(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 668 207**

(51) Int. CI.:

A61K 38/16 (2006.01)
A61K 39/395 (2006.01)
C07K 14/525 (2006.01)
A61P 35/00 (2006.01)
A61P 3/10 (2006.01)
A61P 29/00 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **03.05.2011 PCT/US2011/035056**

(87) Fecha y número de publicación internacional: **10.11.2011 WO11140135**

(96) Fecha de presentación y número de la solicitud europea: **03.05.2011 E 11778207 (8)**

(97) Fecha y número de publicación de la concesión europea: **28.02.2018 EP 2566496**

(54) Título: **Descubrimiento innovador de composiciones terapéuticas, de diagnóstico y de anticuerpos relacionadas con fragmentos de proteínas de metionil-ARNt sintetasas**

(30) Prioridad:

04.05.2010 US 331282 P
03.05.2010 US 330596 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
17.05.2018

(73) Titular/es:

ATYR PHARMA, INC. (50.0%)
3545 John Hopkins Court, Suite 250
San Diego, CA 92121, US y
PANGU BIOPHARMA LIMITED (50.0%)

(72) Inventor/es:

GREENE, LESLIE, ANN;
CHIANG, KYLE, P.;
HONG, FEI;
VASSEROT, ALAIN, P.;
LO, WING-SZE;
WATKINS, JEFFRY, D.;
MENDLEIN, JOHN, D. y
QUINN, CHERYL, L.

(74) Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

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

Descubrimiento innovador de composiciones terapéuticas, de diagnóstico y de anticuerpos relacionadas con fragmentos de proteínas de metionil-ARNt sintetasas

Referencia a solicitudes relacionadas

La presente solicitud reivindica el beneficio de prioridad según 35 U.S.C. § 119(e) de la solicitud de patente provisional de Estados Unidos N.º 61/330.596 presentada el 3 de mayo de 2010 y la solicitud de patente provisional de Estados Unidos N.º 61/331.282 presentada el 4 de mayo de 2010.

Campo técnico

La presente invención se refiere en general a composiciones que comprenden fragmentos de proteínas de nueva identificación de aminoacil-ARNt sintetasas y otras proteínas, polinucleótidos que los codifican y complementos de los mismos, agentes relacionados y métodos de uso de los mismos en diagnóstico, descubrimiento de fármacos, investigación y aplicaciones terapéuticas.

Antecedentes

Durante más de cuatro décadas, se ha considerado a las aminoacil-ARNt sintetasas (AARS) proteínas constitutivas esenciales que catalizan la aminoacilación de moléculas de ARNt como parte de la decodificación de la información genética durante el proceso de traducción de proteínas. Las AARS se han estudiado exhaustivamente a este respecto, y muchas de sus secuencias de longitud completa se han clonado para análisis de secuencias y para proporcionar una fuente rica de experimentación bioquímica. Algunos fragmentos de AARS y otras proteínas, sin embargo, poseen actividades inesperadas no asociadas con la aminoacilación, incluyendo actividades de señalización extracelular que modulan rutas más allá de la traducción de proteínas. En general, estas actividades inesperadas no se han observado en el contexto de las secuencias de proteínas de longitud completa o parentales; en su lugar, se han observado después de la retirada o resección de fragmentos proteicos de AARS de sus secuencias parentales, o expresando y purificando suficientemente secuencias de AARS de fragmentos y después ensayando con respecto a nuevas actividades no relacionadas con sintetasa.

Se conocen varios fragmentos de MetRS de bases de datos en línea, por ejemplo los números de referencia UPI0000E00904, UPI0000E4C913 y A6MJU3 en la base de datos UniProt y el número de referencia AI368577 en la base de datos de EMBL.

Aunque las secuencias de longitud completa de AARS se han conocido durante algún tiempo, no se ha realizado un análisis experimental sistemático para dilucidar dichos fragmentos de proteínas AARS, o fragmentos proteicos de proteínas relacionadas o asociadas, o para evaluar el papel potencial de las proteínas de AARS de longitud completa para nuevas actividades biológicas fuera del contexto de la síntesis de aminoácidos. En partes de la presente memoria descriptiva, dichos fragmentos proteicos de AARS, dominios de AARS o variantes de corte y empalme alternativo de AARS se denominan en el presente documento "resectinas". En su contexto más amplio, el término "resectina" se refiere a una parte de una proteína que se ha escindido o restringido (bien por medio de proteólisis, corte y empalme alternativo, mutagénesis o ingeniería genética recombinante) a partir del contexto de su secuencia de proteínas parental o de longitud completa nativa, que con frecuencia enmascara de otro modo sus nuevas actividades biológicas. De forma similar, no se ha realizado ningún análisis experimental sistemático para explorar el uso de dichas resectinas como agentes bioterapéuticos, agentes de diagnóstico o dianas farmacológicas en el tratamiento de diversas afecciones médicas, o su asociación potencial a enfermedades humanas. Como genes constitutivos esenciales con una función conocida en los mamíferos que es crítica para la vida, las AARS no se han considerado dianas farmacológicas en mamíferos, ni se han analizado por secuenciación genómica convencional, bioinformática o esfuerzos similares para identificar resectinas que tengan actividades no sintetasa. De forma similar la dirección de los esfuerzos de investigación bioquímica convencionales se ha alejado de la caracterización de las propiedades biológicas de resectinas AARS y su relevancia terapéutica y de diagnóstico potencial, principalmente debido al papel previamente entendido de sus AARS parentales de longitud completa correspondientes.

Breve descripción de los dibujos

La **Figura 1** muestra la estructura de dominio de metionil aminoacil ARNt sintetasa superpuesta con las posiciones relativas y tamaños de los polipéptidos de MetRS N-terminales identificados mostrados esquemáticamente. La Figura 1A representa fragmentos identificados a partir de análisis de espectrometría de masas, la Figura 1B representa los fragmentos identificados a partir de secuenciación profunda de transcriptomas, y la Figura 1C representa fragmentos identificados a partir de análisis bioinformáticos.

La **Figura 2** muestra la estructura de dominio de la metionil aminoacil ARNt sintetasa superpuesta con las posiciones relativas y tamaños de los polipéptidos de MetRS C-terminales mostrados esquemáticamente. La Figura 2A representa fragmentos identificados a partir de análisis de espectrometría de masas. La Figura 2B representa los fragmentos identificados a partir de secuenciación profunda de transcritomas y la Figura 2C

representa fragmentos identificados a partir de análisis bioinformático.

Breve sumario de la invención

La invención se define por las reivindicaciones adjuntas.

Las realizaciones de la presente divulgación se refieren en general al descubrimiento de fragmentos proteicos de metionil ARNt sintetasa (MetRS), que poseen actividades biológicas no canónicas, tales como actividades de señalización extracelular, y/u otras características de relevancia terapéutica y de diagnóstico. Las AARS son elementos universales y esenciales de la maquinaria de síntesis de proteínas hallada en todos los organismos, pero las AARS humanas y sus proteínas asociadas tienen variantes resecadas de origen natural, con potentes actividades de señalización celular que contribuyen al funcionamiento normal de seres humanos. Las actividades de estos fragmentos proteicos son distintas de las actividades de síntesis de proteínas habitualmente conocidas para AARS, y la presente divulgación incluye el descubrimiento y desarrollo de estas proteínas resecadas como nuevos agentes bioterapéuticos, nuevos reactivos de investigación de descubrimiento y como nuevos antígenos/dianas para biologías y agentes de diagnóstico dirigidos que pueden usarse para potencialmente tratar o diagnosticar una amplia diversidad de enfermedades humanas, tales como enfermedades o trastornos inflamatorios, hematológicos, neurodegenerativos, autoinmunitarios, hematopoyéticos, cardiovasculares y metabólicos.

El fragmento o los fragmentos proteicos de MetRS de la presente divulgación pueden por lo tanto denominarse "resectinas", o como alternativa "apendacrinas". Como se ha indicado anteriormente, el término "resectina" deriva del proceso de escindir o resecar un fragmento proteico de MetRS dado del contexto de su secuencia de MetRS parental de longitud completa, que normalmente enmascara sus actividades no canónicas. En ciertos casos, los fragmentos proteicos de MetRS y polinucleótidos de la presente invención se identificaron mediante la aparición de este proceso de resección, bien de origen natural (por ejemplo, proteolítico, variante de corte y empalme), inducido artificialmente o predicho. El término "apendacrina" deriva de una combinación de "añadir" (del latín-appendere) y "separar" o "distinguir" (del griego -crines), y también refleja la separación de uno o más dominios adjuntos de los fragmentos proteicos de MetRS de sus secuencias de MetRS de longitud completa o parentales correspondientes.

Aunque se ha mostrado previamente que algunos fragmentos de AARS tienen actividades no sintetasa, la expresión, el aislamiento, la purificación y la caracterización de dichos fragmentos para utilidad bioterapéutica, de descubrimiento o de diagnóstico está limitada y los expertos en la materia no apreciarían fácilmente que dichas actividades se asocien con cada miembro de la familia completa de AARS, o con fragmentos alternativos. Aquí, se utilizó el enfoque metódico para descubrir y verificar fragmentos proteicos de AARS para las 20 AARS mitocondriales y las 20 AARS citosólicas (y proteínas asociadas) para el descubrimiento bioterapéutico y utilidad de diagnóstico. Por ejemplo, algunos de los fragmentos proteicos de MetRS presentes y polinucleótidos que los codifican se identifican de muestras biológicas usando espectrometría de masas (EM), principalmente para identificar fragmentos proteolíticos, y se identificaron otros por técnicas de secuenciación profunda, principalmente para identificar variantes de corte y empalme. Otro fragmento u otros fragmentos proteicos de MetRS se identifican usando predicciones por ordenador de secuencias de aminoácidos, tales como comparando computacionalmente sintetasas de seres humanos y organismos inferiores junto con demarcaciones clave (por ejemplo, sitios de proteasa); este enfoque utilizó análisis de secuencias de la MetRS de longitud completa basándose en criterios específicos para distinguir fragmentos proteolíticos y dominios funcionales que poseen actividades biológicas no canónicas.

Las nuevas resectinas de las AARS son inesperadas, y su expresión diferencial también es inesperada. Se ven resecciones específicas en diferentes tratamientos (por ejemplo, de células cultivadas en medios con o sin suero), en diferentes estadios de crecimiento (por ejemplo, cerebro adulto frente a cerebro fetal) y para diferentes tipos tisulares (por ejemplo, páncreas frente a hígado). El patrón de expresión no es igual para todas las aminoacil ARNt sintetasas a pesar del hecho de que las funciones canónicas para todas las aminoacil ARNt sintetasas son necesarias en las mismas localizaciones celulares y en cantidades relativamente proporcionales. No se esperaría que los niveles de una actividad aminoacil ARNt sintetasa aumentaran sin un aumento en las cantidades de otras actividades de aminoacil ARNt sintetasa al mismo tiempo. Los datos de espectrometría de masas y secuenciación profunda indican que las resectinas de aminoacil ARNt sintetasa no tienen diversos niveles y aparecen en diferentes sitios y en diferentes estadios.

Además, los fragmentos proteicos de AARS pueden expresarse y purificarse hasta una pureza suficientemente alta para discernir sus propiedades biológicas. Previamente, los fragmentos no eran con frecuencia de suficiente pureza, plegamiento y estabilidad para permitir la caracterización biológica apropiada de actividades no sintetasa. Se usan ensayos basados en células, por ejemplo, junto con resectinas suficientemente puras, estables, solubles y plegadas para revelar sus actividades bioterapéuticas, de descubrimiento o de diagnóstico importantes.

En particular, las realizaciones de la presente divulgación se refieren a fragmentos proteicos de metionil ARNt sintetasas, agentes relacionados y composiciones de utilidad bioterapéutica, de descubrimiento o de diagnóstico, y métodos de uso de los mismos. Las composiciones de la presente divulgación son útiles en diversas aplicaciones de diagnóstico, de descubrimiento de fármacos y terapéuticas, como se describe en el presente documento.

Preferentemente, las proteínas de MetRS y fragmentos se purifican y almacenan en condiciones adecuadas en la medida requerida para dichos usos bioterapéuticos, de descubrimiento o de diagnóstico.

Ciertas realizaciones incluyen composiciones que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado de al menos aproximadamente 50 a aproximadamente 100 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137, y tiene una solubilidad de al menos aproximadamente 5 mg/ml, y en el que la composición tiene una pureza de al menos aproximadamente 95 % en proteínas, y menos de aproximadamente 10 UE/mg de proteína endotoxina. En un aspecto, la composición es una composición terapéutica. En realizaciones específicas, la composición está sustancialmente sin suero. En algunas realizaciones el fragmento proteico de MetRS comprende una actividad no canónica. En algunas realizaciones, la actividad biológica no canónica se selecciona de modulación de señalización extracelular, modulación de proliferación celular, modulación de diferenciación celular, modulación de la transcripción génica, modulación de la producción o actividad de citocinas, modulación de la actividad receptora de citocinas y modulación de la inflamación. En algunas realizaciones, el fragmento proteico de MetRS tiene una CE_{50} de menos de aproximadamente 1 nM, aproximadamente 5 nM, aproximadamente 10 nM, aproximadamente 50 nM, aproximadamente 100 nM o aproximadamente 200 nM para una actividad biológica no canónica basada en células.

En ciertas realizaciones, el fragmento proteico de MetRS se fusiona con un polipéptido heterólogo. En algunas realizaciones, la proteína de fusión de MetRS conserva sustancialmente una actividad no canónica del fragmento proteico de MetRS. En algunas realizaciones, la proteína de fusión MetRS suprime una actividad no canónica del fragmento proteico de MetRS. En algunas realizaciones, el polipéptido heterólogo se une al extremo N terminal del fragmento proteico de MetRS. En algunas realizaciones, el polipéptido heterólogo se une al extremo C terminal del fragmento proteico de MetRS. En un aspecto de cualquiera de estas realizaciones el polipéptido heterólogo se selecciona del grupo que consiste en marcadores de purificación, marcadores epitópicos, secuencias de dirección, péptidos señal, secuencias de translocación de membrana y modificadores de PK.

En ciertas realizaciones, la composición comprende un fragmento proteico de MetRS a una concentración de al menos aproximadamente 10 mg/ml. En ciertas realizaciones, la composición comprende un fragmento proteico de MetRS que es al menos 90 % monodisperso. En ciertas realizaciones la composición comprende menos de aproximadamente 3 % de proteínas agregadas de alto peso molecular. En ciertas realizaciones la composición muestra menos del 3 % de agregación cuando se almacena a una concentración de al menos aproximadamente 10 mg/ml en PBS durante una semana a 4 ºC. En ciertas realizaciones la composición muestra menos de 3 % de agregación cuando se almacena a una concentración de al menos 10 mg/ml en PBS durante una semana a temperatura ambiente.

Se describen en el presente documento diversos ensayos para medir dichas características de las resectinas y pueden usarse para definir aspectos de la invención. En ciertos aspectos, estas características serán preferibles para la utilidad bioterapéutica de los fragmentos proteicos de MetRS descritos en el presente documento.

Ciertas realizaciones incluyen composiciones que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado de al menos 50 aminoácidos que difieren de una secuencia de aminoácidos expuesta en SEQ ID NO: 137 por sustitución, deleción y/o adición de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 aminoácidos, en el que el fragmento proteico alterado sustancialmente conserva una actividad no canónica de la proteína ninalterada, o tiene un fenotipo negativo dominante en relación con la actividad no canónica, en el que el fragmento proteico tiene una solubilidad de al menos aproximadamente 5 mg/ml, y en el que la composición tiene una pureza de al menos aproximadamente 95 % en proteínas y menos de aproximadamente 10 UE/mg de proteínas de endotoxina. En realizaciones específicas, la composición está sustancialmente sin suero.

Otras realizaciones incluyen composiciones que comprenden un anticuerpo aislado que se une específicamente con un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado como se expone en SEQ ID NO: 137, en el que la afinidad del anticuerpo por el fragmento proteico de MetRS es aproximadamente 10 por mayor que su afinidad por un polipéptido de MetRS de longitud completa correspondiente. Uno de los aspectos sorprendentes de la presente invención incluye ciertas resectinas que poseen "nuevas" superficies accesibles a anticuerpo u otros productos biológicos dirigidos, mientras que la MetRS de longitud completa "oculta" o cubre estas superficies con otras secuencias o dominios adyacentes. El proceso de resecado también puede crear mayor accesibilidad acuosa para revelar actividades biológicas previamente no identificadas. Algunas realizaciones incluyen composiciones, que comprenden un anticuerpo aislado que se une específicamente con un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado como se expone en SEQ ID NO: 137, en el que el anticuerpo tiene una afinidad de al menos aproximadamente 10 nM por el fragmento proteico de MetRS, y una afinidad de al menos aproximadamente 100 nM por un polipéptido de MetRS de longitud completa correspondiente. En algunas realizaciones, el anticuerpo se une a un epítopo localizado dentro de un punto de unión de corte y empalme único de polipéptido de MetRS como se expone en cualquiera de SEQ ID NO: 137, o con una secuencia de aminoácidos en dirección C-terminal de este sitio de corte y empalme. En ciertas realizaciones, el anticuerpo antagoniza la actividad no canónica del fragmento proteico de MetRS. Dichos antagonistas pueden unirse opcionalmente con la MetRS de longitud completa o parental correspondiente.

Otros aspectos se refieren a sistemas de bioensayo, que comprenden un fragmento proteico de metionil ARNt

sintetasa (MetRS) sustancialmente puro de al menos 50 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137 y un compañero de unión que se une al fragmento proteico de MetRS. En un aspecto, el compañero de unión se selecciona del grupo que consiste en una proteína de receptor de superficie celular, ácido nucleico, membrana lipídica, proteína reguladora celular, enzima y factor de transcripción. Opcionalmente, dicho receptor puede ser parte de una célula, preferentemente una célula relevante para la biología revelada de la resectina.

Ciertas realizaciones incluyen composiciones celulares, que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado de al menos 50 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137, y una población modificada técnicamente de células en las que al menos una célula comprende un polinucleótido que codifica dicho fragmento proteico de AARS. En un aspecto, las células pueden crecer en un ambiente sin suero.

También se incluyen sistemas de detección, que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) sustancialmente puro de al menos 50 o 100 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137, una célula que comprende un receptor de superficie celular o una parte extracelular del mismo que se une al fragmento proteico, y una molécula de menos de aproximadamente 2.000 dalton, o un segundo polipéptido, que modula la unión o interacción entre el fragmento proteico de MetRS y el receptor extracelular.

Realizaciones particulares incluyen sistemas de diagnóstico, que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) sustancialmente puro de al menos 50 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137, y una célula que comprende un receptor de superficie celular o una parte extracelular del mismo que se une al fragmento proteico de MetRS, en el que el sistema o la célula comprende una molécula indicadora que permite la detección de un cambio en los niveles o actividad del receptor de superficie celular o parte extracelular del mismo.

Ciertas realizaciones incluyen dispositivos de crecimiento celular, que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado de al menos 50 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137, una población modificada técnicamente de células en la que al menos una célula comprende un polinucleótido que codifica dicho fragmento proteico de MetRS, al menos aproximadamente 10 litros de medio celular sin suero y un recipiente estéril. En realizaciones específicas, las células utilizadas para cualquiera de los métodos o composiciones descritos en este documento pueden crecer en medios sin suero, opcionalmente con un antibiótico y un inductor.

Algunas realizaciones se refieren a agentes antisentido o de interferencia de ARN (ARNi), que comprenden una secuencia que se dirige contra un punto de unión de corte y empalme único de una variante de corte y empalme de MetRS como se expone en SEQ ID NO: 137.

También se incluyen composiciones terapéuticas, que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado de al menos 35 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137, en el que el fragmento proteico se une específicamente con un compañero de unión y tiene una solubilidad de al menos aproximadamente 5 mg/ml, y en el que la composición tiene una pureza de al menos aproximadamente 95 % en proteínas. En algunos aspectos, la composición puede tener menos de 10 UE de endotoxina/mg de proteína

También se incluyen composiciones que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado de al menos 50 aminoácidos que es al menos 80 %, 85 %, 90 %, 95 %, 98 % o 100 % Idéntico a una secuencia de aminoácidos expuesta en SEQ ID NO: 137, en el que el fragmento proteico tiene una solubilidad de al menos aproximadamente 5 mg/ml, y en el que la composición tiene una pureza de al menos aproximadamente 95 % en proteínas y menos de 10 UE de endotoxina/mg de proteína. En cualquiera de estas realizaciones, las composiciones pueden comprender un fragmento proteico de MetRS que es al menos aproximadamente 50 %, aproximadamente 60 %, aproximadamente 70 %, aproximadamente 80 %, aproximadamente 90 % o aproximadamente 95 % monodisperso con respecto a su masa molecular aparente. En otro aspecto de cualquiera de estas realizaciones, las composiciones comprenden menos de aproximadamente 10 % (en proteínas), de proteínas agregadas de alto peso molecular, o menos de aproximadamente 5 % de proteínas agregadas de alto peso molecular, o menos de aproximadamente 4 % de proteínas agregadas de alto peso molecular, o menos de aproximadamente 3 % de proteínas agregadas de alto peso molecular, o menos de 2 % de proteínas agregadas de alto peso molecular, o menos de aproximadamente 1 % de proteínas agregadas de alto peso molecular.

En otro aspecto de cualquiera de estas realizaciones, las composiciones muestran menos de aproximadamente 10 % de agregación cuando se almacenan a una concentración de al menos 10 mg/ml en PBS durante una semana a 4 ºC, o menos de aproximadamente 5 % de agregación cuando se almacena a una concentración de al menos 10 mg/ml en PBS durante una semana a 4 ºC, o menos de aproximadamente 3 % de agregación cuando se almacena a una concentración de al menos 10 mg/ml en PBS durante una semana a 4 ºC, o menos de aproximadamente 2 % de agregación cuando se almacena a una concentración de al menos 10 mg/ml en PBS durante una semana a 4 ºC, o

menos de aproximadamente 1 % de agregación cuando se almacena a una concentración de al menos 10 mg/ml en PBS durante una semana a 4 ºC.

Ciertas realizaciones incluyen composiciones que comprenden un fragmento proteico de metionil ARNt sintetasa (MetRS) sustancialmente puro de al menos 50 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137, y al menos un resto unido de forma covalente o no covalente con el mismo. En algunas realizaciones, el resto es un marcador detectable. En algunas realizaciones, el resto es un polímero soluble en agua. En algunas realizaciones, el resto es PEG. En un aspecto de cualquiera de estas realizaciones, el resto se une al extremo N terminal del fragmento proteico. En un aspecto de cualquiera de estas realizaciones, el resto se une al extremo C terminal del fragmento proteico.

Las realizaciones particulares incluyen composiciones, que comprenden un sustrato sólido unido a un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado de al menos 50 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO: 137, o un fragmento biológicamente activo o variante del mismo, en el que el fragmento proteico tiene una solubilidad de al menos aproximadamente 5 mg/ml, y la composición tiene una pureza de al menos aproximadamente 95 % en proteínas.

También se incluyen composiciones, que comprenden un agente de unión que se une específicamente con un fragmento proteico de metionil ARNt sintetasa (MetRS) aislado como se expone en SEQ ID NO: 137, en el que el agente de unión tiene una afinidad de al menos aproximadamente 1 nM por el fragmento proteico. En un aspecto, el agente de unión se une a un epítopo localizado dentro de un punto de unión de corte y empalme único de polipéptido de MetRS como se expone en SEQ ID NO: 137, o con una secuencia de aminoácidos en dirección C-terminal de este sitio de corte y empalme. En algunas realizaciones, el agente de unión antagoniza una actividad no canónica del polipéptido de MetRS.

Ciertas realizaciones incluyen polipéptidos de metionil ARNt sintetasa (MetRS) aislados, que comprenden una secuencia de aminoácidos de un fragmento proteico de MetRS como se describe en el presente documento, una secuencia de aminoácidos codificada por un polinucleótido de MetRS como se describe en el presente documento, o una variante o fragmento del mismo. Ciertos polipéptidos de MetRS comprenden una secuencia de aminoácidos que es al menos 80 %, 85 %, 90 %, 95 %, 98 % o 100 % idéntica a una secuencia de referencia de MetRS como se desvela en SEQ ID NO: 137. Ciertos polipéptidos de MetRS consisten esencialmente en una secuencia de aminoácidos que es al menos 80 %, 85 %, 90 %, 95 %, 98 % o 100 % idéntica a una secuencia de referencia de MetRS como se desvela en SEQ ID NO: 137. En ciertas realizaciones, el polipéptido comprende una actividad biológica no canónica. En realizaciones específicas, la actividad biológica no canónica se selecciona de modulación de la señalización celular (por ejemplo, señalización extracelular), modulación de la proliferación celular, modulación de la migración celular, modulación de la diferenciación celular, modulación de la apoptosis o muerte celular, modulación de la angiogénesis, modulación de la unión celular, modulación del metabolismo celular, modulación de la captación celular, modulación de la transcripción génica, o secreción, modulación de la producción o actividad de citocinas, modulación de la actividad receptora de citocinas y modulación de la inflamación.

Otros aspectos incluyen anticuerpos y otros agentes de unión que muestran especificidad de unión por un polipéptido de MetRS aislado como se describe en el presente documento, un compañero de unión del polipéptido de MetRS, o el complejo de ambos. En algunas realizaciones, la afinidad del anticuerpo o agente de unión por el polipéptido de MetRS es aproximadamente 10 veces más fuerte que su afinidad por un polipéptido de MetRS de longitud completa correspondiente. En realizaciones específicas, el agente de unión se selecciona de un péptido, peptidomimético, una adnectina, un aptámero, y una molécula pequeña. En ciertas realizaciones, el anticuerpo o agente de unión antagoniza una actividad no canónica del polipéptido MetRS. En otras realizaciones, el anticuerpo o agente de unión es agonista de una actividad no canónica del polipéptido de MetRS.

Ciertas realizaciones incluyen polinucleótidos de metionil ARNt sintetasa (MetRS) aislados, que comprenden una secuencia de nucleótidos de un polinucleótido de MetRS como se describe en el presente documento, una secuencia de nucleótidos que codifica un fragmento proteico de MetRS como se describe en el presente documento, o una variante, un fragmento o un complemento de los mismos. Ciertos polinucleótidos de MetRS comprenden una secuencia de nucleótidos que es al menos 80 %, 85 %, 90 %, 95 %, 98 % o 100 % idéntica a un polinucleótido de referencia de MetRS, o un complemento del mismo, como se desvela en SEQ ID NO: 137. En algunas realizaciones, la secuencia de nucleótidos tiene codones optimizados para expresión bacteriana.

Polinucleótidos de MetRS específicos consisten esencialmente en una secuencia de nucleótidos que es al menos 80 %, 85 %, 90 %, 95 %, 98 % o 100 % idéntica a un polinucleótido de referencia de MetRS, o un complemento del mismo, como se desvela en SEQ ID NO: 137. Otros polinucleótidos de MetRS comprenden o consisten esencialmente en una secuencia de nucleótidos que hibrida específicamente con un polinucleótido de referencia de MetRS, como se desvela en SEQ ID NO: 137. En ciertas realizaciones, el polinucleótido se selecciona de un cebador, una sonda y un oligonucleótido antisentido. En las realizaciones específicas, el cebador, la sonda o el oligonucleótido antisentido se dirige a un punto de unión de corte y empalme específico o único y/o la secuencia 3' de este sitio de corte y empalme dentro de un polinucleótido de MetRS.
Ciertas realizaciones incluyen métodos para determinar la presencia o los niveles de un fragmento proteico de

MetRS en una muestra, que comprenden poner en contacto la muestra con uno o más agentes de unión que se unen específicamente con un fragmento proteico de MetRS como se describe en el presente documento, detectar la presencia o ausencia del agente de unión y determinar de este modo la presencia o los niveles del fragmento proteico de MetRS. Otras realizaciones incluyen métodos para determinar la presencia o los niveles de un fragmento proteico de MetRS en una muestra, que comprenden analizar la muestra con un detector que es capaz de identificar específicamente un fragmento proteico como se describe en el presente documento, y determinar de este modo la presencia o los niveles del fragmento proteico de MetRS. En realizaciones específicas, el detector es un espectrómetro de masas (EM), un citómetro de flujo, un dispositivo de captura de imágenes de proteínas, ensayos inmunoabsorbentes ligados a enzima (ELISA) o una micromatriz de proteínas. Ciertas realizaciones comprenden comparar la presencia o los niveles del fragmento proteico de MetRS con una muestra de control o un valor predeterminado. Ciertas realizaciones comprenden caracterizar el estado de la muestra para distinguirla del control. En realizaciones específicas, la muestra y el control comprenden una célula o un tejido, y el método comprende distinguir entre células o tejidos de diferentes especies, células de diferentes tejidos u órganos, células en diferentes estados del desarrollo celular, células en diferentes estados de diferenciación celular, células en diferentes estados fisiológicos o células sanas y enfermas. Por ejemplo, las resectinas seleccionadas pueden ser más abundantes en condiciones tales como tensión o lesión.

Ciertas realizaciones incluyen métodos de descubrimiento de, y composiciones relacionadas para, identificar un compuesto que se une específicamente con un polipéptido de metionil ARNt sintetasa (MetRS) como se describe en el presente documento, o uno o más de sus compañeros de unión celular, que comprenden a) combinar el polipéptido de MetRS o su compañero de unión celular o ambos con al menos un compuesto de ensayo en condiciones adecuadas, y b) detectar la unión del polipéptido de MetRS o su compañero de unión celular o ambos con el compuesto de ensayo, identificando de este modo un compuesto que se una específicamente con el polipéptido de MetRS o su compañero de unión celular o ambos. En ciertas realizaciones, el compuesto de ensayo es un polipéptido o péptido, un anticuerpo o fragmento de unión del mismo, un peptidomimético, o una molécula pequeña. En ciertas realizaciones, el compuesto de ensayo es agonista de una actividad biológica no canónica del polipéptido de MetRS o su compañero de unión celular. En otras realizaciones, el compuesto de ensayo antagoniza una actividad biológica no canónica del polipéptido de MetRS o su compañero de unión celular. Ciertas realizaciones incluyen un compuesto identificado por el método anterior, tal como un agonista (por ejemplo, molécula pequeña, péptido).

Ciertas realizaciones incluyen métodos para determinar la presencia o los niveles de una secuencia polinucleotídica de una variante de corte y empalme de MetRS en una muestra, que comprende poner en contacto la muestra con uno o más oligonucleótidos que hibridan específicamente con un polinucleótido de MetRS como se describe en el presente documento, detectar la presencia o ausencia de los oligonucleótidos en la muestra y determinar de este modo la presencia o los niveles de la secuencia polinucleotídica de la variante de corte y empalme de MetRS. Otras realizaciones incluyen métodos para determinar la presencia o los niveles de una secuencia polinucleotídica de una variante de corte y empalme de MetRS en una muestra, que comprenden poner en contacto la muestra con al menos dos oligonucleótidos que amplifican específicamente un polinucleótido de MetRS como se describe en el presente documento, realizar una reacción de amplificación, detectar la presencia o ausencia de un producto amplificado, y determinar de este modo la presencia o los niveles de la secuencia polinucleotídica de la variante de corte y empalme de MetRS. En realizaciones específicas, el oligonucleótido o los oligonucleótidos se hibridan específicamente con o amplifican específicamente un punto de unión de corte y empalme que es único de la variante de corte y empalme de MetRS. Ciertas realizaciones incluyen comparar la presencia o los niveles del fragmento de proteína de MetRS o variante de corte y empalme con una muestra de control o un valor predeterminado. Ciertas realizaciones incluyen caracterizar el estado de la muestra para distinguirlo del control. En realizaciones específicas, la muestra y el control comprenden una célula o un tejido, y el método comprende distinguir entre células o tejidos de diferentes especies, células de diferentes tejidos u órganos, células en diferentes estados del desarrollo celular, células en diferentes estados de diferenciación celular o células sanas y enfermas.

Algunas realizaciones incluyen composiciones farmacéuticas, que comprenden un polinucleótido de MetRS descrito en el presente documento, un polipéptido de MetRS descrito en el presente documento, un agente de unión como se describe en el presente documento, o un compuesto identificado por el método anterior o descrito en el presente documento, y un excipiente o vehículo farmacéuticamente aceptable.

Ciertas realizaciones incluyen métodos para modular una actividad celular de una célula, que comprenden poner en contacto la célula con un polinucleótido de MetRS descrito en el presente documento, un polipéptido de MetRS descrito en el presente documento, un agente de unión descrito en el presente documento, un compuesto del método anterior o descrito en el presente documento, o una composición farmacéutica descrita en el presente documento. En realizaciones específicas, la actividad celular se selecciona de proliferación celular, migración celular, diferenciación celular, apoptosis o muerte celular, señalización celular, angiogénesis, unión celular, captación celular, secreción celular, metabolismo, producción o actividad de citocinas, actividad receptora de citocinas, transcripción génica e inflamación. En un aspecto, la célula se selecciona del grupo que consiste en preadipocitos, médula ósea, neutrófilos, células sanguíneas, hepatocitos, astrocitos, células madre mesenquimales y células del músculo esquelético.
En ciertas realizaciones, la célula está en un sujeto. Ciertas realizaciones comprenden tratar el sujeto, en el que el

sujeto tiene una afección asociada con una enfermedad neoplásica, una enfermedad o afección del sistema inmunitario, una enfermedad infecciosa, una enfermedad metabólica, un trastorno inflamatorio, enfermedad neuronal/neurológica, una enfermedad muscular/cardiovascular, una enfermedad asociada con hematopoyesis aberrante, una enfermedad asociada con angiogénesis aberrante o una enfermedad asociada con supervivencia celular aberrante.

También se incluyen procesos para fabricar un compuesto farmacéutico, que comprenden: a) realizar un cribado *in vitro* de uno o más compuestos candidatos en presencia de un fragmento proteico de MetRS de al menos 50 aminoácidos que comprende una secuencia de aminoácidos como se expone en SEQ ID NO:137, para identificar un compuesto que se une específicamente con el fragmento proteico de MetRS; b) realizar un ensayo basado en células o bioquímico o de receptores con el compuesto identificado en la etapa a), para identificar un compuesto que modula una o más actividades no canónicas del fragmento proteico de MetRS; c) opcionalmente evaluar la relación de estructura-actividad (SAR) del compuesto identificado en la etapa b), para correlacionar su estructura con la modulación de la actividad no canónica, y opcionalmente derivatizar el compuesto para alterar su capacidad para modular la actividad no canónica; y d) producir suficientes cantidades del compuesto identificado en la etapa b), o el compuesto derivatizado en la etapa c), para su uso en seres humanos, fabricando de este modo el compuesto farmacéutico.

Otras realizaciones incluyen procesos para fabricar un compuesto farmacéutico, que comprenden: a) realizar un cribado *in vitro* de uno o más compuestos candidatos en presencia de un receptor de superficie celular o una parte extracelular del mismo que se une específicamente con un fragmento proteico de MetRS de SEQ ID NO: 137, para identificar un compuesto que se une específicamente con el receptor de superficie celular o parte extracelular del mismo; b) realizar un ensayo basado en células o bioquímico o de receptores con el compuesto identificado de la etapa a), para identificar un compuesto que modula una o más actividades no canónicas del fragmento proteico de MetRS; c) evaluar opcionalmente la relación de estructura-actividad (SAR) del compuesto identificado en la etapa b), para correlacionar su estructura con la modulación de la actividad no canónica, y opcionalmente derivatizar el compuesto para alterar su capacidad para modular la actividad no canónica; y d) producir suficientes cantidades del compuesto identificado en la etapa b), o el compuesto derivatizado en la etapa c), para su uso en seres humanos, fabricando de este modo el compuesto farmacéutico.

Algunas realizaciones incluyen una composición celular, que comprende una población modificada técnicamente de células en las que al menos una célula comprende un polinucleótido que codifica una proteína de metionil ARNt sintetasa (MetRS) de longitud completa heteróloga, en la que las células son capaces de crecer en un medio sin suero. En un aspecto, la proteína de metionil ARNt sintetasa (MetRS) de longitud completa comprende una purificación heteróloga o un marcador epitópico para facilitar la purificación de un fragmento proteico de MetRS. En otro aspecto, la proteína de metionil ARNt sintetasa (MetRS) de longitud completa comprende un sitio de proteólisis heterólogo para permitir la producción del fragmento proteico de MetRS tras su escisión.

Algunas realizaciones incluyen un método para producir un polipéptido de MetRS como se expone en SEQ ID NO: 137 in situ dentro de una célula, que comprende: i) expresar una proteína de metionil ARNt sintetasa (MetRS) de longitud completa heteróloga dentro de la célula, en la que la célula comprende una proteasa capaz de escindir la proteína de metionil ARNt sintetasa (MetRS) de longitud completa heteróloga para producir el polipéptido de MetRS.

Algunas realizaciones incluyen un método para producir un polipéptido de MetRS como se expone en SEQ ID NO: 137 que comprende poner en contacto una proteína de metionil ARNt sintetasa (MetRS) de longitud completa aislada con una proteasa que es capaz de escindir la proteína de la metionil ARNt sintetasa (MetRS) de longitud completa y producir un polipéptido de MetRS.

Algunas realizaciones incluyen una proteína de metionil ARNt sintetasa (MetRS) de longitud completa modificada técnicamente que comprende un sitio de proteólisis heterólogo para permitir la generación proteolítica de un fragmento proteico de MetRS como se expone en SEQ ID NO: 137.

Algunas realizaciones incluyen una composición, que comprende una proteína de metionil ARNt sintetasa de longitud completa aislada, en la que la composición tiene una pureza de al menos aproximadamente 95 % en proteínas, menos de aproximadamente 10 UE de endotoxina/mg de proteína, y está sustancialmente sin suero. En un aspecto, la proteína de metionil ARNt sintetasa de longitud completa está presente a una concentración de al menos 10 mg/ml y está al menos 90 % monodispersa.

Una realización adicional incluye un método para tratar una enfermedad o un trastorno mediado por la desregulación de la expresión, actividad o localización espaciotemporal de una ARNt sintetasa mediante la administración de un fragmento proteico de MetRS, o un ácido nucleico que codifica el fragmento proteico de ARRS, como se expone en SEQ ID NO: 137. En un aspecto de esta realización, la enfermedad se selecciona de cáncer, neuropatía, diabetes y trastornos inflamatorios.

Descripción detallada de la invención

TABLA DE CONTENIDOS

I. VISIÓN DE CONJUNTO 15
II. DEFINICIONES 16
III. FRAGMENTOS PROTEICOS DE METRS PURIFICADOS Y VARIANTES 28
IV. POLINUCLEÓTIDOS DE METRS 97
V. ANTICUERPOS 110
VI. ALTERNATIVAS A ANTICUERPOS Y OTROS AGENTES DE UNIÓN 115
VII. BIOENSAYOS Y ENSAYOS ANALÍTICOS 119
VIII. SISTEMAS DE EXPRESIÓN Y PURIFICACIÓN 122
IX. MÉTODOS Y COMPOSICIONES DE DIAGNÓSTICO 135
X. AGENTES ANTISENTIDO Y DE ARNI 151
A. AGENTES ANTISENTIDO 152
B. AGENTES DE INTERFERENCIA DE ARN 160
XI. DESCUBRIMIENTO DE FÁRMACOS 168
XII. MÉTODOS DE USO 176
XIII. FORMULACIONES FARMACÉUTICAS, ADMINISTRACIÓN Y KITS 180
XIV. EJEMPLOS 189

I. VISIÓN DE CONJUNTO

La invención se define por las reivindicaciones adjuntas.

La presente invención se dirige, al menos en parte, al descubrimiento de nuevos polipéptidos de MetRS, y métodos para su preparación y uso, que representan la transformación de proteínas de tipo silvestre nativas en nuevas formas que muestran características notablemente diferentes en comparación con los genes de metionil ARNt sintetasa de longitud completa de origen natural. Dichos polipéptidos de MetRS se identificaron basándose en secuencia extensiva y análisis de espectro de masas de metionil ARNt sintetasa expresada en diferentes tejidos, seguido de la producción sistemática y el ensayo de cada polipéptido de MetRS potencial para identificar secuencias proteicas que representan dominios proteicos estables y solubles que muestran nuevas actividades biológicas, y características farmacológicas terapéuticas favorables.

Basándose en este análisis, se han identificado al menos dos familias nuevas de polipéptidos de MetRS derivados de metionil ARNt sintetasa.

En un aspecto, dichos polipéptidos de MetRS derivados de metionil ARN sintetasa comprenden secuencias polipeptídicas que comprenden aproximadamente los aminoácidos 1-197 de metionil ARNt sintetasa.

En un segundo aspecto, dichos polipéptidos de MetRS derivados de metionil ARNt sintetasa comprenden secuencias polipeptídicas que comprenden aproximadamente los aminoácidos 846-900 de metionil ARNt sintetasa.

Estas nuevas familias de polipéptidos de MetRS representan nuevos productos proteicos previamente desconocidos que muestran entre otros i) nueva actividad biológica, ii) características de estabilidad y agregación de proteínas favorables, y iii) la capacidad de expresarse y producirse a alto nivel en sistemas de expresión procariotas, que son características materialmente diferentes no halladas en la proteína de tipo silvestre intacta.

II. DEFINICIONES

A no ser que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que se entiende habitualmente por los expertos habituales en la materia a la que pertenece la invención. Aunque puede usarse cualquier método y material similar o equivalente a los descritos en el presente documento en la práctica o ensayo de la presente invención, se describen métodos y materiales preferidos. Para los fines de la presente invención, los siguientes términos se definen posteriormente.

Los artículos “un” y “una” se usan en el presente documento para hacer referencia a uno o más de uno (es decir, a al menos uno) del objeto gramatical del artículo. Como ejemplo, “un elemento” significa un elemento o más de un elemento.

Por “aproximadamente” se entiende una cantidad, un nivel, un valor, un número, una frecuencia, un porcentaje, una dimensión, un tamaño, una cantidad, un peso o una longitud que varía en hasta 30, 25, 20, 25, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 % con respecto a una cantidad, un nivel, un valor, un número, una frecuencia, un porcentaje, una dimensión, un tamaño, una cantidad, un peso o una longitud de referencia.

Un “agonista” se refiere a una molécula que intensifica o imita una actividad, por ejemplo, una actividad biológica no canónica de una MetRS u otra proteína. Los agonistas pueden incluir proteínas, ácidos nucleicos, carbohidratos,

moléculas pequeñas o cualquier otro compuesto o composición que module la actividad de una MetRS bien interaccionando directamente con la MetRS o con su compañero de unión, o bien actuando sobre componentes de la ruta biológica en la que participa la MetRS. Se incluyen agonistas parciales y completos.

Como se usa en el presente documento, se entiende que el término “aminoácido” significa aminoácidos tanto de origen natural como de origen no natural así como análogos y miméticos de aminoácidos. Los aminoácidos de origen natural incluyen los 20 (L)-aminoácidos utilizados durante la biosíntesis de proteínas así como otros tales como 4-hidroxiprolina, hidroxilisina, desmosina, isodesmosina, homocisteína, citrulina y ornitina, por ejemplo. Los aminoácidos de origen no natural incluyen, por ejemplo, (D)-aminoácidos, norleucina, norvalina, p-fluorofenilalanina, etionina y similares, que conocen los expertos en la materia. Los análogos de aminoácidos incluyen formas modificadas de aminoácidos de origen natural y no natural. Dichas modificaciones pueden incluir, por ejemplo, sustitución o reemplazo de grupos químicos y restos en el aminoácido o por derivatización del aminoácido. Los miméticos de aminoácidos incluyen, por ejemplo, estructuras orgánicas que muestran propiedades funcionalmente similares tales como carga y separación de cargas características del aminoácido de referencia. Por ejemplo, una estructura orgánica que imita a arginina (Arg o R) tendría un resto de carga positiva localizado en espacio molecular similar y que tiene el mismo grado de movilidad que el grupo e-amino de la cadena lateral del aminoácido Arg de origen natural. Los miméticos también incluyen estructuras restringidas para mantener la separación y las interacciones de carga óptimas del aminoácido o de los grupos funcionales de aminoácidos. Los expertos en la materia saben o pueden determinar qué estructuras constituyen análogos de aminoácidos y miméticos de aminoácidos funcionalmente equivalentes.

En ciertos aspectos, el uso de aminoácidos no naturales puede utilizarse para modificar (por ejemplo, aumentar) una actividad no canónica seleccionada de un fragmento proteico de MetRS, o para alterar la semivida *in vivo* o *in vitro* de la proteína. También pueden utilizarse aminoácidos no naturales para facilitar las modificaciones químicas (selectivas) (por ejemplo, pegilación) de una proteína MetRS. Por ejemplo, ciertos aminoácidos no naturales permiten la unión selectiva de polímeros tales como PEG con una proteína dada, y mejoran de este modo sus propiedades farmacocinéticas.

Pueden encontrarse ejemplos específicos de análogos y miméticos de aminoácidos descritos en, por ejemplo, Roberts y Vellaccio, The Peptides: Analysis, Synthesis, Biology, Eds. Gross y Meinhofer, Vol. 5, pág. 341, Academic Press, Inc., Nueva York, N. Y. (1983). Otros ejemplos incluyen aminoácidos peralquilados, particularmente aminoácidos permetilados. Véase, por ejemplo, Combinatorial Chemistry, Eds. Wilson y Czarnik, Ch. 11, pág. 235, John Wiley & Sons Inc., Nueva York, N. Y. (1997). Otros ejemplos más incluyen aminoácidos cuya parte de amida (y, por lo tanto, la cadena principal de amida del péptido resultante) se ha reemplazado, por ejemplo, por un anillo de azúcar, esteroide, benzodiazepina o carbociclo. Véase, por ejemplo, Burger’s Medicinal Chemistry and Drug Discovery, Ed. Manfred E. Wolff, C. 15, págs. 619-620, John Wiley & Sons Inc., Nueva York, N. Y. (1995). Se conocen bien en la técnica métodos para sintetizar péptidos, polipéptidos, peptidomiméticos y proteínas (véanse, por ejemplo, Patente de Estados Unidos n.º 5.420.109, M. Bodanzsky, Principles of Peptide Synthesis (1ª ed. & 2ª rev. ed.), Springer-Verlag, Nueva York, N. Y. (1984 y 1993), véase Capítulo 7; Stewart y Young, Solid Phase Peptide Synthesis, (2ª ed.), Pierce Chemical Co., Rockford, Ill. (1984)). En consecuencia, los polipéptidos de MetRS de la presente invención pueden estar compuestos de aminoácidos de origen natural y de origen no natural así como análogos y miméticos de aminoácidos.

El término “antagonista” se refiere a una molécula que reduce o atenúa una actividad, por ejemplo, una actividad biológica no canónica de una MetRS, u otra proteína. 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 MetRS o su compañero de unión, bien interaccionando directamente con la MetRS o su compañero de unión o bien actuando en componentes de la ruta biológica en la que participa la MetRS. Se incluyen antagonistas parciales y completos.

La expresión “aminoacil ARNt sintetasa” (AARS) se refiere en general a enzimas que en su forma natural o de tipo silvestre son capaces de catalizar la esterificación de un aminoácido específico o su precursor a uno de todos sus ARNt afines compatibles para formar un aminoacil-ARNt. En esta actividad “canónica”, las aminoacil ARNt sintetasas catalizan una reacción de dos etapas: en primer lugar, activan su aminoácido respectivo formando un aminoacil-adenilato, en el que el carboxilo del aminoácido se une al fosfato alfa de ATP desplazando pirofosfato, y después, cuando se une el ARNt correcto, el grupo aminoacilo del aminoacil-adenilato se transfiere al OH 2’ o 3’ terminal del ARNt.

Las aminoacil ARNt sintetasas de clase I normalmente tienen dos motivos de secuencia altamente conservados. Estas enzimas aminoacilan en el 2’-OH de un nucleótido de adenosina, y son habitualmente monoméricos o diméricos. Las aminoacil ARNt sintetasas de clase II normalmente tienen tres motivos de secuencia altamente conservados. Estas enzimas aminoacilan en el 3’-OH de la misma adenosina, y son habitualmente diméricas o tetraméricas. Los sitios activos de enzimas de clase II se componen principalmente de una lámina β antiparalela de siete cadenas flanqueada por hélices α. Aunque la fenilalanina ARNt sintetasa es de clase II, aminoacila en el 2’-OH.

