad del compañero de unión en presencia del compuesto de ensayo con la actividad del compañero de unión en ausencia del compuesto de ensayo, en donde un cambio en la actividad del compañero de unión en presencia del compuesto de ensayo es indicativa de un compuesto que modula la actividad del compañero de unión. Típicamente, estas realizaciones y realizaciones relacionadas descritas en la presente memoria incluyen la evaluación de una actividad no canónica seleccionada que está asociada con un polipéptido de AARS o su compañero de unión. Se describen en la presente memoria las condiciones *in vitro* e *in vivo*, tales como las condiciones de cultivo celular.

Ciertas realizaciones descritas en la presente memoria incluyen los métodos de escrutinio de un compuesto para determinar la eficacia como agonista total o parcial de un polipéptido de AARS o un fragmento activo o una variante del mismo, que comprenden a) exponer una muestra que comprende el polipéptido a un compuesto, y b) detectar la actividad agonística en la muestra, típicamente midiendo un incremento en la actividad no canónica del polipéptido de AARS. Ciertos métodos incluyen a) exponer una muestra que comprende un compañero de unión del polipéptido de AARS a un compuesto, y b) detectar la actividad agonística en la muestra, típicamente midiendo un incremento en la actividad no canónica seleccionada del polipéptido de AARS. Ciertas realizaciones de la presente descripción incluyen composiciones que comprenden un compuesto agonista identificado por el método y un portador o excipiente farmacéuticamente aceptables.

Asimismo se describen en la presente memoria los métodos de escrutinio de un compuesto para determinar la eficacia como antagonista total o parcial de un polipéptido de AARS, que comprenden a) exponer una muestra que comprende el polipéptido a un compuesto, y b) detectar la actividad antagónica en la muestra, típicamente midiendo un descenso en la actividad no canónica del polipéptido de AARS. Ciertos métodos incluyen a) exponer una muestra que comprende un compañero de unión del polipéptido de AARS a un compuesto, y b) detectar la actividad antagónica en la muestra, típicamente midiendo un descenso en la actividad no canónica seleccionada del polipéptido de AARS. Ciertas realizaciones descritas en la presente memoria incluyen composiciones que comprenden un compuesto antagónico identificado por el método y un portador o excipiente farmacéuticamente aceptables.

En ciertas realizaciones descritas en la presente memoria, se pueden diseñar sistemas *in vitro* para identificar compuestos capaces de interactuar con o modular una secuencia de referencia de AARS o su compañero de unión. Ciertos compuestos identificados por medio de tales sistemas pueden ser útiles, por ejemplo, en la modulación de la actividad de la ruta, y en la elaboración de los componentes de la propia ruta. Asimismo se pueden utilizar en escrutinios para identificar compuestos que interrumpen las interacciones entre los componentes de la ruta; o pueden interrumpir tales interacciones directamente. Un enfoque ilustrativo implica preparar una mezcla de reacción del polipéptido de AARS y un compuesto de ensayo en condiciones y durante un tiempo suficiente para permitir que los dos interactúen y se unan, formando de ese modo un complejo que puede ser eliminado de y/o detectado en la mezcla de reacción.

Los análisis de escrutinio *in vitro* se pueden llevar a cabo en una variedad de formas. Por ejemplo, un polipéptido de AARS, un compañero de unión celular, o uno o varios compuestos de ensayo se pueden anclar sobre una fase

sólida. En estas realizaciones y en realizaciones relacionadas descritas en la presente memoria, los complejos resultantes se pueden capturar y detectar sobre la fase sólida al final de la reacción. En un ejemplo de semejante método, el polipéptido de AARS y/o su compañero de unión se anclan sobre una superficie sólida, y el compuesto o los compuestos de ensayo, que no están anclados, se pueden marcar, directamente o indirectamente, de manera que se pueda detectar su captura por el componente sobre la superficie sólida. En otros ejemplos, el compuesto o los compuestos de ensayo se anclan a la superficie sólida, y el polipéptido de AARS y/o su compañero de unión, que no están anclados, se marcan o se hacen detectables de algún modo. En ciertas realizaciones descritas en la presente memoria, se pueden utilizar convenientemente placas de microtitulación como fase sólida. El componente anclado (o compuesto de ensayo) puede ser inmovilizado por medio de anclajes no covalentes o covalentes. El anclaje no covalente se puede completar simplemente recubriendo la superficie sólida con una solución de la proteína y secando. Alternativamente, se puede utilizar un anticuerpo inmovilizado, preferiblemente un anticuerpo monoclonal, específico para la proteína que se va a inmovilizar para anclar la proteína a la superficie sólida. Las superficies se pueden preparar de antemano y se pueden almacenar.

Para llevar a cabo un análisis ilustrativo, el componente no inmovilizado se añade típicamente a la superficie recubierta que contiene el componente anclado. Una vez que la reacción se ha completado, los componentes que no han reaccionado se eliminan (p. ej., mediante lavado) en condiciones tales que cualquiera de los complejos específicos formados permanezcan inmovilizados sobre la superficie sólida. La detección de complejos anclados sobre la superficie sólida se puede completar de numerosas maneras. Por ejemplo, cuando el componente previamente no inmovilizado está pre-marcado, la detección de la marca inmovilizada sobre la superficie indica que se habían formado complejos. Cuando el componente previamente no inmovilizado no está pre-marcado, se puede utilizar una marca indirecta para detectar los complejos anclados sobre la superficie; p. ej., utilizando un anticuerpo marcado específico para el componente previamente no inmovilizado (el anticuerpo, a su vez, puede estar marcado directamente o marcado indirectamente con un anticuerpo anti-Ig marcado).

Alternativamente, se puede determinar la presencia o ausencia de unión de un compuesto de ensayo, por ejemplo, utilizando la resonancia de plasmón superficial (SPR) y el cambio en el ángulo de resonancia como un índice, en donde un polipéptido de AARS o un compañero de unión celular es inmovilizado sobre la superficie de un chip sensor disponible en el mercado (p. ej., fabricado por Biacore™) de acuerdo con un método convencional, el compuesto de ensayo se pone en contacto con el mismo, y el chip sensor se ilumina con una luz de una longitud de onda concreta desde un ángulo concreto. La unión de un compuesto de ensayo también se puede medir mediante la detección de la aparición de un pico correspondiente al compuesto de ensayo mediante un método en donde un polipéptido de AARS o un compañero de unión celular son inmovilizados sobre la superficie de un chip de proteína adaptable a un espectrómetro de masas, un compuesto de ensayo se pone en contacto con el mismo, y un método de ionización tal como MALDI-MS, ESI-MS, FAB-MS y similares se combina con un espectrómetro de masas (p. ej., espectrómetro de masas de doble enfoque, espectrómetro de masas de cuadrupolo, espectrómetro de masas de tiempo de vuelo, espectrómetro de masas de transformada de Fourier, espectrómetro de masas de ión ciclotrón y similares).

En ciertas realizaciones descritas en la presente memoria, se pueden utilizar análisis basados en células, análisis basados en vesículas de membrana, o análisis basados en fracciones de membrana para identificar compuestos que modulan las interacciones en la ruta no canónica del polipéptido de AARS seleccionado. Con este fin, se pueden utilizar líneas celulares que expresan un polipéptido de AARS y/o un compañero de unión, o una proteína de fusión que contiene un dominio o un fragmento de tales proteínas (o una combinación de los mismos), o líneas celulares (p. ej., células COS, células CHO, células HEK293, células Hela, etc.) que han sido modificadas genéticamente para expresar tales proteínas o proteínas de fusión. Los compuestos de ensayo que influyen en la actividad no canónica se pueden identificar controlando el cambio (p. ej., un cambio estadísticamente significativo) en esa actividad en comparación con una cantidad de control o predeterminada.

Para las realizaciones de la presente descripción que relacionan los agentes antisentido y ARNi, por ejemplo, también se incluyen los métodos de escrutinio de un compuesto para determinar la eficacia en la alteración de la expresión de una secuencia de polinucleótido de AARS, que comprenden a) exponer una muestra que comprende el polinucleótido de AARS a un compuesto tal como un oligonucleótido antisentido potencial, y b) detectar la alteración de la expresión del polinucleótido de AARS. En ciertos ejemplos no limitantes, se pueden emplear estas realizaciones y realizaciones relacionadas descritas en la presente memoria en análisis basados en células o en análisis de traducción libres de células, de acuerdo con mecanismos rutinarios en la técnica. Asimismo se describen oligonucleótidos antisentido y agentes de ARNi identificados por tales métodos.

También se describe en la presente memoria cualquiera de los métodos anteriores, u otros métodos de escrutinio conocidos en la técnica, que están adaptados para el escrutinio de alto rendimiento (HTS). El HTS típicamente utiliza la automatización para ejecutar un escrutinio de un análisis frente a una biblioteca de compuestos candidato, por ejemplo, un análisis que mide un incremento o un descenso en una actividad no canónica, como se describe en la presente memoria.

Métodos de tratamiento

En otro aspecto, la presente descripción se refiere a métodos de utilización de las composiciones de la presente descripción para el tratamiento de una célula, tejido o sujeto con una composición como se describe en la presente memoria. Las células o tejidos que se pueden modular por medio de la presente descripción son preferiblemente células o tejidos de mamífero, o más preferiblemente células o tejidos de ser humano. Tales células o tejidos pueden proceder de un estado sano o de un estado enfermo.

En determinadas realizaciones de la presente descripción, por ejemplo, se proporcionan métodos para modular actividades celulares terapéuticamente relevantes incluyendo, pero no limitadas a, metabolismo celular, diferenciación celular, proliferación celular, muerte celular, movilización celular, migración celular, transcripción génica, traducción del ARNm, impedancia celular, y similares, que comprenden poner en contacto una célula con una composición como se describe en la presente memoria. Los ejemplos de las células o tejidos que pueden ser modulados incluyen los de la piel (p. ej., dermis, epidermis, capa subcutánea), folículos pilosos, sistema nervioso (p. ej., cerebro, médula espinal, nervios periféricos), sistema auditivo u órganos del equilibrio (p. ej., oído interno, oído medio, oído externo), sistema respiratorio (p. ej., nariz, tráquea, pulmones), tejidos gastroesofágicos, el sistema gastrointestinal (p. ej., boca, esófago, estómago, intestino delgado, intestino grueso, recto), sistema vascular (p. ej., corazón, vasos sanguíneos, y arterias), hígado, vesícula biliar, sistema linfático/inmunitario (p. ej., nódulos linfáticos, folículos linfoides, bazo, timo, médula ósea), sistema uro-genital (p. ej., riñones, uréter, vejiga, uretra, cuello uterino, trompas de Falopio, ovarios, útero, vulva, próstata, glándulas bulbouretrales, epidídimo, próstata, vesículas seminales, testículos), sistema músculo-esquelético (p. ej., músculos esqueléticos, músculos lisos, hueso, cartílago, tendones, ligamentos), tejido adiposo, mamas, y el sistema endocrino (p. ej., hipotálamo, pituitaria, tiroides, páncreas, glándulas adrenales). Por consiguiente, las composiciones se pueden emplear en el tratamiento esencialmente de cualquier célula o tejido o sujeto que se beneficiaría de una o más de tales actividades ilustrativas, o uno o más de tales células o tejidos ilustrativos.

Las composiciones descritas en la presente memoria también se pueden utilizar en cualquiera de varios contextos terapéuticos incluyendo, por ejemplo, los relacionados con el tratamiento o la prevención de enfermedades neoplásicas, enfermedades del sistema inmunitario (p. ej., enfermedades autoinmunitarias e inflamación), enfermedades infecciosas, enfermedades metabólicas, enfermedades neuronales/neurológicas, enfermedades musculares/cardiovasculares, enfermedades asociadas con hematopoyesis aberrante, enfermedades asociadas con angiogénesis aberrante, enfermedades asociadas con supervivencia celular aberrante, y otras.

Por ejemplo, en ciertas realizaciones ilustrativas descritas en la presente memoria, las composiciones de la presente descripción se pueden utilizar para modular la angiogénesis, p. ej., a través de la modulación de la proliferación y/o la señalización de las células endoteliales. La proliferación y/o la señalización de células endoteliales se pueden controlar utilizando una línea celular apropiada (p. ej., células de pulmón endoteliales microvasculares humanas (HMVEC-L) y células endoteliales de vena umbilical humana (HUVEC)), y utilizando un análisis apropiado (p. ej., análisis de migración de células endoteliales, análisis de proliferación de células endoteliales, análisis de formación de tubo, análisis de tapón de matrigel, etc.), muchos de los cuales son conocidos y se encuentran disponibles en la técnica.

Por lo tanto, en realizaciones relacionadas descritas en la presente memoria, las composiciones de la presente descripción se pueden emplear en el tratamiento esencialmente de cualquier células o tejido o sujeto que se beneficiaría de la modulación de la angiogénesis. Por ejemplo, en algunas realizaciones descritas en la presente memoria, se puede poner en contacto una célula o tejido o sujeto que está experimentando o es susceptible de angiogénesis (p. ej., una afección angiogénica) con una composición de la presente descripción para inhibir una afección angiogénica. En otras realizaciones descritas en la presente memoria, se puede poner en contacto una célula o tejido que está experimentando o es susceptible de angiogénesis insuficiente (p. ej., una afección angiostática) con una composición apropiada de la presente descripción con el fin de interferir en la actividad angiostática y/o promover la angiogénesis.

Los ejemplos ilustrativos de las afecciones angiogénicas incluyen, pero no se limitan a, degeneración macular relacionada con la edad (AMD), cáncer (tanto sólido como hematológico), anomalías del desarrollo (organogénesis), ceguera por diabetes, endometriosis, neovascularización ocular, psoriasis, artritis reumatoide (RA), y decoloraciones de la piel (p. ej., hemangioma, nevus flammeus o nevus simplex). Los ejemplos de las afecciones anti-angiogénicas incluyen, pero no se limitan a, enfermedad cardiovascular, restenosis, lesión tisular tras reperusión de tejido isquémico o insuficiencia cardíaca, inflamación crónica y curación de heridas.

Las composiciones de la presente descripción también pueden ser útiles como inmunomoduladores para el tratamiento de indicaciones pro-inflamatorias mediante la modulación de las células que median, bien directamente bien indirectamente, enfermedades, afecciones y trastornos autoinmunitarios y/o inflamatorios. La utilidad de las composiciones de la presente descripción como inmunomoduladores se puede controlar utilizando cualquiera de los numerosos mecanismos conocidos y disponibles en la técnica incluyendo, por ejemplo, análisis de migración (p. ej., utilizando leucocitos o linfocitos) o análisis de viabilidad celular (p. ej., utilizando células B, células T, monocitos o células NK). En ciertas realizaciones descritas en la presente memoria, las composiciones de la presente descripción pueden modular las citoquinas inflamatorias, tales como TNF- α . En ciertas realizaciones descritas en la

presente memoria, las composiciones de la presente descripción pueden modular la quimiotaxis de células inmunitarias, tales como monocitos.

"Inflamación" se refiere generalmente a la respuesta biológica de tejidos a estímulos dañinos, tales como patógenos, células dañadas (por ejemplo, heridas), e irritantes. El término "respuesta inflamatoria" se refiere a los mecanismos específicos por los que la inflamación se consigue y regula, incluyendo, meramente como ilustración, activación o migración de células inmunes, producción de citoquinas, vasodilatación, incluyendo liberación de quinina, fibrinolisis, y coagulación, entre otros descritos en la presente memoria y conocidos en la técnica. Idealmente, la inflamación es un intento protector del cuerpo tanto para eliminar el estímulo injurioso como para iniciar el proceso de cicatrización para el tejido o tejidos afectados. En ausencia de inflamación las heridas e infecciones nunca cicatrizarían, creando una situación en la que la destrucción progresiva del tejido amenazaría la supervivencia. Por otra parte, la inflamación excesiva o crónica puede estar asociada con una variedad de enfermedades, tal como fiebre del heno, aterosclerosis, y artritis reumatoide, entre otras descritas en la presente memoria y conocidas en la técnica.

Las composiciones de la presente descripción pueden modular la inflamación aguda, la inflamación crónica, o ambas. Ciertas realizaciones de la presente descripción se refieren al aumento de inflamación aguda o de las respuestas inflamatorias agudas, y determinadas realizaciones descritas en la presente memoria se refieren a la reducción de la inflamación crónica o de las respuestas inflamatorias crónicas. Dependiendo de las necesidades del sujeto, determinadas realizaciones descritas en la presente memoria se refieren a la reducción de la inflamación aguda o de las respuestas inflamatorias, y determinadas realizaciones descritas en la presente memoria se refieren a la reducción de la inflamación crónica o de las respuestas inflamatorias crónicas.

La inflamación aguda se refiere a la respuesta inicial del cuerpo a estímulos presumiblemente dañinos e implica un movimiento incrementado de plasma y leucocitos desde la sangre a los tejidos dañados. Es un proceso a corto plazo, empezando típicamente en minutos u horas y finalizando después de la eliminación del estímulo injurioso. La inflamación aguda puede caracterizarse por una cualquiera o más de enrojecimiento, calor incrementado, hinchazón, dolor, y pérdida de función. El enrojecimiento y calor se deben principalmente a un flujo sanguíneo incrementado a temperatura corporal central al sitio inflamado, la hinchazón está causada por la acumulación de fluido, el dolor se debe típicamente a la liberación de agentes químicos que estimulan las terminaciones nerviosas, y la pérdida de función tiene múltiples causas.

Las respuestas inflamatorias agudas se inician principalmente por células inmunitarias locales, tales como macrófagos residentes, células dendríticas, histiocitos, células de Kupffer y mastocitos. En el inicio de una infección, quemadura, u otras lesiones, estas células experimentan activación y liberan mediadores inflamatorios responsables de los signos clínicos de la inflamación, tales como aminas vasoactivas y eicosanoides. La vasodilatación y su flujo sanguíneo incrementado resultante causan el enrojecimiento y aumento de calor. El aumento de permeabilidad de los vasos sanguíneos da como resultado una exudación o extravasación de proteínas plasmáticas y fluido en el tejido, lo que crea la hinchazón. Determinados mediadores liberados tales como bradiquinina incrementan la sensibilidad al dolor, y alteran los vasos sanguíneos para permitir la migración o extravasación de leucocitos, tales como neutrófilos, que migran típicamente a lo largo de un gradiente quimiotáctico creado por las células inmunitarias locales.

Las respuestas inflamatorias agudas también incluyen uno o más sistemas de cascada bioquímicos acelulares, que consisten en un producto modulado de proteínas plasmáticas preformadas, que actúan en paralelo para iniciar y propagar la respuesta inflamatoria. Estos sistemas incluyen el sistema del complemento, que se activa principalmente por bacterias, y los sistemas de coagulación y fibrinolisis, que se activan principalmente por necrosis, tal como el tipo de daño tisular que es causado por determinadas infecciones, quemaduras, u otros traumas. Por consiguiente, las composiciones de la descripción se pueden utilizar para modular la inflamación aguda, o cualquiera de una o más de las respuestas inflamatorias agudas individuales.

La inflamación crónica, una respuesta inflamatoria prolongada y retrasada, se caracteriza por un desplazamiento progresivo en el tipo de células que están presentes en el sitio de la inflamación, y da lugar frecuentemente a la destrucción y cicatrización simultánea o casi simultánea del tejido a partir del proceso inflamatorio. A nivel celular, las respuestas inflamatorias crónicas implican una variedad de células inmunitarias tales como monocitos, macrófagos, linfocitos, células plasmáticas, y fibroblastos, aunque a diferencia de la inflamación aguda, que está mediada principalmente por granulocitos, la inflamación crónica está mediada principalmente por células mononucleares tales como monocitos y linfocitos. La inflamación crónica también implica una variedad de mediadores inflamatorios, tales como IFN- γ y otras citoquinas, factores de crecimiento, especies de oxígeno reactivas, y enzimas hidrolíticas. La inflamación crónica puede durar durante muchos meses o años, y puede dar como resultado la destrucción de tejido y fibrosis no deseada.

Los signos clínicos de la inflamación crónica dependen de la duración de la enfermedad, lesiones inflamatorias, causa y área anatómica afectada. (Véase, por ejemplo, Kumar et al., Robbins Basic Pathology-8^a Ed., 2009 Elsevier, Londres; Miller, LM, Pathology Lecture Notes, Atlantic Veterinary College, Charlottetown, PEI, Canadá). La inflamación crónica está asociada con una variedad de afecciones o enfermedades patológicas, incluyendo, por ejemplo, alergias, enfermedad de Alzheimer, anemia, estenosis de la válvula aórtica, artritis tal como artritis reumatoide y osteoartritis, cáncer, fallo cardíaco congestivo, fibromialgia, fibrosis, ataque cardíaco, fallo renal, lupus,

pancreatitis, ictus, complicaciones quirúrgicas, enfermedad pulmonar inflamatoria, enfermedad inflamatoria intestinal, aterosclerosis, y psoriasis, entre otras descritas en la presente memoria y conocidas en la técnica. Por lo tanto, las composiciones proporcionada en la presente memoria se pueden utilizar para tratar o gestionar la inflamación crónica, modular cualquiera de las una o más respuestas inflamatorias crónicas individuales, o tratar cualquiera de las una o más enfermedades o afecciones asociadas con la inflamación crónica.

Las composiciones de la presente descripción también pueden modular la inflamación proliferativa, un proceso inflamatorio caracterizado por un incremento en el número de células tisulares. Éstas pueden englobar afecciones de la piel tales como psoriasis, seborrea o eccema, o también puede pensarse en ellas en términos de cánceres y crecimientos anormales especialmente a la vista de la evidencia que se acumula basada en métodos moleculares más eficientes para documentar incluso inflamación crónica de grado bajo.

Los criterios para evaluar los signos y síntomas de las afecciones inflamatorias y otras, incluyendo el propósito de realizar un diagnóstico diferencial y también para controlar tratamientos tales como la determinación de si se ha administrado una dosis terapéuticamente eficaz en el curso del tratamiento, por ejemplo, la determinación de la mejoría de acuerdo con criterios clínicos aceptados, serán evidentes para los expertos en la técnica y se ilustran por las enseñanzas, por ejemplo, de Berkow et al., eds., The Merck Manual, 16a edición, Merck y Co., Rahway, N.J., 1992; Goodman et al., eds., Goodman y Gilman's The Pharmacological Basis of Therapeutics, 10a edición, Pergamon Press, Inc., Elmsford, N.Y., (2001); Avery's Drug Treatment: Principles and Practice of Clinical Pharmacology and Therapeutics, 3^a edición, ADIS Press, Ltd., Williams y Wilkins, Baltimore, MD. (1987); Ebadi, Pharmacology, Little, Brown y Co., Boston, (1985); Osolci al., eds., Remington's Pharmaceutical Sciences, 18a edición, Mack Publishing Co., Easton, PA (1990); Katzung, Basic and Clinical Pharmacology, Appleton y Lange, Norwalk, CT (1992).

Las enfermedades, trastornos o afecciones del sistema inmunitario ilustrativos que se pueden tratar de acuerdo con la presente invención incluyen, pero no están limitados a, inmunodeficiencias primarias, trombocitopenia mediada por inmunidad, síndrome de Kawasaki, trasplante de médula ósea (por ejemplo, trasplante reciente de médula ósea en adultos o niños), leucemia linfocítica crónica de células B, infección por VIH (por ejemplo, infección por VIH de adulto o pediátrica), polineuropatía desmielinizante inflamatoria crónica, púrpura post-transfusión, y similares.

Como se ha observado anteriormente, ciertas realizaciones descritas en la presente memoria pueden emplear las composiciones de la presente descripción para incrementar la inflamación o incrementar la respuesta inmunitaria. Por ejemplo, dependiendo de las necesidades del sujeto, ciertas realizaciones descritas en la presente memoria pueden incrementar la inflamación aguda o incrementar las respuestas inflamatorias agudas o ambas. Ciertas realizaciones descritas en la presente memoria pueden incrementar la inflamación crónica o las respuestas inflamatorias crónicas o ambas. Ciertas realizaciones descritas en la presente memoria pueden incrementar la inflamación tanto aguda como crónica. Ciertas realizaciones descritas en la presente memoria pueden incrementar la inflamación local o sistémica o ambas.

En ciertas realizaciones descritas en la presente memoria, se pueden utilizar las composiciones para tratar o gestionar inmunodeficiencias, incluyendo inmunodeficiencias primarias e inmunodeficiencias secundarias, en las que el organismo puede no montar una respuesta inflamatoria adecuada. Los ejemplos de las inmunodeficiencias primarias incluyen diversas afecciones genéticas recesivas y ligadas a X tales como inmunodeficiencias de células T y de células B, incluyendo las inmunodeficiencias de células T y de células B combinadas, deficiencias de anticuerpos, síndromes bien definidos, enfermedades de desregulación inmunitaria, trastornos de fagocitos, trastornos innatos de la inmunidad, y deficiencias del complemento.

Los ejemplos de las inmunodeficiencias de células T y células B incluyen deficiencias T-/B+ tales como deficiencia de γc , deficiencia de JAK3, deficiencia de la cadena α del receptor de interleuquina 7, deficiencia de CD45, deficiencia de CD3 δ /CD3 ϵ ; y deficiencias T-/B- tales como deficiencia de RAG 1/2, deficiencia de DCLRE1C, deficiencia de adenosina desaminasa (ADA), disgénesis reticular. Otros ejemplos incluyen el síndrome de Omenn, la deficiencia de ADN ligasa de tipo IV, la deficiencia de ligando de CD40, deficiencia de CD40, deficiencia de purina nucleósido fosforilasa (PNP), deficiencia de MHC clase II, deficiencia de CD3 γ , deficiencia de CD8, deficiencia de ZAP-70, deficiencia de TAP-1/2, y deficiencia de hélice alada.

Los ejemplos de las deficiencias de anticuerpos incluyen agammaglobulinemia ligada a X (deficiencia de btk, o agammaglobulinemia de Bruton), deficiencia de cadena pesada μ , deficiencia de I-5, deficiencia de Ig α , deficiencia de BLNK, timoma con inmunodeficiencia, inmunodeficiencia variable común (CVID), deficiencia de ICOS, deficiencia de CD19, deficiencia de TACI (TNFRSF13B), y deficiencia de receptor BAFF. Los ejemplos adicionales incluyen deficiencia de AID, deficiencia de UNG, deleciones de cadena pesada, deficiencia de cadena kappa, deficiencias de la subclase IgG aislada, deficiencia de la subclase IgA con IgG, deficiencia de inmunoglobulina A selectiva, e hipogammaglobulinemia transitoria de la infancia (THI).

Los ejemplos de los "síndrome bien definidos" incluyen síndrome de Wiskott-Aldrich, ataxia telangiectasia, síndrome de tipo ataxia, síndrome de rotura de Nijmegen, síndrome de Bloom, síndrome de DiGeorge, displasias immuno-óseas tales como hipoplasia de cartilago-cabello, síndrome de Schimke, síndrome de Hermansky-Pudlak tipo 2, síndrome de Hiper-IgE, candidiasis mucocutánea crónica.

Los ejemplos de las enfermedades por desregulación inmunitaria incluyen inmunodeficiencia con hipopigmentación o albinismo tal como síndrome de Chediak-Higashi y síndrome de Griscelli tipo 2, linfocitosis hemofagocítica familiar tal como deficiencia de perforina, deficiencia de MUNC13D, y deficiencia de sintaxina 11, síndrome linfoproliferativo ligado a X, síndrome linfoproliferativo autoinmunitario tal como tipo 1a (defectos en CD95), tipo 1b (defectos en el ligando Fas), tipo 2a (defectos en CASP10), y tipo 2b (defectos en CASP8), poliendocrinopatía autoinmunitaria con candidiasis y distrofia ectodérmica, y síndrome de enteropatía poliendocrinopatía e inmunodesregulación ligado a X. Adicionalmente, las enfermedades que afectan a la médula ósea pueden dar como resultado leucocitos anómalos o escasos, tal como en la leucopenia. La leucopenia puede ser inducida por ciertas infecciones y enfermedades, incluyendo infección viral, infección por *Rickettsia*, algunos protozoos, tuberculosis, y ciertos cánceres.

Los ejemplos de los trastornos de los fagocitos incluyen la neutropenia congénita grave tal como la deficiencia de ELA2 (p. ej., mielodisplasia), deficiencia de GF11 (con linfopenia T/B) o deficiencia de G-CSFR (G-CSF no responsiva), síndrome de Kostmann, neutropenia cíclica, neutropenia/mielodisplasia ligada a X, deficiencia de la adherencia de leucocitos tipos 1, 2 y 3, deficiencia de RAC2, deficiencia de β -actina, periodontitis juvenil localizada, síndrome de Papillon-Lefèvre, deficiencia de gránulos específicos, síndrome de Shwachman-Diamond, enfermedad granulomatosa crónica, incluyendo formas ligadas a X y autosómicas, deficiencia de glucosa-6-fosfato deshidrogenasa de neutrófilos, deficiencia de la cadena β 1 de IL-12 e IL-23, deficiencia de IL-12p40, deficiencia de receptor 1 de interferon γ , deficiencia de receptor 2 de interferon γ , y deficiencia de STAT1.