Los polipéptidos de AARS incluyen fuentes de formas mitocondriales y citoplásmicas de tirosil ARNt sintetasa

(TyrRS), una triptofanil ARNt sintetasa (TrpRS), una glutaminil ARNt sintetasa (GlnRS), una glicil ARNt sintetasa (GlyRS), una histidil ARNt sintetasa (HisRS), una seril ARNt sintetasa (SerRS), una fenilalanil ARNt sintetasa (PheRS), una alanil ARNt sintetasa (AlaRS), una asparaginil ARNt sintetasa (AsnRS), un aspartil ARNt sintetasa (AspRS), una cisteinil ARNt sintetasa (CysRS), una glutamil ARNt sintetasa (GluRS), un prolil ARNt sintetasa (ProRS), una arginil ARNt sintetasa (ArgRS), un isoleucil ARNt sintetasa (IleRS), un leucil ARNt sintetasa (LeuRS), una lisil ARNt sintetasa (LysRS), una treonil ARNt sintetasa (ThrRS), un metionil ARNt sintetasa (MetRS) o una valil ARNt sintetasa (ValRS). Las secuencias de tipo silvestre o parentales de estos polipéptidos de AARS se conocen en la técnica.

Por “secuencia codificante” se entiende cualquier secuencia de ácido nucleico que contribuya al código para el producto polipeptídico de un gen. Por el contrario, la expresión “secuencia no codificante” se refiere a cualquier secuencia de ácido nucleico que no contribuya al código para el producto polipeptídico de un gen.

A lo largo de la presente memoria descriptiva, a no ser que el contexto requiera otra cosa, se entenderá que las palabras “comprender”, “comprende” y “comprendiendo” implican la inclusión de una etapa o elemento o grupo de etapas o elementos indicados pero no la exclusión de cualquier otra etapa o elemento o grupo de etapas o elementos.

Por “que consiste en” se entiende incluyendo, y limitado a, lo que esté a continuación de la frase “que consiste en”. Por lo tanto, la frase “que consiste en” indica que los elementos enumerados se requieren o son obligatorios, y que no pueden estar presentes otros elementos. Por “que consiste esencialmente en” se entiende que incluye cualquier elemento enumerado después de la frase, y limitado a otros elementos que no interfieren con o contribuyen a la actividad o acción especificada en la divulgación para los elementos enumerados. Por lo tanto, la frase “que consistente esencialmente en” indica que los elementos enumerados se requieren o son obligatorios, pero que otros elementos son opcionales y pueden estar o no presentes dependiendo de si afectan o no materialmente a la actividad o acción de los elementos enumerados.

La indicación “sin endotoxinas” o “sustancialmente sin endotoxinas” se refiere en general a composiciones, disolventes y/o vasos que contienen como máximo cantidades traza (por ejemplo, cantidades que no tienen efectos fisiológicos clínicamente adversos para un sujeto) de endotoxina, y preferentemente cantidades indetectables de endotoxina. Las endotoxinas son toxinas asociadas con ciertas bacterias, normalmente bacterias gram negativas, aunque pueden encontrarse endotoxinas en bacterias gram positivas, tales como *Listeria monocytogenes*. Las endotoxinas más prevalentes son lipopolisacáridos (LPS) o lipo-oligosacáridos (LOS) hallados en la membrana externa de diversas bacterias gram negativas, y que representan un elemento patógeno central en la capacidad de estas bacterias para provocar enfermedad. Cantidades pequeñas de endotoxina en seres humanos pueden producir fiebre, una reducción de la tensión arterial, y activación de inflamación y coagulación, entre otros efectos fisiológicos adversos.

Por lo tanto, en la producción farmacéutica de polipéptidos de MetRS, es con frecuencia deseable retirar la mayoría de o todas las trazas de endotoxina de productos farmacológicos y/o recipientes farmacológicos, porque incluso cantidades pequeñas pueden provocar efectos adversos en seres humanos. Puede usarse un horno de despirogenización para este fin, ya que se requieren normalmente temperaturas de más de 300 ºC para degradar la mayoría de endotoxinas. Por ejemplo, basándose en el material de envasado primario tal como jeringas o viales, la combinación de una temperatura de vidrio de 250 ºC y un tiempo de mantenimiento de 30 minutos es con frecuencia suficiente para conseguir una reducción 3 log de los niveles de endotoxina. Se contemplan otros métodos para retirar endotoxinas, incluyendo, por ejemplo, métodos de cromatografía y filtración, como se describe en el presente documento y se conoce en la técnica. También se incluyen métodos para producir polipéptidos de MetRS en y aislarlos de células eucariotas tales como células de mamífero para reducir, si no eliminar, el riesgo de que estén presentes endotoxinas en una composición de la invención. Se prefieren métodos para producir polipéptidos de MetRS en y aislarlos de células sin suero. Dichas composiciones que comprenden polipéptidos de MetRS representan nuevas formulaciones que muestran características biológicas y terapéuticas nuevas y novedosas no halladas en composiciones polipeptídicas de MetRS contaminadas con suero o endotoxina que tienen el potencial de unirse a y alterar las propiedades biológicas novedosas de los polipéptidos de MetRS.

Pueden detectarse endotoxinas usando técnicas rutinarias conocidas en este campo. Por ejemplo, el ensayo de Lisado de Amebocitos de *Limulus*, que utiliza sangre del cangrejo herradura, es un ensayo muy sensible para detectar la presencia de endotoxina, y los reactivos, kits e instrumentación para la detección de endotoxinas basándose en este ensayo están disponibles en el mercado, por ejemplo del Grupo Lonza. En este ensayo, niveles muy bajos de LPS pueden provocar coagulación detectable del lisado de *Limulus* debido a una potente cascada enzimática que amplifica esta reacción. Las endotoxinas también pueden cuantificarse por ensayo inmunoabsorbente ligado a enzimas (ELISA). Para ser sustancialmente sin endotoxinas, los niveles de endotoxinas pueden ser menos de aproximadamente 0,001, 0,005, 0,01, 0,02, 0,03, 0,04, 0,05, 0,06, 0,08, 0,09, 0,1, 0,5, 1,0, 1,5, 2, 2,5, 3, 4, 5, 6, 7, 8, 9 o 10 UE/mg de proteína. normalmente, 1 ng de lipopolisacárido (LPS) corresponde a aproximadamente 1-10 UE.

En ciertas realizaciones, la “pureza” de cualquier agente dado (por ejemplo, fragmento proteico de MetRS) en una

composición puede definirse específicamente. Por ejemplo, ciertas composiciones pueden comprender un agente que es al menos 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 o 100 % puro, incluyendo todos los decimales entre medias, como se mide, por ejemplo, y en ningún modo limitante, mediante cromatografía líquida de alta presión (HPLC), una forma bien conocida de cromatografía en columna usada frecuentemente en bioquímica y química analítica para separar, identificar y cuantificar compuestos.

Como se usa en el presente documento, los términos “función” y “funcional” y similares se refieren a una función biológica, enzimática o terapéutica.

Por “gen” se entiende una unidad de herencia que puede ocupar un locus específico en un cromosoma y consiste en secuencias reguladoras de la transcripción y/o de la traducción y/o una región codificante y/o secuencias no traducidas (es decir, intrones, secuencias no traducidas 5’ y 3’).

“Homología” se refiere al porcentaje del número de aminoácidos que son idénticos o constituyen substituciones conservativas. La homología puede determinarse usando programas de comparación de secuencias tales como GAP (Deveraux *et al*., 1984, Nucleic Acids Research 12, 387-395). De esta manera podrían compararse secuencias de una longitud similar o sustancialmente diferente a las citadas en el presente documento por inserción de huecos en el alineamiento, determinándose dichos huecos, por ejemplo, por el algoritmo de comparación usado por GAP.

La expresión “célula hospedadora” incluye una célula individual o un cultivo celular que puede ser o ha sido un receptor de cualquier vector o vectores recombinantes, polinucleótido aislado o polipéptido de la invención. Las células hospedadoras incluyen descendencia de una única célula hospedadora, y la descendencia puede no ser necesariamente completamente idéntica (en morfología o en complemento de ADN total) a la célula parental original debido a mutación y/o cambio natural, accidental o deliberado. Una célula hospedadora incluye células transfectadas o infectadas *in vivo* o *in vitro* con un vector recombinante o un polinucleótido de la invención. Una célula hospedadora que comprende un vector recombinante de la invención es una célula hospedadora recombinante.

Por “aislado” se entiende material que está sustancialmente o esencialmente libre de componentes que normalmente lo acompañan en su estado nativo. Por ejemplo, un “polinucleótido aislado”, como se usa en el presente documento, incluye un polinucleótido que se ha purificado a partir de las secuencias que lo flanquean en su estado de origen natural, por ejemplo, un fragmento de ADN que se ha retirado de las secuencias que normalmente están adyacentes al fragmento. Como alternativa, un “péptido aislado” o un “polipéptido aislado” y similares, como se usa en el presente documento, incluye el aislamiento y/o la purificación *in vitro* de una molécula peptídica o polipeptídica de su ambiente celular natural, y de asociación con otros componentes de la célula; es decir, no se asocia significativamente con sustancias *in vivo*.

El término “ARNm” o denominado en ocasiones “transcritos de ARNm” como se usa en el presente documento, incluye, pero sin limitación, transcrito o transcritos de preARNm, intermedios de procesamiento de transcritos, ARNm maduros listos para traducción y transcritos del gen o los genes, o ácidos nucleicos derivados del transcrito o los transcritos de ARNm. El procesamiento de transcritos puede incluir corte y empalme, edición y degradación. Como se usa en el presente documento, un ácido nucleico derivado de un transcrito de ARNm se refiere a un ácido nucleico para cuya síntesis el transcrito de ARNm o una subsecuencia del mismo ha actuado en última instancia como un molde. Un ADNc inverso transcrito a partir de un ARNm, un ARN transcrito a partir de ese ADNc, un ADN amplificado del ADNc, un ARN transcrito del ADN amplificado, etc., derivan todos del transcrito de ARNm y la detección de dichos productos derivados es indicativa de la presencia y/o abundancia del transcrito original en una muestra. Por lo tanto, las muestras derivadas de ARNm incluyen, pero sin limitación, transcritos de ARNm del gen o los genes, ADNc inverso transcrito del ARNm, ARNc transcrito del ADNc, ADN amplificado de los genes, ARN transcrito de ADN amplificado y similares.

La actividad “no canónica” como se usa en el presente documento, se refiere en general a i) una nueva actividad poseída por un polipéptido de MetRS de la invención que no posee en ningún grado significativo la proteína parental de longitud completa nativa intacta, o ii) una actividad que poseía la proteína parental de longitud completa nativa intacta, en la que el polipéptido de MetRS muestra una actividad específica significativamente mayor (es decir al menos 20 % mayor) en comparación con la proteína parental de longitud completa nativa intacta o muestra la actividad en un nuevo contexto; por ejemplo aislando la actividad de otras actividades poseídas por la proteína parental de longitud completa nativa intacta. En el caso de polipéptidos de MetRS, los ejemplos no limitantes de actividades no canónicas incluyen señalización extracelular, unión a ARN, unión a aminoácidos, modulación de la proliferación celular, modulación de la migración celular, modulación de la diferenciación celular (por ejemplo, hematopoyesis, neurogénesis, miogénesis, osteogénesis y adipogénesis), modulación de la transcripción génica, modulación de la apoptosis u otras formas de muerte celular, modulación de la señalización celular, modulación de la captación celular o secreción, modulación de la angiogénesis, modulación de la unión celular, modulación del metabolismo celular, modulación de la producción o actividad de citocinas, modulación de la actividad receptora de citocinas, modulación de la inflamación y similares.

La expresión “concentración eficaz semimáxima” o “CE_{50}” se refiere a la concentración de un fragmento proteico de MetRS, anticuerpo u otro agente descrito en el presente documento a la que induce una respuesta a medio camino

entre la línea basal y el máximo después de algún tiempo de exposición especificado; la CE_{50} de una curva de respuesta a dosis graduada representa por lo tanto la concentración de un compuesto a la que se observa 50 % de su efecto máximo. En ciertas realizaciones, la CE_{50} de un agente proporcionado en el presente documento se indica en relación con una actividad “no canónica”, como se ha indicado anteriormente. CE_{50} también representa la concentración en plasma requerida para obtener el 50 % de un efecto máximo *in vivo*. De forma similar, la “CE_{90}” se refiere a la concentración de un agente o composición a la que se observa el 90 % de su efecto máximo. La “CE_{90}” puede calcularse a partir de la “CE_{50}” y la pendiente de Hill, o puede determinarse a partir de los datos directamente, usando conocimiento rutinario en la técnica. En algunas realizaciones, la CE_{50} de un fragmento proteico de MetRS, anticuerpo u otro agente es menor de aproximadamente 0,01, 0,05, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, 60, 70, 80, 90 o 100 nM. Preferentemente, la composición bioterapéutica tendrá un valor de CE_{50} de aproximadamente 1 nM o menos.

El término “modular” incluye “aumentar” o “estimular”, así como “reducir” o “disminuir”, normalmente en una cantidad estadísticamente significativa o una fisiológicamente significativa en comparación con un control. En consecuencia un “modulador” puede ser un agonista, un antagonista, o cualquier mezcla de los mismos dependiendo de las condiciones usadas. Una cantidad “aumentada” o “potenciada” es normalmente una cantidad “estadísticamente significativa”, y puede incluir un aumento que es 1,1, 1,2, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30 o más veces (por ejemplo, 500, 1000 veces) (incluyendo todos los números enteros y puntos decimales entre medias y por encima de 1, por ejemplo, 1,5, 1,6, 1,7, 1,8, etc.) la cantidad producida por ausencia de composición (la ausencia de un agente o compuesto) o una composición de control. Una cantidad “disminuida” o reducida es normalmente una cantidad “estadísticamente significativa”, y puede incluir una disminución de 1 %, 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 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 por ausencia de composición (la ausencia de un agente o compuesto) o una composición de control, incluyendo todos los números enteros entre medias. Como un ejemplo no limitante, un control en la comparación de actividades canónicas y no canónicas podría incluir el fragmento de interés de proteína MetRS en comparación con su MetRS de longitud completa correspondiente, o un fragmento de MetRS que tiene actividad canónica comparable a su MetRS de longitud completa correspondiente. Se describen en el presente documento otros ejemplos de cantidades “estadísticamente significativas”.

Por “obtenerse de” se entiende que una muestra tal como, por ejemplo, un extracto polinucleotídico o extracto polipeptídico se aísla de, o deriva de, una fuente particular del sujeto. Por ejemplo, el extracto puede obtenerse a partir de un tejido o un fluido biológico aislado directamente del sujeto. “Derivado” u “obtenido de” también puede referirse a la fuente de una secuencia polipeptídica o polinucleotídica. Por ejemplo, una secuencia de MetRS de la presente invención puede “derivar” de la información de secuencia de un fragmento proteolítico de MetRS o variante de corte y empalme de MetRS, o una parte del mismo, bien de origen natural o bien generado artificialmente, y puede por lo tanto comprender, consistir esencialmente en o consistir en esa secuencia.

Los términos “polipéptido” y “proteína” se usan indistintamente en el presente documento para hacer referencia a un polímero de restos de aminoácidos y a variantes y análogos sintéticos y de origen natural de los mismos. Por lo tanto, estos términos se aplican a polímeros de aminoácidos en los que uno o más restos de aminoácidos son aminoácidos de origen no natural sintéticos, tales como un análogo químico de un aminoácido de origen natural correspondiente, así como a polímeros de aminoácidos de origen natural y derivados químicos de origen natural de los mismos. Dichos derivados incluyen, por ejemplo, modificaciones postraduccionales y productos de degradación que incluyen variantes de piroglutamilo, isoaspartilo, proteolíticas, fosforiladas, glucosiladas, oxidadas, isomerizadas y desaminadas del fragmento de referencia de MetRS.

Las indicaciones “identidad de secuencia” o, por ejemplo, que comprende una “secuencia 50 % idéntica a”, como se usa en el presente documento, se refieren a la medida en que las secuencias son idénticas nucleótido a nucleótido o aminoácido a aminoácido sobre una ventana de comparación. Por lo tanto, un “porcentaje de identidad de secuencia” puede calcularse comparando dos secuencias alineadas de forma óptima sobre la ventana de comparación, determinando el número de posiciones en las que la base de ácido nucleico idéntica (por ejemplo A, T, C, G, I) o el resto de aminoácido idéntico (por ejemplo Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys y Met) aparece en ambas secuencias para producir el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación (es decir, el tamaño de la ventana) y multiplicando el resultado por 100 para producir el porcentaje de identidad de secuencia.

Los términos usados para describir relaciones de secuencia entre dos o más polinucleótidos o polipéptidos incluyen “secuencia de referencia”, “ventana de comparación”, “identidad de secuencia”, “porcentaje de identidad de secuencia” e “identidad sustancial”. Una “secuencia de referencia” es de al menos 12, pero frecuentemente de 15 a 18 y con frecuencia de al menos 25 unidades monoméricas, incluyendo restos de nucleótidos y aminoácidos, de longitud. Debido a que dos polinucleótidos pueden comprender cada uno (1) una secuencia (es decir, solamente una parte de la secuencia polinucleotídica completa) que es similar entre los dos polinucleótidos, y (2) una secuencia que es divergente entre los dos polinucleótidos, las comparaciones de secuencias entre dos (o más) polinucleótidos se realizan normalmente comparando secuencias de los dos polinucleótidos sobre una “ventana de comparación” para

identificar y comparar regiones locales de similitud de secuencia. Una “ventana de comparación” se refiere a un segmento conceptual de al menos 6 posiciones contiguas, habitualmente de aproximadamente 50 a aproximadamente 100, más habitualmente de aproximadamente 100 a aproximadamente 150, en el que se compara una secuencia con una secuencia de referencia del mismo número de posiciones contiguas después de alinearse óptimamente las dos secuencias. La ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) de aproximadamente 20 % o menos en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para alineamiento óptimo de las dos secuencias. El alineamiento óptimo de secuencias para alinear una ventana de comparación puede realizarse por implementaciones computarizadas de algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Paquete de Software de Wisconsin Genetics Versión 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, EE.UU.) o mediante inspección y puede seleccionarse el mejor alineamiento (es decir, que da como resultado el mayor porcentaje de homología sobre la ventana de comparación) generado por cualquiera de los diversos métodos. También puede hacerse referencia a la familia de programas BLAST como por ejemplo se desvela en Altschul *et al*., 1997, Nucl. Acids Res. 25: 3389. Puede encontrarse discusión detallada del análisis de secuencia en la Unidad 19.3 de Ausubel *et al*., “Current Protocols in Molecular Biology”, John Wiley & Sons Inc, 1994-1998, Capítulo 15.

Se realizan cálculos de similitud de secuencia o identidad de secuencia entre secuencias (los términos se usan indistintamente en el presente documento) de la siguiente manera. Para determinar el porcentaje de identidad de dos secuencias de aminoácidos, o de dos secuencias de ácido nucleico, las secuencias se alinean para fines de comparación óptima (por ejemplo, pueden introducirse huecos en una o ambas de una primera y una segunda secuencia de aminoácidos o ácido nucleico para alineamiento óptimo y las secuencias no homólogas pueden descartarse para fines de comparación). En ciertas realizaciones, la longitud de una secuencia de referencia alineada para fines de comparación es de al menos 30 %, preferentemente al menos 40 %, más preferentemente al menos 50 %, 60 %, e incluso más preferentemente al menos 70 %, 80 %, 90 %, 100 % de la longitud de la secuencia de referencia. Los restos de aminoácidos o nucleótidos en posiciones de aminoácidos o posiciones de nucleótidos correspondientes se comparan después. Cuando una posición en la primera secuencia está ocupada por el mismo resto de aminoácido o nucleótido que la posición correspondiente en la segunda secuencia, entonces las moléculas son idénticas en esa posición.

El porcentaje de identidad entre las dos secuencias está en función del número de posiciones idénticas compartido por las secuencias, teniendo en cuenta el número de huecos, y la longitud de cada hueco, que es necesario introducir para alineamiento óptimo de las dos secuencias.

La comparación de secuencias y determinación del porcentaje de identidad entre dos secuencias puede conseguirse usando un algoritmo matemático. En una realización preferida, el porcentaje de identidad entre dos secuencias de aminoácidos se determina usando el algoritmo de Needleman y Wunsch, (1970, J. Mol. Biol. 48: 444-453) que se ha incorporado en el programa GAP en el paquete de software GCG (disponible en http://www.gcg.com), usando una matriz Blossum 62 o una matriz PAM250, y un peso de hueco de 16, 14, 12, 10, 8, 6 o 4 y un peso de longitud de 1, 2, 3, 4, 5 o 6. En otra realización preferida más, el porcentaje de identidad entre dos secuencias de nucleótidos se determina usando el programa GAP en el paquete de software GCG (disponible en http://www.gcg.com), usando una matriz NWSgapdna.CMP y un peso de hueco de 40, 50, 60, 70 u 80 y un peso de longitud de 1, 2, 3, 4, 5 o 6. Un conjunto particularmente preferido de parámetros (y el que debería usarse a no ser que se especifique de otro modo) son una matriz de puntuación Blossum 62 con una penalización de hueco de 12, una penalización de extensión de hueco de 4 y una penalización de hueco de desplazamiento de fase de 5.

El porcentaje de identidad entre dos secuencias de aminoácidos o nucleótidos puede determinarse usando el algoritmo de E. Meyers y W. Miller (1989, Cabios, 4: 11-17) que se ha incorporado al programa ALIGN (versión 2.0), usando una tabla de restos de pesos PAM120, una penalización de longitud de hueco de 12 y una penalización de hueco de 4.

Las secuencias de ácido nucleico y proteínas descritas en el presente documento pueden usarse como una “secuencia de consulta” para realizar una búsqueda frente a bases de datos públicas para, por ejemplo, identificar otros miembros de la familia o secuencias relacionadas. Dichas búsquedas pueden realizarse usando los programas NBLAST y XBLAST (versión 2.0) de Altschul *et al*., (1990, J. Mol. Biol. 215: 403-10). Pueden realizarse búsquedas de nucleótidos de BLAST con el programa NBLAST, puntuación = 100, longitud de palabra = 12 para obtener secuencias de nucleótidos homólogas de moléculas de ácido nucleico de la invención. Pueden realizarse búsquedas de proteínas BLAST con el programa XBLAST, puntuación = 50, longitud de palabra = 3 para obtener secuencias de aminoácidos homólogas de moléculas proteicas de la invención. Para obtener alineamientos con huecos para fines de comparación, puede utilizarse BLAST con huecos como se describe en Altschul *et al*., (1997, Nucleic Acids Res, 25: 3389-3402). Cuando se utilizan programas BLAST y BLAST con huecos, pueden usarse los parámetros por defecto de los programas respectivos (por ejemplo, XBLAST y NBLAST).

El término “solubilidad” se refiere a la propiedad de un agente proporcionado en el presente documento para disolverse en un disolvente líquido y formar una solución homogénea. La solubilidad se expresa normalmente como una concentración, bien por masa de soluto por unidad de volumen de disolvente (g de soluto por kg de disolvente, g por dl (100 ml), mg/ml, etc.), molaridad, molalidad, fracción molar u otras descripciones similares de concentración.

La cantidad en equilibrio máxima de soluto que puede disolverse por cantidad de disolvente es la solubilidad de ese soluto en ese disolvente en las condiciones especificadas, incluyendo temperatura, presión, pH y la naturaleza del disolvente. En ciertas realizaciones, la solubilidad se mide a pH fisiológico. En ciertas realizaciones, la solubilidad se mide en agua o un tampón fisiológico tal como PBS. En ciertas realizaciones, la solubilidad se mide en un fluido biológico (disolvente) tal como sangre o suero. En ciertas realizaciones, la temperatura puede ser aproximadamente temperatura ambiente (por ejemplo, aproximadamente 20, 21, 22, 23, 24, 25 ºC) o aproximadamente temperatura corporal (37 ºC). En ciertas realizaciones, un agente tal como un fragmento proteico de MetRS tiene una solubilidad de al menos aproximadamente 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25 o 30 mg/ml a temperatura ambiente o a 37 ºC.

Un “punto de unión de corte y empalme” como se usa en el presente documento incluye la región en un transcrito de ARNm maduro o el polipéptido codificado en el que el extremo 3’ de un primer exón se une al extremo 5’ de un segundo exón. El tamaño de la región puede variar, y puede incluir 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, 100 o más (incluyendo todos los números enteros entre medias) restos de nucleótidos o aminoácidos en uno de los lados de los restos exactos donde el extremo 3’ de un exón se une al extremo 5’ de otro exón. Un “exón” se refiere a una secuencia de ácido nucleico que está representada en la forma madura de una molécula de ARN después de haberse retirado por corte y empalme en cis partes de un ARN precursor (intrones) o haberse ligado por corte y empalme en trans dos o más moléculas de ARN precursoras. La molécula de ARN madura puede ser un ARN mensajero o una forma funcional de un ARN no codificante tal como ARNr o ARNt. Dependiendo del contexto, un exón puede referirse a la secuencia en el ADN o su transcrito de ARN. Un “intrón” se refiere a una región de ácido nucleico no codificante dentro de un gen, que no se traduce a una proteína. Se transcriben secciones intrónicas no codificantes a ARNm precursor (preARNm) y algunos otros ARN (tales como ARN no codificantes largos) y posteriormente se retiran por corte y empalme durante el procesamiento a ARN maduro.

Una “variante de corte y empalme” se refiere a un ARNm maduro y su proteína codificada que se producen por corte y empalme alternativo, un proceso por el que los exones del ARN (un transcrito génico primario o preARNm) se reconectan de múltiples maneras durante el corte y empalme de ARN. Los ARNm diferentes resultantes pueden traducirse a diferentes isoformas proteicas, lo que permite que un único gen codifique múltiples proteínas.

Un “sujeto”, como se usa en el presente documento, incluye cualquier animal que muestre un síntoma, o esté en riesgo de mostrar un síntoma, que puede tratarse o diagnosticarse con un polinucleótido o polipéptido de MetRS de la invención. También se incluyen sujetos para los que es deseable perfilar niveles de polipéptidos y/o polinucleótidos de MetRS de la invención, para fines de diagnóstico u otros. 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 gato o perro). Se incluyen primates no humanos y, preferentemente, pacientes humanos.

“Tratamiento” o “tratar”, como se usa en el presente documento, incluye cualquier efecto deseable en los síntomas o la patología de una enfermedad o afección que puede efectuarse por las actividades no canónicas de un polinucleótido o polipéptido de MetRS, como se describe en el presente documento, y pueden incluir incluso cambios mínimos o mejoras en uno o más marcadores medibles de la enfermedad o afección que se trate. También se incluyen tratamientos que están relacionados con terapias no MetRS, en los que una secuencia de MetRS descrita en el presente documento proporciona un marcador clínico de tratamiento. “Tratamiento” o “tratar” no indican necesariamente la erradicación completa o cura de la enfermedad o afección, o síntomas asociados con la misma. El sujeto que recibe este tratamiento es cualquier sujeto que lo necesite. Los marcadores a modo de ejemplo de mejora clínica resultarán evidentes para los expertos en la materia.

La práctica de la presente invención empleará, a no ser que se indique específicamente lo contrario, métodos convencionales de biología molecular y técnicas de ADN recombinante dentro de la experiencia de la técnica, muchos de los cuales se describen posteriormente con el fin de ilustrar. Dichas técnicas se explican completamente en la bibliografía. Véase, por ejemplo, Sambrook, *et al*., Molecular Cloning: A Laboratory Manual (3ª Edición, 2000); DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.); Oligonucleotide Synthesis (N. Gait, ed., 1984); Oligonucleotide Synthesis: Methods and Applications (P. Herdewijn, ed., 2004); Nucleic Acid Hybridization (B. Hames y S. Higgins, eds., 1985); Nucleic Acid Hybridization: Modern Applications (Buzdin y Lukyanov, eds., 2009); Transcription and Translation (B. Hames y S. Higgins, eds., 1984); Animal Cell Culture (R. Freshney, ed., 1986); Freshney, R. I. (2005) Culture of Animal Cells, a Manual of Basic Techniques, 5ª Ed. Hoboken NJ, John Wiley & Sons; B. Perbal, A Practical Guide to Molecular Cloning (3ª Edición 2010); Farrell, R., RNA Methodologies: A Laboratory Guide for Isolation and Characterization (3ª Edición 2005), Methods of Enzymology: DNA Structure Part A: Synthesis and Physical Analysis of DNA Methods in Enzymology, Academic Press; Using Antibodies: A Laboratory Manual: Portable Protocol NO. I por Edward Harlow, David Lane, Ed Harlow (1999, Cold Spring Harbor Laboratory Press, ISBN 0-87969-544-7); Antibodies: A Laboratory Manual by Ed Harlow (Editor), David Lane (Editor) (1988, Cold Spring Harbor Laboratory Press, ISBN 0-87969-3,4-2), 1855. Handbook of Drug Screening, editado por Ramakrishna Seethala, Prabhavathi B. Fernandes (2001, Nueva York, N. Y., Marcel Dekker, ISBN 0-8247-0562-9); y Lab Ref: A Handbook of Recipes, Reagents, and Other Reference Tools for Use at the Bench, Edited Jane Roskams y Linda Rodgers, (2002, Cold Spring Harbor Laboratory, ISBN 0-87969-630-3).

III. FRAGMENTOS PROTEICOS DE METRS PURIFICADOS Y VARIANTES PARA PRODUCTOS TERAPÉUTICOS Y OTRAS APLICACIONES

Sorprendentemente, y a diferencia de sus secuencias parentales de longitud completa que se conocen solamente por sus actividades de aminoacilación, se ha descubierto que los fragmentos de MetRS poseen actividades biológicas importantes para aplicaciones bioterapéuticas, de descubrimiento y de diagnóstico. Las realizaciones de la presente invención incluyen por lo tanto proteínas de longitud completa, isoformas de proteínas maduras y fragmentos proteicos de metionil ARNt sintetasas (MetRS), además de variantes biológicamente activas y fragmentos de las mismas. En ciertas realizaciones, las proteínas y fragmentos pueden surgir mediante proteólisis endógena, proteólisis *in vitro*, variación de corte y empalme o predicción por ordenador, entre otros mecanismos.

Los fragmentos proteicos de MetRS descritos en el presente documento, y variantes de los mismos, pueden poseer al menos una actividad biológica “no canónica”. El fragmento o los fragmentos proteicos de MetRS de la presente invención también se indican en el presente documento como “polipéptidos de MetRS” o “polipéptidos de referencia de MetRS”. En ciertas realizaciones, los polipéptidos de MetRS proporcionados en el presente documento comprenden o consisten esencialmente en todas o una parte de las “secuencia o secuencias de referencia” de polipéptidos de MetRS como se expone en SEQ ID NO: 137, que representan la secuencia o las secuencias de aminoácidos de diversos fragmentos de metionil ARNt sintetasas. Las secuencias proteicas de MetRS de ratón y de seres humanos están altamente relacionadas, difiriendo normalmente en no más de unos pocos aminoácidos dentro de una secuencia completa, un dominio particular o un fragmento proteico particular.

Polipéptidos de MetRS N terminales: (Tablas 1, 2 y 3)

Tabla 1A
Polipéptidos de MetRS identificados por EM

Nombre	Tipo / especie / Restos	Secuencias de Aminoácidos y Ácido Nucleico	SEQ. ID. NO.
MetRS^{N1}	Proteína / Humana / 1-231	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGNY LFSTSAICRYFFLLSGWEQDDLTNQWLEW EATELQPALSAALYYLVVQGKKGEDVLGS VRRALTHIDHSLSRQNCPFLAGETESLADI VLWGALYPLLQDPAYLPEELSALHSWFQT LSTQEPCQRAAETVLKQQGVLALRPYLQK QPQPSPAEGRAVTNEPEEEELATLSEE	SEQ. ID. NO.12
MetRS^{N1}	ADN / Humana /	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGAT TGTGTGGTCCCGTTCCTGACCCGGCCTAA GGTCCCTGTCTTGCAGCTGGATAGCGGC AACTACCTCTTCTCCACTAGTGCAATCTG CCGATATTTTTTTTGTTATCTGGCTGGG AGCAAGATGACCTCACTAACCAGTGGCT GGAATGGGAAGCGACAGAGCTGCAGCC AGCTTTGTCTGCTGCCCTGTACTATTTAG TGGTCCAAGGCAAGAAGGGGGAAGATG TTCTTGGTTCAGTGCGGAGAGCCCTGACT CACATTGACCACAGCTTGAGTCGTCAGA ACTGTCCTTTCCTGGCTGGGGAGACAGA ATCTCTAGCCGACATTGTTTTGTGGGGAG	SEQ. ID. NO.13

		CCCTATACCCATTACTGCAAGATCCCGCC TACCTCCCTGAGGAGCTGAGTGCCCTGC ACAGCTGGTTCCAGACACTGAGTACCCA GGAACCATGTCAGCGAGCTGCAGAGACT GTACTGAAACAGCAAGGTGTCCTGGCTC TCCGGCCTTACCTCCAAAAGCAGCCCCA GCCCAGCCCCGCTGAGGGAAGGGCTGTC ACCAATGAGCCTGAGGAGGAGGAGCTG GCTACCCTATCTGAGGAG	

Tabla 1B
Péptidos de enlace inferidos y péptidos detectados por espectrometría de masas de MetRS^{N1}

Tipo / especie	Secuencia	SEQ. ID. NO.
Proteína / ratón	LFVSEGSPGSLPVLAAAAR	SEQ. ID. NO.14
Proteína / ratón	GRAELLISTVGPEECVVPFLTR	SEQ. ID. NO.15
Proteína / ratón	VPVLQLDSGNYLFSASAICR	SEQ. ID. NO.16
Proteína / ratón	YFFLLCGWEQDDLTNQWLEWEATELQPVLSAALHCL VVQGKKGEDILGPLRRVLTHIDHSLSRQNCPFLAGDTE SLADIVLWGALYPLLQDPAYLPEELGALQSWFQTLST QEPCQRAAETVLKQQGVLALRLYLQK	SEQ. ID. NO.17
Proteína / ratón	QPQPQPPPPEGR	SEQ. ID. NO.18

Tabla 1C
Secuencias concatenadas de MetRS^{N1} basadas en péptidos detectados por espectrometría de masas

Tipo / especie	Secuencia	SEQ. ID. NO.
Proteína / ratón	LFVSEGSPGSLPVLAAAARARGRAELLISTVGPEECVV PFLTRPKVPVLQLDSGNYLFSASAICRYFFLLCGWEQD DLTNQWLEWEATELQPVLSAALHCLVVQGKKGEDIL	SEQ. ID. NO.19
	GPLRRVLTHIDHSLSRQNCPFLAGDTESLADIVLWGAL YPLLQDPAYLPEELGALQSWFQTLSTQEPCQRAAETV LKQQGVLALRLYLQKQPQPQPPPPEGR	

Tabla 2
Polipéptidos de MetRS y transcritos alternativos identificados por Secuenciación Profunda

Nombre	Tipo / especie / Restos	Secuencias de Aminoácidos y Ácido Nucleico	SEQ. ID. NO.
MetRS^{N3}	Proteína / Humana / 1-67 + 48 aa	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYSMSGLMPLLAICPSQPTTQ TSGRDGGRTQSKWTCISSWPKTMFLSIA	SEQ. ID. NO.20

MetRS^{N3}	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGGTATTCTATGTCTGGTTTGATG CCACTATTGGCTATCTGTCCATCACAGC CAACTACACAGACCAGTGGGAGAGATG GTGGAAGAACCCAGAGCAAGTGGACCT GTATCAGTTCATGGCCAAAGACAATGTT CCTTTCCATAGCTTAG	SEQ. ID. NO.21
		CAGAGCGAGCTTATGCGACTGCCCAGT GGCAAAGTGGGCTTCAGCAAGGCCATG TCCAACAAGTGGATTCGGGTGGACAAG AGTGCGGCTGACGGGCCCCGGGTGTTC CGAGTGGACTCTGAAGACCTACTGGGT GAGCAAAGGCAGTGCCTTTAG	
MetRS^{N4}	Proteína / Humana / 1-364 + 63 aa	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GERNVLITSALPYVNNVPHLGNIIGCVLSA DVFARYSRLRQWNTLYLCGTDEYGTATE TKALEEGLTPQEICDKYHIIHADIYRWFNI SFDIFGRTTTPQQTKSLSVKSADHALWCS RASTCFWTCLSWRSDWRSGWGGHCLAV	SEQ. ID. NO.22
		TGHPMPSLSPVLGFGMASSHAA	

$MetRS^{N4}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAAAGGAATGTGCTC ATCACCAGTGCCCTCCCTTACGTCAACA ATGTCCCCCACCTTGGGAACATCATTGG TTGTGTGCTCAGTGCCGATGTCTTTGCC AGGTACTCTCGCCTCCGCCAGTGGAACA CCCTCTATCTGTGTGGGACAGATGAGTA TGGTACAGCAACAGAGACCAAGGCTCT GGAGGAGGGACTAACCCCCCAGGAGAT CTGCGACAAGTACCACATCATCCATGCT GACATCTACCGCTGGTTTAACATTTCGT TTGATATTTTTGGTCGCACCACCACTCC ACAGCAGACCAAAAGCCTCAGTGTAAA GTCTGCCGATCATGCCCTGTGGTGCAGT	SEQ. ID. NO.23
		CGAGCCAGCACCTGTTTCTGGACCTGCC TAAGCTGGAGAAGCGACTGGAGGAGTG GTTGGGGAGGACATTGCCTGGCAGTGA CTGGACACCCAATGCCCAGTTTATCACC CGTTCTTGGCTTCGGGATGGCCTCAAGC CACGCTGCATAA	

MetRS^{N5}	Proteína / Humana / 1-513 + 546-900	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GERNVLITSALPYVNNVPHLGNIIGCVLSA DVFARYSRLRQWNTLYLCGTDEYGTATE TKALEEGLTPQEICDKYHIIHADIYRWFNI SFDIFGRTTTPQQTKITQDIFQQLLKRGFV LQDTVEQLRCEHCARFLADRFVEGVCPFC GYEEARGDQCDKCGKLINAVELKKPQCK VCRSCPVVQSSQHLFLDLPKLEKRLEEWL GRTLPGSDWTPNAQFITRSWLRDGLKPRC ITRDLKWGTPVPLEGFEDKVDLYQFMAK DNVPFHSLVFPCSALGAEDNYTLVSHLIA TEYLNYEDGKFSKSRGVGVFGDMAQDTG IPADIWRFYLLYIRPEGQDSAFSWTDLLLK NNSELLNNLGNFINRAGMFVSKFFGGYVP EMVLTPDDQRLLAHVTLELQHYHQLLEK VRIRDALRSILTISRHGNQYIQVNEPWKRI KGSEADRQRAGTVTGLAVNIAALLSVML QPYMPTVSATIQAQLQLPPPACSILLTNFL CTLPAGHQIGTVSPLFQKLENDQIESLRQR FGGGQAKTSPKPAVVETVTTAKPQQIQAL MDEVTKQGNIVRELKAQKADKNEVAAE VAKLLDLKKQLAVAEGKPPEAPKGKKKK	SEQ. ID. NO.24

$MetRS^{N5}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAAAGGAATGTGCTC ATCACCAGTGCCCTCCCTTACGTCAACA ATGTCCCCCACCTTGGGAACATCATTGG TTGTGTGCTCAGTGCCGATGTCTTTGCC AGGTACTCTCGCCTCCGCCAGTGGAACA CCCTCTATCTGTGTGGGACAGATGAGTA TGGTACAGCAACAGAGACCAAGGCTCT GGAGGAGGGACTAACCCCCCAGGAGAT CTGCGACAAGTACCACATCATCCATGCT GACATCTACCGCTGGTTTAACATTTCGT TTGATATTTTTGGTCGCACCACCACTCC ACAGCAGACCAAAATCACCCAGGACAT TTTCCAGCAGTTGCTGAAACGAGGTTTT GTGCTGCAAGATACTGTGGAGCAACTG CGATGTGAGCACTGTGCTCGCTTCCTGG CTGACCGCTTCGTGGAGGGCGTGTGTCC CTTCTGTGGCTATGAGGAGGCTCGGGGT GACCAGTGTGACAAGTGTGGCAAGCTC ATCAATGCTGTCGAGCTTAAGAAGCCTC AGTGTAAAGTCTGCCGATCATGCCCTGT	SEQ. ID. NO.25

		GGTGCAGTCGAGCCAGCACCTGTTTCTG GACCTGCCTAAGCTGGAGAAGCGACTG GAGGAGTGGTTGGGGAGGACATTGCCT GGCAGTGACTGGACACCCAATGCCCAG TTTATCACCCGTTCTTGGCTTCGGGATG GCCTCAAGCCACGCTGCATAACCCGAG ACCTCAAATGGGGAACCCCTGTACCCTT AGAAGGTTTTGAAGACAAGGTGGACCT GTATCAGTTCATGGCCAAAGACAATGTT CCTTTCCATAGCTTAGTCTTTCCTTGCTC AGCCCTAGGAGCTGAGGATAACTATAC CTTGGTCAGCCACCTCATTGCTACAGAG TACCTGAACTATGAGGATGGGAAATTCT CTAAGAGCCGCGGTGTGGGAGTGTTTG GGGACATGGCCCAGGACACGGGGATCC CTGCTGACATCTGGCGCTTCTATCTGCT GTACATTCGGCCTGAGGGCCAGGACAG TGCTTTCTCCTGGACGGACCTGCTGCTG AAGAATAATTCTGAGCTGCTTAACAACC TGGGCAACTTCATCAACAGAGCTGGGA TGTTTGTGTCTAAGTTCTTTGGGGGCTA TGTGCCTGAGATGGTGCTCACCCCTGAT GATCAGCGCCTGCTGGCCCATGTCACCC TGGAGCTCCAGCACTATCACCAGCTACT TGAGAAGGTTCGGATCCGGGATGCCTTG CGCAGTATCCTCACCATATCTCGACATG GCAACCAATATATTCAGGTGAATGAGC CCTGGAAGCGGATTAAAGGCAGTGAGG CTGACAGGCAACGGGCAGGAACAGTGA CTGGCTTGGCAGTGAATATAGCTGCCTT GCTCTCTGTCATGCTTCAGCCTTACATG CCCACGGTTAGTGCCACAATCCAGGCCC AGCTGCAGCTCCCACCTCCAGCCTGCAG TATCCTGCTGACAAACTTCCTGTGTACC TTACCAGCAGGACACCAGATTGGCACA GTCAGTCCCTTGTTCCAAAAATTGGAAA ATGACCAGATTGAAAGTTTAAGGCAGC GCTTTGGAGGGGGCCAGGCAAAAACGT CCCCGAAGCCAGCAGTTGTAGAGACTG TTACAACAGCCAAGCCACAGCAGATAC AAGCGCTGATGGATGAAGTGACAAAAC AAGGAAACATTGTCCGAGAACTGAAAG	
		CACAAAAGGCAGACAAGAACGAGGTTG CTGCGGAGGTGGCGAAACTCTTGGATCT AAAGAAACAGTTGGCTGTAGCTGAGGG GAAACCCCCTGAAGCCCCTAAAGGCAA GAAGAAAAAGTAA	

$MetRS^{N6}$	Proteína / Humana / 1-545 + 2 aa	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GERNVLITSALPYVNNVPHLGNIIGCVLSA DVFARYSRLRQWNTLYLCGTDEYGTATE TKALEEGLTPQEICDKYHIIHADIYRWFNI SFDIFGRTTTPQQTKITQDIFQQLLKRGFV LQDTVEQLRCEHCARFLADRFVEGVCPFC GYEEARGDQCDKCGKLINAVELKKPQCK VCRSCPVVQSSQHLFLDLPKLEKRLEEWL GRTLPGSDWTPNAQFITRSWLRDGLKPRC ITRDLKWGTPVPLEGFEDKVFYVWFDATI GYLSITANYTDQWERWWKNPEQST	SEQ. ID. NO.26
$MetRS^{N6}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG	SEQ. ID. NO.27

		TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAAAGGAATGTGCTC ATCACCAGTGCCCTCCCTTACGTCAACA ATGTCCCCCACCTTGGGAACATCATTGG TTGTGTGCTCAGTGCCGATGTCTTTGCC AGGTACTCTCGCCTCCGCCAGTGGAACA CCCTCTATCTGTGTGGGACAGATGAGTA TGGTACAGCAACAGAGACCAAGGCTCT GGAGGAGGGACTAACCCCCCAGGAGAT CTGCGACAAGTACCACATCATCCATGCT GACATCTACCGCTGGTTTAACATTTCGT TTGATATTTTTGGTCGCACCACCACTCC ACAGCAGACCAAAATCACCCAGGACAT TTTCCAGCAGTTGCTGAAACGAGGTTTT GTGCTGCAAGATACTGTGGAGCAACTG CGATGTGAGCACTGTGCTCGCTTCCTGG CTGACCGCTTCGTGGAGGGCGTGTGTCC CTTCTGTGGCTATGAGGAGGCTCGGGGT GACCAGTGTGACAAGTGTGGCAAGCTC ATCAATGCTGTCGAGCTTAAGAAGCCTC AGTGTAAAGTCTGCCGATCATGCCCTGT GGTGCAGTCGAGCCAGCACCTGTTTCTG GACCTGCCTAAGCTGGAGAAGCGACTG GAGGAGTGGTTGGGGAGGACATTGCCT GGCAGTGACTGGACACCCAATGCCCAG TTTATCACCCGTTCTTGGCTTCGGGATG GCCTCAAGCCACGCTGCATAACCCGAG ACCTCAAATGGGGAACCCCTGTACCCTT AGAAGGTTTTGAAGACAAGGTATTCTAT GTCTGGTTTGATGCCACTATTGGCTATC TGTCCATCACAGCCAACTACACAGACCA GTGGGAGAGATGGTGGAAGAACCCAGA GCAAAGTACCTGA	

$MetRS^{N7}$	Proteína / Humana / 1-735 + 15 aa	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GERNVLITSALPYVNNVPHLGNIIGCVLSA DVFARYSRLRQWNTLYLCGTDEYGTATE TKALEEGLTPQEICDKYHIIHADIYRWFNI SFDIFGRTTTPQQTKITQDIFQQLLKRGFV LQDTVEQLRCEHCARFLADRFVEGVCPFC GYEEARGDQCDKCGKLINAVELKKPQCK VCRSCPVVQSSQHLFLDLPKLEKRLEEWL GRTLPGSDWTPNAQFITRSWLRDGLKPRC ITRDLKWGTPVPLEGFEDKVFYVWFDATI GYLSITANYTDQWERWWKNPEQVDLYQ FMAKDNVPFHSLVFPCSALGAEDNYTLVS HLIATEYLNYEDGKFSKSRGVGVFGDMA QDTGIPADIWRFYLLYIRPEGQDSAFSWT DLLLKNNSELLNNLGNFINRAGMFVSKFF GGYVPEMVLTPDDQRLLAHVTLELQHYH QLLEKVRIRDALRSILTISRHGNQYIQVNE PWKRIKGSEADRSVPCSKNWKMTRLKV	SEQ. ID. NO.28
$MetRS^{N7}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG	SEQ. ID. NO.29

		TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAAAGGAATGTGCTC ATCACCAGTGCCCTCCCTTACGTCAACA ATGTCCCCCACCTTGGGAACATCATTGG TTGTGTGCTCAGTGCCGATGTCTTTGCC AGGTACTCTCGCCTCCGCCAGTGGAACA CCCTCTATCTGTGTGGGACAGATGAGTA TGGTACAGCAACAGAGACCAAGGCTCT GGAGGAGGGACTAACCCCCCAGGAGAT CTGCGACAAGTACCACATCATCCATGCT GACATCTACCGCTGGTTTAACATTTCGT TTGATATTTTTGGTCGCACCACCACTCC ACAGCAGACCAAAATCACCCAGGACAT TTTCCAGCAGTTGCTGAAACGAGGTTTT GTGCTGCAAGATACTGTGGAGCAACTG CGATGTGAGCACTGTGCTCGCTTCCTGG CTGACCGCTTCGTGGAGGGCGTGTGTCC CTTCTGTGGCTATGAGGAGGCTCGGGGT GACCAGTGTGACAAGTGTGGCAAGCTC ATCAATGCTGTCGAGCTTAAGAAGCCTC AGTGTAAAGTCTGCCGATCATGCCCTGT GGTGCAGTCGAGCCAGCACCTGTTTCTG GACCTGCCTAAGCTGGAGAAGCGACTG GAGGAGTGGTTGGGGAGGACATTGCCT GGCAGTGACTGGACACCCAATGCCCAG TTTATCACCCGTTCTTGGCTTCGGGATG GCCTCAAGCCACGCTGCATAACCCGAG ACCTCAAATGGGGAACCCCTGTACCCTT AGAAGGTTTTGAAGACAAGGTATTCTAT GTCTGGTTTGATGCCACTATTGGCTATC TGTCCATCACAGCCAACTACACAGACCA	