Los ejemplos de las deficiencias de inmunidad innata incluyen displasia ectodérmica hipohidrótica tal como deficiencia de NEMO y deficiencia de IKBA, deficiencia de IRAK-4, síndrome de WHIM (verrugas, hipogammaglobulinemia, infecciones, mielocatexis), y epidermodisplasia verruciformis. Los ejemplos de las deficiencias del complemento y las afecciones asociadas ilustrativas incluyen deficiencia de C1 q (p. ej., síndrome de tipo lupus, enfermedad reumatoide, infecciones), deficiencia de C1r, deficiencia de C4, deficiencia de C2 (p. ej., síndrome de tipo lupus, vasculitis, polimiositis, infecciones piogénicas), deficiencia de C3 (p. ej., infecciones piogénicas recurrentes), deficiencia de C5 (p. ej., infecciones neisseriales), deficiencia de C6, deficiencia de C7 (p. ej., vasculitis), deficiencia de C8a y C8b, deficiencia de C9 (p. ej., infecciones neisseriales), deficiencia de inhibidor de C1 (p. ej., angioedema hereditario), deficiencia de Factor I (infecciones piogénicas), deficiencia de Factor H (p. ej., síndrome hemolítico-urémico, glomerulonefritis membranoproliferativa), deficiencia de Factor D (p. ej., infecciones neisseriales), deficiencia de Properdina (p. ej., infecciones neisseriales), deficiencia de MBP (p. ej., infecciones piogénicas), y deficiencia de MASP2.

Las deficiencias inmunitarias primarias se pueden diagnosticar de acuerdo con mecanismos rutinarios en la técnica. Los ensayos de diagnóstico ilustrativos incluyen, sin limitación, la realización de recuentos de los diferentes tipos de células mononucleares en la sangre (p. ej., linfocitos y monocitos, incluyendo linfocitos, diferentes grupos de linfocitos B tales como linfocitos CD19+, CD20+, y CD21+, células asesinas naturales, y monocitos positivos para CD15+), la medición de la presencia de marcadores de activación (p. ej., HLA-DR, CD25, CD80), la realización de ensayos para determinar la función de células T por ejemplo ensayos cutáneos para la hipersensibilidad de tipo retardado, respuestas celulares a mitógenos y células alogénicas, producción de citoquinas por las células, realización de ensayos para determinar la función de las células B por ejemplo identificando anticuerpos para inmunizaciones rutinarias e infecciones comúnmente adquiridas y cuantificando subclases de IgG, y la realización de ensayos de la función de los fagocitos, por ejemplo midiendo la reducción de cloruro de nitroazul de tetrazolio, y la realización de análisis de quimiotaxis y actividad bactericida. Por lo tanto los polipéptidos de AARS se pueden utilizar para estimular o mantener la inflamación aguda o las respuestas inflamatorias agudas en sujetos con una inmunodeficiencia primaria, como se describe en la presente memoria y es conocido en la técnica.

Los ejemplos de las causas de las inmunodeficiencias secundarias incluyen malnutrición, envejecimiento, y medicaciones (p. ej., quimioterapia, fármacos anti-reumáticos modificadores de la enfermedad, fármacos inmunosupresores después de transplantes de órganos, glucocorticoides). Otras causas incluyen diversos cánceres, incluyendo cánceres de médula ósea y células sanguíneas (p. ej., leucemia, linfoma, mieloma múltiple), y ciertas infecciones crónicas, tales como el síndrome de inmunodeficiencia adquirida (SIDA), causado por el virus de la inmunodeficiencia humana (VIH). Los polipéptidos de AARS se pueden utilizar para estimular o mantener la inflamación aguda o las respuestas inflamatorias agudas en sujetos con una inmunodeficiencia, como se describe en la presente memoria y es conocido en la técnica. Los polipéptidos de AARS también se pueden utilizar para estimular o mantener la inflamación crónica o las respuestas inflamatorias crónicas en sujetos con una inmunodeficiencia secundaria, como se describe en la presente memoria y es conocido en la técnica.

Además, las enfermedades, trastornos y afecciones adicionales incluyen síndrome de Guillain-Barre, anemia (por ejemplo, anemia asociada con parvovirus B19, pacientes con mieloma múltiple estable que presentan alto riesgo de infección (por ejemplo, infección recurrente), anemia hemolítica autoinmunitaria (por ejemplo, anemia hemolítica autoinmunitaria de tipo caliente), trombocitopenia (por ejemplo, trombocitopenia neonatal), y neutropenia mediada por inmunidad, trasplante (por ejemplo, receptores citomegalovirus (CMV) negativos de órganos CMV positivos), hipogammaglobulinemia (por ejemplo, neonatos hipogammaglobulinémicos con factor de riesgo para infección o morbilidad), epilepsia (por ejemplo, epilepsia intratable), síndromes vasculíticos sistémicos, miastenia grave (por ejemplo, descompensación en miastenia grave), dermatomiositis, y polimiositis.

de hormonas, incluyendo, pero no limitados a cáncer de colon, tumores cardiacos, cáncer pancreático, melanoma, retinoblastoma, glioblastoma, cáncer de pulmón, cáncer intestinal, cáncer testicular, cáncer de estómago, neuroblastoma, mixoma, mioma, linfoma, endotelioma, osteoblastoma, osteoclastoma, osteosarcoma, condrosarcoma, adenoma, cáncer de mama, cáncer de próstata, sarcoma de Kaposi y cáncer de ovario); trastornos autoinmunitarios (tales como, esclerosis múltiple, síndrome de Sjogren, enfermedad de Graves, tiroiditis de Hashimoto, diabetes autoinmunitaria, cirrosis biliar, enfermedad de Behcet, enfermedad de Crohn, polimiositis, lupus eritematoso sistémico y glomerulonefritis relacionada con inmunidad, gastritis autoinmunitaria, púrpura trombocitopénica autoinmunitaria, y artritis reumatoide) e infecciones virales (tales como virus de herpes, virus de viruela y adenovirus), inflamación, enfermedad de injerto frente a anfitrión (aguda y/o crónica), rechazo de injerto agudo, y rechazo de injerto crónico.

Las enfermedades o afecciones ilustrativas adicionales asociadas con la supervivencia celular incrementada incluyen, pero no están limitadas a, progreso y/o metástasis de malignidades y trastornos relacionados tales como leucemia (incluyendo leucemias agudas (por ejemplo, leucemia linfocítica aguda, leucemia mielocítica aguda, incluyendo mieloblástica, promielocítica, mielomonocítica, monocítica, y eritroleucemia)) y leucemias crónicas (por ejemplo, leucemia mielocítica crónica (granulocítica) y leucemia linfocítica crónica), síndrome mielodisplásico, policitemia vera, linfomas (por ejemplo enfermedad de Hodgkin y enfermedad no de Hodgkin), mieloma múltiple, macroglobulinemia de Waldenstrom, enfermedades de cadena pesada, y tumores sólidos incluyendo, pero no limitados a, sarcomas y carcinomas tales como fibrosarcoma, mixosarcoma, liposarcoma, condrosarcoma, sarcoma osteogénico, cordoma, angiosarcoma, endoteliosarcoma, linfangiosarcoma, linfangioendoteliosarcoma, sinovioma, mesotelioma, tumor de Ewing, leiomiomasarcoma, rabdomiosarcoma, carcinoma de colon, cáncer pancreático, cáncer de mama, cáncer de ovario, cáncer de próstata, carcinoma de células escamosas, carcinoma de células basales, adenocarcinoma, carcinoma de glándula sudorípara, carcinoma de glándula sebácea, carcinoma papilar, adenocarcinomas papilares, cistadenocarcinoma, carcinoma medular, carcinoma broncogénico, carcinoma de células renales, hepatoma, carcinoma del conducto biliar, coriocarcinoma, seminoma, carcinoma embrionario, tumor de Wilm, cáncer cervical, tumor testicular, carcinoma de pulmón, carcinoma de pulmón de células pequeñas, carcinoma de vejiga, carcinoma epitelial, glioma, astrocitoma, meduloblastoma, craneofaringioma, ependimoma, pinealoma, hemangioblastoma, neuroma acústico, oligodendroglioma, menangioma, melanoma, neuroblastoma, y retinoblastoma.

Las enfermedades ilustrativas asociadas con el incremento de la apoptosis incluyen, pero no están limitadas a, SIDA (tal como nefropatía inducida por VIH y encefalitis por VIH), trastornos neurodegenerativos (tal como enfermedad de Alzheimer, enfermedad de Parkinson, esclerosis lateral amiotrófica, retinitis pigmentosa, degeneración cerebelar y tumor cerebral o enfermedad asociada con priones), trastornos autoinmunitarios tales como esclerosis múltiple, síndrome de Sjogren, enfermedad de Graves, tiroiditis de Hashimoto, diabetes autoinmunitaria, cirrosis biliar, enfermedad de Behcet, enfermedad de Crohn, polimiositis, lupus eritematoso sistémico, glomerulonefritis relacionada con inmunidad, gastritis autoinmunitaria, púrpura trombocitopénica, y artritis reumatoide, síndromes mielodisplásicos (tales como anemia aplásica), enfermedad de injerto frente a anfitrión (aguda y/o crónica), daño isquémico (tal como el causado por infarto de miocardio, ictus y daño por reperfusión), daño o enfermedad hepática (por ejemplo, daño hepático relacionado con hepatitis, cirrosis, daño por isquemia/reperfusión, colestosis (daño en el conducto biliar) y cáncer de hígado), enfermedad hepática inducida por toxinas (tal como la causada por alcohol), choque séptico, colitis ulcerosa, caquexia, y anorexia.

En otras realizaciones adicionales más descritas en la presente memoria, se pueden utilizar las composiciones de la presente descripción en el tratamiento de enfermedades o trastornos neuronales/neurológicos, incluyendo los ejemplos ilustrativos de éstos enfermedad de Parkinson, enfermedad de Alzheimer, enfermedad de Pick, enfermedad de Creutzfeldt-Jacob, corea de Huntington, hemiplejia alternante, esclerosis lateral amiotrófica, ataxia, parálisis cerebral, síndrome de fatiga crónica, síndromes de dolor crónico, anomalías neurológicas congénitas, enfermedades de los nervios craneales, delirio, demencia, enfermedades desmielinizantes, disautonomía, epilepsia, dolores de cabeza, enfermedad de Huntington, hidrocefalia, meningitis, trastornos del movimiento, enfermedades musculares, neoplasmas del sistema nervioso, síndromes neurocutáneos, enfermedades neurodegenerativas, síndromes de neurotoxicidad, trastornos de motilidad ocular, trastornos del sistema nervioso periférico, trastornos de la pituitaria, porencefalia, síndrome de Rett, trastornos del sueño, trastornos de la médula espinal, ictus, corea de sydenham, síndrome de tourette, trauma y daños del sistema nervioso, etc.

Además, las realizaciones adicionales descritas en la presente memoria se refieren al uso de las composiciones de la presente descripción en el tratamiento de trastornos metabólicos tales como adrenoleucodistrofia, enfermedad de Krabbe (leucodistrofia de células globoides), leucodistrofia metacromática, enfermedad de Alexander, enfermedad de Canavan (leucodistrofia esponjiforme), enfermedad de Pelizaeus-Merzbacher, síndrome de Cockayne, enfermedad de Hurler, síndrome de Lowe, enfermedad de Leigh, enfermedad de Wilson, enfermedad de Hallervorden-Spatz, enfermedad de Tay-Sachs, etc. La utilidad de las composiciones de la presente descripción para modular procesos metabólicos puede verificarse usando cualquiera de una variedad de mecanismos conocidos y disponibles en la técnica incluyendo, por ejemplo, ensayos que miden la lipogénesis de adipocitos o lipólisis de adipocitos.

Ejemplos**EJEMPLO COMPARATIVO 1****Identificación del motivo estructural dentro de un fragmento de polipéptido de GLYRS que tiene actividad de señalización celular**

5 Como se muestra en las FIG. 2A-2C, se descubrió inesperadamente que un fragmento de glicil-ARNt sintetasa (GlyRS) humana referido como G6, y los correspondientes residuos 214-439 de GlyRS humana (SEQ ID NO: 1), era eficaz al inducir quimiotaxis de monocitos THP-1 dependiente de GPCR. El análisis adicional de este fragmento activo reveló la presencia de un motivo estructural, que comprendía 3 láminas β antiparalelas flanqueadas por una hélice α en cada extremo, que se cree que es responsable de la actividad quimioquina observada. Se identificó que el motivo estructural estaba contenido en los residuos de aminoácido 345-438 de la secuencia de proteína GlyRS humana mostrada en el SEQ ID NO: 1.

Ejemplo 2**Identificación del motivo estructural en otras proteínas de AARS de clase II**

15 El motivo estructural identificado anteriormente en el Ejemplo 1 también se encontró en regiones similares de otras AARS de Clase II humanas. Estas incluían AspRS, HisRS, ThrRS, GluProRS y SerRS. Esta región de las proteínas AARS de Clase II está altamente conservada estructuralmente entre los miembros de esta clase de AARS indicando que la región contiene la misma arquitectura estructural básica, pero no está bien conservada en identidad de secuencia de los residuos, creando el potencial de presentar varias secuencias diferentes que pueden tener diferentes actividades de señalización en esta arquitectura. (p. ej., FIG. 4, adaptada de O'Donoghue y Luthey-Schulten, Microbiology and Molecular Biology Reviews, Dic 2003)

20 Se identificó que el motivo estructural estaba contenido en los residuos de aminoácido 367-448 de la secuencia de proteína AspRS humana mostrada en el SEQ ID NO: 2; en los residuos de aminoácido 294-372 de la proteína HisRS humana mostrada en el SEQ ID NO: 3; en los residuos de aminoácido 469-586 of the la proteína ThRS humana mostrada en el SEQ ID NO: 4; en los residuos de aminoácido 1171-1253 de la proteína GluProRS humana mostrada en el SEQ ID NO: 5; en los residuos de aminoácido 325-410 de la proteína SerRS humana mostrada en el SEQ ID NO: 6; en los residuos de aminoácido 380-449 de la proteína PheRS humana mostrada en el SEQ ID NO: 7; en los residuos de aminoácido 425-523 de la proteína LysRS humana mostrada en el SEQ ID NO: 8; en los residuos de aminoácido 416-494 de la proteína AsnRS humana mostrada en el SEQ ID NO: 9; y en los residuos de aminoácido 148-258 de la proteína AlaRS humana mostrada en el SEQ ID NO: 10.

EJEMPLO COMPARATIVO 3**Proteínas no AARS ilustrativas que contienen el motivo estructural**

35 Un análisis adicional identificó la presencia de este mismo motivo estructural que comprende 3 láminas β antiparalelas y dos hélices α en varias proteínas distintas de AARS (FIG 5). Por ejemplo, el motivo se identificó en la tiorredoxina humana, contenido en los residuos 20-105 de la secuencia de tiorredoxina humana del SEQ ID NO: 11, así como también en los residuos 20-84 de Trx80, que es una forma trucada, secretada de tiorredoxina que contiene los 84 residuos de aminoácido N-terminales.

El motivo también se identificó en la proteína factor inhibidor de macrófagos humana, contenido en los residuos 1-90 de la secuencia de tiorredoxina humana del SEQ ID NO: 12.

40 El motivo también se identificó en la proteína de isoforma B de peroxirredoxina 5 humana, contenido en los fragmentos no contiguos correspondientes a los residuos 32-68 y 125-161 de la secuencia de la isoforma B de la peroxirredoxina 5 humana del SEQ ID NO: 13.