		GTGGGAGAGATGGTGGAAGAACCCAGA GCAAGTGGACCTGTATCAGTTCATGGCC AAAGACAATGTTCCTTTCCATAGCTTAG TCTTTCCTTGCTCAGCCCTAGGAGCTGA GGATAACTATACCTTGGTCAGCCACCTC ATTGCTACAGAGTACCTGAACTATGAGG ATGGGAAATTCTCTAAGAGCCGCGGTGT GGGAGTGTTTGGGGACATGGCCCAGGA CACGGGGATCCCTGCTGACATCTGGCGC TTCTATCTGCTGTACATTCGGCCTGAGG GCCAGGACAGTGCTTTCTCCTGGACGGA CCTGCTGCTGAAGAATAATTCTGAGCTG CTTAACAACCTGGGCAACTTCATCAACA GAGCTGGGATGTTTGTGTCTAAGTTCTT TGGGGGCTATGTGCCTGAGATGGTGCTC ACCCCTGATGATCAGCGCCTGCTGGCCC ATGTCACCCTGGAGCTCCAGCACTATCA CCAGCTACTTGAGAAGGTTCGGATCCGG GATGCCTTGCGCAGTATCCTCACCATAT CTCGACATGGCAACCAATATATTCAGGT GAATGAGCCCTGGAAGCGGATTAAAGG CAGTGAGGCTGACAGGTCAGTCCCTTGT TCCAAAAATTGGAAAATGACCAGATTG AAAGTTTAA	
MetRS^{N8}	Proteína / Humana / 1-821 + 853-900	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GERNVLITSALPYVNNVPHLGNIIGCVLSA DVFARYSRLRQWNTLYLCGTDEYGTATE TKALEEGLTPQEICDKYHIIHADIYRWFNI SFDIFGRTTTPQQTKITQDIFQQLLKRGFV LQDTVEQLRCEHCARFLADRFVEGVCPFC GYEEARGDQCDKCGKLINAVELKKPQCK VCRSCPVVQSSQHLFLDLPKLEKRLEEWL GRTLPGSDWTPNAQFITRSWLRDGLKPRC ITRDLKWGTPVPLEGFEDKVFYVWFDATI	SEQ. ID. NO.30

		GYLSITANYTDQWERWWKNPEQVDLYQ FMAKDNVPFHSLVFPCSALGAEDNYTLVS HLIATEYLNYEDGKFSKSRGVGVFGDMA QDTGIPADIWRFYLLYIRPEGQDSAFSWT DLLLKNNSELLNNLGNFINRAGMFVSKFF GGYVPEMVLTPDDQRLLAHVTLELQHYH QLLEKVRIRDALRSILTISRHGNQYIQVNE PWKRIKGSEADRQRAGTVTGLAVNIAAL LSVMLQPYMPTVSATIQAQLQLPPPACSIL LTNFLCTLPAGHQIGTVSPLFQKLENDQIE SLRQRFGGGQGNIVRELKAQKADKNEVA AEVAKLLDLKKQLAVAEGKPPEAPKGKK KK	
MetRS^{N8}	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAAAGGAATGTGCTC	SEQ. ID. NO.31

		ATCACCAGTGCCCTCCCTTACGTCAACA ATGTCCCCCACCTTGGGAACATCATTGG TTGTGTGCTCAGTGCCGATGTCTTTGCC AGGTACTCTCGCCTCCGCCAGTGGAACA CCCTCTATCTGTGTGGGACAGATGAGTA TGGTACAGCAACAGAGACCAAGGCTCT GGAGGAGGGACTAACCCCCCAGGAGAT CTGCGACAAGTACCACATCATCCATGCT GACATCTACCGCTGGTTTAACATTTCGT TTGATATTTTTGGTCGCACCACCACTCC ACAGCAGACCAAAATCACCCAGGACAT TTTCCAGCAGTTGCTGAAACGAGGTTTT GTGCTGCAAGATACTGTGGAGCAACTG CGATGTGAGCACTGTGCTCGCTTCCTGG CTGACCGCTTCGTGGAGGGCGTGTGTCC CTTCTGTGGCTATGAGGAGGCTCGGGGT GACCAGTGTGACAAGTGTGGCAAGCTC ATCAATGCTGTCGAGCTTAAGAAGCCTC AGTGTAAAGTCTGCCGATCATGCCCTGT GGTGCAGTCGAGCCAGCACCTGTTTCTG GACCTGCCTAAGCTGGAGAAGCGACTG GAGGAGTGGTTGGGGAGGACATTGCCT GGCAGTGACTGGACACCCAATGCCCAG TTTATCACCCGTTCTTGGCTTCGGGATG GCCTCAAGCCACGCTGCATAACCCGAG ACCTCAAATGGGGAACCCCTGTACCCTT AGAAGGTTTTGAAGACAAGGTATTCTAT GTCTGGTTTGATGCCACTATTGGCTATC TGTCCATCACAGCCAACTACACAGACCA GTGGGAGAGATGGTGGAAGAACCCAGA GCAAGTGGACCTGTATCAGTTCATGGCC AAAGACAATGTTCCTTTCCATAGCTTAG TCTTTCCTTGCTCAGCCCTAGGAGCTGA GGATAACTATACCTTGGTCAGCCACCTC ATTGCTACAGAGTACCTGAACTATGAGG ATGGGAAATTCTCTAAGAGCCGCGGTGT GGGAGTGTTTGGGGACATGGCCCAGGA CACGGGGATCCCTGCTGACATCTGGCGC TTCTATCTGCTGTACATTCGGCCTGAGG GCCAGGACAGTGCTTTCTCCTGGACGGA CCTGCTGCTGAAGAATAATTCTGAGCTG CTTAACAACCTGGGCAACTTCATCAACA	

		GAGCTGGGATGTTTGTGTCTAAGTTCTT TGGGGGCTATGTGCCTGAGATGGTGCTC ACCCCTGATGATCAGCGCCTGCTGGCCC ATGTCACCCTGGAGCTCCAGCACTATCA CCAGCTACTTGAGAAGGTTCGGATCCGG GATGCCTTGCGCAGTATCCTCACCATAT CTCGACATGGCAACCAATATATTCAGGT GAATGAGCCCTGGAAGCGGATTAAAGG CAGTGAGGCTGACAGGCAACGGGCAGG AACAGTGACTGGCTTGGCAGTGAATAT AGCTGCCTTGCTCTCTGTCATGCTTCAG CCTTACATGCCCACGGTTAGTGCCACAA TCCAGGCCCAGCTGCAGCTCCCACCTCC AGCCTGCAGTATCCTGCTGACAAACTTC CTGTGTACCTTACCAGCAGGACACCAGA TTGGCACAGTCAGTCCCTTGTTCCAAAA ATTGGAAAATGACCAGATTGAAAGTTT AAGGCAGCGCTTTGGAGGGGGCCAGGG AAACATTGTCCGAGAACTGAAAGCACA AAAGGCAGACAAGAACGAGGTTGCTGC GGAGGTGGCGAAACTCTTGGATCTAAA GAAACAGTTGGCTGTAGCTGAGGGGAA ACCCCCTGAAGCCCCTAAAGGCAAGAA GAAAAAGTAA	
MetRS^{N12}	Proteína / Humana / 1-67 + 47 aa	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRGAECPAQLVPDTEYPGTMS ASCRDCTETARCPGSPALPPKAAPAQPR	SEQ. ID. NO. 32
MetRS^{N12}	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGAGGAGCTGAGTGCCCTGCACA GCTGGTTCCAGACACTGAGTACCCAGG AACCATGTCAGCGAGCTGCAGAGACTG TACTGAAACAGCAAGGTGTCCTGGCTCT CCGGCCTTACCTCCAAAAGCAGCCCCAG CCCAGCCCCGCTGA	SEQ. ID. NO. 33

$MetRS^{N13}$	Proteína / Humana / 1-221 + 55 aa	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEVACGWRKEC AHHQCPPLRQQCPPPWEHHWLCAQCRCL CQVLSPPPVEHPLSVWDR	SEQ. ID. NO. 34
$MetRS^{N13}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG TTGCCTGTGGCTGGAGAAAGGAATGTG CTCATCACCAGTGCCCTCCCTTACGTCA ACAATGTCCCCCACCTTGGGAACATCAT TGGTTGTGTGCTCAGTGCCGATGTCTTT GCCAGGTACTCTCGCCTCCGCCAGTGGA ACACCCTCTATCTGTGTGGGACAGATGA	SEQ. ID. NO. 35

$MetRS^{N13}$	Proteína / Humana / 1-257 + 16 aa	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVWTCI SSWPKTMFLSIA	SEQ. ID. NO. 36
$MetRS^{N14}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTG GACCTGTATCAGTTCATGGCCAAAGACA ATGTTCCTTTCCATAGCTTAG	SEQ. ID. NO. 37

$MetRS^{N15}$	Proteína / Humana / 1-797 + 822-900	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GERNVLITSALPYVNNVPHLGNIIGCVLSA DVFARYSRLRQWNTLYLCGTDEYGTATE TKALEEGLTPQEICDKYHIIHADIYRWFNI SFDIFGRTTTPQQTKITQDIFQQLLKRGFV LQDTVEQLRCEHCARFLADRFVEGVCPFC GYEEARGDQCDKCGKLINAVELKKPQCK VCRSCPVVQSSQHLFLDLPKLEKRLEEWL GRTLPGSDWTPNAQFITRSWLRDGLKPRC ITRDLKWGTPVPLEGFEDKVFYVWFDATI GYLSITANYTDQWERWWKNPEQVDLYQ FMAKDNVPFHSLVFPCSALGAEDNYTLVS HLIATEYLNYEDGKFSKSRGVGVFGDMA QDTGIPADIWRFYLLYIRPEGQDSAFSWT DLLLKNNSELLNNLGNFINRAGMFVSKFF GGYVPEMVLTPDDQRLLAHVTLELQHYH QLLEKVRIRDALRSILTISRHGNQYIQVNE PWKRIKGSEADRQRAGTVTGLAVNIAAL LSVMLQPYMPTVSATIQAQLQLPPPACSIL LTNFLCTLPAGHQIGTAKTSPKPAVVETV TTAKPQQIQALMDEVTKQGNIVRELKAQ KADKNEVAAEVAKLLDLKKQLAVAEGK PPEAPKGKKKK	SEQ. ID. NO. 38
$MetRS^{N15}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG	SEQ. ID. NO. 39

		TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAAAGGAATGTGCTC ATCACCAGTGCCCTCCCTTACGTCAACA ATGTCCCCCACCTTGGGAACATCATTGG TTGTGTGCTCAGTGCCGATGTCTTTGCC AGGTACTCTCGCCTCCGCCAGTGGAACA CCCTCTATCTGTGTGGGACAGATGAGTA TGGTACAGCAACAGAGACCAAGGCTCT GGAGGAGGGACTAACCCCCCAGGAGAT CTGCGACAAGTACCACATCATCCATGCT GACATCTACCGCTGGTTTAACATTTCGT TTGATATTTTTGGTCGCACCACCACTCC ACAGCAGACCAAAATCACCCAGGACAT TTTCCAGCAGTTGCTGAAACGAGGTTTT GTGCTGCAAGATACTGTGGAGCAACTG CGATGTGAGCACTGTGCTCGCTTCCTGG CTGACCGCTTCGTGGAGGGCGTGTGTCC CTTCTGTGGCTATGAGGAGGCTCGGGGT GACCAGTGTGACAAGTGTGGCAAGCTC ATCAATGCTGTCGAGCTTAAGAAGCCTC AGTGTAAAGTCTGCCGATCATGCCCTGT GGTGCAGTCGAGCCAGCACCTGTTTCTG GACCTGCCTAAGCTGGAGAAGCGACTG GAGGAGTGGTTGGGGAGGACATTGCCT GGCAGTGACTGGACACCCAATGCCCAG TTTATCACCCGTTCTTGGCTTCGGGATG GCCTCAAGCCACGCTGCATAACCCGAG ACCTCAAATGGGGAACCCCTGTACCCTT AGAAGGTTTTGAAGACAAGGTATTCTAT GTCTGGTTTGATGCCACTATTGGCTATC TGTCCATCACAGCCAACTACACAGACCA	

		GTGGGAGAGATGGTGGAAGAACCCAGA GCAAGTGGACCTGTATCAGTTCATGGCC AAAGACAATGTTCCTTTCCATAGCTTAG TCTTTCCTTGCTCAGCCCTAGGAGCTGA GGATAACTATACCTTGGTCAGCCACCTC ATTGCTACAGAGTACCTGAACTATGAGG ATGGGAAATTCTCTAAGAGCCGCGGTGT GGGAGTGTTTGGGGACATGGCCCAGGA CACGGGGATCCCTGCTGACATCTGGCGC TTCTATCTGCTGTACATTCGGCCTGAGG GCCAGGACAGTGCTTTCTCCTGGACGGA CCTGCTGCTGAAGAATAATTCTGAGCTG CTTAACAACCTGGGCAACTTCATCAACA GAGCTGGGATGTTTGTGTCTAAGTTCTT TGGGGGCTATGTGCCTGAGATGGTGCTC ACCCCTGATGATCAGCGCCTGCTGGCCC ATGTCACCCTGGAGCTCCAGCACTATCA CCAGCTACTTGAGAAGGTTCGGATCCGG GATGCCTTGCGCAGTATCCTCACCATAT CTCGACATGGCAACCAATATATTCAGGT GAATGAGCCCTGGAAGCGGATTAAAGG CAGTGAGGCTGACAGGCAACGGGCAGG AACAGTGACTGGCTTGGCAGTGAATAT AGCTGCCTTGCTCTCTGTCATGCTTCAG CCTTACATGCCCACGGTTAGTGCCACAA TCCAGGCCCAGCTGCAGCTCCCACCTCC AGCCTGCAGTATCCTGCTGACAAACTTC CTGTGTACCTTACCAGCAGGACACCAGA TTGGCACAGCAAAAACGTCCCCGAAGC CAGCAGTTGTAGAGACTGTTACAACAG CCAAGCCACAGCAGATACAAGCGCTGA TGGATGAAGTGACAAAACAAGGAAACA TTGTCCGAGAACTGAAAGCACAAAAGG CAGACAAGAACGAGGTTGCTGCGGAGG TGGCGAAACTCTTGGATCTAAAGAAAC AGTTGGCTGTAGCTGAGGGGAAACCCC CTGAAGCCCCTAAAGGCAAGAAGAAAA AGTAA	

Tabla 2B
Punto de unión de corte y empalme únicos de polipéptidos de MetRS

Nombre	Tipo / especie	Secuencias de Aminoácidos y Ácido Nucleico en las cercanías del punto de unión de corte y empalme único	SEQ. ID. NO.
M1-AS01	ADN / Humana /	TCTTCTCCACTAGTGCAATCTGCCG|GTA TTCTATGTCTGGTTTGATGCCA	SEQ. ID. NO.40
	Proteína / Humana /	FSTSAICRYSMSGLMP	SEQ. ID. NO.41
M1-AS06	ADN / Humana /	GCACCACCACTCCACAGCAGACCAA|AA GCCTCAGTGTAAAGTCTGCCGAT	SEQ. ID. NO.42
	Proteína / Humana /	TTTPQQTKSLSVKSAD	SEQ. ID. NO.43

M1-AS07	ADN / Humana /	ACCCTTAGAAGGTTTTGAAGACAAG|GT GGACCTGTATCAGTTCATGGCCA	SEQ. ID. NO.44
	Proteína / Humana /	PLEGFEDKVDLYQFMA	SEQ. ID. NO.45
M1-AS08	ADN / Humana /	GAGATGGTGGAAGAACCCAGAGCAA|A GTACCTGAACTATGAGGATGGGAA	SEQ. ID. NO.46
	Proteína / Humana /	RWWKNPEQST	SEQ. ID. NO.47
M1-AS09	ADN / Humana /	GGATTAAAGGCAGTGAGGCTGACAG|GT CAGTCCCTTGTTCCAAAAATTGG	SEQ. ID. NO.48
	Proteína / Humana /	IKGSEADRSVPCSKNW	SEQ. ID. NO.49
M1-AS10	ADN / Humana /	AAGGCAGCGCTTTGGAGGGGGCCAG|GG AAACATTGTCCGAGAACTGAAAG	SEQ. ID. NO.50
	Proteína / Humana /	RQRFGGGQGNIVRELK	SEQ. ID. NO.51
M1-AS13	ADN / Humana /	TCTTCTCCACTAGTGCAATCTGCCG|AGG AGCTGAGTGCCCTGCACAGCTG	SEQ. ID. NO.52
	Proteína / Humana /	FSTSAICRGAECPAQL	SEQ. ID. NO.53
M1-AS14	ADN / Humana /	AAGGGCTGTCACCAATGAGCCTGAG|GT TGCCTGTGGCTGGAGAAAGGAAT	SEQ. ID. NO.54
	Proteína / Humana /	RAVTNEPEVACGWRKE	SEQ. ID. NO.55
M1-AS15	ADN / Humana /	CGCTGCGGCCCCAGCAGAATCCAGT|GT GGACCTGTATCAGTTCATGGCCA	SEQ. ID. NO.56
	Proteína / Humana /	LRPQQNPVWTCISSWP	SEQ. ID. NO.57
M1-AS18	ADN / Humana /	ACCAGCAGGACACCAGATTGGCACA|GC AAAAACGTCCCCGAAGCCAGCAG	SEQ. ID. NO.58
	Proteína / Humana /	PAGHQIGTAKTSPKPA	SEQ. ID. NO.59

Tabla 3
Polipéptidos y ácidos nucleicos de MetRS identificados por Bioinformática

Nombre	Tipo / especie / Restos	Secuencias de Aminoácidos y Ácido Nucleico	SEQ. ID. NO.
$MetRS^{N2}$	Proteína / Humana / 1-263	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GE	SEQ. ID. NO.60

MetRS^{N2}	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAA	SEQ. ID. NO.61
MetRS^{N9}	Proteína / Humana / 1-197	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLA	SEQ. ID. NO.62

MetRS^{N9}	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA	SEQ. ID. NO.63
		GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCT	
MetRS^{N10}	Proteína / Humana / 1-214	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGR	SEQ. ID. NO.64

MetRS^{N10}	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGG	SEQ. ID. NO.65
MetRS^{N11}	Proteína / Humana / 1-221	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE	SEQ. ID. NO.66
		WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPE	

MetRS^{N11}	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAG	SEQ. ID. NO.67

Polipéptidos de MetRS C terminales: (Tablas 4, 5 y 6)

Tabla 4A
Polipéptidos de MetRS identificados por EM

Nombre	Tipo / especie / Restos	Secuencias de Aminoácidos y Ácido Nucleico	SEQ. ID. NO.
MetRS^{C1}	Proteína / Humana / 194-900	GVLALRPYLQKQPQPSPAEGRAVTNEPEE EELATLSEEEIAMAVTAWEKGLESLPPLR PQQNPVLPVAGERNVLITSALPYVNNVPH LGNIIGCVLSADVFARYSRLRQWNTLYLC GTDEYGTATETKALEEGLTPQEICDKYHII	SEQ. ID. NO.101

		HADIYRWFNISFDIFGRTTTPQQTKITQDIF QQLLKRGFVLQDTVEQLRCEHCARFLAD RFVEGVCPFCGYEEARGDQCDKCGKLIN AVELKKPQCKVCRSCPVVQSSQHLFLDLP KLEKRLEEWLGRTLPGSDWTPNAQFITRS WLRDGLKPRCITRDLKWGTPVPLEGFED KVFYVWFDATIGYLSITANYTDQWERW WKNPEQVDLYQFMAKDNVPFHSLVFPCS ALGAEDNYTLVSHLIATEYLNYEDGKFSK SRGVGVFGDMAQDTGIPADIWRFYLLYIR PEGQDSAFSWTDLLLKNNSELLNNLGNFI NRAGMFVSKFFGGYVPEMVLTPDDQRLL AHVTLELQHYHQLLEKVRIRDALRSILTIS RHGNQYIQVNEPWKRIKGSEADRQRAGT VTGLAVNIAALLSVMLQPYMPTVSATIQA QLQLPPPACSILLTNFLCTLPAGHQIGTVSP LFQKLENDQIESLRQRFGGGQAKTSPKPA VVETVTTAKPQQIQALMDEVTKQGNIVR ELKAQKADKNEVAAEVAKLLDLKKQLA VAEGKPPEAPKGKKKK	
MetRS^{C1}	ADN / Humana	GGTGTCCTGGCTCTCCGGCCTTACCTCC AAAAGCAGCCCCAGCCCAGCCCCGCTG AGGGAAGGGCTGTCACCAATGAGCCTG AGGAGGAGGAGCTGGCTACCCTATCTG AGGAGGAGATTGCTATGGCTGTTACTGC TTGGGAGAAGGGCCTAGAAAGTTTGCC CCCGCTGCGGCCCCAGCAGAATCCAGT GTTGCCTGTGGCTGGAGAAAGGAATGT GCTCATCACCAGTGCCCTCCCTTACGTC AACAATGTCCCCCACCTTGGGAACATCA TTGGTTGTGTGCTCAGTGCCGATGTCTT TGCCAGGTACTCTCGCCTCCGCCAGTGG AACACCCTCTATCTGTGTGGGACAGATG AGTATGGTACAGCAACAGAGACCAAGG CTCTGGAGGAGGGACTAACCCCCCAGG AGATCTGCGACAAGTACCACATCATCCA TGCTGACATCTACCGCTGGTTTAACATT TCGTTTGATATTTTTGGTCGCACCACCA CTCCACAGCAGACCAAAATCACCCAGG ACATTTTCCAGCAGTTGCTGAAACGAGG TTTTGTGCTGCAAGATACTGTGGAGCAA CTGCGATGTGAGCACTGTGCTCGCTTCC	SEQ. ID. NO.102

		TGGCTGACCGCTTCGTGGAGGGCGTGTG TCCCTTCTGTGGCTATGAGGAGGCTCGG GGTGACCAGTGTGACAAGTGTGGCAAG CTCATCAATGCTGTCGAGCTTAAGAAGC CTCAGTGTAAAGTCTGCCGATCATGCCC TGTGGTGCAGTCGAGCCAGCACCTGTTT CTGGACCTGCCTAAGCTGGAGAAGCGA CTGGAGGAGTGGTTGGGGAGGACATTG CCTGGCAGTGACTGGACACCCAATGCCC AGTTTATCACCCGTTCTTGGCTTCGGGA TGGCCTCAAGCCACGCTGCATAACCCGA GACCTCAAATGGGGAACCCCTGTACCCT TAGAAGGTTTTGAAGACAAGGTATTCTA TGTCTGGTTTGATGCCACTATTGGCTAT CTGTCCATCACAGCCAACTACACAGACC AGTGGGAGAGATGGTGGAAGAACCCAG AGCAAGTGGACCTGTATCAGTTCATGGC CAAAGACAATGTTCCTTTCCATAGCTTA GTCTTTCCTTGCTCAGCCCTAGGAGCTG AGGATAACTATACCTTGGTCAGCCACCT CATTGCTACAGAGTACCTGAACTATGAG GATGGGAAATTCTCTAAGAGCCGCGGT GTGGGAGTGTTTGGGGACATGGCCCAG GACACGGGGATCCCTGCTGACATCTGGC GCTTCTATCTGCTGTACATTCGGCCTGA GGGCCAGGACAGTGCTTTCTCCTGGACG GACCTGCTGCTGAAGAATAATTCTGAGC TGCTTAACAACCTGGGCAACTTCATCAA CAGAGCTGGGATGTTTGTGTCTAAGTTC TTTGGGGGCTATGTGCCTGAGATGGTGC TCACCCCTGATGATCAGCGCCTGCTGGC CCATGTCACCCTGGAGCTCCAGCACTAT CACCAGCTACTTGAGAAGGTTCGGATCC GGGATGCCTTGCGCAGTATCCTCACCAT ATCTCGACATGGCAACCAATATATTCAG GTGAATGAGCCCTGGAAGCGGATTAAA GGCAGTGAGGCTGACAGGCAACGGGCA GGAACAGTGACTGGCTTGGCAGTGAAT ATAGCTGCCTTGCTCTCTGTCATGCTTC AGCCTTACATGCCCACGGTTAGTGCCAC AATCCAGGCCCAGCTGCAGCTCCCACCT CCAGCCTGCAGTATCCTGCTGACAAACT	

		TCCTGTGTACCTTACCAGCAGGACACCA GATTGGCACAGTCAGTCCCTTGTTCCAA AAATTGGAAAATGACCAGATTGAAAGT TTAAGGCAGCGCTTTGGAGGGGGCCAG GCAAAAACGTCCCCGAAGCCAGCAGTT GTAGAGACTGTTACAACAGCCAAGCCA CAGCAGATACAAGCGCTGATGGATGAA GTGACAAAACAAGGAAACATTGTCCGA GAACTGAAAGCACAAAAGGCAGACAAG AACGAGGTTGCTGCGGAGGTGGCGAAA CTCTTGGATCTAAAGAAACAGTTGGCTG TAGCTGAGGGGAAACCCCCTGAAGCCC CTAAAGGCAAGAAGAAAAAGTAA	

Tabla 4B
Péptidos detectados por espectrometría de masas y péptidos de enlace inferidos de MetRS[C1]

Tipo / especie	Secuencia	SEQ. ID. NO.
Proteína / ratón	TVSNELEEEELATLSEEDIVTAVAAWEK	SEQ. ID. NO. 103
Proteína / ratón	GLESLPPLKLQQHPVLPVPGER	SEQ. ID. NO. 104
Proteína / ratón	NVLITSALPY	SEQ. ID. NO. 105
Proteína / ratón	VNNVPHLGNIIGCVLSADVFAR	SEQ. ID. NO. 106
Proteína / ratón	YCR	SEQ. ID. NO. 107
Proteína / ratón	LRQWNTLYLCGTDEYGTATETK	SEQ. ID. NO. 108
Proteína / ratón	AMEEGLTPR	SEQ. ID. NO. 109
Proteína / ratón	EICDKYHAIHADIYR	SEQ. ID. NO. 110
Proteína / ratón	WFGISFDTFGR	SEQ. ID. NO. 111
Proteína / ratón	TTTPQQTK	SEQ. ID. NO. 112
Proteína / ratón	ITQDIFQR	SEQ. ID. NO. 113
Proteína / ratón	LLTRGFVLRDTVEQLRCERCARFLADR	SEQ. ID. NO. 114
Proteína / ratón	FVEGVCPFCGYEEAR	SEQ. ID. NO. 115
Proteína / ratón	GDQCDRCGKLINAIELKKPQCKICRSCPVVR	SEQ. ID. NO. 116
Proteína / ratón	SSQHLFLDLPK	SEQ. ID. NO. 117
Proteína / ratón	LEKRLEDWLGKTVPGSDWTPNARFIIRSWLRDGLKPR CITRDLKWGTPVPLEGFEDKVFYVWFDATIGYVSITA NYTDQWEKWWKNPEQVDLYQFMAK	SEQ. ID. NO. 118
Proteína / ratón	DNVPFHGLVFPCSVLGAEDNYTLVK	SEQ. ID. NO. 119
Proteína / ratón	HIIATEYLNYEDGK	SEQ. ID. NO. 120
Proteína / ratón	FSKSRGIGVFGDMAKDTGIPADIWRFYLLY	SEQ. ID. NO. 121

Proteína / ratón	IRPEGQDSAFSWTDLLIKNNSELLNNLGNFINR	SEQ. ID. NO. 122
Proteína / ratón	AGMFVSKFFGGCVPEMALTPDDRRLVAHVSWELQH YHQLLEKVRIRDALRSILTISR	SEQ. ID. NO. 123
Proteína / ratón	HGNQYIQVNEPWKR	SEQ. ID. NO. 124
Proteína / ratón	IKGGEMDRQRAGTVTGMAVNMAALLSVMLQPYMPT VSSTIQTQLQLPPAACRILATSFICTLPAGHR	SEQ. ID. NO. 125
Proteína / ratón	IGTVSPLFQK	SEQ. ID. NO. 126
Proteína / ratón	LENDQIENLR	SEQ. ID. NO. 127
Proteína / ratón	QRFGGGQAK	SEQ. ID. NO. 128
Proteína / ratón	GSPKPAAVEAVTAAGSQHIQTLTDEVTK	SEQ. ID. NO. 129
Proteína / ratón	QGNVVRELKAQKADKNQVAAEVAKLLDLKK	SEQ. ID. NO. 130
Proteína / ratón	QLALAEGKPIETPK	SEQ. ID. NO. 131

Tabla 4C
Secuencias concatenadas basadas en péptidos detectados por espectrometría de masas de MetRS^{C1}

Tipo / especie	Secuencia	SEQ. ID. NO.
Proteína / ratón	**TVSNELEEEELATLSEEDIVTAVAAWEKGLESLPPL KLQQHPVLPVPGERNVLITSALPYVNNVPHLGNIIG CVLSADVFARYCRLRQWNTLYLCGTDEYGTATET KAMEEGLTPREICDKYHAIHADIYRWFGISFDTFG** RTTTPQQTK**ITQDIFQR**LLTRGFVLRDTVEQLRCERC ARFLADR**FVEGVCPFCGYEEAR**GDQCDRCGKLINAI ELKKPQCKICRSCPVVR**SSQHLFLDLPK**LEKRLEDWL GKTVPGSDWTPNARFIIRSWLRDGLKPRCITRDLKWG TPVPLEGFEDKVFYVWFDATIGYVSITANYTDQWEK WWKNPEQVDLYQFMAK**DNVPFHGLVFPCSVLGAE DNYTLVKHIIATEYLNYEDGK**FSKSRGIGVFGDMAK DTGIPADIWRFYLLY**IRPEGQDSAFSWTDLLIKNNSE**	SEQ. ID. NO. 132
Proteína / ratón	**LLNNLGNFINR**AGMFVSKFFGGCVPEMALTPDDRRL VAHVSWELQHYHQLLEKVRIRDALRSILTISR**HGNQY IQVNEPWKR**IKGGEMDRQRAGTVTGMAVNMAALLS VMLQPYMPTVSSTIQTQLQLPPAACRILATSFICTLPA GHR**IGTVSPLFQKLENDQIENLR**QRFGGGQAK**GSPK PAAVEAVTAAGSQHIQTLTDEVTK**QGNVVRELKAQ KADKNQVAAEVAKLLDLKK**QLALAEGKPIETPK**	

Tabla 5
Polipéptidos de MetRS y transcritos alternativos identificados por Secuenciación Profunda

Nombre	Tipo / especie / Restos	Secuencias de Aminoácidos y Ácido Nucleico	SEQ. ID. NO.

$MetRS^{C3}$	Proteína / Humana / 552-900	MAKDNVPFHSLVFPCSALGAEDNYTLVS HLIATEYLNYEDGKFSKSRGVGVFGDMA QDTGIPADIWRFYLLYIRPEGQDSAFSWT DLLLKNNSELLNNLGNFINRAGMFVSKFF GGYVPEMVLTPDDQRLLAHVTLELQHYH QLLEKVRIRDALRSILTISRHGNQYIQVNE PWKRIKGSEADRQRAGTVTGLAVNIAAL LSVMLQPYMPTVSATIQAQLQLPPPACSIL LTNFLCTLPAGHQIGTVSPLFQKLENDQIE SLRQRFGGGQAKTSPKPAVVETVTTAKPQ QIQALMDEVTKQGNIVRELKAQKADKNE VAAEVAKLLDLKKQLAVAEGKPPEAPKG KKKK	SEQ. ID. NO.133
$MetRS^{C3}$	ADN / Humana	ATGGCCAAAGACAATGTTCCTTTCCATA GCTTAGTCTTTCCTTGCTCAGCCCTAGG AGCTGAGGATAACTATACCTTGGTCAGC CACCTCATTGCTACAGAGTACCTGAACT ATGAGGATGGGAAATTCTCTAAGAGCC GCGGTGTGGGAGTGTTTGGGGACATGG CCCAGGACACGGGGATCCCTGCTGACA TCTGGCGCTTCTATCTGCTGTACATTCG GCCTGAGGGCCAGGACAGTGCTTTCTCC TGGACGGACCTGCTGCTGAAGAATAATT CTGAGCTGCTTAACAACCTGGGCAACTT CATCAACAGAGCTGGGATGTTTGTGTCT AAGTTCTTTGGGGGCTATGTGCCTGAGA TGGTGCTCACCCCTGATGATCAGCGCCT GCTGGCCCATGTCACCCTGGAGCTCCAG	SEQ. ID. NO.134

		CACTATCACCAGCTACTTGAGAAGGTTC GGATCCGGGATGCCTTGCGCAGTATCCT CACCATATCTCGACATGGCAACCAATAT ATTCAGGTGAATGAGCCCTGGAAGCGG ATTAAAGGCAGTGAGGCTGACAGGCAA CGGGCAGGAACAGTGACTGGCTTGGCA GTGAATATAGCTGCCTTGCTCTCTGTCA TGCTTCAGCCTTACATGCCCACGGTTAG TGCCACAATCCAGGCCCAGCTGCAGCTC CCACCTCCAGCCTGCAGTATCCTGCTGA CAAACTTCCTGTGTACCTTACCAGCAGG ACACCAGATTGGCACAGTCAGTCCCTTG TTCCAAAAATTGGAAAATGACCAGATT GAAAGTTTAAGGCAGCGCTTTGGAGGG GGCCAGGCAAAAACGTCCCCGAAGCCA GCAGTTGTAGAGACTGTTACAACAGCC AAGCCACAGCAGATACAAGCGCTGATG GATGAAGTGACAAAACAAGGAAACATT GTCCGAGAACTGAAAGCACAAAAGGCA GACAAGAACGAGGTTGCTGCGGAGGTG GCGAAACTCTTGGATCTAAAGAAACAG TTGGCTGTAGCTGAGGGGAAACCCCCTG AAGCCCCTAAAGGCAAGAAGAAAAAGT AA	
MetRS^{C4}	Proteína / Humana / 8 aa + 586-900	MVEEPRAKYLNYEDGKFSKSRGVGVFGD MAQDTGIPADIWRFYLLYIRPEGQDSAFS WTDLLLKNNSELLNNLGNFINRAGMFVS KFFGGYVPEMVLTPDDQRLLAHVTLELQ HYHQLLEKVRIRDALRSILTISRHGNQYIQ VNEPWKRIKGSEADRQRAGTVTGLAVNI AALLSVMLQPYMPTVSATIQAQLQLPPPA CSILLTNFLCTLPAGHQIGTVSPLFQKLEN DQIESLRQRFGGGQAKTSPKPAVVETVTT AKPQQIQALMDEVTKQGNIVRELKAQKA DKNEVAAEVAKLLDLKKQLAVAEGKPPE APKGKKKK	SEQ. ID. NO.135
MetRS^{C4}	ADN / Humana	ATGGTGGAAGAACCCAGAGCAAAGTAC CTGAACTATGAGGATGGGAAATTCTCTA AGAGCCGCGGTGTGGGAGTGTTTGGGG ACATGGCCCAGGACACGGGGATCCCTG CTGACATCTGGCGCTTCTATCTGCTGTA CATTCGGCCTGAGGGCCAGGACAGTGC	SEQ. ID. NO.136

		TTTCTCCTGGACGGACCTGCTGCTGAAG AATAATTCTGAGCTGCTTAACAACCTGG GCAACTTCATCAACAGAGCTGGGATGTT TGTGTCTAAGTTCTTTGGGGGCTATGTG CCTGAGATGGTGCTCACCCCTGATGATC AGCGCCTGCTGGCCCATGTCACCCTGGA GCTCCAGCACTATCACCAGCTACTTGAG AAGGTTCGGATCCGGGATGCCTTGCGCA GTATCCTCACCATATCTCGACATGGCAA CCAATATATTCAGGTGAATGAGCCCTGG AAGCGGATTAAAGGCAGTGAGGCTGAC AGGCAACGGGCAGGAACAGTGACTGGC TTGGCAGTGAATATAGCTGCCTTGCTCT CTGTCATGCTTCAGCCTTACATGCCCAC GGTTAGTGCCACAATCCAGGCCCAGCTG CAGCTCCCACCTCCAGCCTGCAGTATCC TGCTGACAAACTTCCTGTGTACCTTACC AGCAGGACACCAGATTGGCACAGTCAG TCCCTTGTTCCAAAAATTGGAAAATGAC CAGATTGAAAGTTTAAGGCAGCGCTTTG GAGGGGGCCAGGCAAAAACGTCCCCGA AGCCAGCAGTTGTAGAGACTGTTACAA CAGCCAAGCCACAGCAGATACAAGCGC TGATGGATGAAGTGACAAAACAAGGAA ACATTGTCCGAGAACTGAAAGCACAAA AGGCAGACAAGAACGAGGTTGCTGCGG AGGTGGCGAAACTCTTGGATCTAAAGA AACAGTTGGCTGTAGCTGAGGGGAAAC CCCCTGAAGCCCCTAAAGGCAAGAAGA AAAAGTAA	
MetRS^{C5}	Proteína / Humana / 846-900	MDEVTKQGNIVRELKAQKADKNEVAAE VAKLLDLKKQLAVAEGKPPEAPKGKKKK	SEQ. ID. NO.137
MetRS^{C5}	ADN / Humana	ATGGATGAAGTGACAAAACAAGGAAAC ATTGTCCGAGAACTGAAAGCACAAAAG GCAGACAAGAACGAGGTTGCTGCGGAG GTGGCGAAACTCTTGGATCTAAAGAAA CAGTTGGCTGTAGCTGAGGGGAAACCC CCTGAAGCCCCTAAAGGCAAGAAGAAA AAGTAA	SEQ. ID. NO.138

$MetRS^{C6}$	Proteína / Humana / 1-36 + 164-900	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEEELSALHSWFQTLSTQEPCQRA AETVLKQQGVLALRPYLQKQPQPSPAEG RAVTNEPEEEELATLSEEEIAMAVTAWEK GLESLPPLRPQQNPVLPVAGERNVLITSAL PYVNNVPHLGNIIGCVLSADVFARYSRLR QWNTLYLCGTDEYGTATETKALEEGLTP QEICDKYHIIHADIYRWFNISFDIFGRTTTP QQTKITQDIFQQLLKRGFVLQDTVEQLRC EHCARFLADRFVEGVCPFCGYEEARGDQ CDKCGKLINAVELKKPQCKVCRSCPVVQ SSQHLFLDLPKLEKRLEEWLGRTLPGSDW TPNAQFITRSWLRDGLKPRCITRDLKWGT PVPLEGFEDKVFYVWFDATIGYLSITANY TDQWERWWKNPEQVDLYQFMAKDNVPF HSLVFPCSALGAEDNYTLVSHLIATEYLN YEDGKFSKSRGVGVFGDMAQDTGIPADI WRFYLLYIRPEGQDSAFSWTDLLLKNNSE LLNNLGNFINRAGMFVSKFFGGYVPEMV LTPDDQRLLAHVTLELQHYHQLLEKVRIR DALRSILTISRHGNQYIQVNEPWKRIKGSE ADRQRAGTVTGLAVNIAALLSVMLQPYM PTVSATIQAQLQLPPPACSILLTNFLCTLPA GHQIGTVSPLFQKLENDQIESLRQRFGGG QAKTSPKPAVVETVTTAKPQQIQALMDE VTKQGNIVRELKAQKADKNEVAAEVAKL LDLKKQLAVAEGKPPEAPKGKKKK	SEQ. ID. NO.139
$MetRS^{C6}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA GGAGCTGAGTGCCCTGCACAGCTGGTTC CAGACACTGAGTACCCAGGAACCATGT CAGCGAGCTGCAGAGACTGTACTGAAA CAGCAAGGTGTCCTGGCTCTCCGGCCTT ACCTCCAAAAGCAGCCCCAGCCCAGCC CCGCTGAGGGAAGGGCTGTCACCAATG AGCCTGAGGAGGAGGAGCTGGCTACCC TATCTGAGGAGGAGATTGCTATGGCTGT TACTGCTTGGGAGAAGGGCCTAGAAAG TTTGCCCCCGCTGCGGCCCCAGCAGAAT CCAGTGTTGCCTGTGGCTGGAGAAAGG AATGTGCTCATCACCAGTGCCCTCCCTT ACGTCAACAATGTCCCCCACCTTGGGAA	SEQ. ID. NO.140

		CATCATTGGTTGTGTGCTCAGTGCCGAT GTCTTTGCCAGGTACTCTCGCCTCCGCC AGTGGAACACCCTCTATCTGTGTGGGAC AGATGAGTATGGTACAGCAACAGAGAC CAAGGCTCTGGAGGAGGGACTAACCCC CCAGGAGATCTGCGACAAGTACCACAT CATCCATGCTGACATCTACCGCTGGTTT AACATTTCGTTTGATATTTTTGGTCGCA CCACCACTCCACAGCAGACCAAAATCA CCCAGGACATTTTCCAGCAGTTGCTGAA ACGAGGTTTTGTGCTGCAAGATACTGTG GAGCAACTGCGATGTGAGCACTGTGCTC GCTTCCTGGCTGACCGCTTCGTGGAGGG CGTGTGTCCCTTCTGTGGCTATGAGGAG GCTCGGGGTGACCAGTGTGACAAGTGT GGCAAGCTCATCAATGCTGTCGAGCTTA AGAAGCCTCAGTGTAAAGTCTGCCGATC ATGCCCTGTGGTGCAGTCGAGCCAGCAC CTGTTTCTGGACCTGCCTAAGCTGGAGA AGCGACTGGAGGAGTGGTTGGGGAGGA CATTGCCTGGCAGTGACTGGACACCCAA TGCCCAGTTTATCACCCGTTCTTGGCTTC GGGATGGCCTCAAGCCACGCTGCATAA CCCGAGACCTCAAATGGGGAACCCCTG TACCCTTAGAAGGTTTTGAAGACAAGGT ATTCTATGTCTGGTTTGATGCCACTATT GGCTATCTGTCCATCACAGCCAACTACA CAGACCAGTGGGAGAGATGGTGGAAGA ACCCAGAGCAAGTGGACCTGTATCAGTT CATGGCCAAAGACAATGTTCCTTTCCAT AGCTTAGTCTTTCCTTGCTCAGCCCTAG GAGCTGAGGATAACTATACCTTGGTCAG CCACCTCATTGCTACAGAGTACCTGAAC TATGAGGATGGGAAATTCTCTAAGAGC CGCGGTGTGGGAGTGTTTGGGGACATG GCCCAGGACACGGGGATCCCTGCTGAC ATCTGGCGCTTCTATCTGCTGTACATTC GGCCTGAGGGCCAGGACAGTGCTTTCTC CTGGACGGACCTGCTGCTGAAGAATAA TTCTGAGCTGCTTAACAACCTGGGCAAC TTCATCAACAGAGCTGGGATGTTTGTGT CTAAGTTCTTTGGGGGCTATGTGCCTGA	

		GATGGTGCTCACCCCTGATGATCAGCGC CTGCTGGCCCATGTCACCCTGGAGCTCC AGCACTATCACCAGCTACTTGAGAAGGT TCGGATCCGGGATGCCTTGCGCAGTATC CTCACCATATCTCGACATGGCAACCAAT ATATTCAGGTGAATGAGCCCTGGAAGC GGATTAAAGGCAGTGAGGCTGACAGGC AACGGGCAGGAACAGTGACTGGCTTGG CAGTGAATATAGCTGCCTTGCTCTCTGT CATGCTTCAGCCTTACATGCCCACGGTT AGTGCCACAATCCAGGCCCAGCTGCAG CTCCCACCTCCAGCCTGCAGTATCCTGC TGACAAACTTCCTGTGTACCTTACCAGC AGGACACCAGATTGGCACAGTCAGTCC CTTGTTCCAAAAATTGGAAAATGACCAG ATTGAAAGTTTAAGGCAGCGCTTTGGAG GGGGCCAGGCAAAAACGTCCCCGAAGC CAGCAGTTGTAGAGACTGTTACAACAG CCAAGCCACAGCAGATACAAGCGCTGA TGGATGAAGTGACAAAACAAGGAAACA TTGTCCGAGAACTGAAAGCACAAAAGG CAGACAAGAACGAGGTTGCTGCGGAGG TGGCGAAACTCTTGGATCTAAAGAAAC AGTTGGCTGTAGCTGAGGGGAAACCCC CTGAAGCCCCTAAAGGCAAGAAGAAAA AGTAA	
MetRS^{C7}	Proteína / Humana / 235-900	MAVTAWEKGLESLPPLRPQQNPVLPVAG ERNVLITSALPYVNNVPHLGNIIGCVLSAD VFARYSRLRQWNTLYLCGTDEYGTATET KALEEGLTPQEICDKYHIIHADIYRWFNIS FDIFGRTTTPQQTKITQDIFQQLLKRGFVL QDTVEQLRCEHCARFLADRFVEGVCPFC GYEEARGDQCDKCGKLINAVELKKPQCK VCRSCPVVQSSQHLFLDLPKLEKRLEEWL GRTLPGSDWTPNAQFITRSWLRDGLKPRC ITRDLKWGTPVPLEGFEDKVFYVWFDATI GYLSITANYTDQWERWWKNPEQVDLYQ FMAKDNVPFHSLVFPCSALGAEDNYTLVS HLIATEYLNYEDGKFSKSRGVGVFGDMA QDTGIPADIWRFYLLYIRPEGQDSAFSWT DLLLKNNSELLNNLGNFINRAGMFVSKFF GGYVPEMVLTPDDQRLLAHVTLELQHYH	SEQ. ID. NO.141

		QLLEKVRIRDALRSILTISRHGNQYIQVNE PWKRIKGSEADRQRAGTVTGLAVNIAAL LSVMLQPYMPTVSATIQAQLQLPPPACSIL LTNFLCTLPAGHQIGTVSPLFQKLENDQIE SLRQRFGGGQAKTSPKPAVVETVTTAKPQ QIQALMDEVTKQGNIVRELKAQKADKNE VAAEVAKLLDLKKQLAVAEGKPPEAPKG KKKK	
MetRS^{C7}	ADN / Humana	ATGGCTGTTACTGCTTGGGAGAAGGGCC TAGAAAGTTTGCCCCCGCTGCGGCCCCA GCAGAATCCAGTGTTGCCTGTGGCTGGA GAAAGGAATGTGCTCATCACCAGTGCC CTCCCTTACGTCAACAATGTCCCCCACC TTGGGAACATCATTGGTTGTGTGCTCAG TGCCGATGTCTTTGCCAGGTACTCTCGC CTCCGCCAGTGGAACACCCTCTATCTGT GTGGGACAGATGAGTATGGTACAGCAA CAGAGACCAAGGCTCTGGAGGAGGGAC TAACCCCCCAGGAGATCTGCGACAAGT ACCACATCATCCATGCTGACATCTACCG CTGGTTTAACATTTCGTTTGATATTTTTG GTCGCACCACCACTCCACAGCAGACCA AAATCACCCAGGACATTTTCCAGCAGTT GCTGAAACGAGGTTTTGTGCTGCAAGAT ACTGTGGAGCAACTGCGATGTGAGCAC TGTGCTCGCTTCCTGGCTGACCGCTTCG TGGAGGGCGTGTGTCCCTTCTGTGGCTA TGAGGAGGCTCGGGGTGACCAGTGTGA CAAGTGTGGCAAGCTCATCAATGCTGTC GAGCTTAAGAAGCCTCAGTGTAAAGTCT GCCGATCATGCCCTGTGGTGCAGTCGAG CCAGCACCTGTTTCTGGACCTGCCTAAG CTGGAGAAGCGACTGGAGGAGTGGTTG GGGAGGACATTGCCTGGCAGTGACTGG ACACCCAATGCCCAGTTTATCACCCGTT CTTGGCTTCGGGATGGCCTCAAGCCACG CTGCATAACCCGAGACCTCAAATGGGG AACCCCTGTACCCTTAGAAGGTTTTGAA GACAAGGTATTCTATGTCTGGTTTGATG CCACTATTGGCTATCTGTCCATCACAGC CAACTACACAGACCAGTGGGAGAGATG GTGGAAGAACCCAGAGCAAGTGGACCT	SEQ. ID. NO.142

		GTATCAGTTCATGGCCAAAGACAATGTT CCTTTCCATAGCTTAGTCTTTCCTTGCTC AGCCCTAGGAGCTGAGGATAACTATAC CTTGGTCAGCCACCTCATTGCTACAGAG TACCTGAACTATGAGGATGGGAAATTCT CTAAGAGCCGCGGTGTGGGAGTGTTTG GGGACATGGCCCAGGACACGGGGATCC CTGCTGACATCTGGCGCTTCTATCTGCT GTACATTCGGCCTGAGGGCCAGGACAG TGCTTTCTCCTGGACGGACCTGCTGCTG AAGAATAATTCTGAGCTGCTTAACAACC TGGGCAACTTCATCAACAGAGCTGGGA TGTTTGTGTCTAAGTTCTTTGGGGGCTA TGTGCCTGAGATGGTGCTCACCCCTGAT GATCAGCGCCTGCTGGCCCATGTCACCC TGGAGCTCCAGCACTATCACCAGCTACT TGAGAAGGTTCGGATCCGGGATGCCTTG CGCAGTATCCTCACCATATCTCGACATG GCAACCAATATATTCAGGTGAATGAGC CCTGGAAGCGGATTAAAGGCAGTGAGG CTGACAGGCAACGGGCAGGAACAGTGA CTGGCTTGGCAGTGAATATAGCTGCCTT GCTCTCTGTCATGCTTCAGCCTTACATG CCCACGGTTAGTGCCACAATCCAGGCCC AGCTGCAGCTCCCACCTCCAGCCTGCAG TATCCTGCTGACAAACTTCCTGTGTACC TTACCAGCAGGACACCAGATTGGCACA GTCAGTCCCTTGTTCCAAAAATTGGAAA ATGACCAGATTGAAAGTTTAAGGCAGC GCTTTGGAGGGGGCCAGGCAAAAACGT CCCCGAAGCCAGCAGTTGTAGAGACTG TTACAACAGCCAAGCCACAGCAGATAC AAGCGCTGATGGATGAAGTGACAAAAC AAGGAAACATTGTCCGAGAACTGAAAG CACAAAAGGCAGACAAGAACGAGGTTG CTGCGGAGGTGGCGAAACTCTTGGATCT AAAGAAACAGTTGGCTGTAGCTGAGGG GAAACCCCCTGAAGCCCCTAAAGGCAA GAAGAAAAAGTAA	

$MetRS^{C8}$	Proteína / Humana / 4 aa + 258-900	MSLRLPVAGERNVLITSALPYVNNVPHLG NIIGCVLSADVFARYSRLRQWNTLYLCGT DEYGTATETKALEEGLTPQEICDKYHIIHA DIYRWFNISFDIFGRTTTPQQTKITQDIFQQ LLKRGFVLQDTVEQLRCEHCARFLADRF VEGVCPFCGYEEARGDQCDKCGKLINAV ELKKPQCKVCRSCPVVQSSQHLFLDLPKL EKRLEEWLGRTLPGSDWTPNAQFITRSWL RDGLKPRCITRDLKWGTPVPLEGFEDKVF YVWFDATIGYLSITANYTDQWERWWKN PEQVDLYQFMAKDNVPFHSLVFPCSALG AEDNYTLVSHLIATEYLNYEDGKFSKSRG VGVFGDMAQDTGIPADIWRFYLLYIRPEG QDSAFSWTDLLLKNNSELLNNLGNFINRA GMFVSKFFGGYVPEMVLTPDDQRLLAHV TLELQHYHQLLEKVRIRDALRSILTISRHG NQYIQVNEPWKRIKGSEADRQRAGTVTG LAVNIAALLSVMLQPYMPTVSATIQAQLQ LPPPACSILLTNFLCTLPAGHQIGTVSPLFQ KLENDQIESLRQRFGGGQAKTSPKPAVVE TVTTAKPQQIQALMDEVTKQGNIVRELK AQKADKNEVAAEVAKLLDLKKQLAVAE GKPPEAPKGKKKK	SEQ. ID. NO.143
$MetRS^{C8}$	ADN / Humana	ATGAGCCTGAGGTTGCCTGTGGCTGGAG AAAGGAATGTGCTCATCACCAGTGCCCT CCCTTACGTCAACAATGTCCCCCACCTT GGGAACATCATTGGTTGTGTGCTCAGTG CCGATGTCTTTGCCAGGTACTCTCGCCT CCGCCAGTGGAACACCCTCTATCTGTGT GGGACAGATGAGTATGGTACAGCAACA GAGACCAAGGCTCTGGAGGAGGGACTA ACCCCCCAGGAGATCTGCGACAAGTAC CACATCATCCATGCTGACATCTACCGCT GGTTTAACATTTCGTTTGATATTTTTGGT CGCACCACCACTCCACAGCAGACCAAA ATCACCCAGGACATTTTCCAGCAGTTGC TGAAACGAGGTTTTGTGCTGCAAGATAC TGTGGAGCAACTGCGATGTGAGCACTGT GCTCGCTTCCTGGCTGACCGCTTCGTGG AGGGCGTGTGTCCCTTCTGTGGCTATGA GGAGGCTCGGGGTGACCAGTGTGACAA GTGTGGCAAGCTCATCAATGCTGTCGAG CTTAAGAAGCCTCAGTGTAAAGTCTGCC GATCATGCCCTGTGGTGCAGTCGAGCCA GCACCTGTTTCTGGACCTGCCTAAGCTG	SEQ. ID. NO.144

		GAGAAGCGACTGGAGGAGTGGTTGGGG AGGACATTGCCTGGCAGTGACTGGACA CCCAATGCCCAGTTTATCACCCGTTCTT GGCTTCGGGATGGCCTCAAGCCACGCTG CATAACCCGAGACCTCAAATGGGGAAC CCCTGTACCCTTAGAAGGTTTTGAAGAC AAGGTATTCTATGTCTGGTTTGATGCCA CTATTGGCTATCTGTCCATCACAGCCAA CTACACAGACCAGTGGGAGAGATGGTG GAAGAACCCAGAGCAAGTGGACCTGTA TCAGTTCATGGCCAAAGACAATGTTCCT TTCCATAGCTTAGTCTTTCCTTGCTCAGC CCTAGGAGCTGAGGATAACTATACCTTG GTCAGCCACCTCATTGCTACAGAGTACC TGAACTATGAGGATGGGAAATTCTCTAA GAGCCGCGGTGTGGGAGTGTTTGGGGA CATGGCCCAGGACACGGGGATCCCTGC TGACATCTGGCGCTTCTATCTGCTGTAC ATTCGGCCTGAGGGCCAGGACAGTGCTT TCTCCTGGACGGACCTGCTGCTGAAGAA TAATTCTGAGCTGCTTAACAACCTGGGC AACTTCATCAACAGAGCTGGGATGTTTG TGTCTAAGTTCTTTGGGGGCTATGTGCC TGAGATGGTGCTCACCCCTGATGATCAG CGCCTGCTGGCCCATGTCACCCTGGAGC TCCAGCACTATCACCAGCTACTTGAGAA GGTTCGGATCCGGGATGCCTTGCGCAGT ATCCTCACCATATCTCGACATGGCAACC AATATATTCAGGTGAATGAGCCCTGGA AGCGGATTAAAGGCAGTGAGGCTGACA GGCAACGGGCAGGAACAGTGACTGGCT TGGCAGTGAATATAGCTGCCTTGCTCTC TGTCATGCTTCAGCCTTACATGCCCACG GTTAGTGCCACAATCCAGGCCCAGCTGC AGCTCCCACCTCCAGCCTGCAGTATCCT GCTGACAAACTTCCTGTGTACCTTACCA GCAGGACACCAGATTGGCACAGTCAGT CCCTTGTTCCAAAAATTGGAAAATGACC AGATTGAAAGTTTAAGGCAGCGCTTTGG AGGGGGCCAGGCAAAAACGTCCCCGAA GCCAGCAGTTGTAGAGACTGTTACAAC AGCCAAGCCACAGCAGATACAAGCGCT	
		GATGGATGAAGTGACAAAACAAGGAAA CATTGTCCGAGAACTGAAAGCACAAAA GGCAGACAAGAACGAGGTTGCTGCGGA GGTGGCGAAACTCTTGGATCTAAAGAA ACAGTTGGCTGTAGCTGAGGGGAAACC CCCTGAAGCCCCTAAAGGCAAGAAGAA AAAGTAA	

MetRS^{C9}	Proteína / Humana / 1-296 + 365-900	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GERNVLITSALPYVNNVPHLGNIIGCVLSA DVFARITQDIFQQLLKRGFVLQDTVEQLR CEHCARFLADRFVEGVCPFCGYEEARGD QCDKCGKLINAVELKKPQCKVCRSCPVV QSSQHLFLDLPKLEKRLEEWLGRTLPGSD WTPNAQFITRSWLRDGLKPRCITRDLKW GTPVPLEGFEDKVFYVWFDATIGYLSITA NYTDQWERWWKNPEQVDLYQFMAKDN VPFHSLVFPCSALGAEDNYTLVSHLIATEY LNYEDGKFSKSRGVGVFGDMAQDTGIPA DIWRFYLLYIRPEGQDSAFSWTDLLLKNN SELLNNLGNFINRAGMFVSKFFGGYVPEM VLTPDDQRLLAHVTLELQHYHQLLEKVRI RDALRSILTISRHGNQYIQVNEPWKRIKGS EADRQRAGTVTGLAVNIAALLSVMLQPY MPTVSATIQAQLQLPPPACSILLTNFLCTL PAGHQIGTVSPLFQKLENDQIESLRQRFGG GQAKTSPKPAVVETVTTAKPQQIQALMD EVTKQGNIVRELKAQKADKNEVAAEVAK LLDLKKQLAVAEGKPPEAPKGKKKK	SEQ. ID. NO.145
MetRS^{C9}	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG	SEQ. ID. NO.146

		GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAAAGGAATGTGCTC ATCACCAGTGCCCTCCCTTACGTCAACA ATGTCCCCCACCTTGGGAACATCATTGG TTGTGTGCTCAGTGCCGATGTCTTTGCC AGAATCACCCAGGACATTTTCCAGCAGT TGCTGAAACGAGGTTTTGTGCTGCAAGA TACTGTGGAGCAACTGCGATGTGAGCA CTGTGCTCGCTTCCTGGCTGACCGCTTC GTGGAGGGCGTGTGTCCCTTCTGTGGCT ATGAGGAGGCTCGGGGTGACCAGTGTG ACAAGTGTGGCAAGCTCATCAATGCTGT CGAGCTTAAGAAGCCTCAGTGTAAAGT CTGCCGATCATGCCCTGTGGTGCAGTCG AGCCAGCACCTGTTTCTGGACCTGCCTA AGCTGGAGAAGCGACTGGAGGAGTGGT TGGGGAGGACATTGCCTGGCAGTGACT GGACACCCAATGCCCAGTTTATCACCCG TTCTTGGCTTCGGGATGGCCTCAAGCCA CGCTGCATAACCCGAGACCTCAAATGG GGAACCCCTGTACCCTTAGAAGGTTTTG	

		AAGACAAGGTATTCTATGTCTGGTTTGA TGCCACTATTGGCTATCTGTCCATCACA GCCAACTACACAGACCAGTGGGAGAGA TGGTGGAAGAACCCAGAGCAAGTGGAC CTGTATCAGTTCATGGCCAAAGACAATG TTCCTTTCCATAGCTTAGTCTTTCCTTGC TCAGCCCTAGGAGCTGAGGATAACTAT ACCTTGGTCAGCCACCTCATTGCTACAG AGTACCTGAACTATGAGGATGGGAAAT TCTCTAAGAGCCGCGGTGTGGGAGTGTT TGGGGACATGGCCCAGGACACGGGGAT CCCTGCTGACATCTGGCGCTTCTATCTG CTGTACATTCGGCCTGAGGGCCAGGAC AGTGCTTTCTCCTGGACGGACCTGCTGC TGAAGAATAATTCTGAGCTGCTTAACAA CCTGGGCAACTTCATCAACAGAGCTGG GATGTTTGTGTCTAAGTTCTTTGGGGGC TATGTGCCTGAGATGGTGCTCACCCCTG ATGATCAGCGCCTGCTGGCCCATGTCAC CCTGGAGCTCCAGCACTATCACCAGCTA CTTGAGAAGGTTCGGATCCGGGATGCCT TGCGCAGTATCCTCACCATATCTCGACA TGGCAACCAATATATTCAGGTGAATGA GCCCTGGAAGCGGATTAAAGGCAGTGA GGCTGACAGGCAACGGGCAGGAACAGT GACTGGCTTGGCAGTGAATATAGCTGCC TTGCTCTCTGTCATGCTTCAGCCTTACAT GCCCACGGTTAGTGCCACAATCCAGGCC CAGCTGCAGCTCCCACCTCCAGCCTGCA GTATCCTGCTGACAAACTTCCTGTGTAC CTTACCAGCAGGACACCAGATTGGCAC AGTCAGTCCCTTGTTCCAAAAATTGGAA AATGACCAGATTGAAAGTTTAAGGCAG CGCTTTGGAGGGGGCCAGGCAAAAACG TCCCCGAAGCCAGCAGTTGTAGAGACT GTTACAACAGCCAAGCCACAGCAGATA CAAGCGCTGATGGATGAAGTGACAAAA CAAGGAAACATTGTCCGAGAACTGAAA GCACAAAAGGCAGACAAGAACGAGGTT GCTGCGGAGGTGGCGAAACTCTTGGAT CTAAAGAAACAGTTGGCTGTAGCTGAG GGGAAACCCCCTGAAGCCCCTAAAGGC	
		AAGAAGAAAAAGTAA	

$MetRS^{C10}$	Proteína / Humana / 1-431 + 457-900	MRLFVSDGVPGCLPVLAAAGRARGRAEV LISTVGPEDCVVPFLTRPKVPVLQLDSGN YLFSTSAICRYFFLLSGWEQDDLTNQWLE WEATELQPALSAALYYLVVQGKKGEDVL GSVRRALTHIDHSLSRQNCPFLAGETESLA DIVLWGALYPLLQDPAYLPEELSALHSWF QTLSTQEPCQRAAETVLKQQGVLALRPYL QKQPQPSPAEGRAVTNEPEEEELATLSEEE IAMAVTAWEKGLESLPPLRPQQNPVLPVA GERNVLITSALPYVNNVPHLGNIIGCVLSA DVFARYSRLRQWNTLYLCGTDEYGTATE TKALEEGLTPQEICDKYHIIHADIYRWFNI SFDIFGRTTTPQQTKITQDIFQQLLKRGFV LQDTVEQLRCEHCARFLADRFVEGVCPFC GYEEARGDQCDKCGKLINAVELKLEKRL EEWLGRTLPGSDWTPNAQFITRSWLRDG LKPRCITRDLKWGTPVPLEGFEDKVFYV WFDATIGYLSITANYTDQWERWWKNPEQ VDLYQFMAKDNVPFHSLVFPCSALGAED NYTLVSHLIATEYLNYEDGKFSKSRGVGV FGDMAQDTGIPADIWRFYLLYIRPEGQDS AFSWTDLLLKNNSELLNNLGNFINRAGMF VSKFFGGYVPEMVLTPDDQRLLAHVTLE LQHYHQLLEKVRIRDALRSILTISRHGNQY IQVNEPWKRIKGSEADRQRAGTVTGLAV NIAALLSVMLQPYMPTVSATIQAQLQLPP PACSILLTNFLCTLPAGHQIGTVSPLFQKL ENDQIESLRQRFGGGQAKTSPKPAVVETV TTAKPQQIQALMDEVTKQGNIVRELKAQ KADKNEVAAEVAKLLDLKKQLAVAEGK PPEAPKGKKKK	SEQ. ID. NO.147
$MetRS^{C10}$	ADN / Humana	ATGAGACTGTTCGTGAGTGATGGCGTCC CGGGTTGCTTGCCGGTGCTGGCCGCCGC CGGGAGAGCCCGGGGCAGAGCAGAGGT GCTCATCAGCACTGTAGGCCCGGAAGA TTGTGTGGTCCCGTTCCTGACCCGGCCT AAGGTCCCTGTCTTGCAGCTGGATAGCG GCAACTACCTCTTCTCCACTAGTGCAAT CTGCCGATATTTTTTTTTGTTATCTGGCT GGGAGCAAGATGACCTCACTAACCAGT GGCTGGAATGGGAAGCGACAGAGCTGC	SEQ. ID. NO.148

		AGCCAGCTTTGTCTGCTGCCCTGTACTA TTTAGTGGTCCAAGGCAAGAAGGGGGA AGATGTTCTTGGTTCAGTGCGGAGAGCC CTGACTCACATTGACCACAGCTTGAGTC GTCAGAACTGTCCTTTCCTGGCTGGGGA GACAGAATCTCTAGCCGACATTGTTTTG TGGGGAGCCCTATACCCATTACTGCAAG ATCCCGCCTACCTCCCTGAGGAGCTGAG TGCCCTGCACAGCTGGTTCCAGACACTG AGTACCCAGGAACCATGTCAGCGAGCT GCAGAGACTGTACTGAAACAGCAAGGT GTCCTGGCTCTCCGGCCTTACCTCCAAA AGCAGCCCCAGCCCAGCCCCGCTGAGG GAAGGGCTGTCACCAATGAGCCTGAGG AGGAGGAGCTGGCTACCCTATCTGAGG AGGAGATTGCTATGGCTGTTACTGCTTG GGAGAAGGGCCTAGAAAGTTTGCCCCC GCTGCGGCCCCAGCAGAATCCAGTGTTG CCTGTGGCTGGAGAAAGGAATGTGCTC ATCACCAGTGCCCTCCCTTACGTCAACA ATGTCCCCCACCTTGGGAACATCATTGG TTGTGTGCTCAGTGCCGATGTCTTTGCC AGGTACTCTCGCCTCCGCCAGTGGAACA CCCTCTATCTGTGTGGGACAGATGAGTA TGGTACAGCAACAGAGACCAAGGCTCT GGAGGAGGGACTAACCCCCCAGGAGAT CTGCGACAAGTACCACATCATCCATGCT GACATCTACCGCTGGTTTAACATTTCGT TTGATATTTTTGGTCGCACCACCACTCC ACAGCAGACCAAAATCACCCAGGACAT TTTCCAGCAGTTGCTGAAACGAGGTTTT GTGCTGCAAGATACTGTGGAGCAACTG CGATGTGAGCACTGTGCTCGCTTCCTGG CTGACCGCTTCGTGGAGGGCGTGTGTCC CTTCTGTGGCTATGAGGAGGCTCGGGGT GACCAGTGTGACAAGTGTGGCAAGCTC ATCAATGCTGTCGAGCTTAAGCTGGAGA AGCGACTGGAGGAGTGGTTGGGGAGGA CATTGCCTGGCAGTGACTGGACACCCAA TGCCCAGTTTATCACCCGTTCTTGGCTTC GGGATGGCCTCAAGCCACGCTGCATAA CCCGAGACCTCAAATGGGGAACCCCTG	

		TACCCTTAGAAGGTTTTGAAGACAAGGT ATTCTATGTCTGGTTTGATGCCACTATT GGCTATCTGTCCATCACAGCCAACTACA CAGACCAGTGGGAGAGATGGTGGAAGA ACCCAGAGCAAGTGGACCTGTATCAGTT CATGGCCAAAGACAATGTTCCTTTCCAT AGCTTAGTCTTTCCTTGCTCAGCCCTAG GAGCTGAGGATAACTATACCTTGGTCAG CCACCTCATTGCTACAGAGTACCTGAAC TATGAGGATGGGAAATTCTCTAAGAGC CGCGGTGTGGGAGTGTTTGGGGACATG GCCCAGGACACGGGGATCCCTGCTGAC ATCTGGCGCTTCTATCTGCTGTACATTC GGCCTGAGGGCCAGGACAGTGCTTTCTC CTGGACGGACCTGCTGCTGAAGAATAA TTCTGAGCTGCTTAACAACCTGGGCAAC TTCATCAACAGAGCTGGGATGTTTGTGT CTAAGTTCTTTGGGGGCTATGTGCCTGA GATGGTGCTCACCCCTGATGATCAGCGC CTGCTGGCCCATGTCACCCTGGAGCTCC AGCACTATCACCAGCTACTTGAGAAGGT TCGGATCCGGGATGCCTTGCGCAGTATC CTCACCATATCTCGACATGGCAACCAAT ATATTCAGGTGAATGAGCCCTGGAAGC GGATTAAAGGCAGTGAGGCTGACAGGC AACGGGCAGGAACAGTGACTGGCTTGG CAGTGAATATAGCTGCCTTGCTCTCTGT CATGCTTCAGCCTTACATGCCCACGGTT AGTGCCACAATCCAGGCCCAGCTGCAG CTCCCACCTCCAGCCTGCAGTATCCTGC TGACAAACTTCCTGTGTACCTTACCAGC AGGACACCAGATTGGCACAGTCAGTCC CTTGTTCCAAAAATTGGAAAATGACCAG ATTGAAAGTTTAAGGCAGCGCTTTGGAG GGGGCCAGGCAAAAACGTCCCCGAAGC CAGCAGTTGTAGAGACTGTTACAACAG CCAAGCCACAGCAGATACAAGCGCTGA TGGATGAAGTGACAAAACAAGGAAACA TTGTCCGAGAACTGAAAGCACAAAAGG CAGACAAGAACGAGGTTGCTGCGGAGG TGGCGAAACTCTTGGATCTAAAGAAAC AGTTGGCTGTAGCTGAGGGGAAACCCC	
		CTGAAGCCCCTAAAGGCAAGAAGAAAA AGTAA	

Tabla 5B Puntos de unión de corte y empalme únicos de polipéptidos de MetRS			
Nombre	**Tipo / especie**	**Secuencias de Aminoácidos y Ácido Nucleico en las cercanías del punto de unión de corte y empalme único**	**SEQ. ID. NO.**
M1-AS01	ADN / Humana /	TCTTCTCCACTAGTGCAATCTGCCG|GTA TTCTATGTCTGGTTTGATGCCA	SEQ. ID. NO.149
	Proteína / Humana /	N/D	
M1-AS06	ADN / Humana /	GCACCACCACTCCACAGCAGACCAA|AA GCCTCAGTGTAAAGTCTGCCGAT	SEQ. ID. NO.150
	Proteína / Humana /	N/D	
M1-AS08	ADN / Humana /	GAGATGGTGGAAGAACCCAGAGCAA|A GTACCTGAACTATGAGGATGGGAA	SEQ. ID. NO.151
	Proteína / Humana /	MVEEPRAKYLNYEDG	SEQ. ID. NO.152
M1-AS09	ADN / Humana /	GGATTAAAGGCAGTGAGGCTGACAG|GT CAGTCCCTTGTTCCAAAAATTGG	SEQ. ID. NO.153
	Proteína / Humana /	N/D	
M1-AS12	ADN / Humana /	CTCATCAGCACTGTAGGCCCGGAAG|AG GAGCTGAGTGCCCTGCACAGCTG	SEQ. ID. NO.154
	Proteína / Humana /	LISTVGPEEELSALHS	SEQ. ID. NO.155
M1-AS13	ADN / Humana /	TCTTCTCCACTAGTGCAATCTGCCG|AGG AGCTGAGTGCCCTGCACAGCTG	SEQ. ID. NO.156
	Proteína / Humana /	N/D	
M1-AS14	ADN / Humana /	AAGGGCTGTCACCAATGAGCCTGAG|GT TGCCTGTGGCTGGAGAAAGGAAT	SEQ. ID. NO.157
	Proteína / Humana /	GLSPMSLRLPVAGERN	SEQ. ID. NO.158
M1-AS16	ADN / Humana /	TGCTCAGTGCCGATGTCTTTGCCAG|AAT CACCCAGGACATTTTCCAGCAG	SEQ. ID. NO.159
	Proteína / Humana /	LSADVFARITQDIFQQ	SEQ. ID. NO.160
M1-AS17	ADN / Humana /	GCTCATCAATGCTGTCGAGCTTAAG|CTG GAGAAGCGACTGGAGGAGTGGT	SEQ. ID. NO.161
	Proteína / Humana /	LINAVELKLEKRLEEW	SEQ. ID. NO.162

Tabla 6 Polipéptidos y ácidos nucleicos de MetRS identificados por Bioinformática			
Nombre	**Tipo / especie / Restos**	**Secuencias de Aminoácidos y Ácido Nucleico**	**SEQ. ID. NO.**
$MetRS^{C2}$	Proteína / Humana / 773-900	LPPPACSILLTNFLCTLPAGHQIGTVSPLF QKLENDQIESLRQRFGGGQAKTSPKPAV VETVTTAKPQQIQALMDEVTKQGNIVRE LKAQKADKNEVAAEVAKLLDLKKQLAV AEGKPPEAPKGKKKK	SEQ. ID. NO.163
$MetRS^{C2}$	ADN / Humana	CTCCCACCTCCAGCCTGCAGTATCCTGC TGACAAACTTCCTGTGTACCTTACCAGC AGGACACCAGATTGGCACAGTCAGTCC CTTGTTCCAAAAATTGGAAAATGACCA GATTGAAAGTTTAAGGCAGCGCTTTGG AGGGGGCCAGGCAAAAACGTCCCCGAA GCCAGCAGTTGTAGAGACTGTTACAAC AGCCAAGCCACAGCAGATACAAGCGCT GATGGATGAAGTGACAAAACAAGGAA ACATTGTCCGAGAACTGAAAGCACAAA AGGCAGACAAGAACGAGGTTGCTGCGG AGGTGGCGAAACTCTTGGATCTAAAGA AACAGTTGGCTGTAGCTGAGGGGAAAC CCCCTGAAGCCCCTAAAGGCAAGAAGA	SEQ. ID. NO.164
		AAAAGTAA	

"Fragmentos proteicos", o la secuencia de aminoácidos de fragmentos proteicos, tales como fragmentos proteolíticos o fragmentos de variantes de corte y empalme, pueden caracterizarse, identificarse o derivarse de acuerdo con diversas técnicas. Por ejemplo, pueden identificarse variantes de corte y empalme por técnicas tales como secuenciación profunda (véase, por ejemplo, Xing *et al*., ARN 14: 1470-1479, 2008, y Zhang *et al*., Genome Research, 17:503-509, 2007). Como ejemplo adicional, pueden identificarse fragmentos proteicos tales como fragmentos proteolíticos *in vitro*, tal como incubando polipéptidos de MetRS de longitud completa u otros con proteasas seleccionadas, o pueden identificarse de forma endógena (por ejemplo, *in vivo*). En ciertas realizaciones, pueden generarse o identificarse fragmentos proteicos tales como fragmentos proteolíticos endógenos, por ejemplo, expresando de forma recombinante polipéptidos de MetRS de longitud completa u otros en un microorganismo seleccionado o una célula eucariota que se ha modificado para contener una o más proteasas seleccionadas o que contiene de forma natural una o más proteasas que son capaces de actuar en un polipéptido de MetRS seleccionado, y aislar y caracterizar los fragmentos proteicos producidos de forma endógena del mismo.

En ciertas realizaciones, los fragmentos proteicos tales como fragmentos proteolíticos endógenos (por ejemplo, de origen natural) pueden generarse o identificarse, por ejemplo, de diversas fracciones celulares (por ejemplo, citosólicas, de membrana, nucleares) y/o medio de cultivo de diversos tipos celulares, incluyendo, por ejemplo, células inmunitarias tales como monocitos, células dendríticas, macrófagos (por ejemplo, macrófagos RAW 264.7), neutrófilos, eosinófilos, basófilos y linfocitos, tales como linfocitos B y linfocitos T (por ejemplo, linfocitos auxiliares CD4+ y citolíticos CD8+), incluyendo linfocitos T primarios y líneas de linfocitos T, tales como linfocitos T de Jurkat, así como linfocitos citolíticos naturales (NK).

En ciertas realizaciones, pueden identificarse fragmentos proteicos tales como fragmentos proteolíticos endógenos, generados de cualquier modo, por técnicas tales como espectrometría de masas, o técnicas equivalentes. Una vez que se ha generado o identificado un fragmento proteico *in vitro* o identificado de forma endógena, este puede mapearse o secuenciarse y, por ejemplo, clonarse en un vector de expresión para producción recombinante, o producirse de forma sintética.

Puede usarse una amplia diversidad de proteasas para producir, identificar, derivar o caracterizar la secuencia de fragmentos proteicos de MetRS tales como fragmentos proteolíticos. En general, las proteasas se clasifican habitualmente según tres criterios principales: (i) la reacción catalizada, (ii) la naturaleza química del sitio catalítico y (iii) la relación evolutiva, como se revela por la estructura. Los ejemplos generales de proteasas o proteinasas, como se clasifican por mecanismo de catálisis, incluyen proteasas aspárticas, serina proteasas, cisteína proteasas y metaloproteasas.

La mayoría de proteasas aspárticas pertenecen a la familia de pepsina. Esta familia incluye enzimas digestivas tales como pepsina y quimosina, así como catepsinas D lisosómicas y enzimas de procesamiento tales como renina, y ciertas proteasas fúngicas (por ejemplo, penicillopepsina, rhizopuspepsina, endotiapepsina). Una segunda familia de proteasas aspárticas incluye proteinasas víricas tales como la proteasa del virus del SIDA (VIH), también denominada retropepsina.

Las serinas proteasas incluyen dos familias distintas. En primer lugar, la familia de quimotripsina, que incluye las enzimas de mamíferos tales como quimotripsina, tripsina, elastasa y calicreína, y en segundo lugar, la familia de subtilisina, que incluye las enzimas bacterianas tales como subtilisina. La estructura tridimensional general entre estas dos familias es diferente, pero tienen la misma geometría de sitio activo, y la catálisis se produce por el mismo mecanismo. Las serina proteasas muestran especificidades de sustrato diferentes, diferencias que se relacionan principalmente con sustituciones de aminoácidos en los diversos subsitios enzimáticos (sitios de interacción con restos de sustrato). Algunas serina proteasas tienen un sitio de interacción extendido con el sustrato mientras que otras tienen una especificidad que está restringida al resto del sustrato P1.

La familia de cisteína proteasa incluye las proteasas vegetales tales como papaína, actinidina y bromelina, varias catepsinas lisosómicas de mamífero, las calpaínas citosólicas (activadas por calcio), así como varias proteasas parasitarias (por ejemplo, *Trypanosoma, Schistosoma*). Papaína es el arquetipo y el miembro mejor estudiado de la familia. La reciente determinación de la estructura por rayos X de la enzima conversora de interleucina-1-beta ha revelado un nuevo tipo de plegamiento para las cisteína proteasas.

Las metaloproteasas son una de las clases más antiguas de proteasas, halladas en bacterias, hongos y organismos superiores. Difieren ampliamente en sus secuencias y sus estructuras tridimensionales, pero la gran mayoría de enzimas contienen un átomo de zinc que es catalíticamente activo. En algunos casos, el zinc puede reemplazarse por otro metal tal como cobalto o níquel sin pérdida de la actividad proteolítica. La termolisina bacteriana se ha caracterizado bien y su estructura cristalográfica indica que el zinc está unido con dos histidinas y un ácido glutámico. Muchas metaloproteasas contienen el motivo de secuencia HEXXH, que proporciona dos ligandos de histidina para el zinc. El tercer ligando es un ácido glutámico (termolisina, neprilisina, alanil aminopeptidasa) o una histidina (astacina, serralisina).

Las proteasas ilustrativas incluyen, por ejemplo, acromopeptidasa, aminopeptidasa, ancrod, enzima conversora de angiotensina, bromelaína, calpaína, calpaína I, calpaína II, carboxipeptidasa A, carboxipeptidasa B, carboxipeptidasa G, carboxipeptidasa P, carboxipeptidasa W, carboxipeptidasa Y, caspasa 1, caspasa 2, caspasa 3, caspasa 4, caspasa 5, caspasa 6, caspasa 7, caspasa 8, caspasa 9, caspasa 10, caspasa 11, caspasa 12, caspasa 13, catepsina B, catepsina C, catepsina D, catepsina E , catepsina G, catepsina H, catepsina L, quimopapaína, quimasa, quimotripsina, clostripaína, colagenasa, complemento C1r, complemento C1s, factor de complemento D, factor de complemento I, cucumisina, dipeptidil peptidasa IV, elastasa (leucocito), elastasa (pancreática), endoproteinasa Arg-C, endoproteinasa Asp-N, endoproteinasa Glu-C, endoproteinasa Lys-C, enteroquinasa, factor Xa, ficina, furina, granzima A, granzima B, proteasa del VIH, IGasa, calicreína tisular, leucina aminopeptidasa (general), leucina aminopeptidasa (citosol), leucina aminopeptidasa (microsómica), metaloproteasa de matriz, metionina aminopeptidasa, neutrasa, papaína, pepsina, plasmina, prolidasa, pronasa E, antígeno específico de próstata, proteasa alcalófila de *Streptomyces griseus*, proteasa *de Aspergillus*, proteasa de *Aspergillus saitoi*, proteasa de *Aspergillus sojae*, proteasa (*B. licheniformis*) (alcalina o alcalasa), proteasa de *Bacillus polymyxa*, proteasa de *Bacillus sp*, proteasa de *Rhizopus sp*., proteasa S, proteasomas, proteinasa de *Aspergillus oryzae*, proteinasa 3, proteinasa A, proteinasa K, proteína C, piroglutamato aminopeptidasa, renina, estreptoquinasa, subtilisina, termolisina, trombina, activador de plasminógeno tisular, tripsina, triptasa y uroquinasa.

Ciertas realizaciones se refieren a polipéptidos de MetRS aislados, que comprenden, que consisten esencialmente en, o que consisten en secuencias de aminoácidos que se han derivado de fragmentos polipeptídicos de MetRS de origen natural, endógenos y composiciones farmacéuticas que comprenden dichos fragmentos, y métodos de uso de los mismos. Estas realizaciones y otras relacionadas pueden generarse o identificarse *in vivo*, *ex vivo* y/o *in vitro*. En ciertas realizaciones *in vitro* preferidas, se generan o identifican fragmentos proteolíticos de MetRS incubando un polipéptido de MetRS, tal como un polipéptido de MetRS de longitud completa, con una o más proteasas humanas aisladas, principalmente las proteasas que son endógenas o naturales de seres humanos, tales como elastasa y otras descritas en el presente documento y conocidas en la técnica. Otras realizaciones se refieren a polipéptidos de MetRS aislados, que comprenden, consisten esencialmente en o consisten en secuencias de aminoácidos que se han derivado de variantes de corte y empalme de MetRS de origen natural, endógenas, y composiciones farmacéuticas que comprenden dichos fragmentos, y métodos de uso de los mismos. Esencialmente, un fragmento proteico de MetRS puede aislarse de muestras que se han expuesto a proteasas, bien *in vivo* o bien *in vitro*.

En ciertas realizaciones, pueden identificarse fragmentos proteicos de MetRS mediante técnicas tales como espectrometría de masas, o técnicas equivalentes. Únicamente como ilustración y no como limitación, en ciertas realizaciones los proteomas de diversos tipos celulares, tejidos o fluidos corporales de diversos estados fisiológicos (por ejemplo, hipoxia, dieta, edad, enfermedad) o fracciones de los mismos pueden separarse por SDS-PAGE 1D y los carriles de los geles cortarse en bandas a intervalos fijos; después de lo cual las bandas pueden digerirse opcionalmente con una proteasa apropiada, tal como tripsina, para liberar los péptidos, que pueden después

analizarse por CL-EM/EM de fase inversa 1D. Los datos proteómicos resultantes pueden integrarse en los denominados peptógrafos que representan, en el panel izquierdo, la cobertura de secuencia para una proteína dada en la dimensión horizontal (del extremo N a C terminal, de izquierda a derecha) frente a migración en SDS-PAGE en dimensión vertical (peso molecular de alto a bajo, de arriba a abajo). Los fragmentos peptídicos específicos pueden después secuenciarse o mapearse. En ciertas realizaciones, el fragmento de referencia de MetRS puede caracterizarse por su peso molecular único, en comparación, por ejemplo, con el peso molecular de la MetRS de longitud completa correspondiente.

Como se indica anteriormente, las realizaciones de la presente invención incluyen los polipéptidos MetRS expuestos en SEQ ID NO: 137. También se incluyen "variantes" de los polipéptidos de referencia de MetRS. La indicación "variante" polipeptídica se refiere a polipéptidos que se distinguen de un polipéptido de MetRS de referencia mediante la adición, deleción y/o sustitución de al menos un resto de aminoácido, y que normalmente conservan (por ejemplo, imitan) o modulan (por ejemplo, antagonizan) una o más actividades no canónicas de un polipéptido de MetRS de referencia.

Además, las Metionil ARNt sintetasas humanas incluyen varios cientos de formas polimórficas altamente relacionadas, y se sabe en la técnica que estas son al menos parcialmente funcionalmente intercambiables. Sería por lo tanto un asunto rutinario seleccionar una variante de origen natural de Metionil ARNt sintetasa, incluyendo, por ejemplo, las formas polimórficas de un único nucleótido enumeradas en la **Tabla A** para crear un polipéptido MetRS que contiene uno o más cambios de aminoácidos basados en la secuencia de cualquiera de los homólogos, ortólogos e isoformas de origen natural de ser humano así como otras especies de Metionil ARNt sintetasa.

Tabla A
SNP de MetionilMetionil ARNt sintetasa humana

Número de Referencia de GenBank	Cambio de Nucleótidos	Número de Referencia de GenBank	Cambio de Nucleótidos
rs118120417	A/C	rs36085497	A/G
rs118042640	C/T	rs36052671	-/A
rs117948353	A/G	rs36019863	-/C
rs117914586	A/G	rs35924774	-/A
rs117912819	G/T	rs35843015	C/T
rs117895855	C/T	rs35786533	-IT
rs117833843	C/T	rs35674919	-IT
rs117762340	A/G	rs35515685	-/A
rs117659381	A/C	rs35405310	A/G
rs117633211	A/G	rs35354985	-/C
rs117476973	A/G	rs35152353	-/T
rs117420860	C/T	rs35066896	-/G
rs117385528	C/T	rs34943124	G/T
rs117328675	C/T	rs34537592	-/T
rs117316797	A/G	rs34526594	-/A
rs117101415	C/T	rs34351056	-IT
rs117093278	C/T	rs34251115	-/A
rs116975692	A/T	rs34238800	-IT
rs116920370	C/T	rs34010387	-U
rs116899349	G/T	rs28382855	C/T
rs116795882	A/C	rs28382854	G/T
rs116714288	A/G	rs17119735	A/G
rs116260314	A/G	rs12829556	C/T
rs116178189	A/C	rs12579188	C/T
rs116119105	C/T	rs12427340	A/T
rs115823928	C/G	rs12425095	A/G
rs115762355	A/T	rs12372044	A/G
rs115447727	A/G	rs12372043	A/G
rs115402894	C/T	rs12371589	A/T
rs115111847	A/T	rs12367179	G/T
rs114933325	C/T	rs12366545	A/T
rs114852246	A/C	rs12315660	A/G
rs114790812	C/T	rs12298311	A/G
rs114718314	A/T	rs11831454	A/G
rs114665682	A/G	rs11831009	G/T
rs114609270	A/G	rs11616111	C/T
rs114325763	C/T	rs11614017	C/T
rs114312802	A/G	rs11547020	C/T

rs113957069	C/T	rs11540811	G/T
rs113917592	C/T	rs11540809	C/G
rs113908259	C/T	rs11540808	A/T
rs113808165	A/G	rs11172245	C/T
rs113796766	A/T	rs11172243	A/T
rs113773319	C/G	rs11172241	C/T
rs113617531	C/G	rs11172240	C/T
rs113549256	C/T	rs11172239	C/T
rs113274678	C/T	rs11172237	A/T
rs113263363	-/C	rs11172236	A/T
rs113171815	A/G	rs11172234	C/G
rs113090086	C/G	rs11172233	A/G
rs112917609	C/T	rs11172232	A/G
rs112829410	A/T	rs11172231	A/G
rs112799168	A/G	rs10783829	A/G
rs112714462	A/G	rs10715375	-/A
rs112710318	-/A	rs71084758	-IT
rs112687648	-/GA	rs68188746	-ITITT
rs112678779	C/T	rs68097611	-IT
rs112660326	A/G	rs67410914	-ITG
rs112404047	A/G	rs67291515	-/A
rs112327724	G/T	rs61935708	A/G
rs112196139	C/T	rs61935705	A/T
rs112116601	A/C	rs61935704	C/T
rs111995146	C/T	rs61935703	G/T
rs111987764	A/G	rs61935702	G/T
rs111969677	A/G	rs61935701	C/T
rs111812045	A/T	rs61349055	C/T
rs111647559	A/C	rs61159732	A/T
rs111609806	A/T	rs60709374	A/T
rs111515147	C/T	rs60164085	-/T
rs111451812	A/G	rs60038237	-ITT
rs111321301	C/T	rs59457163	C/T
rs111306654	-/G	rs59160934	-/TTT
rs111287174	C/T	rs59108082	G/T
rs111257978	A/C	rs59072876	A/T
rs111226219	C/G	rs58998083	G/T
rs80291028	G/T	rs58844705	A/T
rs80290696	G/T	rs58533497	A/T
rs79885244	G/T	rs58353609	A/T
rs79870226	A/G	rs58338213	-/T
rs79791981	C/G	rs58223781	-IT
rs79748949	C/T	rs58057273	-IGT
rs79580523	A/T	rs57976382	-/ATATA
rs79531790	C/T	rs57959685	A/T
rs79354820	C/T	rs57934778	-/AATA
rs78854130	A/C	rs57598635	A/T
rs78617675	C/T	rs57586332	A/T
rs78563683	C/T	rs57440582	A/T
rs78509000	A/C	rs57411748	C/T
rs78451666	A/G	rs57329666	C/T
rs78442993	C/T	rs57276054	-IT
rs78377626	A/G	rs56986351	A/T
rs78220036	C/T	rs56354627	C/G
rs78069812	A/G	rs56198826	C/T
rs78027605	A/G	rs56189503	-/A
rs77981382	C/T	rs56126868	-/TAC
rs77902929	-/TAGTT	rs56071532	-/AA
rs77631229	C/G	rs56060087	-/CA
rs77551014	C/T	rs56032642	A/G
rs77389838	A/G	rs55985221	-/A
rs77298381	A/G	rs55964032	-/G
rs77159901	A/G	rs55901364	A/T

rs76780434	-/AAA	rs55826438	-/T
rs76763017	C/G	rs55813010	A/G
rs76608716	A/G	rs55784218	-/TAT
rs76535211	C/G	rs55757029	-/TA
rs76515218	C/T	rs55742554	-/TATA
rs76477870	A/G	rs55727879	A/T
rs76399088	A/G	rs55702220	-/A
rs76305181	G/T	rs9737891	A/G
rs76150936	C/G	rs7969683	C/T
rs76119157	A/G	rs7485148	G/T
rs75719980	A/G	rs7314803	A/T
rs75683176	A/G	rs7304550	A/G
rs75617785	G/T	rs7300617	A/C
rs75549553	A/G	rs7134758	C/T
rs75425333	C/T	rs5029303	C/T
rs75308443	A/T	rs4760273	C/T
rs75040517	A/G	rs4583018	C/T
rs74978137	A/T	rs2943696	A/G
rs74974479	C/T	rs2928382	C/T
rs74921288	C/T	rs2672571	C/T
rs74875570	C/T	rs2620680	A/G
rs74832951	A/G	rs2620676	A/T
rs74683277	C/T	rs2600004	A/G
rs74625799	C/T	rs2599998	A/G
rs74520477	C/G	rs2459050	C/T
rs73344102	A/G	rs2305390	A/G
rs73344100	G/T	rs2290297	A/G
rs73344088	A/G	rs2278846	C/T
rs73344082	C/T	rs1871547	A/T
rs72386117	-/TT	rs1709874	C/T
rs72343151	-/CA	rs1709871	A/C
rs72306101	-/AT	rs1679366	A/G
rs72264062	-/AT	rs1679365	A/G
rs72260515	-/TATA	rs1678537	A/G
rs72229554	-/AT	rs 1678510	G/T
rs72141714	-/AT	rs1669294	C/T
rs72047411	-/TG	rs1669293	A/C
rs72039971	-/AA	rs1669292	A/C
rs71858271	-/A	rs1669291	A/C
rs71847544	-/CA	rs 1629410	C/G
rs71770989	-/GA	rs1299737	C/T
rs71622721	C/T	rs1284606	A/G
rs71448526	-IT	rs1284467	C/T
rs71446621	A/G	rs1284465	A/G
rs71445351	C/G	rs1054519	A/C
rs71445350	C/T	rs 1054403	A/C/G
rs71445349	A/G	rs899653	C/T
rs71445348	A/G	rs796546	C/T
rs71432309	A/G	rs776039	A/G
rs71432308	A/G	rs776038	G/T
rs71360927	A/G	rs747341	C/G
rs71317711	C/T	rs747340	C/G
rs71280719	-IT	rs747339	C/T
rs71280718	-IT	rs537161	C/T
rs71084766	-IT	rs508904	A/G
rs71084765	-/A	rs496245	C/G
rs71084764	-/G	rs474738	C/T
rs71084763	-/A	rs474544	C/T
rs71084762	-/A		
rs71084761	-IT		
rs71084759	-/A		
Número de Referencia de GenBank	Cambio de Nucleótidos		

rs76311414	-/ATATATATTTTTTAATTTTAATTTT
rs55784614	-/TTATAATATATA
rs57783219	-/TAGAGGGTAGAGGGTAGAGG
rs58931225	-/CACACACACA
rs67907939	-/TATATAT
rs59488863	-/TATATAT
rs71276496	/GAGGTTGGGGGGTCAGCCCCCCGCCCGGCCAGCTGC CCCGTCCGGGAGGGAGGTTGAGGGGTCAGCCCCCCGC CCGGCCAGCCGCCCCCGCCCGGGAGGGAGGTTGGGGG GTCAGCCCCCCGCCCGGCCAGCCGCCCCGTCCGGGAG GGAGGTTGGGGGGTCAGCCCCCCGCCCGGCCAGCCGC CCCGTCCGGGAGGGAGGTGGGGGGGTCAGCCCCCCCG CCCGGCCAGCCGCCCCGCCCGGGAGG
rs71084760	-/TATATATTA
rs71280717	-/TCTACC
rs71759599	(DELECIÓN GRANDE)
rs71658612	(DELECIÓN GRANDE)
rs71931724	-/ACACACAC
rs72270156	-/CACACACACACACACACACACA

En ciertas realizaciones, una variante polipeptídica se distingue de un polipéptido de referencia en una o más sustituciones, que pueden ser conservativas o no conservativas, como se describe en el presente documento y se conoce en la técnica. En ciertas realizaciones, la variante polipeptídica comprende sustituciones conservativas y, a este respecto, se entiende bien en la técnica que algunos aminoácidos pueden cambiarse a otros con propiedades ampliamente similares sin cambiar la naturaleza de la actividad del polipéptido.

En ciertas realizaciones, una variante polipeptídica incluye una secuencia de aminoácidos que tiene al menos aproximadamente 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o más identidad de secuencia o similitud con una secuencia correspondiente de un polipéptido de referencia de MetRS, como se describe en el presente documento, y conserva sustancialmente la actividad no canónica de ese polipéptido de referencia. También se incluyen secuencias que difieren de las secuencias de MetRS de referencia por la adición, deleción o sustitución de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150 o más aminoácidos pero que conservan las propiedades del polipéptido de MetRS de referencia. En ciertas realizaciones, las adiciones o deleciones de aminoácidos suceden en el extremo C-terminal y/o en el extremo N-terminal del polipéptido de referencia de MetRS. En ciertas realizaciones, las adiciones de aminoácidos incluyen 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50 o más restos de tipo silvestre (es decir, del polipéptido de MetRS de longitud completa correspondiente) que están próximos al extremo C-terminal y/o el extremo N-terminal del polipéptido de referencia de MetRS.

En ciertas realizaciones, los polipéptidos variantes difieren de las secuencias de referencia de MetRS correspondientes en al menos un 1 %, pero menos del 20 %, 15 %, 10 % o 5 % de los restos. (Si esta comparación requiere alineamiento, las secuencias deberían alinearse para similitud máxima. Las secuencias “omitidas” de deleciones o inserciones, o desapareamientos, se consideran diferencias). Las diferencias son, convenientemente, las diferencias o cambios en un resto no esencial o una sustitución conservativa. En ciertas realizaciones, el peso molecular de un polipéptido de MetRS variante difiere del polipéptido de referencia de MetRS en aproximadamente 2 %, 3 %, 4 %, 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 4 %, 5 %, 6 %, 7 %, 8 % %, 11 %, 12 %, 13 %, 14 %, 15 %, 16 %, 17 %, 18 %, 19 %, 20 % o más.