De este modo, escrutando proteínas distintas de AARS para determinar la presencia del motivo estructural identificado, la presente descripción proporciona un medio para identificar fragmentos de proteínas que tienen actividades biológicas no canónicas de relevancia terapéutica.

Ejemplo 4**Identificación del motivo estructural dentro de un fragmento de polipéptido de SERRS que tiene actividad de señalización celular**

50 Como se muestra en la FIG. 7, se descubrió inesperadamente que un fragmento de seril-ARNt sintetasa (SerRS) humana referido como S3, y correspondiente a los residuos 253-484 de la SerRS humana completa (SEQ ID NO: 6), era eficaz en la unión a monocitos y células B humanos, estimulando la secreción de TNF- α desde una línea celular monocítica humana, y señalizando a través de un receptor 2 de tipo Toll. El análisis adicional de este fragmento activo reveló la presencia de un motivo estructural, que comprende 3 láminas β antiparalelas flanqueadas por una hélice α en cada extremo, que se cree que es responsable de la actividad observada. Se identificó que el

motivo estructural estaba contenido en los residuos de aminoácido 325-410 de la secuencia de proteína SerRS humana mostrada en el SEQ ID NO: 6.

5 Para la unión de S3, se purificaron células mononucleares de sangre periférica (PBMC) a partir de un donante de sangre sano y se resuspendieron en RPMI con FBS al 10% a 1×10^7 por ml. Después se trataron 1×10^6 células en 100 μ l de medio con PBS o S3 500 nM. Después de 45 minutos las células se lavaron dos veces con 1 mL de tampón de tinción (PBS + FBS al 2%) y se tiñeron con un anticuerpo α V5-FITC de Invitrogen (Núm. de Catálogo R96325), CD19 para las células B (Núm. de Catálogo BD 555413), en tampón de tinción con 200 μ g/ml de IgG humana durante 30 minutos. Las células se lavaron a continuación dos veces con 1 mL de tampón de tinción, se resuspendieron y se analizaron mediante FACS. Los resultados se muestran en la FIG. 7A, que muestra la unión de S3 a monocitos humanos, y la FIG. 7B, que muestra la unión de S3 a células B humanas, en comparación con un control negativo de PBS.

15 Para la secreción de TNF α , se cultivaron células THP en medio RPMI con FBS al 10% y BME 0,5 mM. Se trataron un millón de células con PBS, S3 (100 nM), o Lipopolisacárido (LPS, 1 μ g/mL). Después de 4 horas, el sobrenadante celular se recogió por centrifugación a 2000xg durante 10 min y se evaluó en un ELISA de TNF- α (R&D Systems; Núm. de Cat. DTA00C) mediante las directrices del kit. Los resultados se muestran en la FIG. 7C, que muestra la inducción de secreción de TNF- α a partir de la línea celular monocítica humana (THP-1), en comparación con un control positivo de LPS y un control negativo de PBS.

20 Para la señalización a través del receptor 2 de tipo Toll, se trataron células HEK-Blue 2 (Invivogen células Núm. hb2) transfectadas establemente con receptor 2 de tipo Toll de ratón (mTLR2) unido a un gen informador de fosfatasa alcalina de NF κ B durante la noche con dosis variables de proteína S3. El día siguiente, el medio se eliminó y se analizó durante cuatro horas para determinar la fosfatasa alcalina secretada utilizando QUANTI-Blue (Invivogen Núm. rep-qb). Como se muestra en la Figura 8, el tratamiento logró una fuerte secreción de fosfatasa alcalina de una manera dependiente de la dosis, indicando la activación del receptor y su ruta de señalización de NF κ B asociada.

25 En otro experimento, se pre-trataron las mismas células con 5 μ g/ml de MAb mTLR2 (Invivogen Núm. mab-mtlr2) (marcado (+) TLR-2) o no se trataron (S3) durante 30 minutos, seguido de tratamiento durante la noche con S3 250 nM. Al día siguiente el medio se retiró y se analizó durante cuatro horas para determinar la fosfatasa alcalina secretada utilizando QUANTI-Blue (Invivogen Núm. rep-qb). Como se muestra en la Figura 9, la activación del receptor 2 de tipo Toll era inhibida por el pre-tratamiento con un anticuerpo monoclonal que bloquea la unión a mTLR2.

30

Listado de secuencias

<110> ATYR PHARMA, INC. Greene, Leslie Ann

5 <120> Motivos estructurales de polipéptidos asociados con la actividad de señalización celular

<130> 120161.411PC

<140> PCT

10 <141> 26-02-2010

<150> US 61/156.370

<151> 27-02-2009

15 <160> 13

<170> FastSEQ para Windows Versión 4.0

<210> 1

20 <211> 685

<212> PRT

<213> Homo sapiens

<400> 1

25

```

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

```

275 280 285
 Val Arg Glu Phe Thr Met Ala Glu Ile Glu His Phe Val Asp Pro Ser
 290 295 300
 Glu Lys Asp His Pro Lys Phe Gln Asn Val Ala Asp Leu His Leu Tyr
 305 310 315 320
 Leu Tyr Ser Ala Lys Ala Gln Val Ser Gly Gln Ser Ala Arg Lys Met
 325 330 335
 Arg Leu Gly Asp Ala Val Glu Gln Gly Val Ile Asn Asn Thr Val Leu
 340 345 350
 Gly Tyr Phe Ile Gly Arg Ile Tyr Leu Tyr Leu Thr Lys Val Gly Ile
 355 360 365
 Ser Pro Asp Lys Leu Arg Phe Arg Gln His Met Glu Asn Glu Met Ala
 370 375 380
 His Tyr Ala Cys Asp Cys Trp Asp Ala Glu Ser Lys Thr Ser Tyr Gly
 385 390 395 400
 Trp Ile Glu Ile Val Gly Cys Ala Asp Arg Ser Cys Tyr Asp Leu Ser
 405 410 415
 Cys His Ala Arg Ala Thr Lys Val Pro Leu Val Ala Glu Lys Pro Leu
 420 425 430
 Lys Glu Pro Lys Thr Val Asn Val Val Gln Phe Glu Pro Ser Lys Gly
 435 440 445
 Ala Ile Gly Lys Ala Tyr Lys Lys Asp Ala Lys Leu Val Met Glu Tyr
 450 455 460
 Leu Ala Ile Cys Asp Glu Cys Tyr Ile Thr Glu Met Glu Met Leu Leu
 465 470 475 480
 Asn Glu Lys Gly Glu Phe Thr Ile Glu Thr Glu Gly Lys Thr Phe Gln
 485 490 495
 Leu Thr Lys Asp Met Ile Asn Val Lys Arg Phe Gln Lys Thr Leu Tyr
 500 505 510
 Val Glu Glu Val Val Pro Asn Val Ile Glu Pro Ser Phe Gly Leu Gly
 515 520 525
 Arg Ile Met Tyr Thr Val Phe Glu His Thr Phe His Val Arg Glu Gly
 530 535 540
 Asp Glu Gln Arg Thr Phe Phe Ser Phe Pro Ala Val Val Ala Pro Phe
 545 550 555 560
 Lys Cys Ser Val Leu Pro Leu Ser Gln Asn Gln Glu Phe Met Pro Phe
 565 570 575
 Val Lys Glu Leu Ser Glu Ala Leu Thr Arg His Gly Val Ser His Lys
 580 585 590
 Val Asp Asp Ser Ser Gly Ser Ile Gly Arg Arg Tyr Ala Arg Thr Asp
 595 600 605
 Glu Ile Gly Val Ala Phe Gly Val Thr Ile Asp Phe Asp Thr Val Asn
 610 615 620
 Lys Thr Pro His Thr Ala Thr Leu Arg Asp Arg Asp Ser Met Arg Gln
 625 630 635 640
 Ile Arg Ala Glu Ile Ser Glu Leu Pro Ser Ile Val Gln Asp Leu Ala
 645 650 655
 Asn Gly Asn Ile Thr Trp Ala Asp Val Glu Ala Arg Tyr Pro Leu Phe
 660 665 670
 Glu Gly Gln Glu Thr Gly Lys Lys Glu Thr Ile Glu Glu
 675 680 685

<210> 2
 <211> 501
 5 <212> PRT
 <213> Homo sapiens

<400> 2

Met Pro Ser Ala Ser Ala Ser Arg Lys Ser Gln Glu Lys Pro Arg Glu
 1 5 10 15
 Ile Met Asp Ala Ala Glu Asp Tyr Ala Lys Glu Arg Tyr Gly Ile Ser
 20 25 30

10

ES 2 552 773 T3

Ser Met Ile Gln Ser Gln Glu Lys Pro Asp Arg Val Leu Val Arg Val
35 40 45
Arg Asp Leu Thr Ile Gln Lys Ala Asp Glu Val Val Trp Val Arg Ala
50 55 60
Arg Val His Thr Ser Arg Ala Lys Gly Lys Gln Cys Phe Leu Val Leu
65 70 75 80
Arg Gln Gln Gln Phe Asn Val Gln Ala Leu Val Ala Val Gly Asp His
85 90 95
Ala Ser Lys Gln Met Val Lys Phe Ala Ala Asn Ile Asn Lys Glu Ser
100 105 110
Ile Val Asp Val Glu Gly Val Val Arg Lys Val Asn Gln Lys Ile Gly
115 120 125
Ser Cys Thr Gln Gln Asp Val Glu Leu His Val Gln Lys Ile Tyr Val
130 135 140
Ile Ser Leu Ala Glu Pro Arg Leu Pro Leu Gln Leu Asp Asp Ala Val
145 150 155 160
Arg Pro Glu Ala Glu Gly Glu Glu Glu Gly Arg Ala Thr Val Asn Gln
165 170 175
Asp Thr Arg Leu Asp Asn Arg Val Ile Asp Leu Arg Thr Ser Thr Ser
180 185 190
Gln Ala Val Phe Arg Leu Gln Ser Gly Ile Cys His Leu Phe Arg Glu
195 200 205
Thr Leu Ile Asn Lys Gly Phe Val Glu Ile Gln Thr Pro Lys Ile Ile
210 215 220
Ser Ala Ala Ser Glu Gly Gly Ala Asn Val Phe Thr Val Ser Tyr Phe
225 230 235 240
Lys Asn Asn Ala Tyr Leu Ala Gln Ser Pro Gln Leu Tyr Lys Gln Met
245 250 255
Cys Ile Cys Ala Asp Phe Glu Lys Val Phe Ser Ile Gly Pro Val Phe
260 265 270
Arg Ala Glu Asp Ser Asn Thr His Arg His Leu Thr Glu Phe Val Gly
275 280 285
Leu Asp Ile Glu Met Ala Phe Asn Tyr His Tyr His Glu Val Met Glu
290 295 300
Glu Ile Ala Asp Thr Met Val Gln Ile Phe Lys Gly Leu Gln Glu Arg
305 310 315 320
Phe Gln Thr Glu Ile Gln Thr Val Asn Lys Gln Phe Pro Cys Glu Pro
325 330 335
Phe Lys Phe Leu Glu Pro Thr Leu Arg Leu Glu Tyr Cys Glu Ala Leu
340 345 350
Ala Met Leu Arg Glu Ala Gly Val Glu Met Gly Asp Glu Asp Asp Leu
355 360 365
Ser Thr Pro Asn Glu Lys Leu Leu Gly His Leu Val Lys Glu Lys Tyr
370 375 380
Asp Thr Asp Phe Tyr Ile Leu Asp Lys Tyr Pro Leu Ala Val Arg Pro
385 390 395 400
Phe Tyr Thr Met Pro Asp Pro Arg Asn Pro Lys Gln Ser Asn Ser Tyr
405 410 415
Asp Met Phe Met Arg Gly Glu Glu Ile Leu Ser Gly Ala Gln Arg Ile
420 425 430
His Asp Pro Gln Leu Leu Thr Glu Arg Ala Leu His His Gly Ile Asp
435 440 445
Leu Glu Lys Ile Lys Ala Tyr Ile Asp Ser Phe Arg Phe Gly Ala Pro
450 455 460
Pro His Ala Gly Gly Gly Ile Gly Leu Glu Arg Val Thr Met Leu Phe
465 470 475 480
Leu Gly Leu His Asn Val Arg Gln Thr Ser Met Phe Pro Arg Asp Pro
485 490 495
Lys Arg Leu Thr Pro
500

<210> 3
<211> 509
5 <212> PRT
<213> Homo sapiens

<400> 3

Met Ala Glu Arg Ala Ala Leu Glu Glu Leu Val Lys Leu Gln Gly Glu
 1 5 10
 Arg Val Arg Gly Leu Lys Gln Gln Lys Ala Ser Ala Glu Leu Ile Glu
 20 25 30
 Glu Glu Val Ala Lys Leu Leu Lys Leu Lys Ala Gln Leu Gly Pro Asp
 35 40 45
 Glu Ser Lys Gln Lys Phe Val Leu Lys Thr Pro Lys Gly Thr Arg Asp
 50 55 60
 Tyr Ser Pro Arg Gln Met Ala Val Arg Glu Lys Val Phe Asp Val Ile
 65 70 75 80
 Ile Arg Cys Phe Lys Arg His Gly Ala Glu Val Ile Asp Thr Pro Val
 85 90 95
 Phe Glu Leu Lys Glu Thr Leu Met Gly Lys Tyr Gly Glu Asp Ser Lys
 100 105 110
 Leu Ile Tyr Asp Leu Lys Asp Gln Gly Gly Glu Leu Leu Ser Leu Arg
 115 120 125
 Tyr Asp Leu Thr Val Pro Phe Ala Arg Tyr Leu Ala Met Asn Lys Leu
 130 135 140
 Thr Asn Ile Lys Arg Tyr His Ile Ala Lys Val Tyr Arg Arg Asp Asn
 145 150 155 160
 Pro Ala Met Thr Arg Gly Arg Tyr Arg Glu Phe Tyr Gln Cys Asp Phe
 165 170 175
 Asp Ile Ala Gly Asn Phe Asp Pro Met Ile Pro Asp Ala Glu Cys Leu
 180 185 190
 Lys Ile Met Cys Glu Ile Leu Ser Ser Leu Gln Ile Gly Asp Phe Leu
 195 200 205
 Val Lys Val Asn Asp Arg Arg Ile Leu Asp Gly Met Phe Ala Ile Cys
 210 215 220
 Gly Val Ser Asp Ser Lys Phe Arg Thr Ile Cys Ser Ser Val Asp Lys
 225 230 235 240
 Leu Asp Lys Val Ser Trp Glu Glu Val Lys Asn Glu Met Val Gly Glu
 245 250 255
 Lys Gly Leu Ala Pro Glu Val Ala Asp Arg Ile Gly Asp Tyr Val Gln
 260 265 270
 Gln His Gly Gly Val Ser Leu Val Glu Gln Leu Leu Gln Asp Pro Lys
 275 280 285
 Leu Ser Gln Asn Lys Gln Ala Leu Glu Gly Leu Gly Asp Leu Lys Leu
 290 295 300
 Leu Phe Glu Tyr Leu Thr Leu Phe Gly Ile Asp Asp Lys Ile Ser Phe
 305 310 315 320
 Asp Leu Ser Leu Ala Arg Gly Leu Asp Tyr Thr Gly Val Ile Tyr
 325 330 335
 Glu Ala Val Leu Leu Gln Thr Pro Ala Gln Ala Gly Glu Glu Pro Leu
 340 345 350
 Gly Val Gly Ser Val Ala Ala Gly Gly Arg Tyr Asp Gly Leu Val Gly
 355 360 365
 Met Phe Asp Pro Lys Gly Arg Lys Val Pro Cys Val Gly Leu Ser Ile
 370 375 380
 Gly Val Glu Arg Ile Phe Ser Ile Val Glu Gln Arg Leu Glu Ala Leu
 385 390 395 400
 Glu Glu Lys Ile Arg Thr Thr Glu Thr Gln Val Leu Val Ala Ser Ala
 405 410 415
 Gln Lys Lys Leu Leu Glu Glu Arg Leu Lys Leu Val Ser Glu Leu Trp
 420 425 430
 Asp Ala Gly Ile Lys Ala Glu Leu Leu Tyr Lys Lys Asn Pro Lys Leu
 435 440 445
 Leu Asn Gln Leu Gln Tyr Cys Glu Glu Ala Gly Ile Pro Leu Val Ala
 450 455 460
 Ile Ile Gly Glu Gln Glu Leu Lys Asp Gly Val Ile Lys Leu Arg Ser
 465 470 475 480
 Val Thr Ser Arg Glu Glu Val Asp Val Arg Arg Glu Asp Leu Val Glu
 485 490 495
 Glu Ile Lys Arg Arg Thr Gly Gln Pro Leu Cys Ile Cys
 500 505