También se incluyen "fragmentos" biológicamente activos de los polipéptidos de referencia de MetRS, es decir, fragmentos biológicamente activos de los fragmentos proteicos de MetRS. Los fragmentos biológicamente activos representativos participan en general en una interacción, por ejemplo, una interacción intramolecular o una interacción intermolecular. Una interacción intermolecular puede ser una interacción de unión específica o una interacción enzimática. Una interacción intermolecular puede ser entre un polipéptido de MetRS y un compañero de unión celular, tal como un receptor celular u otra molécula hospedadora que participa en la actividad no canónica del polipéptido de MetRS. En algunas realizaciones, proteínas de MetRS, variantes y fragmentos biológicamente activos de las mismas, se unen a uno o más compañeros de unión celular con una afinidad de al menos aproximadamente 0,01, 0,05, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40 o 50 nM. La afinidad de unión de un fragmento proteico de MetRS por un compañero de unión celular seleccionado, particularmente un compañero de unión que participa en una actividad no canónica, es normalmente más fuerte que la del polipéptido de MetRS de longitud completa correspondiente al fragmento proteico de MetRS, en al menos aproximadamente 1,5x, 2x, 2,5x, 3x, 3,5x, 4x, 4,5x, 5x, 6x, 7x, 8x, 9x, 10x, 15x, 20x, 25x, 30x, 40x, 50x, 60x, 70x, 80x, 90x, 100x, 200x, 300x, 400x, 500x, 600x, 700x, 800x, 900x, 1.000x

o más (incluyendo todos los números enteros entre medias). La afinidad de unión de un fragmento proteico de MetRS por un compañero de unión que participa en al menos una actividad canónica de un MetRS es normalmente más débil que la del polipéptido de MetRS de longitud completa correspondiente al fragmento proteico de MetRS, en al menos aproximadamente 1,5x, 2x, 2,5x, 3x, 3,5x, 4x, 4,5x, 5x, 6x, 7x, 8x, 9x, 10x, 15x, 20x, 25x, 30x, 40x, 50x, 60x, 70x, 80x, 90x, 100x, 200x, 300x, 400x, 500x, 600x, 700x, 800x, 900x, 1.000x o más.

Normalmente, los fragmentos biológicamente activos comprenden un dominio o motivo con al menos una actividad de un polipéptido de referencia de MetRS y pueden incluir uno o más (y en algunos casos todos) de los diversos dominios activos, e incluyen fragmentos que tienen una actividad no canónica. En algunos casos, los fragmentos biológicamente activos de un polipéptido de MetRS tienen una actividad biológica que es única del fragmento truncado, particular, de modo que el polipéptido de MetRS de longitud completa puede no tener esa actividad. En ciertos casos, la actividad biológica puede revelarse separando el fragmento polipeptídico de MetRS biológicamente activo de las otras secuencias polipeptídicas de MetRS de longitud completa, o alterando ciertos restos de la secuencia polipeptídica de tipo silvestre de MetRS de longitud completa para desenmascarar los dominios biológicamente activos.

Un fragmento biológicamente activo de un polipéptido de referencia de MetRS puede ser un fragmento polipeptídico que es, por ejemplo, de 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 450, 500, 550, 600, 650, 700, 750 o más aminoácidos contiguos o no contiguos, incluyendo todos los números enteros (por ejemplo, 101, 102, 103) e intervalos (por ejemplo, 50-100, 50-150, 50-200) entre medias, de las secuencias de aminoácidos expuestas en uno cualquiera de los polipéptidos de referencia MetRS descritos en el presente documento, pero normalmente excluyen la MetRS de longitud completa. En ciertas realizaciones, un fragmento biológicamente activo comprende una secuencia, un dominio o un motivo no relacionado con actividad no canónica. En ciertas realizaciones, la región C-terminal o N-terminal de cualquier polipéptido de referencia de MetRS puede estar truncada en aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650 o 700 o más aminoácidos, o en aproximadamente 10-50, 20-50, 50-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 400-450, 450-500, 500-550, 550-600, 600-650, 650-700 o más aminoácidos, incluyendo todos los números enteros e intervalos entre medias (por ejemplo, 101, 102, 103, 104, 105), siempre que los polipéptidos de MetRS truncados conserven la actividad no canónica del polipéptido de referencia. normalmente, el fragmento biológicamente activo no tiene menos de aproximadamente 1 %, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 25 % o aproximadamente 50 % de una actividad del polipéptido de referencia de MetRS biológicamente activo (es decir, actividad no canónica) del que deriva. Se describen en los Ejemplos métodos ilustrativos para medir las actividades no canónicas.

Como se ha indicado anteriormente, un polipéptido de MetRS puede alterarse de diversas maneras incluyendo sustituciones, deleciones, truncamientos e inserciones de aminoácidos. Se conocen en general en la técnica métodos para dichas manipulaciones. Por ejemplo, pueden prepararse variantes de secuencia de aminoácidos de un polipéptido de referencia de MetRS por mutaciones en el ADN. Se conocen bien en la técnica métodos para mutagénesis y alteraciones de la secuencia de nucleótidos. Véase, por ejemplo, Kunkel (1985, Proc. Natl. Acad. Sci. USA. 82: 488-492), Kunkel *et al*., (1987, Methods in Enzymol, 154: 367-382), Patente de Estados Unidos N.º 4.873.192, Watson, J. D. *et al*., ("Molecular Biology of the Gene", Cuarta Edicióin, Benjamin/Cummings, Menlo Park, Calif., 1987) y las referencias citadas en las mismas. Puede encontrarse orientación con respecto a sustituciones de aminoácidos apropiadas que no afectan a actividad biológica de la proteína en el modelo de Dayhoff *et al*., (1978) Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found., Washington, D.C.).

De forma similar está dentro de la experiencia de la técnica abordar y/o mitigar las preocupaciones de inmunogenicidad si surgen usando un polipéptido de MetRS, por ejemplo, mediante el uso de programas de reconocimiento informático automáticos para identificar epítopos de linfocitos T potenciales, y enfoques de evolución dirigida para Identificar formas menos inmunogénicas.

Se conocen en la técnica métodos para explorar productos génicos de bibliotecas combinatorias preparadas por mutaciones puntuales o truncamiento y para explorar bibliotecas de ADNc con respecto a productos génicos que tengan una propiedad seleccionada. Dichos métodos son adaptables para cribado rápida de las bibliotecas génicas generadas por mutagénesis combinatoria de polipéptidos de MetRS. Puede usarse mutagénesis de conjunto recursiva (REM), una técnica que potencia la frecuencia de mutantes funcionales en las bibliotecas, en combinación con los ensayos de cribado para identificar variantes polipeptídicas de MetRS (Arkin y Yourvan, Proc. Nat. Acad. 89:7811-7815, Delgrave *et al*., (1993) Protein Engineering, 6:327-331). Como se analiza en más detalle posteriormente pueden ser deseables sustituciones conservativas, tales como intercambiar un aminoácido con otro que tenga propiedades similares.

Los polipéptidos de MetRS variantes y/o truncados biológicamente activos pueden contener substituciones de aminoácidos conservativas en diversas localizaciones a lo largo de su secuencia, en comparación con un resto de aminoácido de MetRS de referencia. Adicionalmente, se han secuenciado variantes de origen natural de proteínas de MetRS y se sabe en la técnica que son al menos parcialmente funcionalmente intercambiables. Sería por lo tanto

un asunto rutinario seleccionar una posición de aminoácido para introducir una mutación conservativa o no conservativa en un polipéptido de MetRS basándose en variaciones de secuencia de origen natural entre los homólogos, ortólogos e isoformas de origen natural de proteína de MetRS de ser humano así como otras especies de una proteína de MetRS.

Una "sustitución de aminoácidos conservativa" es una en la que el resto de aminoácido se remplaza con un resto de aminoácido que tiene una cadena lateral similar. Se han definido en la técnica familias de restos de aminoácidos que tienen cadenas laterales similares, que pueden subclasificarse en general de la siguiente manera:

Ácidos: El resto tiene una carga negativa debido a una pérdida del ion H a pH fisiológico y el resto es atraído por solución acuosa para buscar las posiciones en superficie en la conformación de un péptido en el que está contenido cuando el péptido está en un medio acuoso a pH fisiológico. Los aminoácidos que tienen una cadena lateral ácida incluyen ácido glutámico y ácido aspártico.

Básicos: El resto tiene una carga positiva debido a la asociación con ion H a pH fisiológico o dentro de una o dos unidades de pH del mismo (por ejemplo, histidina) y el resto es atraído por solución acuosa para buscar las posiciones en superficie en la conformación de un péptido en el que está contenido cuando el péptido está en medio acuoso a pH fisiológico. Los aminoácidos que tienen una cadena lateral básica incluyen arginina, lisina e histidina.

Con carga: Los restos tienen carga a pH fisiológico y, por lo tanto, incluyen aminoácidos que tienen cadenas laterales ácidas o básicas (es decir, ácido glutámico, ácido aspártico, arginina, lisina e histidina).

Hidrófobos: Los restos no tienen carga a pH fisiológico y el resto es repelido por solución acuosa para buscar las posiciones internas en la conformación de un péptido en el que está contenido cuando el péptido está en un medio acuoso. Los aminoácidos que tienen una cadena lateral hidrófoba incluyen tirosina, valina, isoleucina, leucina, metionina, fenilalanina y triptófano.

Neutros/polares: Los restos no tienen carga a pH fisiológico, pero el resto no es repelido lo suficiente por las soluciones acuosas de modo que buscaría posiciones internas en la conformación de un péptido en el que esté contenido cuando el péptido está en medio acuoso. Los aminoácidos que tienen una cadena lateral neutra/polar incluyen asparagina, glutamina, cisteína, histidina, serina y treonina.

Esta descripción también caracteriza ciertos aminoácidos como "pequeños" ya que sus cadenas laterales no son suficientemente grandes, incluso si carecen de grupos polares, para conferir hidrofobicidad. Con la excepción de prolina, los aminoácidos "pequeños" son los que tienen cuatro carbonos o menos cuando al menos un grupo polar está en la cadena lateral y tres carbonos o menos cuando no. Los aminoácidos que tienen una cadena lateral pequeña incluyen glicina, serina, alanina y treonina. El aminoácido secundario codificado por genes prolina es un caso especial debido a sus efectos conocidos en la conformación secundaria de cadenas peptídicas. La estructura de prolina difiere de todos los otros aminoácidos de origen natural porque su cadena lateral está unida con el nitrógeno en el grupo a-amino, así como el a-carbono. Se conocen en la técnica varias matrices de similitud de aminoácidos (véase, por ejemplo, matriz PAM120 y matriz PAM250 como se desvela por ejemplo en Dayhoff *et al*., 1978, A model of evolutionary change in proteins). Las matrices para determinar las relaciones de distancia en M. O. Dayhoff, (ed.), Atlas of protein sequence and structure, Vol. 5, págs. 345-358, National Biomedical Research Foundation, Washington DC; y por Gonnet *et al*., (Science, 256:14430-1445, 1992), sin embargo, incluyen prolina en el mismo grupo que glicina, serina, alanina y treonina. En consecuencia, para los fines de la presente invención, la prolina se clasifica como un aminoácido "pequeño".

El grado de atracción o repulsión requerido para la clasificación como polar o no polar es arbitrario y, por lo tanto, los aminoácidos contemplados específicamente por la invención se han clasificado como uno o el otro. La mayoría de aminoácidos no nombrados específicamente pueden clasificarse basándose en el comportamiento conocido.

Los restos de aminoácidos pueden subclasificarse adicionalmente como cíclicos o no cíclicos, y aromáticos o no aromáticos, clasificaciones autoexplicativas con respecto a los grupos sustituyentes de cadena lateral de los restos, y como pequeños o grandes. El resto se considera pequeño si contiene un total de cuatro átomos de carbono o menos, incluyendo el carbono carboxilo, siempre que esté presente un sustituyente polar adicional; tres o menos si no. Los restos pequeños son, por supuesto, siempre no aromáticos. Dependiendo de sus propiedades estructurales, los restos de aminoácidos pueden incluirse en dos o más clases. Para los aminoácidos de proteínas de origen natural, la subclasificación de acuerdo con este esquema se presenta en la **Tabla B**.

Tabla B: Subclasificación de Aminoácidos

Subclases	**Aminoácidos**
Ácidos	Ácido aspártico, ácido Glutámico
Básicos	No cíclicos: Arginina, Lisina; Cíclicos: Histidina
Carga	Ácido aspártico, ácido glutámico, arginina, lisina, histidina
Pequeños	Glicina, Serina, Alanina, Treonina, Prolina

Polares/neutros	Asparagina, Histidina, Glutamina, Cisteína, Serina, Treonina
Polares/grandes	Asparagina, Glutamina
Hidrófobos	Tirosina, Valina, Isoleucina, Leucina, Metionina, Fenilalanina, Triptófano
Aromáticos	Triptófano, Tirosina, Fenilalanina
Restos que influyen en la orientación de cadena	Glicina y Prolina

La sustitución de aminoácidos conservativa también incluye agrupamientos basados en cadenas laterales. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tienen cadenas laterales hidroxilo-alifáticas es serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amidas es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Por ejemplo, es razonable esperar que el remplazo de una leucina con una isoleucina o valina, un aspartato con un glutamato, una treonina con una serina o un remplazo similar de un aminoácido con un aminoácido relacionado estructuralmente no tendrá un efecto importante en las propiedades del polipéptido variante resultante. Si un cambio de aminoácido da como resultado un polipéptido de MetRS variante y/o truncado funcional puede determinarse fácilmente ensayando su actividad no canónica, como se describe en el presente documento. Se muestran sustituciones conservativas en la Tabla C bajo el encabezamiento de sustituciones a modo de ejemplo. Se realizan sustituciones de aminoácidos que quedan dentro del alcance de la invención, en general, seleccionando sustituciones que no difieren significativamente en su efecto en el mantenimiento de (a) la estructura de la cadena principal peptídica en el área de la sustitución, (b) la carga o hidrofobicidad de la molécula en el sitio diana, (c) el volumen de la cadena lateral, o (d) la función biológica. Después de introducirse las sustituciones, las variantes se exploran con respecto a actividad biológica.

Tabla C: sustituciones de aminoácidos a modo de ejemplo

Resto Original	**Sustituciones A modo de ejemplo**	**Sustituciones Preferidas**
Ala	Val, Leu, Ile	Val
Arg	Lys, Gln, Asn	Lys
Asn	Gln, His, Lys, Arg	Gln
Asp	Glu	Glu
Cys	Ser	Ser
Gln	Asn, His, Lys,	Asn
Glu	Asp, Lys	Asp
Gly	Pro	Pro
His	Asn, Gln, Lys, Arg	Arg
Ile	Leu, Val, Met, Ala, Phe, Norleu	Leu
Leu	Norleu, Ile, Val, Met, Ala, Phe	Ile
Lys	Arg, Gln, Asn	Arg
Met	Leu, Ile, Phe	Leu
Phe	Leu, Val, Ile, Ala	Leu
Pro	Gly	Gly
Ser	Thr	Thr
Thr	Ser	Ser
Trp	Tyr	Tyr
Tyr	Trp, Phe, Thr, Ser	Phe
Val	Ile, Leu, Met, Phe, Ala, Norleu	Leu

Como alternativa, los aminoácidos similares para realizar sustituciones conservativas pueden agruparse en tres categorías basándose en la identidad de las cadenas laterales. El primer grupo incluye ácido glutámico, ácido aspártico, arginina, lisina, histidina, que tienen todos cadenas laterales con carga; el segundo grupo incluye glicina, serina, treonina, cisteína, tirosina, glutamina, asparagina; y el tercer grupo incluye leucina, isoleucina, valina, alanina, prolina, fenilalanina, triptófano, metileno, como se describe en Zubay, G., Biochemistry, tercera edición, Wm. C. Brown Publishers (1993).

Por lo tanto, un resto de aminoácido no esencial predicho en un polipéptido de MetRS truncado y/o variante se reemplaza normalmente con otro resto de aminoácido de la misma familia de cadena lateral. Como alternativa, pueden introducirse mutaciones aleatoriamente a lo largo de toda o parte de una secuencia codificante de MetRS, tal como por mutagénesis de saturación, y los mutantes resultantes pueden explorarse con respecto a una actividad del polipéptido parental para identificar mutantes que conservan esa actividad. Después de mutagénesis de las secuencias codificantes, el péptido codificado puede expresarse de forma recombinante y puede determinarse la actividad del péptido. Un resto de aminoácido “no esencial” es un resto que puede alterarse de la secuencia de referencia de un polipéptido de realización sin anular o alterar sustancialmente una o más de sus actividades. Convenientemente, la alteración no ha anulado sustancialmente una de estas actividades, por ejemplo, la actividad

es al menos 20 %, 40 %, 60 %, 70 % u 80 % 100 %, 500 %, 1000 % o más de la secuencia de MetRS de referencia. Un resto de aminoácido “esencial” es un resto que, cuando se altera de la secuencia de referencia de un polipéptido de MetRS, da como resultado una anulación de una actividad de la molécula parental de modo que esté presente menos del 20 % de la actividad de referencia. Por ejemplo, dichos restos de aminoácidos esenciales incluyen los que están conservados en polipéptidos de MetRS entre diferentes especies, incluyendo las secuencias que están conservadas en el sitio o los sitios o motivo o motivos de unión activos de polipéptidos de MetRS de diversas fuentes.

En general, se aíslan polipéptidos y polipéptidos de fusión (así como sus polinucleótidos codificantes). Un polipéptido o polinucleótido “aislado” es uno que está retirado de su ambiente original. Por ejemplo, una proteína de origen natural está aislada si se separa de algunos o todos de los materiales coexistentes en el sistema natural. Preferentemente, dichos polipéptidos son al menos aproximadamente 90 % puros, más preferentemente al menos aproximadamente 95 % puros y más preferentemente al menos aproximadamente 99 % puros. Se considera que un polinucleótido está aislado si, por ejemplo, se clona en un vector que no es parte del ambiente natural.

Ciertas realizaciones también abarcan dímeros de polipéptidos de MetRS. Los dímeros pueden incluir, por ejemplo, homodímeros entre dos polipéptidos de MetRS idénticos, heterodímeros entre dos polipéptidos de MetRS diferentes (por ejemplo, un polipéptido de YRS de longitud completa y un polipéptido de YRS truncado; un polipéptido de YRS truncado y un polipéptido de WRS truncado) y/o heterodímeros entre un polipéptido de MetRS y un polipéptido heterólogo. Ciertos heterodímeros, tales como los que están entre un polipéptido de MetRS y un polipéptido heterólogo, pueden ser bifuncionales, como se describe en el presente documento.

También se incluyen monómeros de polipéptidos de MetRS, incluyendo monómeros de polipéptidos de MetRS aislados que no dimerizan sustancialmente con un segundo polipéptido de MetRS, bien debido a una o más sustituciones, truncamientos, deleciones, adiciones, modificaciones químicas o una combinación de estas alteraciones. En ciertas realizaciones, los polipéptidos de MetRS monoméricos poseen actividades biológicas, incluyendo actividades no canónicas, que no poseen complejos polipeptídicos de MetRS diméricos o multiméricos.

Ciertas realizaciones de la presente invención también contemplan el uso de polipéptidos de MetRS modificados, incluyendo modificaciones que mejoraron las características deseadas de un polipéptido de MetRS, como se describe en el presente documento. Las modificaciones de polipéptidos de MetRS de la invención incluyen derivatizaciones químicas y/o enzimáticas en uno o más aminoácidos constituyentes, incluyendo modificaciones de cadena lateral, modificaciones de cadena principal y modificaciones N y C terminales incluyendo acetilación, hidroxilación, metilación, amidación y la unión de restos de carbohidratos o lipídicos, cofactores y similares. Las modificaciones a modo de ejemplo también incluyen pegilación de un polipéptido de MetRS (véase, por ejemplo, Veronese y Harris, Advanced Drug Delivery Reviews 54: 453-456, 2002; y Pasut *et al*., Expert Opinion. Ther. Patents 14(6) 859-894 2004, ambas incorporadas en el presente documento por referencia).

PEG es un polímero bien conocido que tiene las propiedades de solubilidad en agua y en muchos disolventes orgánicos, ausencia de toxicidad y ausencia de inmunogenicidad. También es transparente, incoloro, inodoro y químicamente estable. Por estas y otras razones, el PEG se ha seleccionado como el polímero preferido para unión, pero se ha empleado solamente para fines de ilustración y no de limitación. Pueden obtenerse productos similares con otros polímeros solubles en agua, incluyendo sin limitación, alcohol polivinílico, otros poli(óxidos de alquileno) tales como poli(propilenglicol) y similares, poli(polioles oxietilados) tales como poli(glicerol oxietilado) y similares, carboximetilcelulosa, dextrano, alcohol polivinílico, polivinilpirrolidona, poli-1,3-dioxolano, poli-1,3,6-trioxano, anhídrido maleico/de etileno y poliaminoácidos. Un experto en la materia será capaz de seleccionar el polímero deseado basándose en la dosificación, el tiempo de circulación, la resistencia a proteólisis y otras consideraciones deseadas.

En particular, una amplia diversidad de derivados de PEG están disponibles y son adecuados para su uso en la preparación de conjugados de PEG. Por ejemplo, los reactivos de PEG de NOF Corp. comercializados con el nombre comercial de serie SUNBRIGHT® proporcionan numerosos derivados de PEG, incluyendo metoxipolietilenglicoles y derivados de PEG activados tales como metoxi-PEG aminas, maleimidas, ésteres de N-hidroxisuccinimida y ácidos carboxílicos para acoplamiento por diversos métodos con el extremo N terminal, C terminal o cualquier aminoácido interno del polipéptido de MetRS. La tecnología de PEGilación avanzada de Nektar Therapeutics también ofrece diversas tecnologías de acoplamiento a PEG para mejorar potencialmente la seguridad y eficacia de un producto terapéutico basado en polipéptido de MetRS.

Una búsqueda de patentes, solicitudes de patente publicadas y publicaciones relacionadas también proporcionará a los expertos en la materia que lean la presente divulgación posibles tecnologías de acoplamiento de PEG significativas y derivados de PEG. Por ejemplo, las Patentes de Estados Unidos n.º 6.436.386; 5.932.462; 5.900.461; 5.824.784; y 4.904.584; describen dichas tecnologías y derivados, y métodos para su fabricación.

En ciertos aspectos, puede utilizarse tecnología de ligamiento quimioselectivo para modificar polipéptidos de MetRS de la invención, tal como uniendo polímeros de una manera específica de sitio y controlada. Dicha tecnología se basa normalmente en la incorporación de anclajes quimioselectivos en la cadena principal proteica por medios

químicos o recombinantes, y la posterior modificación con un polímero que porta un enlazador complementario. Como resultado, el proceso de ensamblaje y la estructura covalente del conjugado de polímero-proteína resultante puede controlarse, permitiendo la optimización racional de propiedades farmacológicas, tales como eficacia y propiedades farmacocinéticas (véase, por ejemplo, Kochendoerfer, Current Opinion in Chemical Biology 9: 555-560, 2005).

En otras realizaciones, también se incluyen proteínas de fusión de polipéptido de MetRS con otras proteínas, y estas proteínas de fusión pueden aumentar la actividad biológica, secreción, dirección, vida biológica, capacidad de penetrar las membranas celulares, o la barrera hematoencefálica, o propiedades farmacocinéticas del polipéptido de MetRS. Los ejemplos de proteínas de fusión que mejoran las propiedades farmacocinéticas ("modificadores de PK") incluyen sin limitación, fusiones con albúmina humana (Osborn *et al*.: Eur. J. Pharmacol. 456 (1-3): 149-158, (2002)), dominios Fc de anticuerpo, secuencias de poli Glu o poli Asp, y transferrina. Adicionalmente, la fusión con secuencias polipeptídicas conformacionalmente desordenadas compuestas de los aminoácidos Pro, Ala y Ser ("PASilación") o almidón de hidroxietilo (comercializado con el nombre comercial HESYLATION®) proporciona un modo sencillo de aumentar el volumen hidrodinámico del polipéptido de MetRS. Esta extensión adicional adopta una estructura aleatoria voluminosa que aumenta significativamente el tamaño de la proteína de fusión resultante. Por este medio la eliminación normalmente rápida de polipéptidos de MetRS más pequeños mediante filtración por riñones se retarda en varios órdenes de magnitud. Adicionalmente se ha mostrado que el uso de proteínas de fusión de IgG permite que algunas proteínas de fusión penetren en la barrera hematoencefálica (Fu *et al*., (2010) Brain Res. 1352: 208-13).

Los ejemplos de proteínas de fusión que mejoran la penetración a través de membranas celulares incluyen fusiones con secuencias de translocación de membrana. En este contexto, la expresión "secuencias de translocación de membrana" se refiere a secuencias de aminoácidos de origen natural y sintéticas que tienen capacidad de translocación de membrana a través de una membrana celular. Las secuencias de translocación de membrana representativas incluyen las basadas en las secuencias de translocación de membrana de origen natural derivadas de la proteína Tat y proteína de transcripción homeótica Antennapedia, así como secuencias de translocación de membrana sintéticas basadas completamente o en parte en restos de poli Arginina y Lisina. Las secuencias de translocación de membrana representativas incluyen por ejemplo las desveladas en las siguientes patentes, US5.652.122; US5.670.617; US5.674.980; US5.747.641; US5.804.604; US6.316.003; US7.585.834; US7.312.244; US7.279.502; US7.229.961; US7.169.814; US7.453.011; US7.235.695; US6.982.351; US6.605.115; US7.306.784; US7.306.783; US6.589.503; US6.348.185; US6.881.825; US7.431.915; WO0074701A2; WO2007111993A2; WO2007106554A2; WO02069930A1; WO03049772A2; WO03106491A2; y WO2008063113A1.

Se apreciará que puede interponerse un enlazador molecular flexible (o espaciador) opcionalmente entre, y unir covalentemente, el polipéptido de MetRS y cualquiera de las proteínas de fusión desveladas en el presente documento.

Adicionalmente, en algunas realizaciones, el polipéptido de MetRS puede incluir secuencias señales de secreción de origen natural o sintéticas, derivadas de otras proteínas secretadas bien caracterizadas. En algunas realizaciones dichas proteínas pueden procesarse por escisión proteolítica para formar el polipéptido de MetRS in situ. Dichas proteínas de fusiones incluyen por ejemplo fusiones de polipéptido de MetRS con ubiquitina para proporcionar un nuevo aminoácido N terminal, o el uso de una señal de secreción para mediar en secreción a alto nivel del polipéptido de MetRS en el medio extracelular, o marcadores epitópicos N o C terminales para mejorar la purificación o detección.

Los polipéptidos de MetRS descritos en el presente documento pueden prepararse por cualquier procedimiento adecuado conocido por los expertos en la materia, tal como por técnicas recombinantes. Además de métodos de producción recombinante, los polipéptidos de la invención pueden producirse por síntesis peptídica directa usando técnicas de fase sólida (Merrifield, J. Am. Chem. Soc. 85: 2149-2154 (1963)). Puede realizarse síntesis de proteínas usando técnicas manuales o por automatización. Puede conseguirse síntesis automática, por ejemplo, usando sintetizador peptídico Applied Biosystems 431A (Perkin Elmer). Como alternativa, diversos fragmentos pueden sintetizarse químicamente por separado y combinarse usando métodos químicos para producir la molécula deseada.

IV. POLINUCLEÓTIDOS DE METRS

Las realizaciones de la presente invención incluyen polinucleótidos que codifican uno o más fragmentos proteicos de nueva identificación de una metionil ARNt sintetasa (MetRS), además de complementos, variantes y fragmentos de los mismos. En ciertas realizaciones, un polinucleótido de MetRS codifica todas o una parte de la secuencia o las secuencias de referencia del polipéptido de MetRS como se expone en SEQ ID NO: 137, que representan variantes de corte y empalme, fragmentos proteolíticos u otro tipo de fragmentos de metionil ARNt sintetasa. Ciertas realizaciones incluyen polinucleótidos que codifican polipéptidos o proteínas que comprenden la secuencia de uno o más puntos de unión de corte y empalme de esas variantes de corte y empalme, además de complementos, variantes y fragmentos de las mismas. En ciertas realizaciones, normalmente debido a la naturaleza singular de una variante de corte y empalme de MetRS seleccionada, que combina exones de una manera nueva o excepcional, las secuencias de referencia de polinucleótidos de MetRS comprenden un punto de unión de corte y empalme único o

excepcional. Ciertas realizaciones excluyen un polinucleótido de MetRS de longitud completa correspondiente.

También se incluyen dentro de los polinucleótidos de MetRS de la presente invención cebadores, sondas, oligonucleótidos antisentido y agentes de interferencia de ARN que comprenden todos o una parte de estos polinucleótidos de referencia, que son complementarios de todos o una parte de estos polinucleótidos de referencia, o que hibridan específicamente con estos polinucleótidos de referencia, como se describe en el presente documento.

La expresión "polinucleótido" o "ácido nucleico" como se usa en el presente documento designa ARNm, ARN, ARNc, ADNc o ADN. El término normalmente se refiere a una forma polimérica de nucleótidos de al menos 10 bases de longitud, bien ribonucleótidos o bien desoxinucleótidos o una forma modificada de un tipo de nucleótido. El término incluye formas mono y bicatenarias de ADN. Los términos "ADN" y "polinucleótido" y "ácido nucleico" se refieren a una molécula de ADN que se ha aislado de ADN genómico total de una especie particular. Por lo tanto, un segmento de ADN aislado que codifica un polipéptido se refiere a un segmento de ADN que contiene una o más secuencias codificantes pero aún está sustancialmente aislado de, o purificado de, ADN genómico total de la especie de la que se obtiene el segmento de ADN. También se incluyen polinucleótidos no codificantes (por ejemplo, cebadores, sondas, oligonucleótidos), que no codifican un polipéptido de MetRS. Se incluyen dentro de las expresiones "segmento de ADN" y "polinucleótido" segmentos de ADN y fragmentos más pequeños de dichos segmentos, y también vectores recombinantes, incluyendo, por ejemplo, plásmidos, cósmidos, fagémidos, fagos, virus y similares.

Las secuencias codificantes o no codificantes adicionales pueden, aunque no es necesario, estar presentes dentro de un polinucleótido de la presente invención, y un polinucleótido puede estar, aunque no es necesario que esté, unido a otras moléculas y/o materiales de apoyo. Por lo tanto, los polinucleótidos de la presente invención, independientemente de la longitud de la secuencia codificante en sí misma, pueden combinarse con otras secuencias de ADN, tales como promotores, señales de poliadenilación, sitios de enzimas de restricción adicionales, sitios de clonación múltiple, otros segmentos codificantes, y similares, de modo que su longitud global puede variar considerablemente.

Se contempla por lo tanto que puede emplearse un fragmento polinucleotídico de casi cualquier longitud; estando limitada la longitud preferentemente por la facilidad de preparación y uso en el protocolo de ADN recombinante pretendido. Se incluyen polinucleótidos de 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100, 110, 120. 130, 140, 150, 160, 170, 180, 190, 200, 220, 240, 260, 270, 280, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000 o más (incluyendo todos los números enteros entre medias) bases de longitud, incluyendo cualquier parte o fragmento (por ejemplo, mayor de aproximadamente 6, 7, 8, 9 o 10 nucleótidos de longitud) de un polinucleótido de referencia de MetRS (por ejemplo, número de bases XY, en el que X es aproximadamente 1-3000 o más e Y es aproximadamente 10-3000 o más), o su complemento.

Las realizaciones de la presente invención también incluyen "variantes" de las secuencias polinucleotídicas de referencia de MetRS. Las "variantes" polinucleotídicas pueden contener una o más sustituciones, adiciones, deleciones y/o inserciones en relación con un polinucleótido de referencia. En general, las variantes de una secuencia polinucleotídica de referencia de MetRS pueden tener al menos aproximadamente 30 %, 40 % 50 %, 55 %, 60 %, 65 %, 70 %, generalmente al menos aproximadamente 75 %, 80 %, 85 %, idealmente aproximadamente 90 % a 95 % o más, y más convenientemente aproximadamente 98 % o más identidad de secuencia con esa secuencia de nucleótidos particular como se determina por programas de alineamiento de secuencias descritos en otra parte en el presente documento usando parámetros por defecto. En ciertas realizaciones, las variantes pueden diferir de una secuencia de referencia en aproximadamente 1, 2, 3, 4, 5, 6, 7, 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, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 90, 100 (incluyendo todos los números enteros entre medias) o más bases. En ciertas realizaciones, tales como cuando la variante polinucleotídica codifica un polipéptido de MetRS que tiene una actividad no canónica, la actividad deseada del polipéptido de MetRS no está sustancialmente disminuida en relación con el polipéptido no modificado. El efecto en la actividad del polipéptido codificado puede evaluarse en general como se describe en el presente documento.

Ciertas realizaciones incluyen polinucleótidos que hibridan con una secuencia polinucleotídica de MetRS de referencia, o con sus complementos, en condiciones rigurosas descritas posteriormente. Como se usa en el presente documento, la expresión "hibrida en condiciones de baja rigurosidad, rigurosidad media, alta rigurosidad o muy alta rigurosidad" describe condiciones para hibridación y lavado. Puede encontrarse orientación para realizar reacciones de hibridación en Ausubel *et al*., (1998, mencionado anteriormente), Secciones 6.3.1-6.3.6. Se describen métodos acuosos y no acuosos en esta referencia y puede usarse cualquiera de ellos.

La referencia en el presente documento a condiciones de baja rigurosidad incluye y abarca de al menos aproximadamente 1 % v/v a al menos aproximadamente 15 % v/v de formamida y de al menos aproximadamente 1 M a al menos aproximadamente 2 M de sal para hibridación a 42 ºC y de al menos aproximadamente 1 M a al menos aproximadamente 2 M de sal para lavado a 42 ºC. Las condiciones de baja rigurosidad también pueden

incluir albúmina de suero bovino (BSA) 1 %, EDTA 1 mM, $NaHPO_4$ 0,5 M (pH 7,2), SDS 7 % para hibridación a 65 ºC y (i) SSC 2 x, SDS 0,1 %; o (ii) BSA a 0,5 %, EDTA 1 mM, $NaHPO_4$ 40 mM (pH 7,2), SDS 5 % para lavado a temperatura ambiente. Una realización de condiciones de baja rigurosidad incluye hibridación en cloruro sódico/citrato sódico (SSC) 6 x a aproximadamente 45 ºC, seguido de dos lavados en SSC 0,2 x, SDS 0,1 % al menos a 50 ºC (la temperatura de los lavados puede aumentarse hasta 55 ºC para condiciones de baja rigurosidad).

Las condiciones de rigurosidad media incluyen y abarcan de al menos aproximadamente 16 % v/v a al menos aproximadamente 30 % v/v de formamida y de al menos aproximadamente 0,5 M a al menos aproximadamente 0,9 M de sal para hibridación a 42 ºC, y de al menos aproximadamente 0,1 M a al menos aproximadamente 0,2 M de sal para lavado a 55 ºC. Las condiciones de rigurosidad media también pueden incluir albúmina de suero bovino 1 % (BSA), EDTA 1 mM, $NaHPO_4$ 0,5 M (pH 7,2), SDS 7 % para hibridación a 65 ºC y (i) SSC 2 x, SDS 0,1 %; o (ii) BSA 0,5 %, EDTA 1 mM, $NaHPO_4$ 40 mM (pH 7,2), SDS 5 % para lavado a 60-65 ºC. Una realización de condiciones de rigurosidad media incluye hibridación en SSC 6 x a aproximadamente 45 ºC, seguido de uno o más lavados en SSC 0,2 x, SDS 0,1 % a 60 ºC. Las condiciones de alta rigurosidad incluyen y abarcan de al menos aproximadamente 31 % v/v a al menos aproximadamente 50 % v/v de formamida y aproximadamente 0,01 M a aproximadamente 0,15 M de sal para hibridación a 42 ºC, y aproximadamente 0,01 M a aproximadamente 0,02 M de sal para lavado a 55 ºC.

Las condiciones de alta rigurosidad también pueden incluir BSA 1 %, EDTA 1 mM, $NaHPO_4$ 0,5 M (pH 7,2), SDS 7 % para hibridación a 65 ºC y (i) SSC 0,2 x, SDS 0,1 %; o (ii) BSA 0,5 %, EDTA 1 mM, $NaHPO_4$ 40 mM (pH 7,2), SDS 1 % para lavado a una temperatura de más de 65 ºC. Una realización de condiciones de alta rigurosidad incluye hibridación en SSC 6 x a aproximadamente 45 ºC, seguido de uno o más lavados en SSC 0,2 x, SDS 0,1 % a 65 ºC. Una realización de condiciones de rigurosidad muy alta incluye hibridación en fosfato sódico 0,5 M, SDS 7 % a 65 ºC, seguido de uno o más lavados en SSC 0,2 x, SDS 1 % a 65 ºC.

Otras condiciones de rigurosidad se conocen bien en la técnica y un experto en la materia reconocerá que pueden manipularse diversos factores para optimizar la especificidad de la hibridación. La optimización de la rigurosidad de los lavados finales puede servir para asegurar un alto grado de hibridación. Para ejemplos detallados, véase Ausubel *et al*., mencionado anteriormente en las páginas 2.10.1 a 2.10.16 y Sambrook *et al*. (1989, mencionado anteriormente) en las secciones 1.101 a 1.104.

Aunque se llevan a cabo normalmente lavados rigurosos a temperaturas de aproximadamente 42 ºC a 68 ºC, un experto en la materia apreciará que pueden ser adecuadas otras temperaturas para condiciones rigurosas. La velocidad de hibridación máxima normalmente sucede a aproximadamente 20 ºC a 25 ºC por debajo de la T_m para formación de híbrido de ADN-ADN. Se conoce bien en la técnica que la Tm es la temperatura de fusión, o temperatura a la que dos secuencias polinucleotídicas complementarias se disocian. Los métodos para estimar la T_m se conocen bien en la técnica (véase Ausubel *et al*., mencionado anteriormente en la página 2.10.8).

En general, la T_m de una doble cadena perfectamente coincidente de ADN puede predecirse como una aproximación por la fórmula: $T_m = 81,5 + 16,6 (\log_{10} M) + 0,41 (\%G+C) - 0,63$ (% formamida) - (600/longitud) en la que: M es la concentración de Na^+, preferentemente en el intervalo de 0,01 molar a 0,4 molar; %G+C es la suma de bases de guanosina y citosina como un porcentaje del número total de bases, dentro del intervalo entre 30 % y 75 % de G+C; el % de formamida es el porcentaje de concentración de formamida por volumen; la longitud es el número de pares de bases en el dúplex de ADN. La T_m de un ADN dúplex se reduce en aproximadamente 1 ºC con cada aumento de 1 % en el número de pares de bases desapareados aleatoriamente. El lavado se lleva a cabo en general a T_m -15 ºC para alta rigurosidad o T_m - 30 ºC para rigurosidad moderada.

En un ejemplo de un procedimiento de hibridación, se hibrida una membrana (por ejemplo, una membrana de nitrocelulosa o una membrana de nylon) que contiene ADN inmovilizado durante una noche a 42 ºC en un tampón de hibridación (formamida desionizada al 50 %, SSC 5 x, Solución de Denhardt 5 x (ficoll 0,1 %, polivinilpirrolidona 0,1 % y albúmina de suero bovino 0,1 %), SDS 0,1 % y ADN de esperma de salmón desnaturalizado 200 mg/ml) que contienen una sonda marcada. La membrana se somete después a dos lavados de rigurosidad media secuenciales (es decir, SSC 2 x, SDS 0,1 % durante 15 minutos a 45 ºC, seguido de SSC 2 x, SDS 0,1 % durante 15 minutos a 50 ºC), seguido de dos lavados de mayor rigurosidad secuenciales (es decir, SSC 0,2 x, SDS 0,1 % durante 12 minutos a 55 ºC seguido de SSC 0,2 % y solución de SDS 0,1 % durante 12 minutos a 65-68 ºC.

Como se ha indicado anteriormente, ciertas realizaciones se refieren a polinucleótidos de MetRS que codifican un polipéptido de MetRS. Entre otros usos, estas realizaciones pueden utilizarse para producir de forma recombinante un polipéptido de MetRS deseado o variante del mismo, o para expresar el polipéptido de MetRS en una célula seleccionada o sujeto. Se apreciará por los expertos habituales en la materia que, como resultado de la degradación del código genético, existen muchas secuencias de nucleótidos que codifican un polipéptido como se describe en el presente documento. Algunos de estos polinucleótidos pueden portar homología mínima con la secuencia de nucleótidos de cualquier gen nativo. No obstante, se contemplan específicamente en la presente invención polinucleótidos que varían debido a las diferencias en el uso codónico, por ejemplo polinucleótidos que se optimizan para selección de codones humanos y/o primates.

Por lo tanto, múltiples polinucleótidos pueden codificar los polipéptidos de MetRS de la invención. Además, la secuencia polinucleotídica puede manipularse por diversas razones. Los ejemplos incluyen pero sin limitación la incorporación de codones preferidos para potenciar la expresión del polinucleótido en diversos organismos (véase en general Nakamura *et al*., Nuc. Acid. Res. (2000) 28 (1): 292). Además, pueden incorporar mutaciones silenciosas para introducir, o eliminar sitios de restricción, reducir la densidad de motivos dinucleotídicos de CpG (véase por ejemplo, Kameda *et al*., Biochem. Biophys Res. Commun. (2006) 349 (4): 1269-1277) o reducir la capacidad de secuencias monocatenarias para formar estructuras de tallo-bucle: (véase, por ejemplo, Zuker M., Nucl. Acid Res. (2003); 31 (13): 3406-3415). Además, la expresión en mamíferos puede optimizarse adicionalmente incluyendo una secuencia consenso de Kozak [es decir, (a/g)cc(a/g)ccATGg] en el codón de inicio. Se conocen en la técnica secuencias consenso de Kozak útiles para este fin (Mantyh *et al*., PNAS 92: 2662-2666 (1995), Mantyh *et al*., Prot. Exp. & Purif 6,124 (1995)).

Los polinucleótidos de la presente invención, independientemente de la longitud de la secuencia codificante en sí misma, pueden combinarse con otras secuencias de ADN, tales como promotores, señales de poliadenilación, sitios de enzimas de restricción adicionales, sitios de clonación múltiples otros segmentos codificantes y similares, de modo que su longitud total pueda variar considerablemente. Se contempla por lo tanto que puede emplearse un fragmento polinucleotídico de casi cualquier longitud; estando la longitud total limitada preferentemente por la facilidad de preparación y uso en el protocolo de ADN recombinante pretendido.

Pueden prepararse polinucleótidos y fusiones de los mismos, manipularse y/o expresarse usando cualquiera de diversas técnicas bien establecidas conocidas y disponibles en este campo. Por ejemplo, pueden usarse secuencias polinucleotídicas que codifiquen polipéptidos de la invención o proteínas de fusión o equivalentes funcionales de las mismas, en moléculas de ADN recombinante para dirigir la expresión de un polipéptido de MetRS en células hospedadoras apropiadas. Debido a la degradación 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.

Como entenderán los expertos en la materia, puede ser ventajoso en algunos casos producir secuencias de nucleótidos que codifican polipéptidos que posean codones de origen no natural. Por ejemplo, pueden seleccionarse codones preferidos por un hospedador procariota o eucariota particular para aumentar la tasa de expresión de proteínas o para producir un transcrito de ARN recombinante que tenga propiedades deseables, tales como una semivida que es más larga que la de un transcrito generado de la secuencia natural. Dichos polinucleótidos se denominan habitualmente “de codones optimizados”. Puede utilizarse cualquiera de los polinucleótidos descritos en el presente documento en una forma de codones optimizados. En ciertas realizaciones, un polinucleótido puede tener codones optimizados para su uso en bacterias específicas tales como *E. coli* o levadura tal como *S. cerevisiae* (véase, por ejemplo, Burgess-Brown *et al*., Protein Expr Purif. 59: 94-102, 2008 Ermolaeva MD (2001) Curr. Iss. Mol. Biol. 3 (4) 91-7; Welch *et al*., PLoS ONE 4 (9): e7007 doi: 10.1371/journal.pone.0007002).

Además, las secuencias de polinucleótidos de la presente invención pueden modificarse técnicamente usando métodos generalmente conocidos en este campo para alterar secuencias que codifican polipéptidos por diversas razones, incluyendo pero sin limitación, alteraciones que modifican la clonación, el procesamiento, la expresión y/o la actividad del producto génico.

De acuerdo con otro aspecto de la invención, pueden suministrarse polinucleótidos que codifican polipéptidos de la invención a un sujeto *in vivo*, por ejemplo, usando 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, particularmente un ser humano, con un trastorno o afecciones para las que se busca dicha terapia. El ácido nucleico se introduce en las células diana seleccionadas de una manera tal que el ADN heterólogo se exprese y se produzca un producto terapéutico codificado por el mismo.

Diversos vectores víricos que pueden utilizarse para terapia génica como se enseña en el presente documento incluyen adenovirus, virus del herpes, vaccinia, virus adenoasociado (VAA), o, preferentemente, un virus de ARN tal como un retrovirus. Preferentemente, el vector retrovírico es un derivado de un retrovirus murino o aviar, o es un vector lentivírico. El vector retrovírico preferido es un vector lentivírico. Los ejemplos de vectores retrovíricos en los que puede insertarse un único gen ajeno incluyen, pero sin limitación: virus de leucemia murina de Moloney (MoMuLV), virus de sarcoma murino de Harvey (HaMuSV), virus de tumor mamario murino (MuMTV), VIS, VIB, VIH y virus del sarcoma de Rous (VSR). Varios vectores retrovíricos adicionales pueden incorporar múltiples genes. Todos estos vectores pueden transferir o incorporar un gen para un marcador seleccionable de modo que puedan identificarse y generarse células transducidas. Insertando una secuencia de interés de polipéptido de unión a ADN derivado de dedo de cinc en el vector viral, junto con otro gen que codifica el ligando para un receptor en una célula diana específica, por ejemplo, el vector puede hacerse específico de diana. Pueden hacerse vectores retrovíricos específicos de diana insertando, por ejemplo, un polinucleótido que codifica una proteína (dímero). Puede conseguirse una dirección ilustrativa usando un anticuerpo que se dirija al vector retrovírico. Los expertos en la materia conocerán, o pueden determinar fácilmente, sin experimentación indebida, secuencias polinucleotídicas específicas que pueden insertarse en el genoma retrovírico para permitir el suministro específico de diana del vector retrovírico que contiene el polinucleótido de proteína de unión a nucleótido de dedo de cinc.

Ya que los retrovirus recombinantes son defectuosos, requieren ayuda para producir partículas de vectores infecciosos. Esta ayuda puede proporcionarse, por ejemplo, usando líneas celulares auxiliares que contienen plásmidos que codifican todos los genes estructurales del retrovirus bajo el control de secuencias reguladoras dentro de la LTR. Estos plásmidos carecen de una secuencia de nucleótidos que permita que el mecanismo de empaquetamiento reconozca un transcrito de ARN para encapsulación. Las líneas celulares auxiliares que tienen deleciones en la señal de empaquetamiento incluyen pero sin limitación PSI.2, PA317 y PNA12, por ejemplo. Estas líneas celulares producen viriones vacíos, ya que no se empaqueta ningún genoma. Si se introduce un vector retrovírico en dichas células en las que la señal de empaquetamiento está intacta, pero los genes estructurales se reemplazan por otros genes de interés, el vector puede empaquetarse y producirse el virión del vector. Los viriones de vector producidos por este método pueden después usarse para infectar una línea celular tisular, tal como células NIH 3T3, para producir grandes cantidades de viriones retrovíricos quiméricos.

También pueden usarse técnicas de suministro “no víricas” para terapia génica incluyendo, por ejemplo, complejos de ADN-ligando, complejos de adenovirus-ligando-ADN, inyección directa de ADN, precipitación con $CaPO_4$, técnicas de pistola génica, electroporación, liposomas, lipofección y similares. Cualquiera de estos métodos está ampliamente disponible para un experto en la materia y sería adecuado para su uso en la presente invención. Están disponibles para los expertos en la materia otros métodos adecuados, y debe entenderse que la presente invención puede conseguirse usando cualquiera de los métodos disponibles de transfección. Puede conseguirse lipofección encapsulando una molécula de ADN aislada dentro de una partícula liposómica y poniendo en contacto la partícula liposómica con la membrana celular de la célula diana. Los liposomas son partículas coloidales, que se autoensamblan, en las que una bicapa lipídica, compuesta de moléculas anfifílicas tales como fosfatidilserina o fosfatidilcolina, encapsula una parte del medio circundante de modo que la bicapa lipídica rodea un interior hidrófilo. Pueden construirse liposomas unilamelares o multilamelares de modo que el interior contenga un producto químico, fármaco o, como en la presente invención, una molécula de ADN aislada deseado.

En otro aspecto, pueden usarse polinucleótidos que codifican polipéptidos de la invención para expresar y suministrar un polipéptido de MetRS mediante terapia celular. De acuerdo con otro aspecto, la presente invención incluye una terapia celular para tratar una enfermedad o un trastorno, que comprende administrar a una célula hospedadora que expresa, o es capaz de expresar, un polipéptido de MetRS.

La terapia celular implica la administración de células que se han seleccionado, multiplicado y tratado o alterado farmacológicamente (es decir modificadas genéticamente) fuera del cuerpo (Bordignon, C. *et al*., Cell Therapy: Achievements and Perspectives (1999), Haematologica, 84, págs. 1110-1149). Dichas células hospedadoras incluyen, por ejemplo, células primarias, incluyendo macrófagos, y células madre que se han modificado genéticamente para expresar un polipéptido de MetRS. El objetivo de la terapia celular es reemplazar, reparar o potenciar la función biológica de tejidos u órganos dañados.