<210> 4
 <211> 723
 <212> PRT
 <213> Homo sapiens

5

<400> 4

ES 2 552 773 T3

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

```

385              390              395              400
Val Glu Lys Glu Leu Phe Ala Leu Lys Pro Met Asn Cys Pro Gly His
      405              410              415
Cys Leu Met Phe Asp His Arg Pro Arg Ser Trp Arg Glu Leu Pro Leu
      420              425              430
Arg Leu Ala Asp Phe Gly Val Leu His Arg Asn Glu Leu Ser Gly Ala
      435              440              445
Leu Thr Gly Leu Thr Arg Val Arg Arg Phe Gln Gln Asp Asp Ala His
      450              455              460
Ile Phe Cys Ala Met Glu Gln Ile Glu Asp Glu Ile Lys Gly Cys Leu
465              470              475              480
Asp Phe Leu Arg Thr Val Tyr Ser Val Phe Gly Phe Ser Phe Lys Leu
      485              490              495
Asn Leu Ser Thr Arg Pro Glu Lys Phe Leu Gly Asp Ile Glu Val Trp
      500              505              510
Asp Gln Ala Glu Lys Gln Leu Glu Asn Ser Leu Asn Glu Phe Gly Glu
      515              520              525
Lys Trp Glu Leu Asn Ser Gly Asp Gly Ala Phe Tyr Gly Pro Lys Ile
      530              535              540
Asp Ile Gln Ile Lys Asp Ala Ile Gly Arg Tyr His Gln Cys Ala Thr
545              550              555              560
Ile Gln Leu Asp Phe Gln Leu Pro Ile Arg Phe Asn Leu Thr Tyr Val
      565              570              575
Ser His Asp Gly Asp Asp Lys Lys Arg Pro Val Ile Val His Arg Ala
      580              585              590
Ile Leu Gly Ser Val Glu Arg Met Ile Ala Ile Leu Thr Glu Asn Tyr
      595              600              605
Gly Gly Lys Trp Pro Phe Trp Leu Ser Pro Arg Gln Val Met Val Val
      610              615              620
Pro Val Gly Pro Thr Cys Asp Glu Tyr Ala Gln Lys Val Arg Gln Gln
625              630              635              640
Phe His Asp Ala Lys Phe Met Ala Asp Ile Asp Leu Asp Pro Gly Cys
      645              650              655
Thr Leu Asn Lys Lys Ile Arg Asn Ala Gln Leu Ala Gln Tyr Asn Phe
      660              665              670
Ile Leu Val Val Gly Glu Lys Glu Lys Ile Ser Gly Thr Val Asn Ile
      675              680              685
Arg Thr Arg Asp Asn Lys Val His Gly Glu Arg Thr Ile Ser Glu Thr
      690              695              700
Ile Glu Arg Leu Gln Gln Leu Lys Glu Phe Arg Ser Lys Gln Ala Glu
705              710              715              720
Glu Glu Phe

```

<210> 5
<211> 1512
5 <212> PRT
<213> Homo sapiens

<400> 5

```

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

```

10

ES 2 552 773 T3

Ser Thr Ile Asn Glu Leu Asn His Cys Leu Ser Leu Arg Thr Tyr Leu
100 105 110
Val Gly Asn Ser Leu Ser Leu Ala Asp Leu Cys Val Trp Ala Thr Leu
115 120 125
Lys Gly Asn Ala Ala Trp Gln Glu Gln Leu Lys Gln Lys Lys Ala Pro
130 135 140
Val His Val Lys Arg Trp Phe Gly Phe Leu Glu Ala Gln Gln Ala Phe
145 150 155 160
Gln Ser Val Gly Thr Lys Trp Asp Val Ser Thr Thr Lys Ala Arg Val
165 170 175
Ala Pro Glu Lys Lys Gln Asp Val Gly Lys Phe Val Glu Leu Pro Gly
180 185 190
Ala Glu Met Gly Lys Val Thr Val Arg Phe Pro Pro Glu Ala Ser Gly
195 200 205
Tyr Leu His Ile Gly His Ala Lys Ala Ala Leu Leu Asn Gln His Tyr
210 215 220
Gln Val Asn Phe Lys Gly Lys Leu Ile Met Arg Phe Asp Asp Thr Asn
225 230 235 240
Pro Glu Lys Glu Lys Glu Asp Phe Glu Lys Val Ile Leu Glu Asp Val
245 250 255
Ala Met Leu His Ile Lys Pro Asp Gln Phe Thr Tyr Thr Ser Asp His
260 265 270
Phe Glu Thr Ile Met Lys Tyr Ala Glu Lys Leu Ile Gln Glu Gly Lys
275 280 285
Ala Tyr Val Asp Asp Thr Pro Ala Glu Gln Met Lys Ala Glu Arg Glu
290 295 300
Gln Arg Ile Glu Ser Lys His Arg Lys Asn Pro Ile Glu Lys Asn Leu
305 310 315 320
Gln Met Trp Glu Glu Met Lys Lys Gly Ser Gln Phe Gly His Ser Cys
325 330 335
Cys Leu Arg Ala Lys Ile Asp Met Ser Ser Asn Asn Gly Cys Met Arg
340 345 350
Asp Pro Thr Leu Tyr Arg Cys Lys Ile Gln Pro His Pro Arg Thr Gly
355 360 365
Asn Lys Tyr Asn Val Tyr Pro Thr Tyr Asp Phe Ala Cys Pro Ile Val
370 375 380
Asp Ser Ile Glu Gly Val Thr His Ala Leu Arg Thr Thr Glu Tyr His
385 390 395 400
Asp Arg Asp Glu Gln Phe Tyr Trp Ile Ile Glu Ala Leu Gly Ile Arg
405 410 415
Lys Pro Tyr Ile Trp Glu Tyr Ser Arg Leu Asn Leu Asn Asn Thr Val
420 425 430
Leu Ser Lys Arg Lys Leu Thr Trp Phe Val Asn Glu Gly Leu Val Asp
435 440 445
Gly Trp Asp Asp Pro Arg Phe Pro Thr Val Arg Gly Val Leu Arg Arg
450 455 460
Gly Met Thr Val Glu Gly Leu Lys Gln Phe Ile Ala Ala Gln Gly Ser
465 470 475 480
Ser Arg Ser Val Val Asn Met Glu Trp Asp Lys Ile Trp Ala Phe Asn
485 490 495
Lys Lys Val Ile Asp Pro Val Ala Pro Arg Tyr Val Ala Leu Leu Lys
500 505 510
Lys Glu Val Ile Pro Val Asn Val Pro Glu Ala Gln Glu Glu Met Lys
515 520 525
Glu Val Ala Lys His Pro Lys Asn Pro Glu Val Gly Leu Lys Pro Val
530 535 540
Trp Tyr Ser Pro Lys Val Phe Ile Glu Gly Ala Asp Ala Glu Thr Phe
545 550 555 560
Ser Glu Gly Glu Met Val Thr Phe Ile Asn Trp Gly Asn Leu Asn Ile
565 570 575
Thr Lys Ile His Lys Asn Ala Asp Gly Lys Ile Ile Ser Leu Asp Ala
580 585 590
Lys Phe Asn Leu Glu Asn Lys Asp Tyr Lys Lys Thr Thr Lys Val Thr

ES 2 552 773 T3

	595					600					605				
Trp	Leu	Ala	Glu	Thr	Thr	His	Ala	Leu	Pro	Ile	Pro	Val	Ile	Cys	Val
	610					615					620				
Thr	Tyr	Glu	His	Leu	Ile	Thr	Lys	Pro	Val	Leu	Gly	Lys	Asp	Glu	Asp
	625				630					635					640
Phe	Lys	Gln	Tyr	Val	Asn	Lys	Asn	Ser	Lys	His	Glu	Glu	Leu	Met	Leu
				645					650					655	
Gly	Asp	Pro	Cys	Leu	Lys	Asp	Leu	Lys	Lys	Gly	Asp	Ile	Ile	Gln	Leu
			660					665					670		
Gln	Arg	Arg	Gly	Phe	Phe	Ile	Cys	Asp	Gln	Pro	Tyr	Glu	Pro	Val	Ser
	675					680						685			
Pro	Tyr	Ser	Cys	Lys	Glu	Ala	Pro	Cys	Val	Leu	Ile	Tyr	Ile	Pro	Asp
	690					695					700				
Gly	His	Thr	Lys	Glu	Met	Pro	Thr	Ser	Gly	Ser	Lys	Glu	Lys	Thr	Lys
	705				710				715						720
Val	Glu	Ala	Thr	Lys	Asn	Glu	Thr	Ser	Ala	Pro	Phe	Lys	Glu	Arg	Pro
				725					730					735	
Thr	Pro	Ser	Leu	Asn	Asn	Asn	Cys	Thr	Thr	Ser	Glu	Asp	Ser	Leu	Val
			740					745					750		
Leu	Tyr	Asn	Arg	Val	Ala	Val	Gln	Gly	Asp	Val	Val	Arg	Glu	Leu	Lys
	755						760					765			
Ala	Lys	Lys	Ala	Pro	Lys	Glu	Asp	Val	Asp	Ala	Ala	Val	Lys	Gln	Leu
	770					775				780					
Leu	Ser	Leu	Lys	Ala	Glu	Tyr	Lys	Glu	Lys	Thr	Gly	Gln	Glu	Tyr	Lys
	785				790					795					800
Pro	Gly	Asn	Pro	Pro	Ala	Glu	Ile	Gly	Gln	Asn	Ile	Ser	Ser	Asn	Ser
			805					810						815	
Ser	Ala	Ser	Ile	Leu	Glu	Ser	Lys	Ser	Leu	Tyr	Asp	Glu	Val	Ala	Ala
			820					825					830		
Gln	Gly	Glu	Val	Val	Arg	Lys	Leu	Lys	Ala	Glu	Lys	Ser	Pro	Lys	Ala
	835						840					845			
Lys	Ile	Asn	Glu	Ala	Val	Glu	Cys	Leu	Leu	Ser	Leu	Lys	Ala	Gln	Tyr
	850					855					860				
Lys	Glu	Lys	Thr	Gly	Lys	Glu	Tyr	Ile	Pro	Gly	Gln	Pro	Pro	Leu	Ser
	865				870					875					880
Gln	Ser	Ser	Asp	Ser	Ser	Pro	Thr	Arg	Asn	Ser	Glu	Pro	Ala	Gly	Leu
			885						890					895	
Glu	Thr	Pro	Glu	Ala	Lys	Val	Leu	Phe	Asp	Lys	Val	Ala	Ser	Gln	Gly
			900					905					910		
Glu	Val	Val	Arg	Lys	Leu	Lys	Thr	Glu	Lys	Ala	Pro	Lys	Asp	Gln	Val
	915						920						925		
Asp	Ile	Ala	Val	Gln	Glu	Leu	Leu	Gln	Leu	Lys	Ala	Gln	Tyr	Lys	Ser
	930					935					940				
Leu	Ile	Gly	Val	Glu	Tyr	Lys	Pro	Val	Ser	Ala	Thr	Gly	Ala	Glu	Asp
	945				950					955					960
Lys	Asp	Lys	Lys	Lys	Lys	Glu	Lys	Glu	Asn	Lys	Ser	Glu	Lys	Gln	Asn
			965						970					975	
Lys	Pro	Gln	Lys	Gln	Asn	Asp	Gly	Gln	Arg	Lys	Asp	Pro	Ser	Lys	Asn
			980					985					990		
Gln	Gly	Gly	Gly	Leu	Ser	Ser	Ser	Gly	Ala	Gly	Glu	Gly	Gln	Gly	Pro
	995						1000					1005			
Lys	Lys	Gln	Thr	Arg	Leu	Gly	Leu	Glu	Ala	Lys	Lys	Glu	Glu	Asn	Leu
	1010					1015						1020			
Ala	Asp	Trp	Tyr	Ser	Gln	Val	Ile	Thr	Lys	Ser	Glu	Met	Ile	Glu	Tyr
	1025				1030					1035					1040
His	Asp	Ile	Ser	Gly	Cys	Tyr	Ile	Leu	Arg	Pro	Trp	Ala	Tyr	Ala	Ile
			1045						1050					1055	
Trp	Glu	Ala	Ile	Lys	Asp	Phe	Phe	Asp	Ala	Glu	Ile	Lys	Lys	Leu	Gly
			1060					1065						1070	
Val	Glu	Asn	Cys	Tyr	Phe	Pro	Met	Phe	Val	Ser	Gln	Ser	Ala	Leu	Glu
	1075						1080					1085			
Lys	Glu	Lys	Thr	His	Val	Ala	Asp	Phe	Ala	Pro	Glu	Val	Ala	Trp	Val
	1090					1095						1100			