El uso de células trasplantadas se ha investigado para el tratamiento de numerosos trastornos endocrinos tales como anemia y enanismo, trastornos hematológicos, insuficiencia renal y hepática, deficiencias de la hipófisis y el SNC y diabetes mellitus (Uludag *et al*., Technology of Mammalian Cell Encapsulation (2000), Advanced Drug Delivery Reviews, 42, págs. 29-64). Las células trasplantadas pueden actuar liberando compuestos bioactivos tales como un polipéptido de MetRS de la invención para reemplazar polipéptidos de MetRS endógenos que están ausentes o se producen en cantidades insuficientes en un sistema afectado.

Las realizaciones de la presente invención también incluyen oligonucleótidos, bien para detección, amplificación, terapias antisentido u otro fin. Para estos y fines relacionados, se entiende que el término “oligonucleótido”, “oligo” u “oligómero” abarca un “oligonucleótido” singular así como “oligonucleótidos” plurales, y se refiere a cualquier polímero de dos o más de nucleótidos, nucleósidos, nucleobases o compuestos relacionados usados como un reactivo en los métodos de amplificación de la presente invención, así como métodos de detección posteriores. El oligonucleótido puede ser ADN y/o ARN y/o análogos de los mismos.

El término oligonucleótido no indica necesariamente ninguna función particular del reactivo, sino que se usa de forma genérica para abarcar todos estos reactivos descritos en el presente documento. Un oligonucleótido puede cumplir diversas funciones diferentes, por ejemplo, puede actuar como un cebador si es capaz de hibridar con una cadena complementaria y puede extenderse adicionalmente en presencia de una polimerasa de ácido nucleico, puede proporcionar un promotor si contiene una secuencia reconocida por una ARN polimerasa y permite la transcripción, y puede actuar para evitar la hibridación o impedir la extensión de cebadores si se sitúa y/o modifica de forma apropiada. Un oligonucleótido también puede actuar como una sonda, o un agente antisentido. Un oligonucleótido puede ser prácticamente de cualquier longitud, limitado solamente por su función específica, por ejemplo, en una reacción de amplificación, en la detección de un producto de amplificación de la reacción de amplificación, o en una aplicación de interferencia de ARN o antisentido. Cualquiera de los oligonucleótidos descritos en el presente documento puede usarse como un cebador, una sonda, un oligómero antisentido, o un agente de interferencia de ARN.

El término “cebador” como se usa en el presente documento se refiere a un oligonucleótido monocatenario capaz de actuar como un punto de inicio para síntesis de ADN dirigida a molde en condiciones adecuadas definidas, por

ejemplo, por tampón y temperatura, en presencia de cuatro nucleósido trifosfatos diferentes y un agente para polimerización, tal como una ADN o ARN polimerasa o transcriptasa inversa. La longitud del cebador, en cualquier caso dado, depende, por ejemplo, del uso pretendido del cebador, y generalmente varía de aproximadamente 15 a 30 nucleótidos, aunque pueden usarse cebadores más cortos y más largos. Las moléculas de cebadores cortos requieren en general temperaturas más frías para formar complejos híbridos suficientemente estables con el molde. Un cebador no refleja necesariamente la secuencia exacta del molde pero debe ser suficientemente complementario para hibridar con dicho molde. El sitio de cebador es el área del molde con el que hibrida un cebador. El par de cebadores es un conjunto de cebadores que incluye un cebador cadena arriba 5' que hibrida con el extremo 5' de la secuencia para amplificar y un cebador cadena abajo 3' que hibrida con el complemento del extremo 3' de la secuencia para amplificar.

El término "sonda" como se usa en el presente documento incluye una molécula inmovilizada en superficie o soluble pero capaz de inmovilizarse que puede reconocerse por una diana particular. Véase, por ejemplo, Patente de Estados Unidos n.º 6.582.908 para un ejemplo de matrices que tienen todas las posibles combinaciones de sondas con 10, 12 y más bases. Las sondas y los cebadores como se usa en el presente documento normalmente comprenden al menos 10-15 nucleótidos contiguos de una secuencia conocida. Para potenciar la especificidad, pueden emplearse también cebadores y sondas más largos, tales como sondas y cebadores que comprenden al menos 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 o al menos 150 nucleótidos de una secuencia de referencia de MetRS o su complemento. Las sondas y los cebadores pueden ser considerablemente más largos que estos ejemplos, y se entiende que puede usarse cualquier longitud apoyada por el conocimiento en la técnica y la memoria descriptiva, incluyendo las tablas, figuras y el listado de secuencias.

Se describen en las referencias métodos para preparar y usar sondas y cebadores, por ejemplo Sambrook, J. *et al*. (1989) Molecular Cloning: A Laboratory Manual, 2ª ed., vol. 1-3, Cold Spring Harbor Press, Plainview N. Y.; Ausubel, F. M. *et al*. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, Nueva York N. Y.; Innis, M. *et al*. (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press, San Diego Calif. Pueden derivarse pares de cebadores de PCR de una secuencia conocida, por ejemplo, usando programas informáticos ideados para este fin tales como Primer (Versión 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge Mass.).

Pueden seleccionarse oligonucleótidos para su uso como sondas o cebadores usando software conocido en la técnica. Por ejemplo, el software OLIGO 4.06 es útil para la selección de pares de cebadores de PCR de hasta 100 nucleótidos cada uno y para el análisis de oligonucleótidos y polinucleótidos mayores de hasta 5.000 nucleótidos de una secuencia polinucleotídica de entrada de hasta 32 kilobases. Los programas de selección de cebadores similares han incorporado elementos adicionales para capacidades expandidas. Por ejemplo, el programa de selección de cebadores PrimOU (disponible al público del Centro de Genoma en el Centro Médico Sur Occidental de Texas, Dallas Tex.) es capaz de elegir cebadores específicos a partir de secuencias de megabases y es por lo tanto útil para diseñar cebadores en un ámbito de genoma completo.

El programa de selección de cebadores Primer3 (disponible al público del Centro para la Investigación del Genoma del Instituto Whitehead/MIT, Cambridge Mass.) permite al usuario introducir una "biblioteca de cebadores erróneos", en la que el usuario especifica secuencias para evitar como sitios de unión de cebadores. Primer3 es útil, en particular, para la selección de oligonucleótidos para micromatrices. (El código fuente para estos dos últimos programas de selección de cebadores también puede obtenerse de sus fuentes respectivas y modificarse para cubrir las necesidades específicas del usuario.) El programa PrimeGen (disponible al público del Centro de Recursos del Proyecto de Mapeo Genómico Humano de Reino Unido, Cambridge, Reino Unido) diseña cebadores basándose en múltiples alineamientos de secuencias, permitiendo de este modo la selección de cebadores que hibridan con las regiones más conservadas o menos conservadas de secuencias de ácido nucleico alineadas. Por lo tanto, este programa es útil para la identificación de oligonucleótidos y fragmentos polinucleotídicos tanto únicos como conservados. Los oligonucleótidos y fragmentos polinucleotídicos identificados por cualquiera de los métodos de selección anteriores son útiles en tecnologías de hibridación, por ejemplo, como cebadores de PCR o de secuenciación, elementos de micromatrices, o sondas específicas para identificar polinucleótidos complementarios parcial o completamente en una muestra de ácidos nucleicos. Los métodos de selección de oligonucleótidos no se limitan a los descritos en el presente documento.

En ciertas realizaciones, pueden prepararse oligonucleótidos por síntesis en fase sólida por etapas, empleando métodos detallados en las referencias citadas anteriormente, y posteriormente con respecto a la síntesis de oligonucleótidos que tienen una mezcla o enlaces de cadena principal catiónicos y sin carga. En algunos casos, puede ser deseable añadir restos químicos adicionales al oligonucleótido, por ejemplo, para potenciar la farmacocinética o para facilitar la captura o detección del compuesto. Dicho resto puede unirse covalentemente, normalmente a un extremo del oligómero, de acuerdo con métodos de síntesis convencionales. Por ejemplo, la adición de un resto de polietilenglicol u otro polímero hidrófilo, por ejemplo, uno que tenga 10-100 subunidades monoméricas, puede ser útil para potenciar la solubilidad. Uno o más grupos con carga, por ejemplo, grupos con carga aniónica tales como un ácido orgánico, pueden potenciar la captación celular.

Puede usarse diversas moléculas detectables para hacer a un oligonucleótido, o una proteína detectable, tal como

radioisótopos, fluorocromos, colorantes, enzimas, nanopartículas, marcadores quimioluminiscentes, biotina u otro monómero conocido en la técnica que puede detectarse directamente (por ejemplo, por emisión de luz) o indirectamente (por ejemplo, por unión de un anticuerpo marcado con fluorescencia).

Los radioisótopos proporcionan ejemplos de moléculas detectables que pueden utilizarse en ciertos aspectos de la presente invención. Varios radioisótopos pueden usarse como moléculas detectables para marcar nucleótidos o proteínas, incluyendo, por ejemplo, ^{32}P, ^{33}P, ^{35}S, ^{3}H y ^{125}I. Estos radioisótopos tienen diferentes semividas, tipos de degradación y niveles de energía que pueden adaptarse para coincidir con las necesidades de un protocolo particular. Por ejemplo, ^{3}H es un emisor de baja energía que da como resultado niveles de fondo bajos, sin embargo esta baja energía también da como resultado periodos de tiempo largos para autorradiografía. Están disponibles en el mercado ribonucleótidos, desoxirribonucleótidos y aminoácidos marcados de forma radiactiva. Están disponibles nucleótidos que están marcados de forma radiactiva en el primer, o α, grupo fosfato, o el tercer, o γ, grupo fosfato. Por ejemplo, tanto [α-^{32}P] dATP como [γ-^{32}P] dATP están disponibles en el mercado. Además, también están disponibles en el mercado diferentes actividades específicas para nucleótidos marcados con radiactividad y pueden adaptarse para diferentes protocolos.

Otros ejemplos de moléculas detectables que pueden utilizarse para detectar un oligonucleótido incluyen fluoróforos. Pueden usarse varios fluoróforos para marcar nucleótidos incluyendo, por ejemplo, fluoresceína, tetrametilrodamina, Texas Red, y varios otros (por ejemplo, Haugland, Handbook of Fluorescent Probes, 9ª Ed., 2002, Molec. Probes, Inc., Eugene OR; Haugland, The Handbook: Guide to Fluorescent Probes and Labelling Technologies, 10ª Ed., 2005, Invitrogen, Carlsbad, CA).

Como un ejemplo, los oligonucleótidos pueden marcarse con fluorescencia durante síntesis química, ya que la incorporación de aminas o tioles durante la síntesis de nucleótidos permite la adición de fluoróforos. Están disponibles en el mercado nucleótidos marcados con fluorescencia. Por ejemplo, están disponibles trifosfatos de desoxiuridina y uridina que se conjugan con diez fluoróforos diferentes para abarcar el espectro. También pueden utilizarse colorantes fluorescentes que pueden unirse directamente con nucleótidos como moléculas detectables. Por ejemplo, FAM, JOE, TAMRA y ROX son colorantes fluorescentes reactivos a amina que se han unido a nucleótidos y se usan en secuenciación de ADN automática. Estos nucleótidos marcados con fluoresceína, por ejemplo, ROX-ddATP, ROX-ddCTP, ROX-ddGTP y ROX-ddUTP, están disponibles en el mercado.

También están disponibles moléculas detectables no radiactivas y no fluorescentes. Como se ha indicado anteriormente, la biotina puede unirse directamente con nucleótidos y detectarse por unión específica y de alta afinidad a avidina o estreptavidina que se ha acoplado químicamente con una enzima que cataliza una reacción colorimétrica (tal como fosfatasa, luciferasa o peroxidasa). Los nucleótidos marcados con digoxigenina también pueden usarse de forma similar para detección no isotópica de ácidos nucleicos. Están disponibles en el mercado nucleótidos marcados con digoxigenina y biotinilados.

También pueden usarse para marcar sondas oligonucleotídicas partículas muy pequeñas, denominadas nanopartículas. Estas partículas varían de 1 a 1000 nm de tamaño e incluyen diversas estructuras químicas tales como partículas de oro y plata y puntos cuánticos. Cuando se irradian con luz blanca incidente en ángulo, las nanopartículas de plata u oro que varían de 40 a 120 nm dispersarán luz monocromática con alta intensidad. La longitud de onda de la luz dispersada depende del tamaño de la partícula. De cuatro a cinco partículas diferentes en proximidad estrecha dispersarán cada una luz monocromática, que cuando se superpone proporcionará un color específico, único. Las partículas son fabricadas por compañías tales como Genicon Sciences (Carlsbad, CA). Pueden unirse partículas de plata u oro derivatizadas a una amplia serie de moléculas incluyendo proteínas, anticuerpos, moléculas pequeñas, ligandos de receptores y ácidos nucleicos. Por ejemplo, la superficie de la partícula puede derivatizarse químicamente para permitir la unión a un nucleótido.

Otros tipos de nanopartículas que pueden usarse para detección de una molécula detectable incluyen puntos cuánticos. Los puntos cuánticos son cristales fluorescentes de 1-5 nm de diámetro que son excitables por luz sobre un amplio intervalo de longitudes de onda. Tras excitación por luz que tiene una longitud de onda apropiada, estos cristales emiten luz, tal como luz monocromática, con una longitud de onda dependiente de su composición química y tamaño. Los puntos cuánticos tales como CdSe, ZnSe, InP o InAs poseen propiedades ópticas únicas; estos y similares puntos cuánticos están disponibles de varias fuentes comerciales (por ejemplo, NN-Labs, Fayetteville, AR, Ocean Nanotech, Fayetteville, AR, Nanoco Technologies, Manchester, Reino Unido, Sigma-Aldrich, St. Louis, MO).

Pueden crearse muchas docenas de clases de partículas de acuerdo con el número de clases de tamaños de los cristales de puntos cuánticos. Las clases de tamaños de los cristales se crean 1) por control estrecho de los parámetros de formación de cristales para crear cada clase de tamaño de partícula deseado o 2) mediante la creación de lotes de cristales como parámetros de formación de cristales controlados ligeramente, seguido de clasificación de acuerdo con el tamaño y/o las longitudes de onda de emisión deseados. Dos ejemplos de referencias en las que se incluyen puntos cuánticos con capas epitaxiales de silicio intrínseco de dispositivos de emisión/detección de luz semiconductores son las Patentes de Estados Unidos n.º 5.293.050 y 5.354.707 de Chapple Sokol *et al*.

En ciertas realizaciones, pueden marcarse cebadores o sondas oligonucleotídicos con uno o más colorantes emisores de luz o de otro modo detectables. La luz emitida por los colorantes puede ser luz visible o luz invisible, tal como luz ultravioleta o infrarroja. En realizaciones a modo de ejemplo, el colorante puede ser un colorante de transferencia de energía por resonancia de fluorescencia (TERF); un colorante de xanteno, tal como fluoresceína y rodamina; un colorante que tiene un grupo amino en la posición alfa o beta (tal como un colorante de naftilamina, 1-dimetilaminonaftil-5-sulfonato, 1-anilino-8-naftaleno sulfonato y 2-p-touidinil-6-naftaleno sulfonato); un colorante que tiene 3-fenil-7-isocianatocumarina; una acridina, tal como 9-isotiocianatoacridina y naranja acridina; un pireno, un bensoxadiazol y un estilbeno; un colorante que tiene 3-(ε-carboxipentil)-3'-etil-5,5'-dimetil-oxacarbocianina (CYA); 6-carboxi fluoresceína (FAM); 5 y 6-carboxirrodamina-110 (R110); 6-carboxirrodamina-6G (R6G); N,N,N',N'-tetrametil-6-carboxirrodamina (TAMRA); 6-carboxi-X-rodamina (ROX); 6-carboxi-4',5'-dicloro-2',7'-dimetoxifluoresceína (JOE); ALEXA FLUOR™; Cy2; Texas Red y rodamina red; 6-carboxi-2',4,7,7'-tetraclorofluoresceína (TET); 6-carboxi-2',4,4',5',7,7'-hexaclorofluoresceína (HEX); 5-carboxi-2',4',5',7'-tetraclorofluoresceína (ZOE); NAN; NED; Cy3; Cy3.5; Cy5; Cy5.5; Cy7; y Cy7.5; IR800CW, ICG, Alexa Fluor 350; Alexa Fluor 488; Alexa Fluor 532; Alexa Fluor 546; Alexa Fluor 568; Alexa Fluor 594; Alexa Fluor 647; Alexa Fluor 680 o Alexa Fluor 750.

Los polinucleótidos y oligonucleótidos de MetRS de la presente invención pueden usarse en cualquiera de las composiciones y los métodos terapéuticos, de diagnóstico, de investigación o de descubrimiento de fármacos descritos en el presente documento.

V. ANTICUERPOS

De acuerdo con otro aspecto, la presente invención proporciona además anticuerpos que muestran especificidad de unión por un polipéptido de MetRS, o su compañero de unión celular nativo (es decir receptor celular, lípido, carbohidrato, proteína o compañero de unión de ácido nucleico), o complejo del mismo, y métodos de uso de los mismos. El término anticuerpo incluye las diversas variaciones del mismo, tales como FAB, anticuerpos humanos, anticuerpos humanos modificados, cadenas sencillas, anticuerpos no humanos y otros derivados del pliegue de inmunoglobulina que subyacen en ligandos del sistema inmunitario para antígenos, como se describe en el presente documento y se conoce en la técnica. Pueden usarse anticuerpos en cualquiera de los métodos y composiciones terapéuticos, de diagnóstico, de descubrimiento de fármacos o de expresión/purificación de proteínas y composiciones proporcionadas en el presente documento.

Ciertos anticuerpos de la presente invención difieren de ciertos anticuerpos previamente preparados porque pueden distinguir entre los fragmentos proteicos de MetRS de SEQ ID NO: 137 y su MetRS de longitud completa correspondiente, normalmente uniéndose con mayor afinidad con los fragmentos proteicos de MetRS que con la MetRS de longitud completa correspondiente. En general, dichos anticuerpos pueden unirse a secuencias únicas o estructuras generadas o reveladas por variaciones de corte y empalme, proteólisis u otro procesamiento celular que genere un fragmento proteico de MetRS de la invención (por ejemplo, procesamiento postraduccional, incluyendo pero sin limitación fosforilación y otras modificaciones que cambien la estructura proteica). Por ejemplo, dichos anticuerpos pueden tener especificidad de unión a una o más caras no expuestas al disolvente que se exponen en el fragmento proteico de MetRS pero no en la MetRS de longitud completa, o secuencias que no se encuentran o están de otro modo inaccesibles en la MetRS de longitud completa. Los anticuerpos también pueden unirse a estructuras tridimensionales únicas que resultan de diferencias en el plegamiento entre el fragmento proteico de MetRS y el MetRS de longitud completa. Dichas diferencias en el plegamiento pueden estar localizadas (por ejemplo, en un dominio o una región específico) o globalizadas. Como un ejemplo, el plegamiento de fragmentos proteicos de MetRS puede generar epítopos continuos o discontinuos únicos que no se encuentran en la MetRS correspondiente o parental. Los ejemplos también incluyen anticuerpos que se unen específicamente con extremos N o C terminales generados por variaciones de corte y empalme, proteólisis u otro procesamiento celular; dichos extremos pueden ser únicos en comparación con la MetRS de longitud completa o pueden no estar expuestos a unión a anticuerpo en las versiones de longitud completa debido a que sus extremos están completamente o parcialmente internados en la estructura general de la molécula parental MetRS mayor.

En algunas realizaciones, los anticuerpos proporcionados en el presente documento no forman agregados, tienen una solubilidad deseada, y/o tienen un perfil de inmunogenicidad que es adecuado para su uso en seres humanos, como se describe en el presente documento y se conoce en la técnica. También se incluyen anticuerpos que son adecuados para trabajo de producción, tal como para purificar los fragmentos proteicos de MetRS descritos en el presente documento. Preferentemente, los anticuerpos activos pueden concentrarse hasta al menos aproximadamente 10 mg/ml y opcionalmente formularse para usos bioterapéuticos.

En ciertas realizaciones, los anticuerpos son eficaces para modular una o más de las actividades no canónicas mediadas por un polipéptido de MetRS de la invención. En ciertas realizaciones, por ejemplo, el anticuerpo es uno que se une a un polipéptido de MetRS y/o su compañero de unión, inhibe su capacidad para interaccionar entre sí y/o antagoniza la actividad no canónica del polipéptido de MetRS. En ciertas realizaciones, por ejemplo, el anticuerpo se une al compañero de unión celular de un polipéptido de MetRS, e imita la actividad polipeptídica de MetRS, tal como aumentando o actuando como agonista de la actividad no canónica mediada por el polipéptido de MetRS. En consecuencia, los anticuerpos pueden usarse para diagnosticar, tratar o prevenir enfermedades, trastornos u otras afecciones que están mediados por un polipéptido de MetRS de la invención, tal como

antagonizando o actuando como agonista de su actividad parcial o completamente.

Se dice que un anticuerpo, o fragmento de unión a antígeno del mismo, “se une específicamente”, “se une inmunológicamente” y/o es “inmunológicamente reactivo” a un polipéptido de la invención si reacciona a un nivel detectable (dentro de, por ejemplo, un ensayo de ELISA) con el polipéptido, y no reacciona de forma detectable de una manera estadísticamente significativa con polipéptidos no relacionados en condiciones similares. En ciertos casos, un agente de unión no interacciona significativamente con una versión de longitud completa del polipéptido de MetRS.

La unión inmunológica, como se usa en este contexto, generalmente se refiere a las interacciones no covalentes del tipo que aparecen entre una molécula de inmunoglobulina y un antígeno para el que la inmunoglobulina es específica. La fuerza, o afinidad de unión tal como interacciones de unión inmunológica puede expresarse con respecto a la constante de disociación (K_d) de la interacción, en la que una K_d más pequeña representa una mayor afinidad. Las propiedades de unión inmunológica de polipéptidos seleccionados pueden cuantificarse usando métodos bien conocidos en la técnica. Véase, por ejemplo, Davies *et al*. (1990) Annual Rev. Biochem. 59: 439-473. En ciertas realizaciones ilustrativas, un anticuerpo tiene una afinidad por un fragmento proteico de MetRS de al menos aproximadamente 0,01, 0,05, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40 o 50 nM. En ciertas realizaciones, la afinidad del anticuerpo por un fragmento proteico de MetRS es más fuerte que su afinidad por un polipéptido de MetRS de longitud completa correspondiente, normalmente en aproximadamente 1,5x, 2x, 2,5x, 3x, 3,5x, 4x, 4,5x, 5x, 6x, 7x, 8x, 9x, 10x, 15x, 20x, 25x, 30x, 40x, 50x, 60x, 70x, 80x, 90x, 100x, 200x, 300x, 400x, 500x, 600x, 800x, 900x, 1000x o más (incluyendo todos los números enteros entre medias). En ciertas realizaciones, un anticuerpo tiene una afinidad por una proteína de MetRS de longitud completa correspondiente de al menos aproximadamente 0,05, 0,1, 0,25, 0,5, 0,75, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 µM. En ciertas realizaciones, un anticuerpo se une débilmente o sustancialmente de forma indetectable con una proteína de MetRS de longitud completa.

Un “sitio de unión a antígeno”, o “parte 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 a antígeno se forma por restos de aminoácidos de las regiones variables (“V”) N terminales de las cadenas pesada (“H”) y ligera (“L”). Tres tramos altamente divergentes dentro de las regiones V de las cadenas pesada y ligera se denominan “regiones hipervariables” que se interponen entre tramos flanqueantes más conservados conocidos como “regiones marco conservadas” o “FR”. Por lo tanto el término “FR” se refiere a secuencias de aminoácidos que se encuentran de forma natural entre y adyacentes a regiones hipervariables en inmunoglobulinas. En una molécula de anticuerpo, las tres regiones hipervariables de una cadena ligera y las tres regiones hipervariables de una cadena pesada se disponen entre sí en un espacio tridimensional para formar una superficie de unión a antígeno. La superficie de unión a antígeno es complementaria de la superficie tridimensional de un antígeno unido, y las tres regiones hipervariables de cada una de las cadenas pesada y ligera se denominan “regiones determinantes de complementariedad” o “CDR”.

Pueden prepararse anticuerpos por cualquiera de diversas técnicas conocidas por los expertos habituales en la materia. Véase, por ejemplo, Harlow y Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988. Pueden prepararse anticuerpos monoclonales específicos para un polipéptido de interés, por ejemplo, usando la técnica de Kohler y Milstein, Eur. J. Immunol. 6: 511-519, 1976, y mejoras de la misma. También se incluyen métodos que utilizan animales transgénicos tales como ratones para expresar anticuerpos humanos. Véase, por ejemplo, Neuberger *et al*., Nature Biotechnology 14: 826, 1996; Lonberg *et al*., Handbook of Experimental Pharmacology 113: 49-101, 1994; y Lonberg *et al*., Internal Review of Immunology 13: 65-93, 1995. Los ejemplos particulares incluyen la plataforma VELOCIMMUNE® de REGERNEREX® (véase, por ejemplo Patente de Estados Unidos n.º 6.596.541). También pueden generarse o identificarse anticuerpos mediante el uso de bibliotecas de presentación en fagos o presentación en levadura (véase, por ejemplo, Patente de Estados Unidos n.º 7.244.592, Chao *et al*., Nature Protocols, 1: 755-768, 2006). Los ejemplos no limitantes de bibliotecas disponibles incluyen bibliotecas clonadas o sintéticas, tales como la Biblioteca de Anticuerpos Combinatoria Humana (HuCAL), en las que la diversidad estructural del repertorio de anticuerpos humanos se representa por siete genes de región variable de cadena pesada y siete genes de región variable de cadena ligera. La combinación de estos genes da lugar a 49 marcos conservados en la biblioteca maestra. Superponiendo casetes genéticos altamente variables (CDR = regiones determinantes de complementariedad) en estos marcos conservados, el amplio repertorio de anticuerpos humanos puede reproducirse. También se incluyen bibliotecas humanas diseñadas con fragmentos obtenidos de donantes humanos que codifican una región variable de cadena ligera, una CDR3 de cadena pesada, ADN sintético que codifica diversidad en CDR1 de cadena pesada y ADN sintético que codifica diversidad en CDR2 de cadena pesada. Otras bibliotecas adecuadas para su uso resultarán evidentes para los expertos en la materia. Los polipéptidos de la presente invención pueden usarse en el proceso de purificación, por ejemplo, en una etapa de cromatografía de afinidad.

Un fragmento “Fv” puede producirse por escisión proteolítica preferente de un IgM, y en raras ocasiones molécula de inmunoglobulina IgG o IgA. Los fragmentos Fv, sin embargo, se derivan más habitualmente usando técnicas recombinantes conocidas en este campo. El fragmento Fv incluye un heterodímero V_H::V_L no covalente que incluye

un sitio de unión a antígeno que conserva muchas de las capacidades de reconocimiento y unión a antígeno de la molécula de anticuerpo nativa. Véase, por ejemplo, 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 Fv monocatenario ("sFv") es un heterodímero de V_H::V_L unido covalentemente que se expresa a partir de una fusión génica que incluye genes que codifican V_H y V_L unidos por un enlazador que codifica péptidos. Huston *et al*. (1988) PNAS USA. 85 (16): 5879-5883. Se han descrito varios métodos para diferenciar estructuras químicas para convertir las cadenas polipeptídicas ligeras y pesadas agregadas de forma natural, pero separadas químicamente, de una región V de anticuerpo en una molécula de sFv que se plegará en una estructura tridimensional sustancialmente similar a la estructura de un sitio de unión a antígeno. Véase, por ejemplo, Patentes de Estados Unidos n.º 5.091.513 y 5.132.405, de Huston *et al*.; y Patente de Estados Unidos n.º 4.946.778, de Ladner *et al*.

Cada una de las moléculas descritas anteriormente incluye un conjunto de CDR de una cadena pesada y una cadena ligera, respectivamente interpuestas entre un conjunto de FR de una cadena pesada y una cadena ligera que proporcionan apoyo a las CDR y definen la relación espacial de las CDR entre sí. Como se usa en el presente documento, la expresión "conjunto de CDR" se refiere a las tres regiones hipervariables de una región V de cadena pesada o ligera. Avanzando desde el extremo N terminal 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 ligera. Un polipéptido que comprende una única CDR (por ejemplo, una CDR1, CDR2 o CDR3) se denomina en el presente documento "unidad de reconocimiento molecular". El análisis cristalográfico de varios complejos de anticuerpo-antígeno ha demostrado que los restos de aminoácidos de CDR forman contacto extensivo con antígeno unido, en el que la mayoría del contacto de antígeno extensivo es con la CDR3 de cadena pesada. Por tanto, las unidades de reconocimiento molecular son principalmente responsables de la especificidad de un sitio de unión a antígeno.

Como se usa en el presente documento, la expresión "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 ligera. Algunos restos de FR pueden entrar en contacto con antígeno unido; sin embargo, las FR son principalmente responsables del plegamiento de la región V en el sitio de unión a antígeno, particularmente los restos de FR directamente adyacentes a las CDR. Dentro de las FR, ciertos restos de amino y ciertas características estructurales están muy altamente conservados. A este respecto, todas las secuencias de región V contienen un bucle disulfuro interno de aproximadamente 90 restos de aminoácidos. Cuando las regiones V se pliegan en un sitio de unión, las CDR se presentan como motivos de bucle de proyección que forman una superficie de unión a antígeno. Se reconoce en general que hay regiones estructurales conservadas de FR que influyen en la forma plegada de los bucles de CDR en ciertas estructuras "canónicas" independientemente de la secuencia de aminoácidos de CDR precisa. Además, se sabe que ciertos restos de FR participan en contactos interdominio no covalentes que estabilizan la interacción de las cadenas pesadas y ligeras de anticuerpos.

Ciertas realizaciones incluyen anticuerpo de dominio sencillo (sdAb o "nanocuerpos"), que se refieren a un fragmento de anticuerpo que consiste en un único dominio de anticuerpo variable monomérico (véase, por ejemplo, Patentes de Estados Unidos n.º 5.840.526, 5.874.541, 6.005.079, 6.765.087, 5.800.988; 5.874.541 y 6.015.695). Dichos sdAB normalmente tienen un peso molecular de aproximadamente 12-15 kDa. En ciertos aspectos, sdAB y otras moléculas de anticuerpos pueden derivarse o aislarse de los anticuerpos de cadena pesada únicos de camellos y llamas inmunizados, con frecuencia denominados camélidos. Véase, por ejemplo, Conrath *et al*., JBC. 276: 7346-7350, 2001.

Varias moléculas de anticuerpos "humanizadas" que comprenden un sitio de unión a antígeno derivado de una inmunoglobulina no humana se han descrito, incluyendo anticuerpos quiméricos que tienen regiones V de roedores y sus CDR asociadas fusionadas con dominios constantes humanos (Winter *et al*. (1991) Nature 349: 293-299; Lobuglio *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 una FR de soporte humana antes de la fusión con un dominio constante de anticuerpo humano apropiado (Riechmann *et al*. (1988) Nature 332: 323-327; Verhoeyen *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 sustituidas en superficie de forma recombinante (Publicación de Patente Europea n.º 519.596, publicada el 23 de diciembre de 1992). Estas moléculas "humanizadas" se diseñan para minimizar una respuesta inmunológica no deseada hacia moléculas de anticuerpos antihumanos de roedores que limita la duración y eficacia de aplicaciones terapéuticas de esos restos en receptores humanos. Véase, por ejemplo, Patentes de Estados Unidos n.º 5.530.101; 5.585.089; 5.693.762; 6.180.370; y 7.022.500.

Los anticuerpos de la presente invención pueden usarse en cualquiera de los métodos terapéuticos, de diagnóstico, de descubrimiento de fármacos, de purificación de proteínas y analíticos y composiciones descritas en el presente documento.

VI. ALTERNATIVAS A ANTICUERPOS Y OTROS AGENTES DE UNIÓN

De acuerdo con otro aspecto, la presente invención proporciona además alternativas a anticuerpos u otros agentes de unión, tales como receptores solubles, adnectinas, péptidos, peptidomiméticos, moléculas pequeñas, aptámeros, etc., que muestran especificidad de unión por un polipéptido de MetRS o su compañero de unión celular como se desvela en el presente documento, o con una parte, variante o derivado de los mismos y composiciones y métodos de uso de los mismos. Pueden usarse agentes de unión en cualquiera de los métodos terapéuticos, de diagnóstico, de descubrimiento de fármacos o de expresión/purificación de proteínas, y analíticos y composiciones descritos en el presente documento. Son particularmente útiles agentes de unión basados en productos biológicos tales como adnectinas, receptores solubles, avímeros y trinectinas.

En ciertas realizaciones, dichos agentes de unión son eficaces para modular una o más de las actividades no canónicas mediadas por un polipéptido de MetRS de la invención. En algunas realizaciones, por ejemplo, el agente de unión es uno que se une a un polipéptido de MetRS y/o su compañero de unión, inhibe su capacidad para interaccionar entre sí, y/o antagoniza la actividad no canónica del polipéptido de MetRS. En ciertas realizaciones, por ejemplo, el agente de unión se une al compañero de unión celular de un polipéptido de MetRS, e imita la actividad polipeptídica de MetRS, tal como aumentando o actuando como agonista de la actividad no canónica mediada por el polipéptido de MetRS. En consecuencia, dichos agentes de unión pueden usarse para diagnosticar, tratar o prevenir enfermedades, trastornos u otras afecciones que están mediadas por un polipéptido de MetRS de la invención, tal como antagonizando o actuando como agonista de su actividad parcial o completamente.

Se dice que un agente de unión “se une específicamente” con un polipéptido de MetRS de la invención, o su compañero de unión celular, si reacciona a un nivel detectable (dentro de, por ejemplo, un ensayo de ELISA) con el polipéptido o su compañero de unión celular, y no reacciona de forma detectable de una manera estadísticamente significativa con polipéptidos no relacionados en condiciones similares. En ciertos casos, un agente de unión no interacciona significativamente con una versión de longitud completa del polipéptido de MetRS. En ciertas realizaciones ilustrativas, un agente de unión tiene una afinidad por un fragmento proteico de MetRS o su compañero de unión celular de al menos aproximadamente 0,01, 0,05, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27 , 28, 29, 30, 40 o 50 nM. En ciertas realizaciones, la afinidad del agente de unión por un fragmento proteico de MetRS es más fuerte que su afinidad por un polipéptido de MetRS de longitud completa correspondiente, normalmente en aproximadamente 1,5x, 2x, 2,5x, 3x, 3,5x, 4x, 4,5x, 5x, 6x, 7x, 8x, 9x, 10x, 15x, 20x, 25x, 30x, 40x, 50x, 60x, 70x, 80x, 90x, 100x, 200x, 300x, 400x, 500x, 600x, 700x, 800x, 900x, 1000x o más (incluyendo todos los número enteros entre medias). En ciertas realizaciones, un agente de unión tiene una afinidad por una proteína de MetRS de longitud completa correspondiente de al menos aproximadamente 0,5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 µM.

Como se ha indicado anteriormente, se incluyen “péptidos” como agentes de unión. El término péptido normalmente se refiere a un polímero de restos de aminoácidos y a variantes y análogos sintéticos del mismo. En ciertas realizaciones, 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 (por ejemplo, 5-10, 8-12, 10-15) entre medias, e interaccionan con un polipéptido de MetRS, su compañero de unión celular, o ambos. Los péptidos pueden estar compuestos de aminoácidos de origen natural y/o aminoácidos de origen no natural, como se ha descrito en el presente documento.

Además de los péptidos que consisten solamente en aminoácidos de origen natural, también se proporcionan peptidomiméticos o análogos peptídicos. Los análogos peptídicos se usan habitualmente en la industria farmacéutica como fármacos no peptídicos con propiedades análogas a las del péptido molde. Estos tipos de compuesto no peptídico se denominan “miméticos peptídicos” o “peptidomiméticos” (Luthman, *et al*., A Textbook of Drug Design and Development, 14: 386-406, 2ª Ed., Harwood Academic Publishers (1996); Joachim Grante, Angew. Chem. Int. Ed. Engl., 33: 1699-1720 (1994); Fauchere, J., Adv. Drug Res., 15: 29 (1986); Veber y Freidinger TINS, p. 392 (1985); y Evans, *et al*., J. Med. Chem. 30: 229 (1987)). Un peptidomimético es una molécula que imita la actividad biológica de un péptido pero ya no es de naturaleza química peptídica. Se conocen en la técnica compuestos peptidomiméticos y se describen, por ejemplo, en la Patente de Estados Unidos n.º 6.245.886.

La presente invención también incluye peptoides. Los derivados peptoides de péptidos representan otra forma de péptidos modificados que conservan los determinantes estructurales importantes para actividad biológica, pero eliminan los enlaces peptídicos, confiriendo de este modo resistencia a proteólisis (Simon, *et al*., PNAS USA 89: 9367-9371, 1992). Los peptoides son oligómeros de glicinas N sustituidas. Se han descrito varios grupos N-alquilo, que corresponden cada uno a la cadena lateral de un aminoácido natural. Los peptidomiméticos de la presente invención incluyen compuestos en los que al menos un aminoácido, algunos aminoácidos o todos los restos de aminoácidos se reemplazan por las glicinas N sustituidas correspondientes. Se describen bibliotecas de peptoides, por ejemplo, en la Patente de Estados Unidos n.º 5.811.387.

Un agente de unión también puede incluir una o más moléculas pequeñas. Una “molécula pequeña” se refiere a un compuesto orgánico que es de origen sintético o biológico (biomolécula), pero que normalmente no es un polímero. Los compuestos orgánicos se refieren a una gran clase de compuestos químicos cuyas moléculas contienen

carbono, normalmente excluyendo las que contienen solamente carbonatos, óxidos sencillos de carbono o cianuros. Una “biomolécula” se refiere en general a una molécula orgánica que se produce por un organismo vivo, incluyendo moléculas poliméricas (biopolímeros) tales como péptidos, polisacáridos y ácidos nucleicos también y moléculas pequeñas tales como metabolitos secundarios primarios, lípidos, fosfolípidos, glucolípidos, esteroles, glicerolípidos, vitaminas y hormonas. Un “polímero” se refiere en general a una molécula grande o macromolécula compuesta de unidades estructurales repetidas que se conectan normalmente por enlace químico covalente.

En ciertas realizaciones, una molécula pequeña tiene un peso molecular de menos de 1000-2000 Dalton, normalmente entre aproximadamente 300 y 700 Dalton, e incluyendo aproximadamente 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 500, 650, 600, 750, 700, 850, 800, 950, 1000 o 2000 Dalton. Se describen en otra parte del presente documento bibliotecas de moléculas pequeñas.

Los aptámeros también se incluyen como agentes de unión (véase, por ejemplo, Ellington *et al*., Nature 346, 818-22, 1990, y Tuerk *et al*., Science 249, 505-10, 1990). Los ejemplos de aptámeros incluyeron aptámeros de ácido nucleico (por ejemplo, aptámeros de ADN, aptámeros de ARN) y aptámeros peptídicos. Los aptámeros de ácido nucleico se refieren en general a especies de ácido nucleico que se han modificado técnicamente mediante ciclos repetidos de selección *in vitro* o un método equivalente, tal como SELEX (evolución sistemática de ligandos por enriquecimiento exponencial), para unirse a diversas dianas moleculares tales como moléculas pequeñas, proteínas, ácidos nucleicos e incluso células, tejidos y organismos. Véase, por ejemplo, Patentes de Estados Unidos n.º 6.376.190; y 6.387.620. Por lo tanto, se incluyen aptámeros de ácidos nucleicos que se unen a los polipéptidos de MetRS descritos en el presente documento y/o sus compañeros de unión celulares.

Los aptámeros peptídicos normalmente incluyen un bucle peptídico variable unido en ambos extremos con un armazón proteico, una restricción estructural doble que aumenta normalmente la afinidad de unión del aptámero peptídico hasta niveles comparables con los de un anticuerpo (por ejemplo, en el intervalo nanomolar). En ciertas realizaciones, la longitud del bucle variable puede estar compuesta de aproximadamente 10-20 aminoácidos (incluyendo todos los números enteros entre medias), y el armazón puede incluir cualquier proteína que tenga buena solubilidad y propiedades de compactibilidad. Ciertas realizaciones a modo de ejemplo pueden utilizar la proteína bacteriana Tiorredoxina-A como una proteína de armazón, insertándose el bucle variable dentro del sitio activo reductor (-Cys-Gly-Pro-Cys-bucle en la proteína silvestre), siendo las dos cadenas laterales de cisteínas capaces de formar un enlace disulfuro. Se describen métodos para identificar aptámeros peptídicos, por ejemplo, en la Solicitud de Estados Unidos n.º 2003/0108532. Por lo tanto, se incluyen aptámeros peptídicos que se unen a los polipéptidos de MetRS descritos en el presente documento y/o sus compañeros de unión celulares. Puede realizarse selección de aptámeros peptídicos usando sistemas diferentes conocidos en la técnica, incluyendo el sistema de dos híbridos de levadura.

También se incluyen ADNECTINS™, AVIMERS™, anáfonos y anticalinas que se unen específicamente con un fragmento proteico de MetRS de la invención. ADNECTINS™ se refiere a una clase de productos biológicos dirigidos derivados de fibronectina humana, una proteína extracelular abundante que se une de forma natural con otras proteínas. Véase, por ejemplo, Solicitud de Estados Unidos n.º 2007/0082365; 2008/0139791; y 2008/0220049. ADNECTINS™ normalmente consiste en una cadena principal de fibronectina natural, así como los dominios de dirección múltiple de una parte específica de fibronectina humana. Los dominios de dirección pueden modificarse técnicamente para permitir que una ADNECTIN™ reconozca específicamente una diana de interés terapéutica, tal como un fragmento proteico de MetRS de la invención.

AVIMERS™ se refiere a proteínas de unión multiméricas o péptidos modificados técnicamente usando redistribución de exones *in vitro* y presentación en fagos. Se unen múltiples dominios de unión, dando como resultado mayor afinidad y especificidad en comparación con dominios de inmunoglobulina de epítopos individuales. Véase, por ejemplo, Silverman *et al*., Nature Biotechnology. 23: 1556-1561, 2005; Patente de Estados Unidos n.º 7.166.697; y Solicitudes de Estados Unidos n.º 2004/0175756, 2005/0048512, 2005/0053973, 2005/0089932 y 2005/0221384.

También se incluyen proteínas de repeticiones de anquirina diseñadas (DARPin), que incluyen una clase de proteínas distintas de inmunoglobulinas que pueden ofrecer ventajas frente a anticuerpos para unión a diana en descubrimiento de fármacos y desarrollo de fármacos. Entre otros usos, las DARPin están adaptadas idealmente para captura de imágenes *in vivo* o suministro de toxinas u otras cargas útiles terapéuticas debido a sus propiedades moleculares favorables, incluyendo pequeño tamaño y alta estabilidad. La producción de bajo coste en bacterias y la rápida generación de muchas DARPin específicas de diana hacen al enfoque de DARPin útil para descubrimiento de fármacos. Adicionalmente, pueden generarse fácilmente DARPin en formatos multiespecíficos, ofreciendo el potencial de dirigir una DARPin efectora a un órgano específico o a múltiples receptores diana con una molécula compuesta de varias DARPin. Véase, por ejemplo, Stumpp *et al*., Curr Opin Drug Discov Devel. 10: 153-159, 2007; Solicitud de Estados Unidos n.º 2009/0082274; y documento PCT/EP2001/10454.

Ciertas realizaciones incluyen “monocuerpos”, que utilizan normalmente el 10º dominio de fibronectina de tipo III de fibronectina humana (FNfn10) como un armazón para presentar múltiples bucles de superficie para unión a diana. FNfn10 es una proteína pequeña (94 restos) con una estructura de tipo sándwich β similar al plegamiento de inmunoglobulina. Es altamente estable sin enlaces disulfuro o iones metálicos, y puede expresarse en la forma

plegada correctamente a un alto nivel en bacterias. El armazón FNfn10 es compatible con prácticamente cualquier tecnología de presentación. Véase, por ejemplo, Batori *et al*., Protein Eng. 15: 1015-20, 2002; y Wojcik *et al*., Nat Struct Mol Biol., 2010; y Patente de Estados Unidos n.º 6.673.901.

Las anticalinas se refieren a una clase de miméticos de anticuerpos, que se sintetizan normalmente a partir de lipocalinas humanas, una familia de proteínas de unión a una región en bucle hipervariable soportada por un armazón estructuralmente rígido. Véase, por ejemplo, Solicitud de Estados Unidos n.º 2006/0058510. Las anticalinas normalmente tienen un tamaño de aproximadamente 20 kDa. Las anticalinas pueden caracterizarse por una estructura en barril formada por ocho cadenas β antiparalelas (un armazón de barril β estable) que se conectan por pares con cuatro bucles peptídicos y una hélice a unida. En ciertos aspectos, se realizan desviaciones conformacionales para conseguir unión específica en la región o las regiones de bucle hipervariable. Véase, por ejemplo, Skerra, FEBS J. 275: 2677-83, 2008.

VII. BIOENSAYOS Y ENSAYOS ANALÍTICOS PARA ENSAYOS DE LIBERACIÓN DE FÁRMACOS Y ESPECIFICACIONES, DIAGNÓSTICOS Y REACTIVOS DE PRODUCTOS

También se incluyen bioensayos que se refieren a los fragmentos proteicos de MetRS y agentes relacionados como reactivos terapéuticos y de diagnóstico. Los ejemplos incluyen bioensayos y ensayos analíticos que miden la pureza, actividad biológica, afinidad, solubilidad, pH, los niveles de endotoxinas, entre otros, muchos de los cuales se describen en el presente documento. También se incluyen ensayos que establecen curvas de respuesta a dosis y/o proporcionan una o más bases para comparación entre diferentes lotes de agentes. Las comparaciones de lotes pueden basarse en uno cualquiera o más de caracterización química, caracterización biológica y caracterización clínica. Para agentes proteicos, también se incluyen métodos para evaluar la potencia, estabilidad, farmacocinética e inmunogenicidad de un agente seleccionado. Entre otros usos, estos y otros métodos pueden usarse para ensayos de liberación por lotes de agentes biológicos o químicos, incluyendo los fragmentos proteicos de MetRS, anticuerpos, agentes de unión, polinucleótidos tales como agentes antisentido y vectores y otros descritos en el presente documento.

Ciertas realizaciones incluyen el uso de ensayos de bioafinidad. Dichos ensayos pueden usarse para evaluar la afinidad de unión, por ejemplo, entre un fragmento proteico de MetRS y un compañero de unión celular, o entre un fragmento proteico de MetRS y un anticuerpo. La afinidad de unión también puede medirse entre un fragmento proteico de MetRS y un agente de unión alternativo tal como un compuesto de ensayo candidato o líder (por ejemplo, modulador de molécula pequeña de una MetRS) o entre un compañero de unión celular de MetRS y un compuesto de ensayo candidato o líder. Ciertos ensayos de afinidad de unión a modo de ejemplo pueden utilizar ensayos de ELISA, como se describe en el presente documento y se conoce en la técnica. Ciertos ensayos utilizan cromatografía de unión a receptor de alto rendimiento (véase, por ejemplo, Roswall *et al*., Biologicals., 24: 25-39, 1996). Otros ensayos de afinidad de unión a modo de ejemplo pueden utilizar tecnologías basadas en resonancia de plasmón superficial (RPS). Los ejemplos incluyen tecnologías BIACore, ciertas de las cuales integran tecnología RPS con un sistema microfluídico para supervisar las interacciones moleculares en tiempo real a concentraciones que varían de pM a mM. También se incluyen ensayos KINEXA™, que proporcionan mediciones precisas de especificidad de unión, afinidad de unión y constantes de velocidad/cinética de unión.

Ciertas realizaciones se refieren a inmunoensayos para evaluar u optimizar la inmunogenicidad de agentes proteicos. Los ejemplos incluyen ensayos celulares humanos *ex vivo* y ensayos inmunoenzimáticos *in vitro* para proporcionar información útil sobre el potencial inmunogénico de una proteína terapéutica. Pueden usarse ensayos de respuesta celular *ex vivo*, por ejemplo, para reproducir la cooperación celular entre células presentadoras de antígenos (APC) y linfocitos T, y de este modo medir la activación de linfocitos T después del contacto con una proteína de interés. Ciertos ensayos enzimáticos *in vitro* pueden utilizar una colección de moléculas de HLA-DR recombinantes que abarcan una parte significativa de una población humana relevante, y pueden incluir ensayos inmunoenzimáticos automatizados para ensayar la unión de péptidos (que surgen de la fragmentación de la proteína terapéutica) con las moléculas de HLA-DR. También se incluyen métodos para reducir la inmunogenicidad de una proteína seleccionada, tal como usando estos y métodos relacionados para identificar y después retirar o alterar uno o más epítopos de linfocitos T de un agente proteico.