Thr Arg Ser Gly Lys Thr Glu Leu Ala Glu Pro Ile Ala Ile Arg Pro
 1105 1110 1115 1120
 Thr Ser Glu Thr Val Met Tyr Pro Ala Tyr Ala Lys Trp Val Gln Ser
 1125 1130 1135
 His Arg Asp Leu Pro Ile Lys Leu Asn Gln Trp Cys Asn Val Val Arg
 1140 1145 1150
 Trp Glu Phe Lys His Pro Gln Pro Phe Leu Arg Thr Arg Glu Phe Leu
 1155 1160 1165
 Trp Gln Glu Gly His Ser Ala Phe Ala Thr Met Glu Glu Ala Ala Glu
 1170 1175 1180
 Glu Val Leu Gln Ile Leu Asp Leu Tyr Ala Gln Val Tyr Glu Glu Leu
 1185 1190 1195 1200
 Leu Ala Ile Pro Val Val Lys Gly Arg Lys Thr Glu Lys Glu Lys Phe
 1205 1210 1215
 Ala Gly Gly Asp Tyr Thr Thr Thr Ile Glu Ala Phe Ile Ser Ala Ser
 1220 1225 1230
 Gly Arg Ala Ile Gln Gly Gly Thr Ser His His Leu Gly Gln Asn Phe
 1235 1240 1245
 Ser Lys Met Phe Glu Ile Val Phe Glu Asp Pro Lys Ile Pro Gly Glu
 1250 1255 1260
 Lys Gln Phe Ala Tyr Gln Asn Ser Trp Gly Leu Thr Thr Arg Thr Ile
 1265 1270 1275 1280
 Gly Val Met Thr Met Val His Gly Asp Asn Met Gly Leu Val Leu Pro
 1285 1290 1295
 Pro Arg Val Ala Cys Val Gln Val Val Ile Ile Pro Cys Gly Ile Thr
 1300 1305 1310
 Asn Ala Leu Ser Glu Glu Asp Lys Glu Ala Leu Ile Ala Lys Cys Asn
 1315 1320 1325
 Asp Tyr Arg Arg Arg Leu Leu Ser Val Asn Ile Arg Val Arg Ala Asp
 1330 1335 1340
 Leu Arg Asp Asn Tyr Ser Pro Gly Trp Lys Phe Asn His Trp Glu Leu
 1345 1350 1355 1360
 Lys Gly Val Pro Ile Arg Leu Glu Val Gly Pro Arg Asp Met Lys Ser
 1365 1370 1375
 Cys Gln Phe Val Ala Val Arg Arg Asp Thr Gly Glu Lys Leu Thr Val
 1380 1385 1390
 Ala Glu Asn Glu Ala Glu Thr Lys Leu Gln Ala Ile Leu Glu Asp Ile
 1395 1400 1405
 Gln Val Thr Leu Phe Thr Arg Ala Ser Glu Asp Leu Lys Thr His Met
 1410 1415 1420
 Val Val Ala Asn Thr Met Glu Asp Phe Gln Lys Ile Leu Asp Ser Gly
 1425 1430 1435 1440
 Lys Ile Val Gln Ile Pro Phe Cys Gly Glu Ile Asp Cys Glu Asp Trp
 1445 1450 1455
 Ile Lys Lys Thr Thr Ala Arg Asp Gln Asp Leu Glu Pro Gly Ala Pro
 1460 1465 1470
 Ser Met Gly Ala Lys Ser Leu Cys Ile Pro Phe Lys Pro Leu Cys Glu
 1475 1480 1485
 Leu Gln Pro Gly Ala Lys Cys Val Cys Gly Lys Asn Pro Ala Lys Tyr
 1490 1495 1500
 Tyr Thr Leu Phe Gly Arg Ser Tyr
 1505 1510

<210> 6
 <211> 514
 5 <212> PRT
 <213> Homo sapiens

<400> 6

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

ES 2 552 773 T3

			20					25				30			
Val	Asp	Gln	Leu	Val	Lys	Ala	Asp	Ser	Glu	Trp	Arg	Arg	Cys	Arg	Phe
		35					40					45			
Arg	Ala	Asp	Asn	Leu	Asn	Lys	Leu	Lys	Asn	Leu	Cys	Ser	Lys	Thr	Ile
	50					55					60				
Gly	Glu	Lys	Met	Lys	Lys	Lys	Glu	Pro	Val	Gly	Asp	Asp	Glu	Ser	Val
65					70					75					80
Pro	Glu	Asn	Val	Leu	Ser	Phe	Asp	Asp	Leu	Thr	Ala	Asp	Ala	Leu	Ala
				85					90					95	
Asn	Leu	Lys	Val	Ser	Gln	Ile	Lys	Lys	Val	Arg	Leu	Leu	Ile	Asp	Glu
			100					105					110		
Ala	Ile	Leu	Lys	Cys	Asp	Ala	Glu	Arg	Ile	Lys	Leu	Glu	Ala	Glu	Arg
	115						120					125			
Phe	Glu	Asn	Leu	Arg	Glu	Ile	Gly	Asn	Leu	Leu	His	Pro	Ser	Val	Pro
	130					135					140				
Ile	Ser	Asn	Asp	Glu	Asp	Val	Asp	Asn	Lys	Val	Glu	Arg	Ile	Trp	Gly
145					150					155					160
Asp	Cys	Thr	Val	Arg	Lys	Lys	Tyr	Ser	His	Val	Asp	Leu	Val	Val	Met
				165					170						175
Val	Asp	Gly	Phe	Glu	Gly	Glu	Lys	Gly	Ala	Val	Val	Ala	Gly	Ser	Arg
			180					185					190		
Gly	Tyr	Phe	Leu	Lys	Gly	Val	Leu	Val	Phe	Leu	Glu	Gln	Ala	Leu	Ile
	195					200						205			
Gln	Tyr	Ala	Leu	Arg	Thr	Leu	Gly	Ser	Arg	Gly	Tyr	Ile	Pro	Ile	Tyr
	210					215					220				
Thr	Pro	Phe	Phe	Met	Arg	Lys	Glu	Val	Met	Gln	Glu	Val	Ala	Gln	Leu
225				230						235					240
Ser	Gln	Phe	Asp	Glu	Glu	Leu	Tyr	Lys	Val	Ile	Gly	Lys	Gly	Ser	Glu
				245					250						255
Lys	Ser	Asp	Asp	Asn	Ser	Tyr	Asp	Glu	Lys	Tyr	Leu	Ile	Ala	Thr	Ser
		260						265					270		
Glu	Gln	Pro	Ile	Ala	Ala	Leu	His	Arg	Asp	Glu	Trp	Leu	Arg	Pro	Glu
	275						280					285			
Asp	Leu	Pro	Ile	Lys	Tyr	Ala	Gly	Leu	Ser	Thr	Cys	Phe	Arg	Gln	Glu
	290					295					300				
Val	Gly	Ser	His	Gly	Arg	Asp	Thr	Arg	Gly	Ile	Phe	Arg	Val	His	Gln
305					310					315					320
Phe	Glu	Lys	Ile	Glu	Gln	Phe	Val	Tyr	Ser	Ser	Pro	His	Asp	Asn	Lys
				325					330					335	
Ser	Trp	Glu	Met	Phe	Glu	Glu	Met	Ile	Thr	Thr	Ala	Glu	Glu	Phe	Tyr
		340						345					350		
Gln	Ser	Leu	Gly	Ile	Pro	Tyr	His	Ile	Val	Asn	Ile	Val	Ser	Gly	Ser
	355					360						365			
Leu	Asn	His	Ala	Ala	Ser	Lys	Lys	Leu	Asp	Leu	Glu	Ala	Trp	Phe	Pro
	370					375					380				
Gly	Ser	Gly	Ala	Phe	Arg	Glu	Leu	Val	Ser	Cys	Ser	Asn	Cys	Thr	Asp
385					390					395					400
Tyr	Gln	Ala	Arg	Arg	Leu	Arg	Ile	Arg	Tyr	Gly	Gln	Thr	Lys	Lys	Met
				405					410						415
Met	Asp	Lys	Val	Glu	Phe	Val	His	Met	Leu	Asn	Ala	Thr	Met	Cys	Ala
		420						425					430		
Thr	Thr	Cys	Thr	Ile	Cys	Ala	Ile	Leu	Glu	Asn	Tyr	Gln	Thr	Glu	Lys
	435						440					445			
Gly	Ile	Thr	Val	Pro	Glu	Lys	Leu	Lys	Glu	Phe	Met	Pro	Pro	Gly	Leu
	450					455					460				
Gln	Glu	Leu	Ile	Pro	Phe	Val	Lys	Pro	Ala	Pro	Ile	Glu	Gln	Glu	Pro
465					470					475					480
Ser	Lys	Lys	Gln	Lys	Lys	Gln	His	Glu	Gly	Ser	Lys	Lys	Lys	Ala	Ala
				485					490					495	
Ala	Arg	Asp	Val	Thr	Leu	Glu	Asn	Arg	Leu	Gln	Asn	Met	Glu	Val	Thr
			500					505					510		
Asp	Ala														

<210> 7

<211> 508

5 <212> PRT

<213> Homo sapiens

<400> 7

Met Ala Asp Gly Gln Val Ala Glu Leu Leu Leu Arg Arg Leu Glu Ala
 1 5 10 15
 Ser Asp Gly Gly Leu Asp Ser Ala Glu Leu Ala Ala Glu Leu Gly Met
 20 25 30
 Glu His Gln Ala Val Val Gly Ala Val Lys Ser Leu Gln Ala Leu Gly
 35 40 45
 Glu Val Ile Glu Ala Glu Leu Arg Ser Thr Lys His Trp Glu Leu Thr
 50 55 60
 Ala Glu Gly Glu Glu Ile Ala Arg Glu Gly Ser His Glu Ala Arg Val
 65 70 75 80
 Phe Arg Ser Ile Pro Pro Glu Gly Leu Ala Gln Ser Glu Leu Met Arg
 85 90 95
 Leu Pro Ser Gly Lys Val Gly Phe Ser Lys Ala Met Ser Asn Lys Trp
 100 105 110
 Ile Arg Val Asp Lys Ser Ala Ala Asp Gly Pro Arg Val Phe Arg Val
 115 120 125
 Val Asp Ser Met Glu Asp Glu Val Gln Arg Arg Leu Gln Leu Val Arg
 130 135 140
 Gly Gly Gln Ala Glu Lys Leu Gly Glu Lys Glu Arg Ser Glu Leu Arg
 145 150 155 160
 Lys Arg Lys Leu Leu Ala Glu Val Thr Leu Lys Thr Tyr Trp Val Ser
 165 170 175
 Lys Gly Ser Ala Phe Ser Thr Ser Ile Ser Lys Gln Glu Thr Glu Leu
 180 185 190
 Ser Pro Glu Met Ile Ser Ser Gly Ser Trp Arg Asp Arg Pro Phe Lys
 195 200 205
 Pro Tyr Asn Phe Leu Ala His Gly Val Leu Pro Asp Ser Gly His Leu
 210 215 220
 His Pro Leu Leu Lys Val Arg Ser Gln Phe Arg Gln Ile Phe Leu Glu
 225 230 235 240
 Met Gly Phe Thr Glu Met Pro Thr Asp Asn Phe Ile Glu Ser Ser Phe
 245 250 255
 Trp Asn Phe Asp Ala Leu Phe Gln Pro Gln Gln His Pro Ala Arg Asp
 260 265 270
 Gln His Asp Thr Phe Phe Leu Arg Asp Pro Ala Glu Ala Leu Gln Leu
 275 280 285
 Pro Met Asp Tyr Val Gln Arg Val Lys Arg Thr His Ser Gln Gly Gly
 290 295 300
 Tyr Gly Ser Gln Gly Tyr Lys Tyr Asn Trp Lys Leu Asp Glu Ala Arg
 305 310 315 320
 Lys Asn Leu Leu Arg Thr His Thr Thr Ser Ala Ser Ala Arg Ala Leu
 325 330 335
 Tyr Arg Leu Ala Gln Lys Lys Pro Phe Thr Pro Val Lys Tyr Phe Ser
 340 345 350
 Ile Asp Arg Val Phe Arg Asn Glu Thr Leu Asp Ala Thr His Leu Ala
 355 360 365
 Glu Phe His Gln Ile Glu Gly Val Val Ala Asp His Gly Leu Thr Leu
 370 375 380
 Gly His Leu Met Gly Val Leu Arg Glu Phe Phe Thr Lys Leu Gly Ile
 385 390 395 400
 Thr Gln Leu Arg Phe Lys Pro Ala Tyr Asn Pro Tyr Thr Glu Pro Ser
 405 410 415
 Met Glu Val Phe Ser Tyr His Gln Gly Leu Lys Lys Trp Val Glu Val
 420 425 430
 Gly Asn Ser Gly Val Phe Arg Pro Glu Met Leu Leu Pro Met Gly Leu
 435 440 445
 Pro Glu Asn Val Ser Val Ile Ala Trp Gly Leu Ser Leu Glu Arg Pro
 450 455 460
 Thr Met Ile Lys Tyr Gly Ile Asn Asn Ile Arg Glu Leu Val Gly His
 465 470 475 480
 Lys Val Asn Leu Gln Met Val Tyr Asp Ser Pro Leu Cys Arg Leu Asp
 485 490 495
 Ala Glu Pro Arg Pro Pro Pro Thr Gln Glu Ala Ala
 500 505

<210> 8
 <211> 597
 <212> PRT
 <213> Homo sapiens

5

<400> 8

ES 2 552 773 T3

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

Tyr Lys Val Thr Tyr His Pro Asp Gly Pro Glu Gly Gln Ala Tyr Asp
 370 375 380
 Val Asp Phe Thr Pro Pro Phe Arg Arg Ile Asn Met Val Glu Glu Leu
 385 390 395 400
 Glu Lys Ala Leu Gly Met Lys Leu Pro Glu Thr Asn Leu Phe Glu Thr
 405 410 415
 Glu Glu Thr Arg Lys Ile Leu Asp Asp Ile Cys Val Ala Lys Ala Val
 420 425 430
 Glu Cys Pro Pro Arg Thr Thr Ala Arg Leu Leu Asp Lys Leu Val
 435 440 445
 Gly Glu Phe Leu Glu Val Thr Cys Ile Asn Pro Thr Phe Ile Cys Asp
 450 455 460
 His Pro Gln Ile Met Ser Pro Leu Ala Lys Trp His Arg Ser Lys Glu
 465 470 475 480
 Gly Leu Thr Glu Arg Phe Glu Leu Phe Val Met Lys Lys Glu Ile Cys
 485 490 495
 Asn Ala Tyr Thr Glu Leu Asn Asp Pro Met Arg Gln Arg Gln Leu Phe
 500 505 510
 Glu Glu Gln Ala Lys Ala Lys Ala Ala Gly Asp Asp Glu Ala Met Phe
 515 520 525
 Ile Asp Glu Asn Phe Cys Thr Ala Leu Glu Tyr Gly Leu Pro Pro Thr
 530 535 540
 Ala Gly Trp Gly Met Gly Ile Asp Arg Val Ala Met Phe Leu Thr Asp
 545 550 555 560
 Ser Asn Asn Ile Lys Glu Val Leu Leu Phe Pro Ala Met Lys Pro Glu
 565 570 575
 Asp Lys Lys Glu Asn Val Ala Thr Thr Asp Thr Leu Glu Ser Thr Thr
 580 585 590
 Val Gly Thr Ser Val
 595

<210> 9

<211> 548

5 <212> PRT

<213> Homo sapiens

<400> 9

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

10

ES 2 552 773 T3

Lys His Asn Asp Leu Asp Asp Val Gly Lys Asp Val Tyr His His Thr
 85 90 95
 Phe Phe Glu Met Leu Gly Ser Trp Ser Phe Gly Asp Tyr Phe Lys Glu
 100 105 110
 Leu Ala Cys Lys Met Ala Leu Glu Leu Leu Thr Gln Glu Phe Gly Ile
 115 120 125
 Pro Ile Glu Arg Leu Tyr Val Thr Tyr Phe Gly Gly Asp Glu Ala Ala
 130 135 140
 Gly Leu Glu Ala Asp Leu Glu Cys Lys Gln Ile Trp Gln Asn Leu Gly
 145 150 155 160
 Leu Asp Asp Thr Lys Ile Leu Pro Gly Asn Met Lys Asp Asn Phe Trp
 165 170 175
 Glu Met Gly Asp Thr Gly Pro Cys Gly Pro Cys Ser Glu Ile His Tyr
 180 185 190
 Asp Arg Ile Gly Gly Arg Asp Ala Ala His Leu Val Asn Gln Asp Asp
 195 200 205
 Pro Asn Val Leu Glu Ile Trp Asn Leu Val Phe Ile Gln Tyr Asn Arg
 210 215 220
 Glu Ala Asp Gly Ile Leu Lys Pro Leu Pro Lys Lys Ser Ile Asp Thr
 225 230 235 240
 Gly Met Gly Leu Glu Arg Leu Val Ser Val Leu Gln Asn Lys Met Ser
 245 250 255
 Asn Tyr Asp Thr Asp Leu Phe Val Pro Tyr Phe Glu Ala Ile Gln Lys
 260 265 270
 Gly Thr Gly Ala Arg Pro Tyr Thr Gly Lys Val Gly Ala Glu Asp Ala
 275 280 285
 Asp Gly Ile Asp Met Ala Tyr Arg Val Leu Ala Asp His Ala Arg Thr
 290 295 300
 Ile Thr Val Ala Leu Ala Asp Gly Gly Arg Pro Asp Asn Thr Gly Arg
 305 310 315 320
 Gly Tyr Val Leu Arg Arg Ile Leu Arg Arg Ala Val Arg Tyr Ala His
 325 330 335
 Glu Lys Leu Asn Ala Ser Arg Gly Phe Phe Ala Thr Leu Val Asp Val
 340 345 350
 Val Val Gln Ser Leu Gly Asp Ala Phe Pro Glu Leu Lys Lys Asp Pro
 355 360 365
 Asp Met Val Lys Asp Ile Ile Asn Glu Glu Glu Val Gln Phe Leu Lys
 370 375 380
 Thr Leu Ser Arg Gly Arg Arg Ile Leu Asp Arg Lys Ile Gln Ser Leu
 385 390 395 400
 Gly Asp Ser Lys Thr Ile Pro Gly Asp Thr Ala Trp Leu Leu Tyr Asp
 405 410 415
 Thr Tyr Gly Phe Pro Val Asp Leu Thr Gly Leu Ile Ala Glu Glu Lys
 420 425 430
 Gly Leu Val Val Asp Met Asp Gly Phe Glu Glu Glu Arg Lys Leu Ala
 435 440 445
 Gln Leu Lys Ser Gln Gly Lys Gly Ala Gly Gly Glu Asp Leu Ile Met
 450 455 460
 Leu Asp Ile Tyr Ala Ile Glu Glu Leu Arg Ala Arg Gly Leu Glu Val
 465 470 475 480
 Thr Asp Asp Ser Pro Lys Tyr Asn Tyr His Leu Asp Ser Ser Gly Ser
 485 490 495
 Tyr Val Phe Glu Asn Thr Val Ala Thr Val Met Ala Leu Arg Arg Glu
 500 505 510
 Lys Met Phe Val Glu Glu Val Ser Thr Gly Gln Glu Cys Gly Val Val
 515 520 525
 Leu Asp Lys Thr Cys Phe Tyr Ala Glu Gln Gly Gly Gln Ile Tyr Asp
 530 535 540
 Glu Gly Tyr Leu Val Lys Val Asp Asp Ser Ser Glu Asp Lys Thr Glu
 545 550 555 560
 Phe Thr Val Lys Asn Ala Gln Val Arg Gly Gly Tyr Val Leu His Ile
 565 570 575
 Gly Thr Ile Tyr Gly Asp Leu Lys Val Gly Asp Gln Val Trp Leu Phe

ES 2 552 773 T3

<210> 12
 <211> 115
 <212> PRT
 <213> Homo sapiens

5

<400> 12

```

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

10

<210> 13
 <211> 170
 <212> PRT
 <213> Homo sapiens

15

<400> 13

```

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

20

REIVINDICACIONES

1. Un polipéptido de aminoacil-ARNt sintetasa (AARS) aislado que consiste en 60-300 residuos de aminoácido, en donde el polipéptido de AARS comprende tres hebras β antiparalelas flanqueadas en cada extremo por una hélice α y muestra una actividad de señalización celular, en donde el polipéptido comprende los residuos 325-410 del SEQ ID NO: 6 o un fragmento activo o variante que muestra una identidad de al menos 90% con los residuos 325-410 del SEQ ID NO: 6, y en donde el polipéptido tiene actividad quimioquina.
2. Una proteína de fusión que comprende el polipéptido aislado de la reivindicación 1, en donde dicho polipéptido está unido covalentemente a un polipéptido heterólogo.
3. El polipéptido aislado de la reivindicación 1 o 2, en donde el polipéptido consiste en 60 a 200 residuos de aminoácido.
4. Una composición farmacéutica que comprende un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1-3.
5. Un a) polinucleótido aislado que codifica un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1-3, b) un vector que comprende un polinucleótido aislado de a), o c) una célula anfitriona que comprende un vector de b).
6. Un método de identificación de un fragmento de polipéptido que tiene una actividad de señalización celular, que comprende las etapas de identificar una secuencia de proteína que contiene un motivo estructural formado por tres hebras β antiparalelas flanqueadas en cada extremo por una hélice α , determinar los límites de residuos de aminoácido del motivo estructural dentro de dicha proteína, e identificar de ese modo un fragmento de polipéptido que tiene actividad de señalización celular.
7. El método de la reivindicación 6, en donde el fragmento de polipéptido comprende de 60 a 200 aminoácidos de la proteína.
8. El método de la reivindicación 6, en donde la etapa de identificación de una secuencia de proteína que contiene un motivo estructural formado por tres hebras β antiparalelas flanqueadas en cada extremo por una hélice α se realiza utilizando un método de predicción de la estructura secundaria.
9. El método de la reivindicación 6, en donde la etapa de identificación de una secuencia de proteína que contiene un motivo estructural formado por tres hebras β antiparalelas flanqueadas en cada extremo por una hélice α se realiza utilizando un método de predicción de la estructura secundaria seleccionado del grupo que consiste en PHDsec, NSSP, SOPM, DSC, SSPRED, MultiPredict, PSA, NNPREDICT, APSSP, GOR, HNN, HTMSRAP, Jpred, JUFO, nnPredict, Porter, PredictProtein, Prof, PSIPred, SOPMA, SSpro y DLP-SVM.
10. Una composición para su uso en la modulación de la señalización celular que comprende una cantidad eficaz de:
- a) un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1-3;
- b) un polinucleótido de acuerdo con la reivindicación 5; y/o
- c) un anticuerpo o fragmento de unión a antígeno del mismo que se une específicamente a un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1-3;
11. La composición de la reivindicación 10, para su uso en la inducción de la secreción de TNF- α .
12. La composición de la reivindicación 10, para su uso en la señalización a través de un receptor 2 de tipo Toll.
13. Un método de escrutinio para identificar un modulador de la señalización celular, que comprende las etapas de:
- (a) formar una mezcla de reacción que incluye:
- (i) un componente seleccionado del grupo que consiste en un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1-3; un polinucleótido de acuerdo con la reivindicación 5; y/o un anticuerpo o fragmento de unión a antígeno del mismo que se une específicamente a un polipéptido de acuerdo con una cualquiera de las reivindicaciones 1-3;
- (ii) un compañero de unión, un efector celular y/o un tipo de célula conocidos por unirse a y/o estar modulados por dicho componente; y
- (iii) un compuesto de ensayo; y
- (b) detectar una interacción de dicho componente con el compañero de unión, el efector celular y/o el tipo de célula, en donde un cambio en la interacción en presencia de un compuesto de ensayo, con respecto a la interacción en ausencia del compuesto de ensayo, identifica de ese modo un modulador de la señalización celular.

14. El polipéptido aislado de acuerdo con una cualquiera de las reivindicaciones 1-3 para su uso como medicamento.

5 15. El polipéptido aislado de acuerdo con una cualquiera de las reivindicaciones 1-3 para su uso en el tratamiento o prevención de enfermedades neoplásicas, enfermedades del sistema inmunitario tales como enfermedades autoinmunitarias e inflamación, enfermedades infecciosas, enfermedades metabólicas, enfermedades neuronales/neurológicas, enfermedades musculares/cardiovasculares, enfermedades asociadas con hematopoyesis aberrante, enfermedades asociadas con angiogénesis aberrante y enfermedades asociadas con supervivencia celular aberrante.