También se incluyen ensayos de liberación biológica (por ejemplo, ensayos basados en células) para medir parámetros tales como actividades biológicas específicas, incluyendo actividades biológicas no canónicas, y citotoxicidad. Ciertos ensayos biológicos específicos incluyen, por ejemplo, ensayos basados en células que utilizan un compañero de unión celular (por ejemplo, receptor de superficie celular) de un fragmento proteico de MetRS seleccionado, que se acopla funcionalmente con una lectura, tal como un indicador fluorescente o luminiscente de una actividad biológica no canónica, como se describe en el presente documento. Por ejemplo, las realizaciones específicas incluyen una célula que comprende un receptor de superficie celular o una parte extracelular del mismo que se une a un fragmento proteico de MetRS, en el que la célula comprende un detector o una lectura. También se incluyen ensayos biológicos *in vivo* para caracterizar la farmacocinética de un agente, tal como un polipéptido de MetRS o anticuerpo, que utiliza normalmente ratones modificados técnicamente u otro mamífero (véase, por ejemplo, Lee *et al*., The Journal of Pharmacology, 281:1431-1439, 1997). Los ejemplos de ensayos biológicos basados en citotoxicidad incluyen ensayos de liberación (por ejemplo, ensayos de liberación de cromo o de europio

para medir la apoptosis; véase, por ejemplo, von Zons *et al*., Clin Diagn Lab Immunol. 4:202-207, 1997), entre otros, que pueden evaluar la citotoxicidad de fragmentos proteicos de MetRS, bien para establecer curvas de respuesta a dosis, ensayos discontinuos u otras propiedades relacionadas con la aprobación por diversas agencias reguladoras, tales como la Administración de Fármacos y Alimentos (FDA).

Dichos ensayos pueden usarse, por ejemplo, para desarrollar una curva de respuesta a dosis para un fragmento proteico de MetRS seleccionado u otro agente, y/o para comparar la curva de respuesta a dosis de diferentes lotes de proteínas u otros agentes. Una curva de respuesta a dosis es un gráfico X-Y que relaciona la magnitud de un tensor con la respuesta de un receptor; la respuesta puede ser una respuesta fisiológica o bioquímica, tal como una actividad biológica no canónica en una célula *in vitro* o en una célula o un tejido *in vivo*, una cantidad terapéuticamente eficaz como se mide *in vivo* (por ejemplo, como se mide por CE_{50}), o muerte, bien medida *in vitro* o *in vivo* (por ejemplo, muerte celular, muerte de un organismo). La muerte se indica habitualmente como una DL_{50}, una dosis derivada estadísticamente que es letal para el 50 % de una población modelada, aunque puede indicarse por CL_{01} (dosis letal para el 1 % de la población de ensayo animal), CL_{100} (dosis letal para el 100 % de la población de ensayo animal), o CL_{LO} (menor dosis que provoca letalidad). Casi cualquier efecto deseado o criterio de valoración puede caracterizarse de esta manera.

La dosis medida de una curva de respuesta se representa normalmente en el eje X y la respuesta se representa en el eje Y. Más normalmente, el logaritmo de la dosis se representa en el eje X, generando con más frecuencia una curva sigmoidea con la parte más pronunciada en el medio. El nivel de efecto no observable (NOEL) se refiere a la dosis experimental más baja para la que no se observa ningún efecto medible y la dosis umbral se refiere al primer punto a lo largo de la gráfica que indica una respuesta por encima de cero. Como regla general, los fármacos más fuertes generan curvas de respuesta a dosis más pronunciadas. Para muchos fármacos, los efectos deseados se encuentran a dosis ligeramente mayores que la dosis umbral, con frecuencia porque dosis menores son relativamente ineficaces y dosis mayores conducen a efectos secundarios indeseados. Para curvas de respuesta a dosis generadas *in vivo*, una curva puede caracterizarse por valores tales como µg/kg, mg/kg o g/kg de peso corporal, si se desea.

Para comparaciones por lotes, puede ser útil calcular el coeficiente de variación (CV) entre diferentes curvas de respuesta a dosis de diferentes lotes (por ejemplo, entre diferentes lotes de fragmentos proteicos de MetRS, anticuerpos u otros agentes), en parte porque el CV permite la comparación entre conjuntos de datos con diferentes unidades o diferentes medios. Por ejemplo, en ciertas realizaciones a modo de ejemplo, dos o tres o más lotes diferentes de fragmentos proteicos de MetRS u otros agentes tienen un CV entre ellos de menos de aproximadamente 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 % o 1 % para una curva de 4, 5, 6, 7 u 8 puntos de dosis. En ciertas realizaciones, la curva de respuesta a dosis se mide en un ensayo basado en células, y su lectura se relaciona con un aumento o una reducción en una actividad no canónica seleccionada del fragmento proteico de MetRS. En ciertas realizaciones, la curva de respuesta a dosis se mide en un ensayo de liberación celular o un modelo animal (por ejemplo, un modelo de ratón), y su lectura se relaciona con la muerte celular o muerte animal. Otras variaciones resultarán evidentes para expertos en la materia.

VIII. SISTEMAS DE EXPRESIÓN Y PURIFICACIÓN

Las realizaciones de la presente invención incluyen métodos y composiciones relacionadas para expresar y purificar los fragmentos proteicos de MetRS u otros agentes basados en polipéptidos de la invención. Dichos polipéptidos de MetRS recombinantes pueden prepararse convenientemente usando protocolos convencionales como se describe por ejemplo en Sambrook, *et al*., (1989, mencionado anteriormente), en particular las Secciones 16 y 17; Ausubel *et al*., (1994, mencionado anteriormente), 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. Como un ejemplo general, pueden prepararse polipéptidos de MetRS por un procedimiento que incluye una o más de las etapas de: (a) preparar una construcción que comprende una secuencia polinucleotídica que codifica un polipéptido de MetRS y que está unida operativamente con un elemento regulador; (b) introducir la construcción en una célula hospedadora; (c) cultivar la célula hospedadora para expresar el polipéptido de MetRS; y (d) aislar el polipéptido de MetRS de la célula hospedadora.

Se describen en otra parte del presente documento polinucleótidos de MetRS. Para expresar un polipéptido deseado, puede insertarse 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. Pueden usarse métodos que conocen bien los expertos en la materia para construir vectores de expresión que contienen secuencias que codifican un polipéptido de interés y elementos de control de la transcripción y la traducción apropiados. Estos métodos incluyen técnicas de ADN recombinante *in vitro*, técnicas de síntesis y de 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 diversos sistemas de vector/hospedador de expresión y pueden utilizarse para contener y expresar secuencias polinucleotídicas. Estos incluyen, pero sin limitación, microorganismos tales como bacterias

transformadas con bacteriófago recombinante, plásmido o vectores de expresión de ADN de cósmidos; levadura transformada con vectores de expresión de levadura; sistemas de células de insectos infectados con vectores de expresión de virus (por ejemplo, baculovirus); sistemas de células vegetales transformados con vectores de expresión de virus (por ejemplo, virus del mosaico de la coliflor, CaMV; virus del mosaico del tabaco, TMV) o con vectores de expresión bacteriana (por ejemplo, plásmidos Ti o pBR322); o sistemas de células animales, incluyendo células de mamífero y más específicamente sistemas de células humanas.

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

En sistemas bacterianos, pueden seleccionarse varios vectores de expresión dependiendo del uso pretendido para el polipéptido expresado. Por ejemplo, cuando son necesarias grandes cantidades, pueden usarse vectores que dirijan expresión de alto nivel de proteínas de fusión que se purifican fácilmente. Dichos vectores incluyen, pero sin limitación, los vectores de clonación y expresión de *E. coli* multifuncionales tales como BLUESCRIPT (Stratagene), en los que la secuencia que codifica el polipéptido de interés puede ligarse en el vector en fase con secuencias para la Met amino-terminal y los 7 restos posteriores de β-galactosidasa de modo que se produzca una proteína híbrida; vectores PIN (Van Heeke y Schuster, J. Biol. Chem. 264: 5503 5509 (1989)); y similares. También pueden usarse vectores pGEX (Promega, Madison, Wis.) para expresar polipéptidos ajenos como proteínas de fusión con glutatión S-transferasa (GST). En general, dichas proteínas de fusión son solubles y pueden purificarse fácilmente a partir de células lisadas mediante adsorción en perlas de glutatión-agarosa seguido de elución en presencia de glutatión libre. Pueden diseñarse proteínas realizadas en dichos sistemas para que incluyan heparina, trombina o sitios de corte y empalme de factor XA proteasa de modo que el polipéptido clonado de interés pueda liberarse del resto GST a voluntad.

Ciertas realizaciones pueden emplear sistemas de expresión basados en *E. coli* (véase, por ejemplo, Structural Genomics Consortium *et al*., Nature Methods 5:135-146, 2008). Estas realizaciones y realizaciones relacionadas pueden basarse parcialmente o totalmente en la clonación independiente de ligamiento (LIC) para producir un vector de expresión adecuado. En realizaciones específicas, la expresión proteica puede controlarse por una ARN polimerasa T7 (por ejemplo, serie de vector pET). Estas realizaciones y realizaciones relacionadas pueden utilizar la cepa hospedadora de expresión BL21 (DE3), un lisógeno λDE3 de BL21 que soporta la expresión mediada por T7 y es deficiente en ion y ompT proteasas para estabilidad de proteína diana mejorada. También se incluyen cepas hospedadoras de expresión que portan plásmidos que codifican ARNt poco usados en *E. coli*, tales como las cepas ROSETTA™ (DE3) y Rosetta 2 (DE3). La lisis celular y la manipulación de muestras también pueden mejorarse usando reactivos comercializados con las marcas comerciales nucleasa BENZONASE® y Reactivo de Extracción de Proteínas BUGBUSTER®. Para el cultivo celular, el medio autoinductor puede mejorar la eficacia de muchos sistemas de expresión, incluyendo sistemas de expresión de alto rendimiento. Los medios de este tipo (por ejemplo, sistema de autoinducción OVERNIGHT EXPRESS™) inducen gradualmente expresión de la proteína mediante desplazamiento metabólico sin la adición de agentes inductores artificiales tales como IPTG. Las realizaciones particulares emplean marcadores de hexahistidina (tales como los comercializados con el nombre comercial fusiones HIS·TAG®), seguidos de purificación por cromatografía de afinidad metálica inmovilizada (IMAC), o técnicas relacionadas. En ciertos aspectos, sin embargo, pueden aislarse proteínas de uso clínico de cuerpos de inclusión de *E. coli*, con o sin el uso de marcadores de afinidad (véase, por ejemplo, Shimp *et al*., Proteins Expr Purif. 50:58-67, 2006). Como ejemplo adicional, ciertas realizaciones pueden emplear un sistema de producción de *E. coli* de alto rendimiento inducido por choque frío, porque la sobreexpresión de proteínas en *Escherichia coli* a baja temperatura mejora su solubilidad y estabilidad (véase, por ejemplo, Qing *et al*., Nature Biotechnology. 22:877-882, 2004).

También se incluyen sistemas de fermentación bacteriana de alta densidad. Por ejemplo, el cultivo de alta densidad celular de *Ralstonia eutropha* permite la producción de proteínas a densidades celulares de más de 150 g/l, y la expresión de proteínas recombinantes a títulos que superan 10 g/l.

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*. (mencionado anteriormente) y Grant *et al*., Methods Enzymol. 153:516-544 (1987). También se incluyen sistemas de expresión de *Pichia pandoris* (véase, por ejemplo, Li *et al*., Nature Biotechnology, 24, 210-215, 2006 y Hamilton *et al*., Science, 301:1244, 2003). Ciertas realizaciones incluyen sistemas de levadura que se modifican técnicamente para glucosilar selectivamente proteínas, incluyendo levaduras que tienen rutas de N-glucosilación humanizadas, entre otros

(véase, por ejemplo, Hamilton *et al*., Science 313:1441-1443, 2006; Wildt *et al*., Nature Reviews Microbiol., 3:119-28, 2005, y Gerngross *et al*., Nature- Biotechnology, 22:1409-1414, 2004; Patentes de Estados Unidos N.º 7.629.163; 7.326.681; y 7.029.872). Únicamente como ejemplo, pueden dejarse crecer cultivos de levadura recombinantes en matraces de Fernbach o fermentadores de 15 l, 50 l, 100 l y 200 l, entre otros.

En casos en los que se usan vectores de expresión vegetales, la expresión de secuencias que codifican polipéptidos puede conducirse por cualquiera de varios promotores. Por ejemplo, pueden usarse promotores víricos tales como los promotores 35S y 19S de CaMV solos o en combinación con la secuencia líder omega de TMV (Takamatsu, EMBO J. 6:307-311 (1987)). Como alternativa, pueden usarse promotores vegetales 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)). Estas construcciones pueden introducirse en células vegetales mediante transformación de ADN directa o transformación mediada por patógenos. Dichas técnicas se describen en varias revisiones disponibles en general (véase, por ejemplo, Hobbs in McGraw Hill, Yearbook of Science and Technology, págs. 191-196 (1992)).

También puede usarse un sistema de insectos para expresar un polipéptido del interés. Por ejemplo, en uno de dichos sistemas, se usa el virus de la poliedrosis nuclear de *Autographa californica* (AcNPV) como un vector para expresar genes ajenos en células de *Spodoptera frugiperda* o en células de *Trichoplusia*. Las secuencias que codifican el polipéptido pueden clonarse en una región no esencial del virus, tal como el gen de poliedrina, y colocarse bajo el control promotor de poliedrina. La inserción exitosa de la secuencia que codifica polipéptidos hará al gen de poliedrina inactivo y producirá virus recombinantes que carecen de proteína de envoltura. Los virus recombinantes pueden después usarse para infectar, por ejemplo, células de *S. frugiperda* o células de *Trichoplusia* en las que el polipéptido de interés puede expresarse (Engelhard *et al*., Proc. Nat. Acad. Sci. USA 91:3224-3227 1994)). También se incluyen sistemas de expresión de baculovirus, incluyendo los que utilizan células SF9, SF21 y *T.* ni (véase, por ejemplo, Murphy y Piwnica-Worms, Curr Protoc Proteins Sci., Capítulo 5: Unidad 5.4, 2001). Los sistemas de insectos pueden proporcionar modificaciones postraduccionales que son similares a sistemas de mamífero

En células hospedadoras de mamífero, están disponibles en general varios sistemas de expresión basados en virus. Por ejemplo, en casos en los que se usa un adenovirus como un vector de expresión, pueden ligarse secuencias que codifican un polipéptido de interés en un complejo de transcripción/traducción de adenovirus que consiste en el promotor tardío y la secuencia líder tripartita. Puede utilizarse inserción en una región E1 o E3 no esencial del genoma viral para obtener un virus viable que es capaz de expresar el polipéptido en células hospedadoras infectadas (Logan y Shenk, Proc. Nat. Acad. Sci. USA 81:3655-3659 (1984)). Además, pueden usarse potenciadores de la transcripción, tales como el potenciador del virus de sarcoma de Rous (VSR), para aumentar la expresión en células hospedadoras de mamífero.

Los ejemplos de líneas celulares hospedadoras de mamífero útiles incluyen línea CV1 de riñón de mono transformada por SV40 (COS-7, ATCC CRL 1651); línea de riñón embrionario humano (293 o células 293 subclonadas para crecimiento en cultivo en suspensión, Graham *et al*., J. Gen Virol., 36:59 (1977)); Células de riñón de cría de hámster (BHK, ATCC CCL 10); células de sertoli de ratón (TM4, Mather, Biol. Reprod. 23:243-251 (1980)); células de riñón de mono (CV1 ATCC CCL 70); células de riñón de mono verde africano (VERO-76, ATCC CRL-1587); células de carcinoma de cuello uterino humano (HELA, ATCC CCL 2); células de riñón canino (MDCK, ATCC CCL 34); células de hígado de rata búfalo (BRL 3A, ATCC CRL 1442); células de pulmón humano (W138, ATCC CCL 75); células de hígado humano (Hep G2, HB 8065); tumor mamario de ratón (MMT 060562, ATCC CCL51); células TR1 (Mather *et al*., Annals N.Y. Acad. Sci. 383:44-68 (1982)); células MRC 5; células FS4; y una línea de hepatoma humano (Hep G2). Otras líneas celulares hospedadoras de mamífero incluyen células de ovario de hámster chino (CHO), incluyendo células DHFR-CHO (Urlaub *et al*., PNAS USA 77:4216 (1980)); y líneas celulares de mieloma tales como NSO y Sp2/0. Para una revisión de ciertas líneas celulares hospedadoras de mamífero adecuadas para producción de anticuerpos, véase, por ejemplo, Yazaki y Wu, Methods in Molecular Biology, Vol. 248 (B. K.C Lo, ed., Humana Press, Totowa, N.J., 2003), pp 255-268. Ciertos sistemas de expresión de células de mamífero preferidos incluyen sistemas de expresión basados en células CHO y HEK293. Los sistemas de expresión de mamífero pueden utilizar líneas celulares unidas, por ejemplo, en matraces en T, frascos rotatorios o fábricas celulares, o cultivos en suspensión, por ejemplo, en agitadores de 1 l y 5 l, biorreactores de tanque agitado de 5 l, 14 l, 40 l, 100 l y 200 l, o biorreactores WAVE de 20/50 l y 100/200 l, entre otros conocidos en la técnica.

También se incluye la expresión sin células de proteínas. Estas realizaciones y otras relacionadas normalmente utilizan ARN polimerasa purificada, ribosomas, ARNt y ribonucleótidos; Estos reactivos pueden producirse por extracción de células o de un sistema de expresión basado en células.

También pueden usarse señales de inicio específicas para conseguir traducción más eficaz de secuencias que codifican un polipéptido de interés. Dichas señales incluyen el codón de inicio ATG y secuencias adyacentes. En casos en los que se insertan secuencias que codifican el polipéptido, su codón de inicio y secuencias cadena arriba en el vector de expresión apropiado, pueden no ser necesarias señales de control de la transcripción o traducción adicionales. Sin embargo, en casos en los que solamente se inserta secuencia codificante, o una parte de la misma, deberían proporcionarse señales de control de la traducción exógenas incluyendo el codón de inicio ATG. Además,

el codón de inicio debería estar en la fase de lectura correcta para asegurar la traducción del inserto completo. Los elementos de traducción y codones de inicio exógenos pueden ser de diversos orígenes, tanto naturales como sintéticos. La eficacia de expresión puede potenciarse por la inclusión de potenciadores que son apropiados para el sistema celular particular que se usa, tales como los descritos en la bibliografía (Scharf *et al*., Results Probl. Cell Difer. 20:125-162 1994)).

Además, puede elegirse una cepa de célula hospedadora por su capacidad para modular la expresión de las secuencias insertadas o para procesar la proteína expresada de la manera deseada. Dichas modificaciones del polipéptido incluyen, pero sin limitación, modificaciones postraduccionales tales como acetilación, carboxilación, glucosilación, fosforilación, lipidación y acilación. El procesamiento postraduccional que escinde una forma "prepro" de la proteína también puede usarse para facilitar la inserción, el plegamiento y/o la función correctos. Pueden elegirse diferentes células hospedadoras tales como levadura, CHO, HeLa, MDCK, HEK293 y W138, además de células bacterianas, que tienen o incluso carecen de maquinaria celular específica y mecanismos característicos para dichas actividades postraduccionales, para asegurar la modificación y el procesamiento correctos de la proteína ajena.

Para producción de alto rendimiento, a largo plazo, de proteínas recombinantes, se prefiere en general la expresión estable. Por ejemplo, las líneas celulares que expresan de forma estable un polinucleótido de interés pueden transformarse usando vectores de expresión que pueden contener orígenes víricos de replicación y/o elementos de expresión endógenos y un gen marcador seleccionable en el mismo vector o en uno separado. Después de la introducción del vector, puede permitirse que las células crezcan durante aproximadamente 1-2 días en un medio enriquecido antes de cambiarse a medio selectivo. El fin del marcador seleccionable es conferir resistencia a la selección, y su presencia permite el crecimiento y recuperación de células que expresan con éxito las secuencias introducidas. Puede hacerse que los clones resistentes de células transformadas de forma estable proliferen usando técnicas de cultivo tisular apropiadas para el tipo celular. También puede emplearse producción transitoria, tal como mediante transfección o infección transitoria. Los sistemas de expresión de mamíferos a modo de ejemplo que son adecuados para la producción transitoria incluyen sistemas basados en HEK293 y CHO.

Puede usarse cualquier variedad de sistemas de selección para recuperar líneas celulares transformadas o transducidas. Estos incluyen, pero sin limitación, los genes de la timidina quinasa del virus del herpes simple (Wigler *et al*., Cell 11:223-232 (1977)) y adenina fosforribosiltransferasa (Lowy *et al*., Cell 22:817-823 (1990)) que pueden emplearse en células tk- o aprt-, respectivamente. Además, puede usarse la resistencia a antimetabolitos, antibióticos o herbicidas como la base de la selección; por ejemplo, dhfr que confiere resistencia al metotrexato (Wigler *et al*., Proc. Nat. Acad. Sci. U.S.A. 77:3567-70 (1980)); Npt, que confiere resistencia a los aminoglucósidos, 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, mencionado anteriormente). Se han descrito genes seleccionables adicionales, por ejemplo, trpB, que permiten que las células utilicen indol en lugar de triptófano, o hisD, que permite que las células utilicen histinol en lugar de histidina (Hartman y Mulligan, Proc. Nat. Acad. Sci. U.S.A. 85:8047-51 (1988)). El uso de marcadores visibles ha obtenido popularidad con marcadores tales como proteína verde fluorescente (GFP) y otras proteínas fluorescentes (por ejemplo, RFP, YFP), antocianinas, β-glucuronidasa y su sustrato GUS, y luciferasa y su sustrato luciferina, usándose ampliamente no solamente para identificar transformantes, sino también para cuantificar la cantidad de expresión proteica transitoria o estable atribuible a un sistema de vector específico (véase, por ejemplo, Rhodes *et al*., Methods Mol. Biol. 55:121-131 (1995)).

Las realizaciones de la presente invención también incluyen sistemas de producción de proteínas de alto rendimiento, o sistemas de microproducción. Ciertos aspectos pueden utilizar, por ejemplo, marcadores de fusión de hexa-histidina para expresión y purificación de proteínas en superficies de portaobjetos modificados con quelados metálicos o partículas de Ni MagneHis (véase, por ejemplo, Kwon *et al*., BMC Biotechnol., 9:72, 2009, y Lin *et al*., Methods Mol Biol. 498:129-41, 2009). También se incluyen sistemas de expresión de proteínas sin células de alto rendimiento (véase, por ejemplo, Sitaraman *et al*., Methods Mol Biol. 498:229-44, 2009). Estas realizaciones y realizaciones relacionadas pueden usarse, por ejemplo, para generar micromatrices de fragmento o fragmentos proteicos de MetRS, que después pueden usarse para explorar bibliotecas para identificar agentes que interaccionen con el fragmento o los fragmentos proteicos de MetRS.

Se conocen en la técnica diversos protocolos para detectar y medir la expresión de productos codificados por polinucleótidos, usando agentes de unión o anticuerpos tales como los anticuerpos policlonales o monoclonales específicos del producto. Los ejemplos incluyen ensayo inmunoabsorbente ligado a enzimas (ELISA), inmunotransferencias de Western, radioinmunoensayos (RIA) y separación de células activadas por fluorescencia (FACS). Estos y otros ensayos se describen, entre otros sitios, en Hampton *et al*., Serological Methods, a Laboratory Manual (1990) y Maddox *et al*., J. Exp. Medicina. 158:1211-1216 (1983).

Los expertos en la materia conocen una amplia diversidad de marcadores y técnicas de conjugación y pueden usarse en diversos ensayos de ácido nucleico y aminoácidos. Los medios para producir hibridación marcada o sondas de PCR para detectar secuencias relacionadas con polinucleótidos incluyen oligomarcaje, traslación de muesca, marcaje de extremos o amplificación por PCR usando un nucleótido marcado. Como alternativa, las

secuencias, o cualquier parte de las mismas pueden clonarse en un vector para la producción de una sonda de ARNm. Dichos vectores se conocen en la técnica, están disponibles en el mercado y pueden usarse para sintetizar sondas de ARN *in vitro* mediante la adición de una ARN polimerasa apropiada tal como T7, T3 o SP6 y nucleótidos marcados. Estos procedimientos pueden realizarse usando diversos kits disponibles en el mercado. Las moléculas indicadoras o los marcadores adecuados que pueden usarse incluyen radionúclidos, enzimas, agentes fluorescentes, quimioluminiscentes o cromogénicos así como sustratos, cofactores, inhibidores, partículas magnéticas y similares.

Pueden cultivarse células hospedadoras transformadas con una secuencia polinucleotídica de interés en condiciones adecuadas para la expresión y recuperación de la proteína de cultivo celular. Ciertas realizaciones específicas utilizan sistemas de expresión celular sin suero. Los ejemplos incluyen células HEK293 y células CHO que pueden cultivarse en medio sin suero (véase, por ejemplo, Rosser *et al*., Protein Expr. Purif. 40:237-43, 2005; y Patente de Estados Unidos n.º 6.210.922).

La proteína producida por una célula recombinante puede secretarse o estar contenida intracelularmente dependiendo de la secuencia y/o el vector usado. Como entenderán los expertos en la materia, los vectores de expresión que contienen polinucleótidos de la invención pueden diseñarse para contener secuencias señal que dirijan la secreción del polipéptido codificado a través de una membrana celular procariótica o eucariótica. Otras construcciones recombinantes pueden usarse para unir secuencias que codifican un polipéptido de interés con una secuencia de nucleótidos que codifica un dominio polipeptídico lo que facilitará la purificación y/o detección de proteínas solubles. Los ejemplos de dichos dominios incluyen marcadores epitópicos y de purificación de afinidad escindibles y no escindibles tales como avidina, marcadores FLAG, marcadores de polihistidina (por ejemplo, 6xHis), marcadores de cMyc, marcadores V5, marcadores de glutatión S-transferasa (GST) y otros.

La proteína producida por una célula recombinante puede purificarse y caracterizarse de acuerdo con diversas técnicas conocidas en este campo. Los sistemas a modo de ejemplo para realizar purificación de proteínas y analizar la pureza de proteínas incluyen cromatografía líquida de proteínas rápida (FPLC) (por ejemplo, sistemas de AKTA y Bio-Rad FPCL), cromatografía líquida de alta presión (HPCL) (por ejemplo, Beckman and Waters HPLC). Las químicas a modo de ejemplo para purificación incluyen cromatografía de intercambio iónico (por ejemplo, Q, S), cromatografía de exclusión por tamaño, gradientes salinos, purificación de afinidad (por ejemplo, Ni, Co, FLAG, maltosa, glutatión, proteína A/G), filtración en gel, fase inversa, cromatografía de intercambio iónico cerámica HYPERD® y columnas de interacción hidrófoba (HIC), entre otras conocidas en la técnica. También se incluyen métodos analíticos tales como SDS-PAGE (por ejemplo, tinción de coomassie, tinción de plata), inmunotransferencia, Bradford, y ELISA, que pueden utilizarse durante cualquier etapa del proceso de producción o purificación, normalmente para medir la pureza de la composición proteica.

También se incluyen métodos para concentrar fragmentos proteicos de MetRS, y composición que comprende proteínas solubles concentradas. En diferentes aspectos dichas soluciones concentradas de polipéptidos de MetRS pueden comprender proteínas a una concentración de aproximadamente 5 mg/ml; o aproximadamente 8 mg/ml; o aproximadamente 10 mg/ml; aproximadamente 15 mg/ml; o aproximadamente 20 mg/ml.

En un aspecto, dichas composiciones pueden ser sustancialmente monodispersas, lo que significa que las composiciones de polipéptido de MetRS existen principalmente (es decir, al menos aproximadamente 90 %, o más) en una forma de peso molecular aparente cuando se evalúan por ejemplo mediante cromatografía de exclusión por tamaño, dispersión de luz dinámica o ultracentrifugación analítica.

En otro aspecto, dichas composiciones tienen una pureza (en proteínas) de al menos aproximadamente 90 %, o en algunos aspectos al menos aproximadamente 95 % de pureza, o en algunas realizaciones, al menos 98 % de pureza. La pureza puede determinarse mediante cualquier método analítico rutinario como se conoce en la técnica.

En otro aspecto, dichas composiciones tienen un contenido de agregado de alto peso molecular de menos de aproximadamente 10 %, en comparación con la cantidad total de proteína presente, o en algunas realizaciones dichas composiciones tienen un contenido de agregado de alto peso molecular de menos de aproximadamente 5 %, o en algunos aspectos dichas composiciones tienen un contenido de agregado de alto peso molecular de menos de aproximadamente 3 %, o en algunas realizaciones un contenido de agregado de alto peso molecular de menos de aproximadamente 1 %. El contenido de agregado de alto peso molecular puede determinarse mediante diversas técnicas analíticas incluyendo, por ejemplo, mediante cromatografía de exclusión por tamaño, dispersión de luz dinámica o ultracentrifugación analítica.

En ciertas realizaciones, como se indica en el presente documento, las composiciones polipeptídicas de MetRS tienen un contenido en endotoxinas de menos de aproximadamente 10 UE/mg de polipéptido de MetRS, o menos de aproximadamente 5 UE/mg de polipéptido de MetRS, menos de aproximadamente 3 UE/mg de polipéptido de MetRS, o menos de aproximadamente 1 UE/mg de polipéptido de MetRS.

Los ejemplos de enfoques de concentración contemplados en el presente documento incluyen liofilización, que se emplea normalmente cuando la solución contiene pocos componentes solubles distintos de la proteína de interés. La

liofilización se realiza con frecuencia después del procesamiento por HPLC, y puede retirar la mayoría o todos de los componentes volátiles de la mezcla. También se incluyen técnicas de ultrafiltración, que normalmente emplean una o más membranas permeables selectivas para concentrar una solución proteica. La membrana permite que la atraviesen agua y moléculas pequeñas y retiene la proteína; la solución puede empujarse contra la membrana por una bomba mecánica, presión de gas o centrifugación, entre otras técnicas.

En ciertas realizaciones, los reactivos, fragmentos proteicos de MetRS, o agentes relacionados (por ejemplo, anticuerpos) tienen una pureza de al menos aproximadamente 90 %, como se mide acuerdo con técnicas rutinarias en este campo. En ciertas realizaciones, tales como composiciones de diagnóstico o ciertas composiciones terapéuticas, las composiciones de MetRS de la presente invención tienen una pureza de al menos aproximadamente 95 %. En realizaciones específicas, tales como composiciones terapéuticas o farmacéuticas, las composiciones de MetRS de la presente invención tienen una pureza de al menos aproximadamente 97 %, 98 % o 99 %. En otras realizaciones, tal como cuando se usan como referencia o reactivos de investigación, los fragmentos proteicos de MetRS pueden ser de menor pureza, y pueden tener una pureza de al menos aproximadamente 50 %, 60 %, 70 % u 80 %. La pureza puede medirse en general o en relación con componentes seleccionados, tales como otras proteínas, por ejemplo, pureza en proteínas.

Los fragmentos proteicos de MetRS purificados también pueden caracterizarse de acuerdo con sus características biológicas. Los ejemplos incluyen afinidad de unión o cinética de unión a un ligando seleccionado (por ejemplo, un compañero de unión celular del fragmento proteico de MetRS tal como un receptor de superficie celular o un dominio extracelular del mismo), y la presencia o los niveles de una o más actividades biológicas canónicas o no canónicas como se describe en el presente documento. La afinidad de unión y cinética de unión pueden medirse de acuerdo con diversas técnicas conocidas en este campo, tales como Biacore® y tecnologías relacionadas que utilizan resonancia de plasmón superficial (RPS), un fenómeno óptico que permite la detección de agentes de interacción no marcados en tiempo real. Pueden usarse biosensores basados en RPS en la determinación de la concentración activa, cribado y caracterización con respecto tanto a afinidad como a cinética. La presencia o los niveles de una o más actividades biológicas no canónicas o canónicas pueden medirse de acuerdo con ensayos basados en células, incluyendo los que utilizan un compañero de unión celular (por ejemplo, receptor de superficie celular) de un fragmento proteico de MetRS seleccionado, que está acoplado funcionalmente con una lectura o un indicador, tal como un indicador fluorescente o luminiscente de una actividad biológica no canónica, como se describe en el presente documento.

En ciertas realizaciones, como se ha indicado anteriormente, las composiciones de polipéptidos de MetRS son aproximadamente sustancialmente sin endotoxinas, incluyendo, por ejemplo, aproximadamente 95 % sin endotoxina, preferentemente aproximadamente 99 % sin endotoxinas, y más preferentemente aproximadamente 99,99 % sin endotoxinas. La presencia de endotoxinas puede detectarse de acuerdo con técnicas rutinarias en este campo, como se describe en el presente documento. En realizaciones específicas, las composiciones de MetRS se preparan a partir de una célula eucariótica tal como una célula de mamífero o humana en medio sustancialmente sin suero.

En ciertas realizaciones, las composiciones polipeptídicas de MetRS comprenden menos de aproximadamente 10 % p/p de agregados de alto peso molecular o menos de aproximadamente 5 % p/p de agregados de alto peso molecular o menos de aproximadamente 2 % p/p de agregados de alto peso molecular, o menos de aproximadamente o menos de aproximadamente 1 % p/p de agregados de alto peso molecular.

También se incluyen ensayos analíticos basados en proteínas y métodos, que pueden usarse para evaluar, por ejemplo, la pureza proteica, el tamaño, la solubilidad y el grado de agregación, entre otras características. La pureza proteica puede evaluarse de varias maneras. Por ejemplo, la pureza puede evaluarse basándose en la estructura primaria, estructura de mayor orden, tamaño, carga, hidrofobicidad y glucosilación. Los ejemplos de métodos para evaluar la estructura primaria incluyen secuenciación N- y C-terminal y mapeo de péptidos (véase, por ejemplo, Allen *et al*., Biologicals., 24:255-275, 1996)). Los ejemplos de métodos para evaluar estructura de mayor orden incluyen dicroísmo circular (véase, por ejemplo, Kelly *et al*., Biochim Biophys Acta. 1751:119-139, 2005), espectroscopia fluorescente (véase, por ejemplo, Meagher *et al*., J. Biol. Chem. 273:23283-89, 1998), FT-IR, cinética de intercambio de hidrógeno-deuterio de amida, calorimetría de cribado diferencial, espectroscopia de RMN, inmunorreactividad con anticuerpos conformacionalmente sensibles. También puede evaluarse una estructura de mayor orden en función de varios parámetros tales como pH, temperatura o sales añadidas. Los ejemplos de métodos para evaluar las características proteicas tales como tamaño incluyen ultracentrifugación analítica y HPLC de exclusión por tamaño (SEC-HPLC), y los métodos a modo de ejemplo para medir la carga incluyen cromatografía de intercambio iónico e isoelectroenfoque. La hidrofobicidad puede evaluarse, por ejemplo, mediante HPLC de fase inversa y HPLC de cromatografía de interacción hidrófoba. La glucosilación puede afectar a farmacocinética (por ejemplo, eliminación), conformación o estabilidad, o unión al receptor y función de proteínas, y puede evaluarse, por ejemplo, en espectrometría de masas y espectroscopia de resonancia magnética nuclear (RMN).

Como se ha indicado anteriormente, ciertas realizaciones incluyen el uso de SEC-HPLC para evaluar características proteicas tales como pureza, tamaño (por ejemplo, homogeneidad de tamaño) o grado de agregación y/o para purificar proteínas, entre otros usos. SEC, que también incluye cromatografía de filtración en gel (GFC) y

cromatografía de permeación de gel (GPC), se refiere a un método cromatográfico en el que se separan moléculas en solución en un material poroso basándose en su tamaño, o más específicamente en su volumen hidrodinámico, coeficiente de difusión y/o propiedades de superficie. El proceso se usa en general para separar moléculas biológicas, y para determinar pesos moleculares y distribuciones de pesos moleculares de polímeros. normalmente, una muestra biológica o proteica (tal como un extracto proteico producido de acuerdo con los métodos de expresión de proteínas proporcionados en el presente documento y conocidos en la técnica) se carga en una columna de exclusión por tamaño seleccionada con una fase estacionaria definida (el material poroso), preferentemente en una fase que no interactúa con las proteínas de la muestra. En ciertos aspectos, la fase estacionaria está compuesta de partículas inertes empaquetadas en una matriz tridimensional densa dentro de una columna de vidrio o acero. La fase móvil puede ser agua pura, un tampón acuoso, un disolvente orgánico o una mezcla de los mismos. Las partículas de fase estacionaria normalmente tienen poros y/o canales pequeños que solamente permiten que entren moléculas por debajo de un cierto tamaño. Las partículas grandes se excluyen por tanto de estos poros y canales, y su interacción limitada con la fase estacionaria las conduce a eluir como un pico "totalmente excluido" al comienzo del experimento. Se retiran moléculas más pequeñas que pueden caber por los poros, de la fase móvil que fluye, y el tiempo que pasan inmovilizadas en los poros de fase estacionaria depende, en parte, de lo lejos que penetran en los poros. Su retirada del flujo de fase móvil provoca que tarden más en eluirse de la columna y da como resultado una separación entre las partículas basándose en diferencias en su tamaño. Una columna de exclusión por tamaño dada tiene un intervalo de pesos moleculares que pueden separarse. En general, moléculas mayores que el límite superior no serán atrapadas por la fase estacionaria, moléculas más pequeñas que el límite inferior entrarán completamente en la fase sólida y eluirán como una única banda, y moléculas dentro del intervalo eluirán a diferentes velocidades, definidas por sus propiedades tales como volumen hidrodinámico. Para ejemplos de estos métodos en la práctica con proteínas farmacéuticas, véase Bruner *et al*., Journal of Pharmaceutical and Biomedical Analysis. 15:1929-1935, 1997.

También se analiza la pureza proteica para aplicaciones clínicas, por ejemplo, en Anicetti *et al*. (Trends in Biotechnology 7:342-349, 1989). Técnicas más recientes para analizar la pureza proteica incluyen, sin limitación, el LabChip GXII, una plataforma automática para el análisis rápido de proteínas y ácidos nucleicos, que proporciona análisis de alto rendimiento de título, calibrado y análisis de pureza de las proteínas. En ciertas realizaciones no limitantes, pueden obtenerse proteínas de uso clínico tales como fragmentos proteicos y anticuerpos utilizando una combinación de materiales cromatográficos en al menos dos etapas ortogonales, entre otros métodos (véase, por ejemplo, Therapeutic Proteins: Methods and Protocols. 308, Eds., Smales y James, Humana Press Inc., 2005). normalmente, los agentes proteicos (por ejemplo, fragmentos proteicos de MetRS, anticuerpos, agentes de unión) y otros agentes (por ejemplo, antisentido, ARNi, moléculas pequeñas) están sustancialmente sin endotoxina, como se mide de acuerdo con técnicas conocidas en este campo y descritas en el presente documento.

También se incluyen ensayos de solubilidad de proteínas. Dichos ensayos pueden utilizarse, por ejemplo, para determinar el crecimiento óptimo y las condiciones de purificación para producción recombinante, para optimizar la elección de tampón o tampones, y para optimizar la elección de fragmentos proteicos de MetRS o variantes de los mismos. La solubilidad o agregación puede evaluarse de acuerdo con diversos parámetros, incluyendo temperatura, pH, sales y la presencia o ausencia de otros aditivos. Los ejemplos de ensayos de cribado de solubilidad incluyen, sin limitación, métodos basados en microplacas para medir la solubilidad de proteínas usando turbidez u otra medida como un criterio de valoración, ensayos de alto rendimiento para análisis de la solubilidad de proteínas recombinantes purificadas (véase, por ejemplo, Stenvall *et al*. , Biochim Biophys Acta. 1752:6-10, 2005), ensayos que usan la complementación estructural de una proteína marcadora genética para supervisar y medir el plegamiento y la solubilidad de las proteínas *in vivo* (véase, por ejemplo, Wigley *et al*., Nature Biotechnology 19:131-136, 2001), y cribado electroquímica de solubilidad proteica recombinante en *Escherichia coli* usando microscopía electroquímica de barrido (SECM) (véase, por ejemplo, Nagamine *et al.*, Biotechnology and Bioengineering, 96:1008-1013, 2006), entre otros. Pueden identificarse y seleccionarse fragmentos proteicos de MetRS con solubilidad aumentada (o agregación reducida) de acuerdo con técnicas rutinarias en este campo, incluyendo ensayos *in vivo* sencillos para solubilidad de proteínas (véase, por ejemplo, Maxwell *et al*., Protein Sci. 8:1908-11, 1999).

La solubilidad de agregación de proteínas también puede medirse por técnicas de dispersión de luz dinámica. La agregación es un término general que abarca varios tipos de interacciones o características, incluyendo interacciones y características solubles/insolubles, covalentes/no covalentes, reversibles/irreversibles y nativas/desnaturalizadas. Para productos terapéuticos proteicos, la presencia de agregados se considera normalmente indeseable debido a la preocupación de que los agregados puedan causar una reacción inmunogénica (por ejemplo, agregados pequeños), o puedan provocar acontecimientos adversos tras su administración (por ejemplo, partículas). La dispersión de luz dinámica se refiere a una técnica que puede usarse para determinar el perfil de distribución de tamaños de partículas pequeñas en suspensión o polímeros tales como proteínas en solución. Esta técnica, también denominada espectroscopia de correlación fotónica (PCS) o dispersión de la luz cuasi elástica (QELS), usa luz dispersada para medir la velocidad de difusión de las partículas proteicas. Pueden observarse fluctuaciones en la intensidad de dispersión debido al movimiento browniano de las moléculas y partículas en la solución. Estos datos de movimiento pueden procesarse convencionalmente para derivar una distribución de tamaños para la muestra, en la que el tamaño se proporciona por el radio de Stokes o radio hidrodinámico de la partícula proteica. El tamaño hidrodinámico depende tanto de la masa como de la forma

(conformación). La dispersión dinámica puede detectar la presencia de cantidades muy pequeñas de proteína agregada (<0,01 % en peso), incluso en muestras que contienen un amplio intervalo de masas. También pueden usarse para comparar la estabilidad de diferentes formulaciones, incluyendo, por ejemplo, aplicaciones que se basan en supervisión en tiempo real de cambios a temperaturas elevadas. En consecuencia, ciertas realizaciones incluyen el uso de dispersión de luz dinámica para analizar la solubilidad y/o presencia de agregados en una muestra que contiene un fragmento proteico de MetRS, anticuerpo u otro agente de la invención.

IX. MÉTODOS DE DIAGNÓSTICO Y COMPOSICIONES

Pueden usarse agentes de MetRS tales como fragmentos proteicos de MetRS, polinucleótidos de MetRS, y anticuerpos y otros agentes de unión descritos en el presente documento en ensayos de diagnóstico y composiciones de diagnóstico. Se incluyen métodos y composiciones bioquímicos, histológicos y basados en células, entre otros.

Estas realizaciones y otras relacionadas incluyen la detección de la secuencia o las secuencias polinucleotídicas de MetRS o secuencia o secuencias polipeptídicas de MetRS correspondientes o partes de las mismas de uno o más fragmentos proteicos de MetRS de nueva identificación, también denominados polipéptidos de MetRS. Por ejemplo, ciertos aspectos incluyen la detección de la secuencia o las secuencias polinucleotídicas de MetRS o secuencia o secuencias polipeptídicas correspondientes o partes de las mismas de una o más variantes de corte y empalme de MetRS de nueva identificación y/o uno o más puntos de unión de corte y empalme de esas variantes de corte y empalme. En ciertas realizaciones, el polinucleótido o secuencia o secuencias polipeptídicas correspondientes de al menos uno de los puntos de unión de corte y empalme es único de esa variante de corte y empalme de MetRS particular.

También se incluye la detección directa de fragmentos proteicos de MetRS, incluyendo variantes de corte y empalme, fragmentos proteolíticos y otros. En ciertas realizaciones, la presencia o los niveles de uno o más fragmentos proteicos de MetRS de nueva identificación se asocian o se correlacionan con uno o más tipos celulares o estados celulares. Por lo tanto, la presencia o los niveles de un polipéptido o polinucleótido de MetRS puede usarse para distinguir entre diferentes tipos celulares o diferentes estados celulares. La presencia o los niveles de fragmentos proteicos de MetRS o sus polinucleótidos relacionados pueden detectarse de acuerdo con técnicas de diagnóstico basadas en polinucleótidos y/o polipéptidos, como se describe en el presente documento y se conoce en la técnica.

Ciertos aspectos pueden emplear los fragmentos proteicos de MetRS, anticuerpo, o polinucleótidos de MetRS como parte de un método de diagnóstico complementario, normalmente para evaluar si un sujeto o una población de sujetos responderá favorablemente a un tratamiento médico específico. Por ejemplo, un agente terapéutico de MetRS dado (por ejemplo, fragmento proteico, antisentido, ARNi, anticuerpo, agente de unión) podría identificarse como adecuado para un sujeto o ciertas poblaciones de sujetos basándose en si el sujeto o los sujetos tienen uno o más biomarcadores seleccionados para una enfermedad o afección dada. Los ejemplos de biomarcadores incluyen marcadores de suero/tejido, así como marcadores que pueden identificarse mediante técnicas de captura de imágenes médicas. En ciertas realizaciones, un fragmento proteico de MetRS de origen natural (o su polinucleótido correspondiente) puede proporcionar en sí mismo un biomarcador de suero y/o tisular que puede utilizarse para medir el resultado farmacológico o evaluar la conveniencia del uso farmacológico en un sujeto específico o una población de sujetos específica. En ciertos aspectos, la identificación de una secuencia de referencia polipeptídica o polinucleotídica de MetRS puede incluir caracterizar la expresión diferencial de esa secuencia, bien en un sujeto seleccionado, tejido seleccionado, o de otro modo, como se describe en el presente documento y se conoce en la técnica.

Algunos de los métodos proporcionados en el presente documento se basan en la expresión diferencial de un polipéptido o polinucleótido de MetRS para caracterizar la condición o el estado de una célula, un tejido o un sujeto y para distinguirlo de otra célula, tejido o sujeto. Los ejemplos no limitantes incluyen métodos para detectar la presencia o los niveles de un polipéptido o polinucleótido de MetRS en una muestra biológica para distinguir entre células o tejidos de diferentes especies, células de diferentes tejidos u órganos, estados de desarrollo celular tales como estados neonatal y adulto, de diferenciación celular, condiciones tales como sanos, enfermos y tratados fracciones intracelulares y extracelulares, además de cultivos celulares primarios y otros cultivos celulares, tales como cultivos celulares inmortalizados

La expresión diferencial incluye una diferencia estadísticamente significativa en uno o más niveles de expresión de genes de una secuencia de referencia polinucleotídica o polipeptídica de MetRS en comparación con los niveles de expresión de la misma secuencia en un control apropiado. La diferencia estadísticamente significativa puede relacionarse con un aumento o una reducción en los niveles de expresión, como se mide por niveles de ARN, niveles proteicos, función proteica o cualquier otra medida relevante de la expresión génica tal como las descritas en el presente documento. También se incluye una comparación entre un polinucleótido o polipéptido de MetRS de la invención y una secuencia de MetRS citosólica o mitocondrial de longitud completa o de tipo silvestre, normalmente del mismo tipo o un tipo correspondiente. Puede detectarse expresión diferencial por diversas técnicas en este campo y descritas en el presente documento, incluyendo técnicas basadas en polinucleótidos y polipéptidos, tales

como PCR en tiempo real, hibridación sustractiva, matrices de polinucleótidos y polipéptidos, y otras.
Un resultado se denomina normalmente estadísticamente significativo si es poco probable que haya aparecido al azar. El nivel de significación de un ensayo o resultado está relacionado tradicionalmente con un concepto de ensayo de hipótesis estadística frecuentista. En casos sencillos, la significación estadística puede definirse como la probabilidad de tomar una decisión para rechazar la hipótesis nula cuando la hipótesis nula es de hecho cierta (una decisión conocida como error de Tipo I o "determinación de falso positivo"). Esta decisión se realiza con frecuencia usando el p valor: si el p valor es menor que el nivel de significación, entonces se rechaza la hipótesis nula. Cuanto menor sea el p valor, más significativo será el resultado. También pueden utilizarse factores Bayes para determinar la significación estadística (véase, por ejemplo, Goodman S., Ann Intern Med 130:1005-13, 1999).