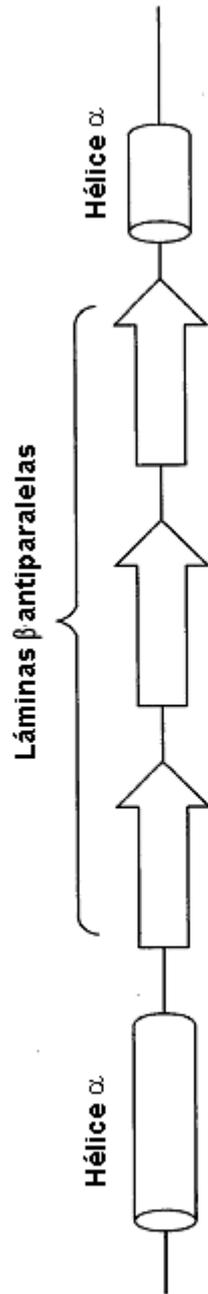


FIG. 1

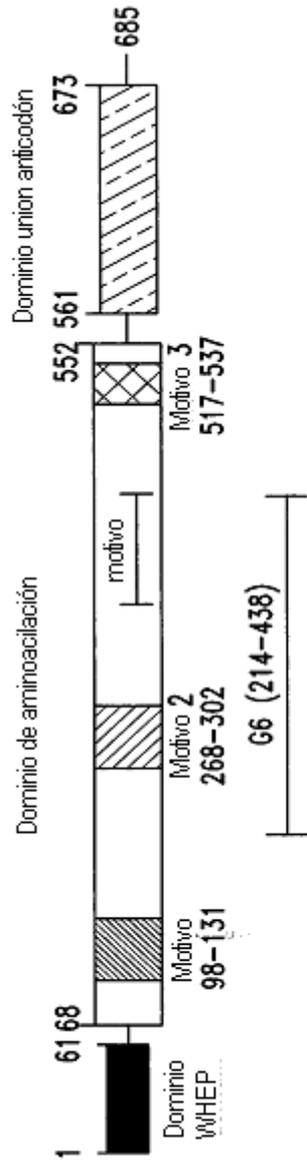


FIG. 2A

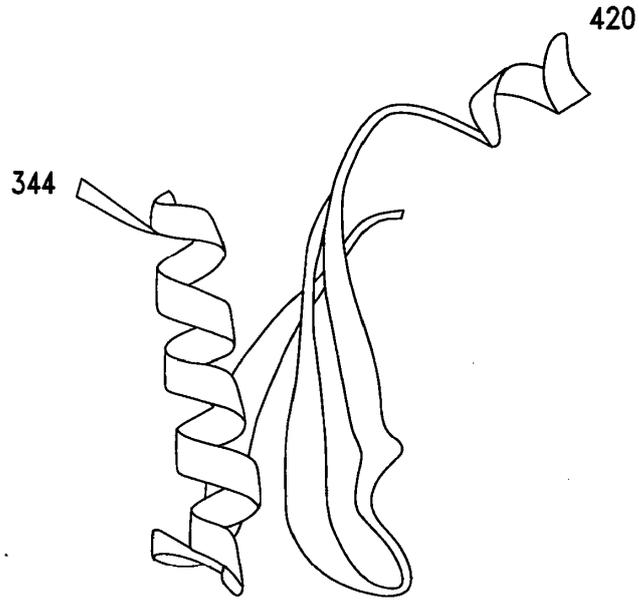


FIG. 2B

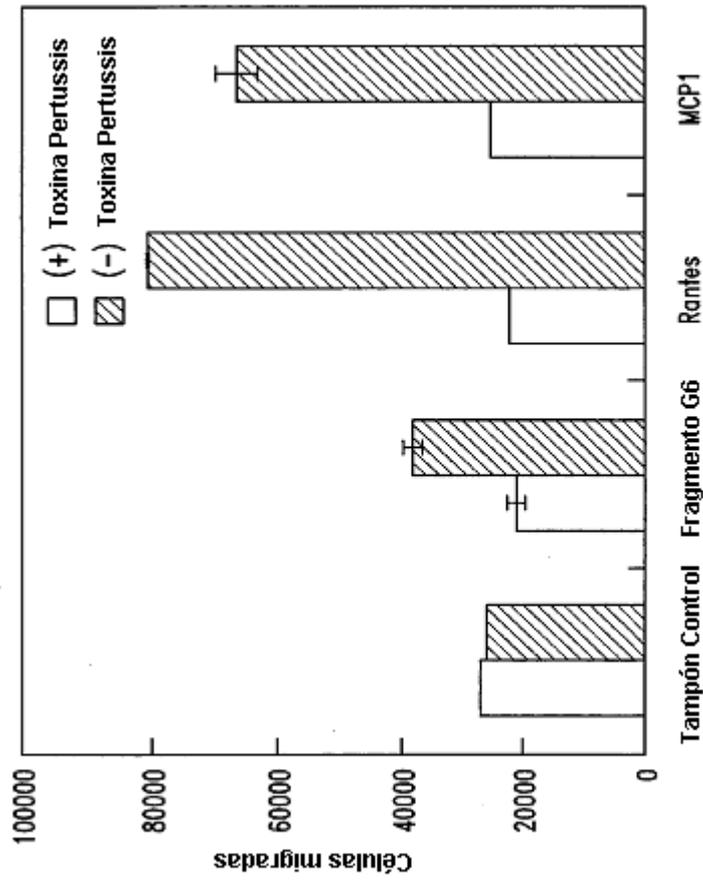


FIG. 2C

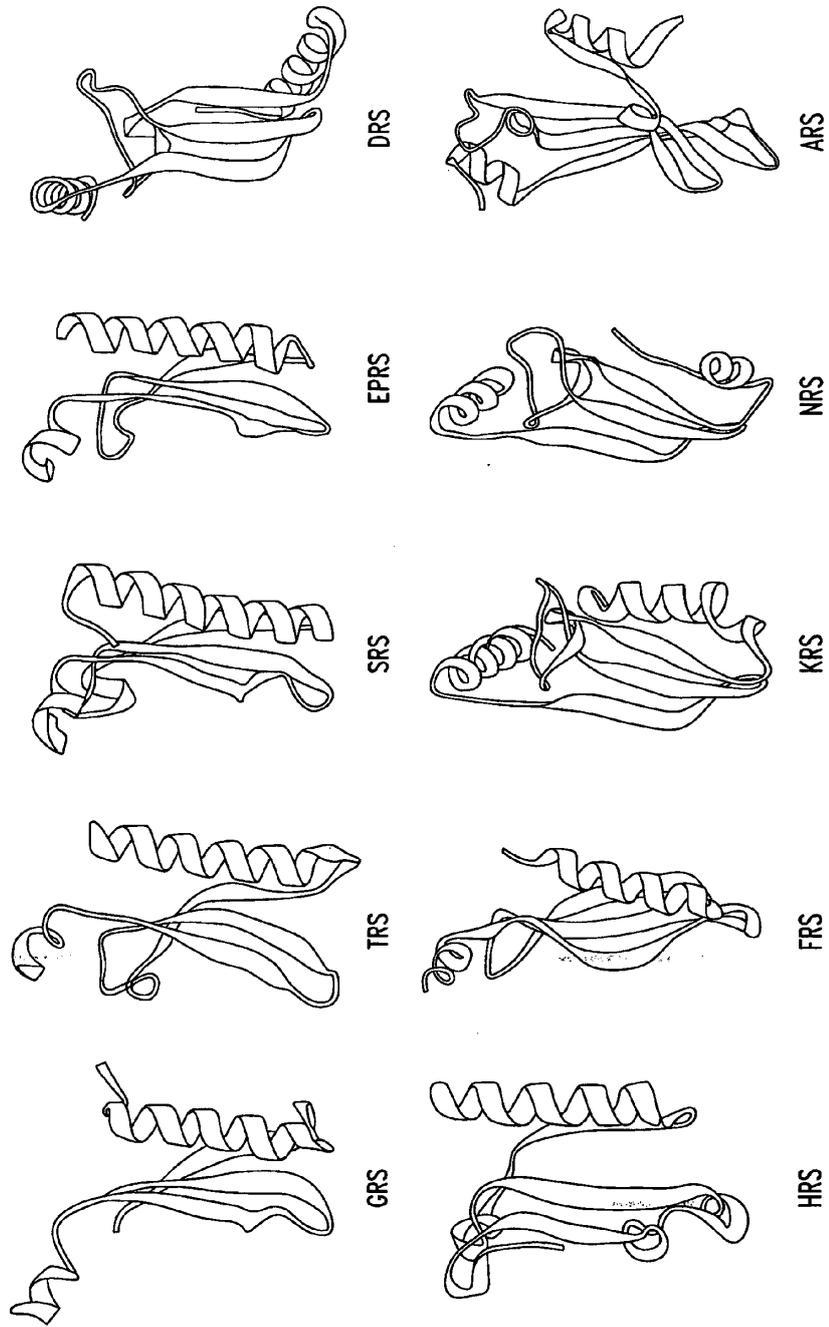


FIG. 3

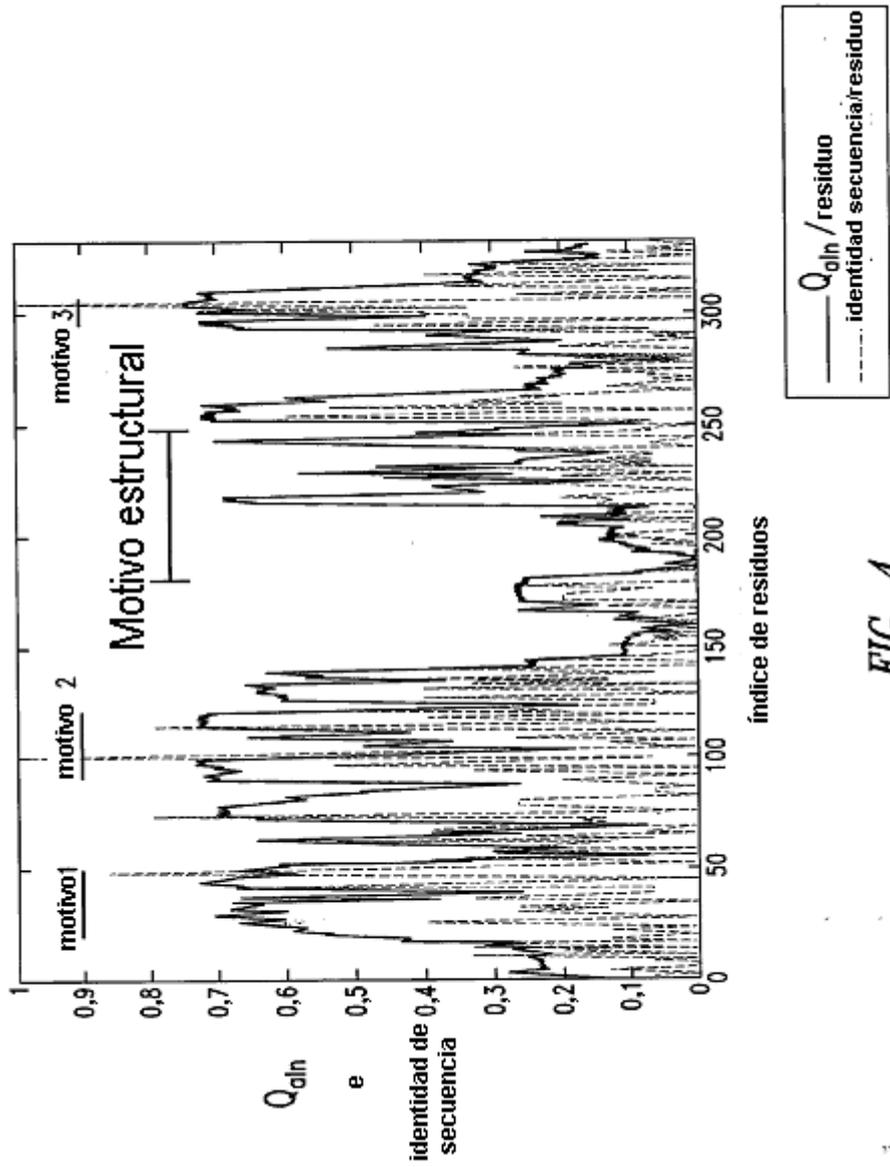


FIG. 4

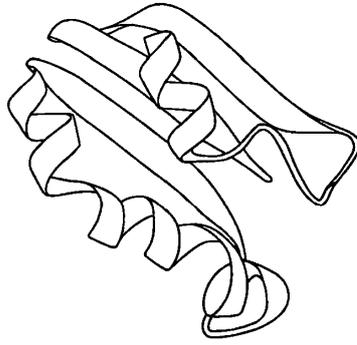


FIG. 5C

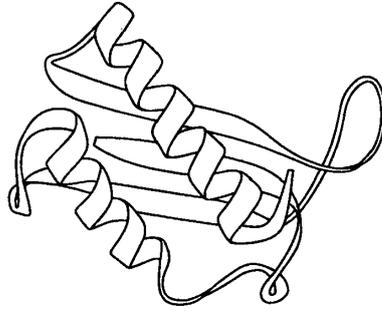


FIG. 5B

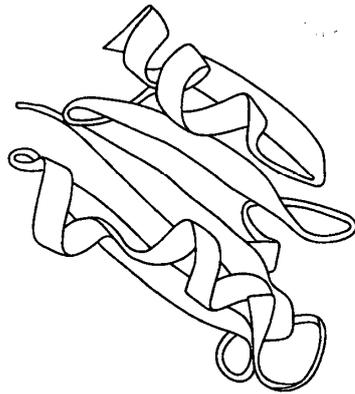


FIG. 5A

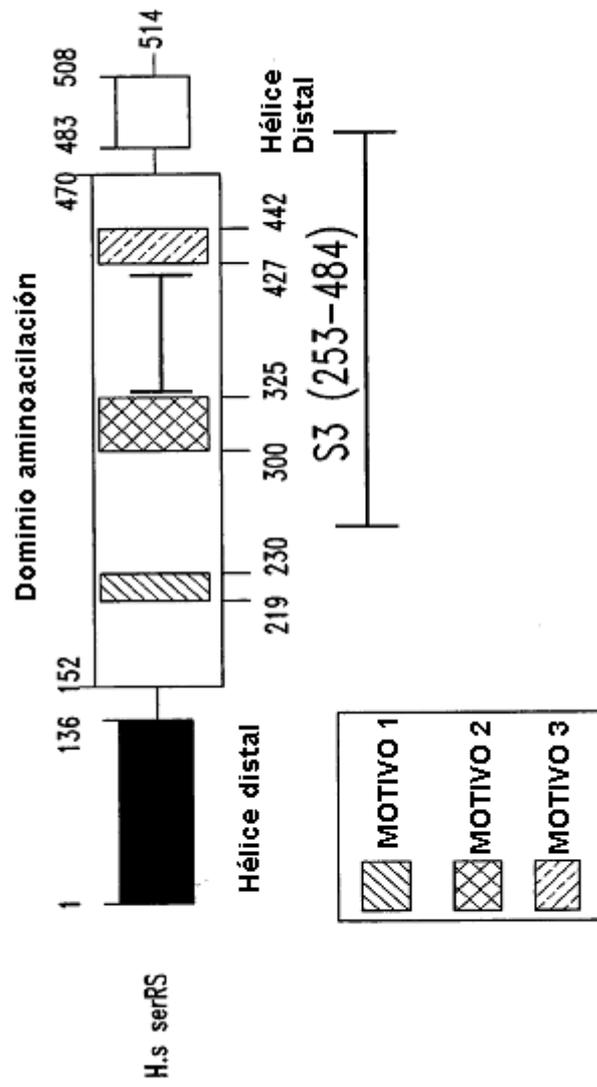


FIG. 6

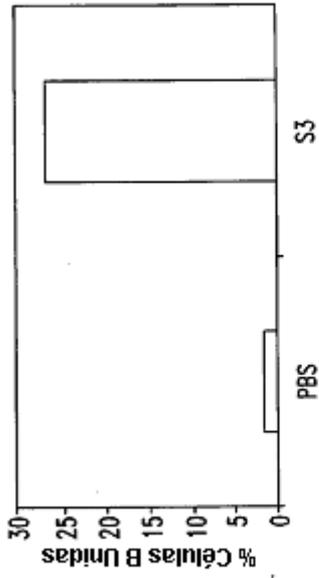


FIG. 7B

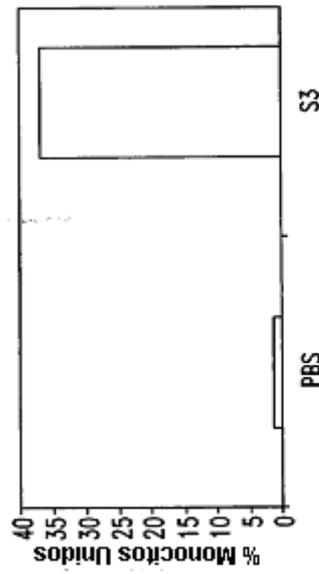


FIG. 7A

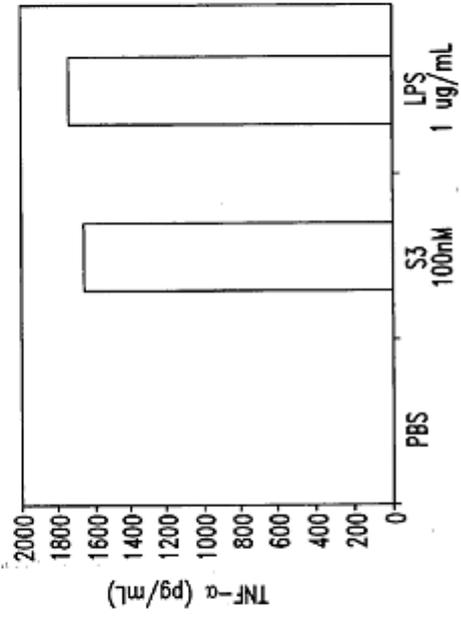


FIG. 7C

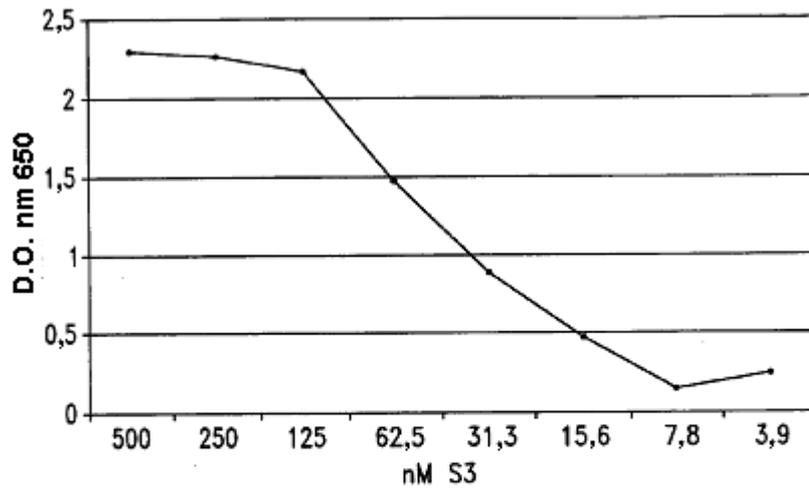


FIG. 8

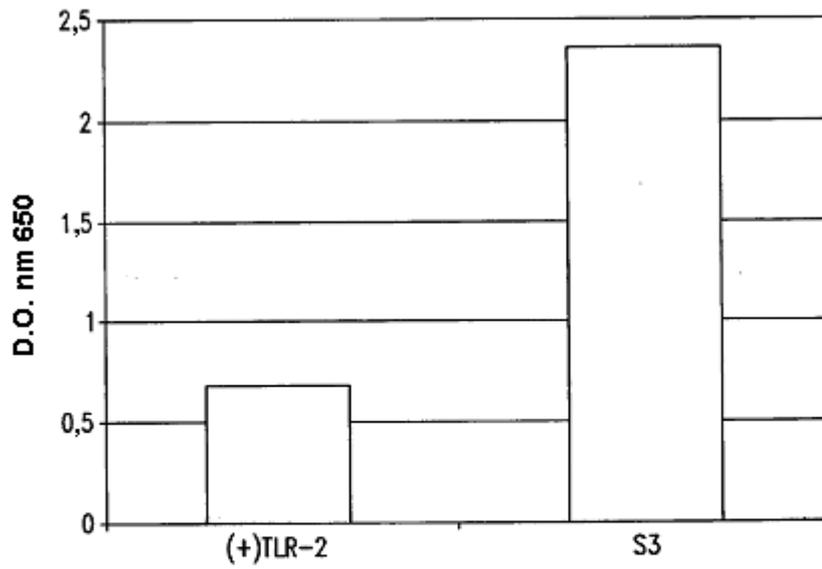


FIG. 9