En casos más complicados, pero prácticamente importantes, el nivel de significación de un ensayo o resultado puede reflejar un análisis en el que la probabilidad de tomar una decisión para rechazar la hipótesis nula cuando la hipótesis nula es de hecho cierta no es mayor que la probabilidad indicada. Este tipo de análisis posibilita las aplicaciones en las que la probabilidad de decidir rechazar puede ser mucho más pequeña que el nivel de significación para algunos conjuntos de suposiciones abarcadas dentro de la hipótesis nula.

En ciertas realizaciones a modo de ejemplo, la expresión diferencial estadísticamente significativa puede incluir situaciones en las que el nivel de expresión de una secuencia de MetRS dada proporciona al menos una diferencia de aproximadamente 1,2X, 1,3X, 1,4X, 1,5X, 1,6X, 1,7X, 1,8X, 1,9X, 2,0X, 2,2X, 2,4X, 2,6X, 2,8X, 3,0X, 4,0X, 5,0X, 6,0X, 7,0X, 8,0X, 9,0X, 10,0X, 15,0X, 20,0X, 50,0 X, 100,0X, o mayor en la expresión (es decir, expresión diferencial que puede ser expresión mayor o menor) en una muestra biológica sospechosa en comparación con un control apropiado, incluyendo todos los números enteros y puntos decimales entre medias (por ejemplo, 1,24X, 1,25X, 2,1X, 2,5X, 60,0X, 75,0X, etc.). En ciertas realizaciones, la expresión diferencial estadísticamente significativa puede incluir situaciones en las que el nivel de expresión de una secuencia de MetRS dada proporciona al menos aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1.000 por ciento (%) o mayor diferencia en expresión (es decir, expresión diferencial que puede ser mayor o menor) en una muestra biológica sospechosa en comparación con un control apropiado, incluyendo todos los números enteros y puntos decimales entre medias.

Como un ejemplo adicional, la expresión diferencial también puede determinarse realizando ensayo Z, es decir, calculando una puntuación Z absoluta, como se describe en el presente documento y se conoce en la técnica (véase Ejemplo 1). El ensayo Z se utiliza normalmente para identificar diferencias significativas entre una media de muestra y una media de población. Por ejemplo, en comparación con una tabla normal convencional (por ejemplo, un tejido de control), a un intervalo de confianza del 95 % (es decir, al nivel de significación del 5 %), una puntuación Z con un valor absoluto mayor de 1,96 indica ausencia de aleatoriedad. Para un intervalo de confianza del 99 %, si el Z absoluto es mayor de 2,58, significa que $p<0,01$ y la diferencia es aún más significativa, la hipótesis nula puede rechazarse con mayor confianza. En estas realizaciones y otras relacionadas, una puntuación Z absoluta de 1,96, 2, 2,58, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más, incluyendo todos los puntos decimales entre medias (por ejemplo,10,1, 10,6, 11,2, etc.), puede proporcionar una fuerte medida de la significación estadística. En ciertas realizaciones, una puntuación Z absoluta mayor de 6 puede proporcionar significación estadística excepcionalmente alta.

Sustancial se refiere de forma similar en general a la falta de una diferencia estadísticamente significativa en los niveles de expresión entre la muestra biológica y el control de referencia. Los ejemplos de niveles de expresión sustancialmente similares pueden incluir situaciones en las que el nivel de expresión de un SSCIGS dado proporciona menos de aproximadamente una diferencia de 0,05X, 0,1X, 0,2X, 0,3X, 0,4X, 0,5X, 0,6X, 0,7X, 0,8X, 0,9X, 1,0X, 1,1X, 1,2X, 1,3X o 1,4X de expresión (es decir, expresión diferencial que puede ser de expresión mayor o menor) en una muestra biológica sospechosa en comparación con una muestra de referencia, incluyendo todos los puntos decimales entre medias (por ejemplo, 0,15X, 0,25X, 0,35X, etc.). En ciertas realizaciones, la expresión diferencial puede incluir situaciones en las que el nivel de expresión de una secuencia de MetRS dada proporciona una diferencia de menos de aproximadamente 0,25, 0,5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50 por ciento (%) de expresión (es decir, expresión diferencial que puede ser mayor o menor) en una muestra biológica sospechosa en comparación con una muestra de referencia, incluyendo todos los puntos decimales entre medias.

En ciertas realizaciones, tal como cuando se usa una micromatriz Affymetrix para medir los niveles de expresión de una secuencia de referencia polipeptídica o polinucleotídica de MetRS, la expresión diferencial también puede determinarse por el valor de expresión medio resumido por el software de Affymetrix Microarray Suite 5 (Affymetrix, Santa Clara, CA), u otro software similar, normalmente con un valor de expresión medio escalado de 1.000.

Las realizaciones de la presente invención incluyen métodos para detectar la presencia o los niveles de una secuencia de referencia polipeptídica o polinucleotídica de MetRS o una parte de la misma para distinguir entre células o tejidos u otra muestra biológica de un organismo o una especie diferente, en los que la presencia o los niveles de esa secuencia se asocia con un organismo o una especie seleccionado. Los ejemplos generales incluyen métodos para distinguir entre seres humanos y cualquier combinación de bacterias, hongos, plantas y otros animales no humanos. Se incluyen dentro de animales métodos para distinguir entre seres humanos y cualquier

combinación de vertebrados e invertebrados, incluyendo vertebrados tales como peces, anfibios, reptiles, aves y mamíferos no humanos, e invertebrados tales como insectos, moluscos, crustáceos y corales. Se incluyen dentro de los mamíferos no humanos métodos para distinguir entre seres humanos y cualquier combinación de mamíferos no humanos de la Orden Afrosoricida, Macroscelidea, Tubulidentata, Hyracoidea, Proboscidea, Sirenia, Cingulata, Pilosa, Scandentia, Dermoptera, Primates, Rodentia, Lagomorpha, Erinaceomorpha, Soricomorpha, Chiroptera, Pholidota, Cetacea, Carnivora, Perissodactyla o Artiodactyla. Se incluyen dentro del Orden Primates monos, simios, gorilas y chimpancés, entre otros conocidos en la técnica. En consecuencia, la presencia o los niveles de una secuencia de referencia polinucleotídica o polipeptídica de MetRS o variante, como se describe en el presente documento, pueden usarse para identificar la fuente de una muestra biológica dada, tal como una célula, un tejido o un órgano, distinguiendo entre cualquier combinación de estos organismos, o distinguiendo entre seres humanos y uno cualquiera o más de estos organismos, tal como un panel de organismos. En ciertas realizaciones, la fuente de una muestra biológica dada también puede determinarse comparando la presencia o los niveles de una secuencia de MetRS o una parte de la misma con un valor predeterminado.

Las realizaciones de la presente invención incluyen métodos para detectar la presencia o los niveles de una secuencia de referencia polinucleotídica o polipeptídica de MetRS o una parte de la misma para distinguir entre células u otras muestras biológicas que se originan de diferentes tejidos u órganos. Los ejemplos no limitantes incluyen métodos para distinguir entre una célula u otra muestra biológica que se origina de cualquier combinación de piel (por ejemplo, dermis, epidermis, capa subcutánea), folículos pilosos, sistema nervioso (por ejemplo, cerebro, médula espinal, nervios periféricos), sistema auditivo u órganos del equilibrio (por ejemplo, oído interno, oído medio, oído externo), sistema respiratorio (por ejemplo, nariz, tráquea, pulmones), tejidos gastroesofágicos, el sistema gastrointestinal (por ejemplo, boca, esófago, estómago, intestino delgado, intestino grueso, recto), sistema vascular (por ejemplo, corazón, vasos sanguíneos y arterias), hígado, vesícula biliar, sistema inmunitario/linfático (por ejemplo, ganglios linfáticos, folículos linfoides, bazo, timo, médula ósea), sistema urogenital (por ejemplo, riñones, uréter, vejiga, uretra, cuello uterino, trompas de Falopio, ovarios, útero, vulva, próstata, glándulas bulbouretrales, epidídimo, próstata, vesículas seminales, testículos), sistema musculoesquelético (por ejemplo, músculos esqueléticos, músculos lisos, hueso, cartílago, tendones, ligamentos), tejido adiposo, tejido mamario y el sistema endocrino (por ejemplo, hipotálamo, hipófisis, tiroides, páncreas, glándulas suprarrenales). Por lo tanto, basándose en la asociación de una secuencia polinucleotídica o polipeptídica de MetRS como se describe en el presente documento, estos métodos pueden usarse para identificar o caracterizar el tejido u órgano del que deriva una célula u otra muestra biológica.

Las realizaciones de la presente invención incluyen métodos para detectar la presencia o los niveles de una secuencia de referencia polinucleotídica o polipeptídica de MetRS o una parte de la misma para distinguir entre o caracterizar el estado de desarrollo o diferenciación de la célula. También se incluyen métodos para diferenciar entre células germinales, células madre y células somáticas. Los ejemplos de estados del desarrollo incluyen neonatal y adulto. Los ejemplos de estados de diferenciación celular incluyen todos los estadios discretos e identificables entre una célula totipotente, una célula pluripotente, una célula madre progenitora multipotente y una célula madura, completamente diferenciada.

Una célula totipotente tiene potencial total, normalmente surge durante la reproducción sexual y asexual, e incluye esporas y cigotos, aunque en ciertos casos las células pueden desdiferenciarse y recuperar su totipotencia. Una célula pluripotente incluye una célula madre que tiene el potencial de diferenciarse en cualquiera de las tres capas germinales, incluyendo el endodermo (revestimiento del estómago interior, tracto gastrointestinal, los pulmones), el mesodermo (músculo, hueso, sangre, sistema urogenital) y el ectodermo (tejidos epidérmicos y sistema nervioso). Las células progenitoras multipotentes son normalmente capaces de diferenciarse en un número limitado de tipos tisulares. Los ejemplos de células multipotentes incluyen, sin limitación, células madre hematopoyéticas (células madre adultas) de la médula ósea que dan lugar a células inmunitarias tales como glóbulos rojos, glóbulos blancos y plaquetas, células madre mesenquimales (células madre adultas) de la médula ósea que dan lugar a células del estómago, células grasas y diversos tipos de células óseas, células madre epiteliales (células progenitoras) que dan lugar a los diversos tipos de células de la piel y células satélite musculares (células progenitoras) que contribuyen a tejido muscular diferenciado. En consecuencia, la presencia o los niveles de una secuencia polinucleotídica o polipeptídica de MetRS particular (por ejemplo, punto de unión de corte y empalme de una variante de corte y empalme de MetRS, fragmento proteolítico de MetRS), puede usarse para distinguir entre o caracterizar los estados de diferenciación celulares indicados anteriormente, en comparación con un nivel de control o uno predeterminado.

Las realizaciones de la presente invención incluyen métodos para detectar la presencia o los niveles de una secuencia de referencia polinucleotídica o polipeptídica de MetRS para caracterizar o diagnosticar la afección o una célula, un tejido, un órgano o un sujeto, en el que esa afección puede caracterizarse como sano, enfermo, en riesgo de estar enfermo, o tratado. Para dichos fines de diagnóstico, el término "diagnóstico" o "diagnosticado" incluye identificar la presencia o la naturaleza de una afección patológica, caracterizando el riesgo de desarrollo de dicha afección y/o midiendo el cambio (o ausencia de cambio) de una afección patológica en respuesta a la terapia. Los métodos de diagnóstico pueden diferir en su sensibilidad y especificidad. En ciertas realizaciones, la "sensibilidad" de un ensayo de diagnóstico se refiere al porcentaje de células, tejidos o sujetos enfermos que tienen un resultado de ensayo positivo (porcentaje de "verdaderos positivos"). Las células, los tejidos o los sujetos enfermos no detectados por el ensayo se denominan normalmente "falsos negativos". Las células, los tejidos o los sujetos que no

están enfermos y que tienen resultados de ensayo negativos en el ensayo pueden denominarse "verdaderos negativos". En ciertas realizaciones, la "especificidad" de un ensayo de diagnóstico puede definirse como uno (1) menos la tasa de falso positivo, en la que la tasa de "falso positivo" se define como la proporción de esas muestras o sujetos sin la enfermedad y que tienen resultados de ensayo positivos. Aunque un método de diagnóstico particular puede no proporcionar un diagnóstico definitivo de una afección, es suficiente si el método proporciona una indicación positiva que ayuda en el diagnóstico.

En ciertos casos, la presencia o el riesgo de desarrollar una afección patológica puede diagnosticarse comparando la presencia o los niveles de una o más secuencias de referencia polinucleotídicas o polipeptídicas de MetRS seleccionadas o partes de las mismas que se correlacionan con la afección, bien por niveles aumentados o disminuidos, en comparación con un control adecuado. Un "control adecuado" o "control apropiado" incluye un valor, un nivel, un elemento, una característica, o una propiedad determinados en una célula u otra muestra biológica de un tejido u organismo, por ejemplo, una célula, un tejido o un organismo de control o normal, que muestra, por ejemplo, rasgos normales, tales como la ausencia de la afección. En ciertas realizaciones, un "control adecuado" o "control apropiado" es un valor, nivel, elemento, característica o propiedad predefinido. Otros controles adecuados resultarán evidentes para expertos en la materia. Se describen en otra parte del presente documento ejemplos de enfermedades y afecciones.

Las realizaciones de la presente invención incluyen técnicas de detección basadas en polinucleótidos o ácidos nucleicos de MetRS, que ofrecen ciertas ventajas debido a la sensibilidad de detección. Por lo tanto, ciertas realizaciones se refieren al uso o la detección de polinucleótidos de MetRS como parte de un método o ensayo de diagnóstico. La presencia y/o los niveles de polinucleótidos de MetRS pueden medirse por cualquier método conocido en la técnica, incluyendo ensayos de hibridación tales como transferencia de Northern, reacción en cadena de la polimerasa cuantitativa o cualitativa (PCR), PCR de transcriptasa inversa (RT-PCR) cuantitativa o cualitativa, micromatriz, transferencias puntuales o por ranuras, o hibridación *in situ* tal como hibridación *in situ* fluorescente (FISH), entre otros. Algunos de estos métodos se describen en mayor detalle posteriormente.

Polinucleótidos de MetRS tales como ADN y ARN pueden recogerse y/o generarse de sangre, fluidos biológicos, tejidos, órganos, líneas celulares u otra muestra relevante usando técnicas conocidas en este campo, tales como las descritas en Kingston. (2002 Current Protocols in Molecular Biology, Greene Publ. Assoc. Inc. & John Wiley & Sons, Inc., NY, NY) (véase, por ejemplo como se describe en Nelson *et al*. Proc Natl Acad Sci U S A, 99: 11890-11895, 2002) y en otras partes. Además, diversos kits disponibles en el mercado para construir ARN son útiles para preparar el ARN para usar en la presente invención. El ARN puede construirse a partir de órganos/tejidos/células obtenidos de sujetos sanos normales; sin embargo, esta invención también contempla la construcción de ARN de sujetos enfermos. Ciertas realizaciones contemplan el uso de cualquier tipo de órgano de cualquier tipo de sujeto o animal. Para muestras de ensayo puede obtenerse ARN de un individuo (por ejemplo, cualquier animal, incluyendo mamíferos) con o sin enfermedad visible y de muestras tisulares, fluidos biológicos (por ejemplo, sangre completa) o similares.

En ciertas realizaciones, la amplificación o construcción de secuencias de ADNc puede ser útil para aumentar las capacidades de detección. La presente divulgación, así como la técnica, proporciona el nivel requerido de detalle para realizar dichas tareas. En una realización ejemplar, se usa sangre completa como la fuente de ARN y, en consecuencia, se usan opcionalmente reactivos de estabilización de ARN, tales como tubos PAX, como se describe, por ejemplo, en Thach *et al*., J. Immunol. Métodos. Dec 283 (1-2):269-279, 2003 y Chai *et al*., J. Clin. Lab Anal. 19(5):182-188, 2005. Puede generarse bibliotecas de ADN complementario (ADNc) usando técnicas conocidas en este campo, tales como las descritas en Ausubel *et al*. (2001 Current Protocols in Molecular Biology, Greene Publ. Assoc. Inc. & John Wiley & Sons, Inc., NY, NY); Sambrook *et al*. (1989 Molecular Cloning, Segunda Ed., Cold Spring Harbor Laboratory, Plainview, NY); Maniatis *et al*. (1982 Molecular Cloning, Cold Spring Harbor Laboratory, Plainview, NY) y en otras partes. Además, diversos kits disponibles en el mercado para construir bibliotecas de ADNc son útiles para preparar las bibliotecas de ADNc de la presente invención. Pueden construirse bibliotecas a partir de órganos/tejidos/células obtenidos de sujetos normales, sanos.

Ciertas realizaciones pueden emplear métodos de hibridación para detectar secuencias polinucleotídicas de MetRS. Se han desarrollado bien en la técnica métodos para realizar ensayos de hibridación de polinucleótidos. Los procedimientos y las condiciones de ensayos de hibridación variarán dependiendo de la aplicación y se seleccionan de acuerdo con los métodos de unión generales conocidos incluyendo los implicados en: Maniatis *et al*. Molecular Cloning: A Laboratory Manual (2ª edición, Cold Spring Harbor, N.Y., 1989); Berger y Kimmel Methods in Enzymology, Vol. 152, Guide to Molecular Cloning Techniques (Academic Press, Inc., San Diego, CA, 1987); Young y Davis, PNAS. 80:1194 (1983). Se han descrito métodos y aparatos para llevar a cabo reacciones de hibridación repetidas y controladas en las Patentes de Estados Unidos N.º 5.871.928, 5.874.219, 6.045.996 y 6.386.749, 6.391.623.

Ciertas realizaciones pueden emplear métodos de amplificación de ácido nucleico para detectar secuencias polinucleotídicas de MetRS. La expresión “amplificación” o “amplificación de ácidos nucleicos” se refiere a la producción de múltiples copias de un ácido nucleico diana que contiene al menos una parte de la secuencia de ácido nucleico diana específica pretendida. Las múltiples copias pueden denominarse amplicones o productos de

amplificación. En ciertas realizaciones, la diana amplificada contiene menos que la secuencia génica diana completa (intrones y exones) o una secuencia génica diana expresada (transcrito cortado y empalmado de exones y secuencias no traducidas flanqueantes). Por ejemplo, pueden producirse amplicones específicos amplificando una parte del polinucleótido diana usando cebadores de amplificación que hibridan con, e inician la polimerización desde, posiciones internas del polinucleótido diana. Preferentemente, la parte amplificada contiene una secuencia diana detectable que puede detectarse usando cualquiera de diversos métodos bien conocidos.

“Amplificación selectiva" o "amplificación específica", como se usa en el presente documento, se refiere a la amplificación de una secuencia de ácido nucleico diana de acuerdo con la presente invención en la que la amplificación detectable de la secuencia diana está sustancialmente limitada a amplificación de secuencia diana a la que contribuye una muestra de ácido nucleico de interés que se ensaya y en la que no contribuye una secuencia de ácido nucleico diana a la que contribuye alguna otra fuente de muestra, por ejemplo, contaminación presente en reactivos usados durante reacciones de amplificación o en el ambiente en el que se realizan reacciones de amplificación.

La expresión "condiciones de amplificación" se refiere a condiciones que permiten la amplificación de ácido nucleico de acuerdo con la presente invención. Las condiciones de amplificación pueden, en algunas realizaciones, ser menos rigurosas que "condiciones de hibridación rigurosas" como se describen en el presente documento. Los oligonucleótidos usados en las reacciones de amplificación de la presente invención hibridan con sus dianas pretendidas en condiciones de amplificación, pero pueden hibridar o no en condiciones de hibridación rigurosas. Por otro lado, las sondas de detección de la presente invención normalmente hibridan en condiciones de hibridación rigurosas. Las condiciones aceptables para llevar a cabo amplificaciones de ácidos nucleicos de acuerdo con la presente invención pueden determinarse fácilmente por alguien que tenga habilidad habitual en la técnica dependiendo del método particular de amplificación empleado.

Muchos métodos bien conocidos de amplificación de ácidos nucleicos requieren termociclación para de forma alterna desnaturalizar ácidos nucleicos bicatenarios e hibridar cebadores; sin embargo, otros métodos bien conocidos de amplificación de ácido nucleico son isotérmicos. La reacción en cadena de la polimerasa (Patente de Estados Unidos N.º 4.683.195, 4.683.202, 4.800.159, 4.965.188), habitualmente denominada PCR, usa múltiples ciclos de desnaturalización, hibridación de pares de cebadores con cadenas opuestas y extensión de cebadores para aumentar exponencialmente los números de copias de la secuencia diana. En una variación denominada RT-PCR, se usa la transcriptasa inversa (RT) para preparar un ADN complementario (ADNc) de ARNm, y el ADNc se amplifica después por PCR para producir múltiples copias de ADN.

Como se ha indicado anteriormente, el término "PCR" se refiere a múltiples ciclos de amplificación que amplifican selectivamente una especie de ácido nucleico diana. Se incluyen PCR cuantitativa (qPCR), PCR en tiempo real, PCR de transcripción inversa (RT-PCR) y PCR de transcripción inversa cuantitativa (qRT-PCR) que están bien descritas en la técnica. El término "pPCR" se refiere a reacción en cadena de la polimerasa cuantitativa, y el término "qRT-PCR" se refiere a reacción en cadena de la polimerasa de transcripción inversa cuantitativa. Pueden usarse qPCR y qRT-PCR para amplificar y cuantificar simultáneamente una molécula de ADNc diana. Permite tanto la detección como la cuantificación de una secuencia específica en un grupo de ADNc, tal como un gen o transcrito de MetRS seleccionado.

La expresión "PCR en tiempo real" puede usar colorante de unión a ADN para unir todos los ADN bicatenarios (bc) en PCR, provocando fluorescencia del colorante. Un aumento en el producto de ADN durante la PCR conduce por tanto a un aumento de la intensidad de fluorescencia y se mide en cada ciclo, permitiendo de este modo cuantificar las concentraciones de ADN. Sin embargo, se unirán colorantes ADNbc tales como SYBR Green a todos los productos de PCR de ADNbc. La fluorescencia se detecta y se mide en el termociclador de PCR en tiempo real, y su aumento geométrico correspondiente a aumento exponencial del producto se usa para determinar el ciclo umbral ("Ct") en cada reacción.

La expresión "Puntuación de Ct" se refiere al número de ciclo umbral, que es el ciclo en el que la amplificación por PCR ha superado un nivel umbral. Si hay una cantidad mayor de ARNm para un gen particular en una muestra, cruzará el umbral antes que un gen poco expresado ya que hay más ARN de partida para amplificar. Por lo tanto, una puntuación de Ct baja indica alta expresión génica en una muestra y una puntuación de Ct alta es indicativa de la expresión génica baja.

Ciertas realizaciones pueden emplear la reacción en cadena de la ligasa (Weiss, Science 254:1292, 1991), habitualmente denominada LCR, que usa dos conjuntos de oligonucleótidos de ADN complementarios que hibridan con regiones adyacentes del ácido nucleico diana. Los oligonucleótidos de ADN se unen covalentemente por una ADN ligasa en ciclos repetidos de desnaturalización térmica, hibridación y ligamiento para producir un producto oligonucleotídico ligado bicatenario detectable.

Otro método es amplificación por desplazamiento de cadena (Walker, G. *et al*., 1992, Proc. Nat. Acad. Sci. USA 89: 392-396; Patentes de Estados Unidos 5.270.184 y 5.455.166), denominado habitualmente SDA , que usa ciclos de pares de hibridación de secuencias de cebadores con cadenas opuestas de una secuencia diana, extensión de

cebador en presencia de un dNTPaS para producir un producto de extensión de cebadores hemifosforotioado de doble cadena, corte mediado por endonucleasa de un sitio de reconocimiento de endonucleasa de restricción hemimodificado y extensión de cebadores mediada por polimerasa desde el extremo 3' del corte para desplazar una cadena existente y producir una cadena para el siguiente ciclo de hibridación de cebadores, corte y desplazamiento de cadena, dando como resultado amplificación geométrica del producto. La SDA termófila (tSDA) utiliza endonucleasas termófilas y polimerasas a mayores temperaturas esencialmente en el mismo método (Patente Europea N.º 0 684 315).

Otros métodos de amplificación incluyen, por ejemplo: amplificación basada en secuencias de ácido nucleico (Patente de Estados Unidos N.º 5.130.238), denominada habitualmente NASBA; una que usa una ARN replicasa para amplificar la molécula sonda en sí misma (Lizardi, P. *et al*., 1988, BioTechnol., 6:1197-1202), habitualmente denominada Qß replicasa; un método de amplificación basado en transcripción (Kwoh, D. *et al*., 1989, Proc. Nat. Acad. Sci. USA 86:1173-1177); replicación de secuencia autosostenida (Guatelli, J. *et al*., 1990, Proc. Nat. Acad. Sci. USA 87:1874-1878); y amplificación mediada por transcripción (Patentes de Estados Unidos N.º 5.480.784 y 5.399.491), denominada habitualmente TMA. Para análisis adicional de métodos de amplificación conocidos, véase Pershing, David H., 1993, "In Vitro Nucleic Acid Amplification Techniques" en Diagnostic Medical Microbiology: Principles and Applications (Persing *et al*., Eds.), págs. 51-87 (American Society for Microbiology, Washington, DC).

Los sistemas de amplificación basados en transcripción ilustrativos de la presente invención incluyen TMA, que emplea una ARN polimerasa para producir múltiples transcritos de ARN de una región diana (Patentes de Estados Unidos N.º 5.480.784 y 5.399.491). TMA usa un "promotor-cebador" que hibrida con un ácido nucleico diana en presencia de una transcriptasa inversa y un ARN polimerasa para formar un promotor bicatenario a partir del que la ARN polimerasa produce transcritos de ARN. Esos transcritos pueden convertirse en moldes para ciclos adicionales de TMA en presencia de un segundo cebador capaz de hibridar con los transcritos de ARN. A diferencia de PCR, LCR u otros métodos que requieren desnaturalización térmica, TMA es un método isotérmico que usa una actividad de RNasa H para digerir la cadena de ARN de un híbrido de ARN:ADN, haciendo de este modo la cadena de ADN disponible para la hibridación con un cebador o promotor-cebador. En general, se usa la actividad RNasa H asociada con la transcriptasa inversa proporcionada para amplificación.

En un método de TMA ilustrativo, un cebador de amplificación es un promotor-cebador oligonucleotídico que comprende una secuencia promotora que se hace funcional cuando es bicatenaria, localizada 5' de una secuencia de unión a diana, que es capaz de hibridar con un sitio de unión de un ARN diana en una localización 3' de la secuencia para amplificar. Un promotor-cebador puede denominarse "cebador T7" cuando es específico para reconocimiento por ARN polimerasa T7. En ciertas circunstancias, el extremo 3' de un promotor-cebador, o una subpoblación de dichos promotores-cebadores, puede modificarse para bloquear o reducir la extensión del cebador. A partir de un promotor-cebador no modificado, la transcriptasa inversa crea una copia de ADNc del ARN diana, mientras que la actividad RNasa H degrada el ARN diana. Un segundo cebador de amplificación se une después con el ADNc. Este cebador puede denominarse "cebador no T7" para distinguirlo de un "cebador T7". A partir de este segundo cebador de amplificación, la transcriptasa inversa crea otra cadena de ADN, dando como resultado un ADN bicatenario con un promotor funcional en un extremo. Cuando es bicatenaria, la secuencia de promotor es capaz de unirse a una ARN polimerasa para comenzar la transcripción de la secuencia diana con la que se hibrida el promotor-cebador. Una ARN polimerasa usa esta secuencia promotora para producir múltiples transcritos de ARN (es decir, amplicones), generalmente de aproximadamente 100 a 1.000 copias. Cada amplicón de nueva síntesis puede hibridar con el segundo cebador de amplificación. La transcriptasa inversa puede después crear una copia de ADN, mientras que la actividad RNasa H degrada el ARN de esta doble cadena ARN:ADN. El promotor-cebador puede después unirse al ADN de nueva síntesis, permitiendo que la transcriptasa inversa cree un ADN bicatenario, a partir del que el ARN polimerasa produce múltiples amplicones. Por lo tanto, puede conseguirse una amplificación isotérmica de miles de millones de veces usando dos cebadores de amplificación.

En ciertas realizaciones, pueden usarse otras técnicas para evaluar transcritos de ARN de los transcritos de una biblioteca de ADNc particular, incluyendo análisis de micromatrices (Han, M., *et al*., Nat Biotechnol, 19:631-635, 2001; Bao, P., *et al*., Anal Chem, 74:1792-1797, 2002; Schena *et al*., Proc. Natl. Acad. Sci. USA 93:10614-19, 1996; y Heller *et al*., Proc. Natl. Acad. Sci. USA 94:2150-55, 1997) y SAGE (análisis en serie de expresión génica). Como MPSS, SAGE es digital y puede generar un gran número de secuencias identificativas. (Véase, por ejemplo, Velculescu, V. E., *et al*., Trends Genet, 16:423-425., 2000; Tuteja R. y Tuteja N. Bioessays. Ago 2004; 26(8):916-22), aunque órdenes de magnitud menores de lo que está disponible de técnicas tales como MPSS.

En ciertas realizaciones, el término "micromatriz" incluye una "micromatriz de ácidos nucleicos" que tiene una pluralidad de ácidos nucleicos unidos a sustrato, siendo detectable por separado la hibridación con cada uno de la pluralidad de ácidos nucleicos unidos. El sustrato puede ser sólido o poroso, planar o no planar, unitario o distribuido. Las micromatrices de ácido nucleico incluyen todos los dispositivos denominados así en Schena (ed.), DNA Microarrays: A Practical Approach (Practical Approach Series), Oxford University Press (1999); Nature Genet. 21(1) (supl.): 1-60 (1999); Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing CompanyBioTechniques Books Division (2000). Las micromatrices de ácido nucleico pueden incluir una pluralidad de ácidos nucleicos unidos a sustrato en los que la pluralidad de ácidos nucleicos están dispuestos en una pluralidad de perlas, en lugar de un sustrato planar unitario, como se describe, por ejemplo, en Brenner *et al*., Proc. Natl. Acad.

Sci. USA 97 (4): 1665-1670 (2000). Pueden encontrarse ejemplos de micromatrices de ácido nucleico en las Patentes de EE.UU. N.º 6.391.623, 6.383.754, 6.383.749, 6.380.377, 6.379.897, 6.376.191, 6.372.431, 6.351.712, 6.344.316, 6.316.193, 6.312.906, 6.309.828, 6.309.824, 6.306.643, 6.300.063, 6.287.850, 6.284.497, 6.284.465. 6.280.954, 6.262.216, 6.251.601, 6.245.518, 6.263.287, 6.251.601, 6.238.866, 6.228.575, 6.214.587, 6.203.989, 6.171.797, 6.103.474, 6.083.726, 6.054.274, 6.040.138, 6.083.726, 6.004.755, 6.001.309, 5.958.342, 5.952.180, 5.936.731, 5.843.655, 5.814.454, 5.837.196, 5.436.327, 5.412.087 y 5.405.783.

Loa ejemplos adicionales incluyen matrices de ácido nucleico que están disponibles en el mercado de Affymetrix (Santa Clara, CA) con el nombre comercial GENECHIP™. Se proporcionan métodos ilustrativos adicionales de fabricación y uso de matrices, por ejemplo, en las Patentes de Estados Unidos N.º 7.028.629; 7.011.949; 7.011.945; 6.936.419; 6.927.032; 6.924.103; 6.921.642; y 6.818.394.

La presente invención, en relación con matrices y micromatrices, también contempla muchos usos para polímeros unidos a sustratos sólidos. Estos usos incluyen supervisión, perfil, cribado de bibliotecas, genotipación y diagnóstico de expresión génica. Los métodos de supervisión y perfilación de expresión génica y métodos útiles para supervisión y perfilación de expresión génica se muestran en las Patentes de Estados Unidos N.º 5.800.992, 6.013.449, 6.020.135, 6.033.860, 6.040.138, 6.177.248 y 6.309.822. La genotipación y usos de la misma se muestran en N.º de Serie de Estados Unidos 10/442.021, 10/013.598 (Solicitud de Estados Unidos N.º 2003/0036069), y Patentes de Estados Unidos. N.º 5.925.525, 6.268.141, 5.856.092, 6.267.152, 6.300.063, 6.525.185, 6.632.611, 5.858.659, 6.284.460, 6.361.947, 6.368.799, 6.673.579 y 6.333.179. Otros métodos de amplificación, marcaje y análisis de ácidos nucleico que pueden usarse en combinación con los métodos desvelados en el presente documento se incluyen en las Patentes de EE.UU. N.º 5.871.928, 5.902.723, 6.045.996, 5.541.061 y 6.197.506.

Como resultará evidente para los expertos en la materia, ciertas realizaciones pueden emplear oligonucleótidos, tales como cebadores o sondas, para amplificación o detección, como se describe en el presente documento. Pueden producirse oligonucleótidos de una secuencia y estructura química definida por técnicas conocidas por los expertos habituales en la materia, tal como por síntesis química o bioquímica, y por expresión *in vitro* o *in vivo* de moléculas de ácido nucleico recombinantes, por ejemplo, vectores bacterianos o víricos. En ciertas realizaciones, un oligonucleótido no consiste solamente en ADN cromosómico de tipo silvestre o los productos de transcripción *in vivo* del mismo.

Los oligonucleótidos o cebadores pueden modificarse de cualquier manera, siempre que una modificación dada sea compatible con la función deseada de un oligonucleótido dado. Un experto en la materia puede determinar fácilmente si una modificación dada es adecuada o deseada para cualquier oligonucleótido dado de la presente invención. Se describen en más detalle en otra parte del presente documento oligonucleótidos de MetRS relevantes.

Aunque el diseño y la secuencia de los oligonucleótidos dependen de su función como se describe en el presente documento, tienen en cuenta en general varias variables. Entre las más relevantes están: longitud, temperatura de fusión (Tm), especificidad, complementariedad con otros oligonucleótidos en el sistema, contenido de G/C, tramos de polipirimidina (T, C) o polipurina (A, G) y la secuencia del extremo 3'. El control de estas y otras variables es un aspecto convencional y bien conocido del diseño de oligonucleótidos, y diversos programas informáticos están fácilmente disponibles para explorar grandes números de oligonucleótidos potenciales para encontrar los óptimos.

Ciertas realizaciones incluyen por lo tanto métodos para detectar un polinucleótido de MetRS diana en una muestra, comprendiendo el polinucleótido la secuencia de un polinucleótido de MetRS de referencia, como se describe en el presente documento, que comprende: a) hibridar la muestra con una sonda que comprende una secuencia complementaria del polinucleótido diana en la muestra, y cuya sonda hibrida específicamente con dicho polinucleótido diana, en condiciones en las que se forma un complejo de hibridación entre dicha sonda y dicho polinucleótido diana o fragmentos de los mismos, y b) detectar la presencia o ausencia de dicho complejo de hibridación, y opcionalmente, si está presente, la cantidad del mismo. También se incluyen métodos para detectar un polinucleótido de MetRS diana en una muestra, comprendiendo el polinucleótido la secuencia de un polinucleótido de MetRS de referencia, como se describe en el presente documento, que comprende: a) amplificar el polinucleótido diana o fragmento del mismo, y b) detectar la presencia o ausencia de dicho polinucleótido diana amplificado o fragmento del mismo y, opcionalmente, si está presente, la cantidad del mismo. Realizaciones específicas se refieren a la detección de variantes de corte y corte y empalme de MetRS, tal como detectando un punto de unión de corte y empalme único de la variante de corte y empalme, bien mediante hibridación, amplificación u otro método de detección.

Las realizaciones de la presente invención incluyen diversas técnicas de detección basadas en polipéptidos de MetRS, incluyendo técnicas de detección basadas en anticuerpos. Se incluye en estas realizaciones el uso de polipéptidos de MetRS para generar anticuerpos u otros agentes de unión, que después pueden usarse en métodos de diagnóstico y composiciones para detectar o cuantificar polipéptidos de MetRS seleccionados en una célula u otra muestra biológica, normalmente de un sujeto.

Ciertas realizaciones pueden emplear metodologías convencionales y detectores tales como transferencia de Western e inmunoprecipitación, ensayos inmunoabsorbentes ligados a enzimas (ELISA), citometría de flujo y

ensayos de inmunofluorescencia (IFA), que utilizan un dispositivo de captura de imágenes. Estos métodos bien conocidos utilizan normalmente uno o más anticuerpos monoclonales o policlonales, como se describe en el presente documento que se unen específicamente con un polipéptido de MetRS seleccionado de la invención, o una región única de ese polipéptido de MetRS, y en general no se unen significativamente con otros polipéptidos de MetRS, tales como un polipéptido de MetRS de longitud completa. En ciertas realizaciones, la región única del polipéptido de MetRS puede representar una estructura tridimensional única poseída por un fragmento proteico de nueva identificación de un MetRS.

Ciertas realizaciones pueden emplear "matrices", tales como "micromatrices". En ciertas realizaciones, una "micromatriz" también puede denominarse "micromatriz de péptido" o "micromatriz de proteína" que tiene una colección unida a sustrato o una pluralidad de polipéptidos, siendo detectable por separado la unión a cada uno de la pluralidad de polipéptidos unidos. Como alternativa la micromatriz de péptidos puede tener una pluralidad de agentes de unión, incluyendo pero sin limitación anticuerpos monoclonales, anticuerpos policlonales, agentes de unión de presentación de fagos, agentes de unión híbridos de levadura 2, y aptámeros, que pueden detectar específicamente la unión de los polipéptidos de MetRS descritos en el presente documento. La matriz puede basarse en la detección de autoanticuerpos de estos polipéptidos de MetRS, como se describe, por ejemplo, en Robinson *et al*., Nature Medicine 8 (3):295-301 (2002). Pueden encontrarse ejemplos de matrices de péptidos en los documentos WO 02/31463, WO 02/25288, WO 01/94946, WO 01/88162, WO 01/68671, WO 01/57259, WO 00/61806, WO 00/54046, WO 00/47774, WO 99/40434, WO 99/39210 y WO 97/42507 y Patentes de Estados Unidos N.º 6.268.210, 5.766.960 y 5.143.854.

Ciertas realizaciones pueden emplear EM u otros métodos basados en peso molecular para detectar de forma diagnóstica secuencias polipeptídicas de MetRS. La espectrometría de masas (EM) se refiere en general a una técnica analítica para determinar la composición elemental de una muestra o molécula. También puede usarse EM para determinar las estructuras químicas de moléculas, tales como péptidos y otros compuestos químicos.

Generalmente, el principio de EM consiste en ionizar compuestos químicos para generar moléculas con carga o fragmentos moleculares, y después medir sus relaciones de masa frente a carga. En un procedimiento de EM ilustrativo: una muestra se carga en el instrumento de EM y se somete a vaporización, los componentes de la muestra se ionizan por uno de diversos métodos (por ejemplo, impactando en ellos con un haz de electrones), lo que da como resultado la formación de partículas con carga positiva, los iones positivos se aceleran después por un campo magnético, se realizan cálculos sobre la relación de masa con respecto a carga (m/z) de las partículas basándose en los detalles del movimiento de los iones a medida que transitan a través de los campos electromagnéticos y detección de los iones, que en una etapa anterior se clasificaron según m/z.

Un instrumento ilustrativo de EM tiene tres módulos: una fuente de iones, que convierte moléculas de la muestra de fase gaseosa en iones (o, en el caso de ionización de electropulverización, mueven iones que existen en solución a la fase de gas); un analizador de masas, que clasifica los iones por sus masas aplicando campos electromagnéticos; y un detector, que mide el valor de una cantidad indicadora y de este modo proporciona datos para calcular las abundancias de cada ion presente.

La técnica EM tiene usos tanto cualitativos como cuantitativos, incluyendo identificar compuestos desconocidos, determinar la composición isotópica de elementos en una molécula y determinar de la estructura de un compuesto observando su fragmentación. Otros usos incluyen cuantificar la cantidad de un compuesto en una muestra o estudiar los fundamentos de química de iones de fase gaseosa (la química de iones y neutros en un vacío). Se incluyen cromatografía de gases-espectrometría de masas (CG/EM o CG-EM), espectrometría de masas por cromatografía líquida (CL/EM o CL-EM) y espectrometría de movilidad de iones/espectrometría de masas (EMI/EM o EMMI). Como consecuencia, las técnicas EM pueden usarse de acuerdo con cualquiera de los métodos proporcionados en el presente documento para medir la presencia o los niveles de un polipéptido de MetRS de la invención en una muestra biológica, y comparar esos niveles con una muestra de control o un valor predeterminado.

Ciertas realizaciones pueden emplear clasificación celular o visualización celular o dispositivos/técnicas de captura de imágenes para detectar o cuantificar la presencia o los niveles de polinucleótidos o polipéptidos de MetRS. Los ejemplos incluyen citometría de flujo o FACS, análisis de inmunofluorescencia (IFA), y técnicas de hibridación *in situ*, tales como hibridación *in situ* fluorescente (FISH).

Ciertas realizaciones pueden emplear métodos de biología convencionales, software y sistemas para fines de diagnóstico. Los productos de software informático de la invención normalmente incluyen soporte informático de lectura que tiene instrucciones ejecutables por ordenador para realizar las etapas lógicas del método de la invención. El soporte informático de lectura adecuado incluye disquete, CD-ROM/DVD/DVD-ROM, disco duro, memoria flash, ROM/RAM cintas magnéticas, etc. Las instrucciones ejecutables por ordenador pueden escribirse en un lenguaje informático adecuado o en combinación de varios lenguajes. Se describen métodos de biología computacional básicos, por ejemplo, en Setubal y Meidanis *et al.*, Introduction to Computational Biology Methods (PWS Publishing Company, Boston, 1997); Salzberg, Searles, Kasif, (Ed.), Computational Methods in Molecular Biology, (Elsevier, Ámsterdam, 1998); Rashidi y Buehler, Bioinformatics Basics: Application in Biological Science and Medicine (CRC Press, Londres, 2000) y Ouelette and Bzevanis Bioinformatics: A Practical Guide for Analysis of Gene and Proteins

(Wiley & Sons, Inc., 2ª ed., 2001). Véase Patente de Estados Unidos N.º 6.420.108.

Ciertas realizaciones pueden emplear diversos productos de programas informáticos y software para diversos fines, tales como diseño de sondas, control de datos, análisis y operación de instrumentos. Véase, Patentes de Estados Unidos N.º 5.593.839, 5.795.716, 5.733.729, 5.974.164, 6.066.454, 6.090.555, 6.185.561, 6.188.783, 6.223.127, 6.229.911 y 6.308.170.

El ensayo de toma de muestras de genoma completo (WGSA) se describe, por ejemplo, en Kennedy *et al*., Nat. Biotech. 21, 1233-1237 (2003), Matsuzaki *et al*., Gen Res. 14: 414-425, (2004), y Matsuzaki *et al*., Nature Methods 1: 109-111 (2004). Se describen algoritmos para uso con ensayos de mapeo, por ejemplo, en Liu *et al*., Bioinformatics. 19:2397-2403 (2003) y Di *et al*. Bioinformática. 21:1958 (2005). Se desvelan métodos adicionales relacionados con WGSA y matrices útiles para WGSA y aplicaciones de WGSA, por ejemplo, en las Solicitudes de Patentes de Estados Unidos N.º 60/676.058 presentada el 29 de abril de 2005, 60/616.273 presentada el 5 de octubre de 2004, 10/912.445, 11/044.831, 10/442.021, 10/650.332 y 10/463.991. Se describen estudios de asociación de todo el genoma usando ensayos de mapeo, por ejemplo, en Hu *et al*., Cancer Res.; 65 (7):2542-6 (2005), Mitra *et al.*, Cancer Res., 64 (21):8116-25 (2004), Butcher *et al.*, Hum Mol Genet., 1415:25 (2005), y Klein *et al*., Science. 308 (5720):385-9 (2005).

Adicionalmente, ciertas realizaciones pueden incluir métodos para proporcionar información genética sobre redes tales como Internet como se muestra, por ejemplo, en las Solicitudes de Estados Unidos N.º 10/197.621, 10/063.559 (Número de Publicación de Estados Unidos 2002/0183936), 10/065.856, 10/065.868, 10/328.818, 10/328.872, 10/423.403 y 60/482.389.

X. AGENTES ANTISENTIDO Y DE ARNI

Las realizaciones de la presente invención también incluyen oligonucleótidos antisentido y agentes de ARNi que se dirigen a las secuencias polinucleotídicas de MetRS, y métodos de uso de los mismos para reducir la expresión de un transcrito y/o fragmento proteico de MetRS seleccionado. Ciertas realizaciones se refieren a dirigirse a uno o más puntos de unión de corte y empalme (con frecuencia únicos) que generan una variante de corte y empalme, fragmento proteico de MetRS de la presente invención. También se incluyen métodos de inhibición antisentido o ARNi que se dirigen a ciertas formas de corte y empalme, bien para estimular o bien para disuadir del corte y empalme de un fragmento proteico seleccionado. En ciertas realizaciones preferidas, los puntos de unión de corte y empalme que generan los fragmentos proteicos de MetRS se sobreexpresan con respecto a tejidos particulares, y son únicos de esa variante de corte y empalme. En estas realizaciones y otras relacionadas, dichas variantes de corte y empalme no son la única fuente de actividad de MetRS citosólica en el tipo celular diana. Por ejemplo, ciertas variantes de corte y empalme para dirigir pueden representar aproximadamente 10 % a 50 % del número de copias totales de las variantes de corte y empalme de ARN de MetRS en una célula o un tejido dado, y preferentemente aproximadamente 1-10 % del número de copias totales de las variantes de corte y empalme de ARN de MetRS en una célula o un tejido dado. También pueden dirigirse variantes de corte y empalme que son aproximadamente <1 % del número de copias totales de las variantes de corte y empalme de ARN de MetRS en una célula o un tejido dado.

En ciertas realizaciones, el agente antisentido o ARNi no se dirige a la proteína de longitud completa, porque dichas proteínas de longitud completa son responsables de una etapa clave en la síntesis de proteínas, y por lo tanto evita la letalidad que con frecuencia resulta de deleciones de MetRS de tipo silvestre. Ciertos métodos descritos en el presente documento pueden usarse por lo tanto para evitar efectos indeseados tales como toxicidades en tratamientos tanto crónicos como agudos, y para modular selectivamente las actividades no canónicas del fragmento proteico de MetRS. Sin embargo, ciertas realizaciones pueden dirigirse genéricamente a secuencias de MetRS, incluyendo secuencias de MetRS de longitud completa, tal como para destruir o modificar sustancialmente la fisiología celular de una célula o un tejido diana.

En ciertas realizaciones, la variante de corte y empalme de MetRS para dirigir posee una actividad biológica no canónica. En algunas realizaciones, la variante de corte y empalme de MetRS tiene actividad de MetRS canónica reducida o indetectable, y el método antisentido o relacionado con ARNi modula más específicamente su actividad no canónica. En ciertas realizaciones, los agentes antisentido o relacionados con ARNi pueden combinarse con un enfoque de suministro dirigido o local para reducir los efectos no deseados sistémicos a células o tejidos no diana. Entre otros descritos en el presente documento, las células o los tejidos a modo de ejemplo que podrían de este modo ser diana incluyen células cancerosas y células para tejidos que se prestan a dirección localizada, tales como tumores o epitelio mediante aplicación tópica.

A. Agentes antisentido

Las expresiones "oligómero antisentido" o "compuesto antisentido" u "oligonucleótido antisentido" se usan indistintamente y se refieren a una secuencia de subunidades cíclicas, que portan cada una un resto de formación de pares de bases, unidas por enlaces intersubunitarios que permiten que los restos de formación de pares de bases hibriden con una secuencia diana en un ácido nucleico (normalmente un ARN) por formación de pares de bases de Watson-Crick, para formar un heterodúplex de ácido nucleico:oligómero dentro de la secuencia diana, y

normalmente evitan de este modo la traducción de ese ARN. También se incluyen métodos de uso de los mismos para modular la expresión de un transcrito de MetRS seleccionado, tal como una variante de corte y empalme o un fragmento proteolítico, y/o su polipéptido correspondiente.

Los oligonucleótidos antisentido pueden contener entre aproximadamente 8 y 40 subunidades, normalmente aproximadamente 8-25 subunidades, y preferentemente aproximadamente 12 a 25 subunidades. En ciertas realizaciones, los oligonucleótidos pueden tener complementariedad de secuencia exacta con la secuencia diana o casi complementariedad, como se define posteriormente. En ciertas realizaciones, el grado de complementariedad entre la secuencia de dirección diana y antisentido es suficiente para formar un dúplex estable. La región de complementariedad de los oligómeros antisentido con la secuencia de ARN diana puede ser tan corta como de 8-11 bases, pero es preferentemente de 12-15 bases o más, por ejemplo, 12-20 bases o 12-25 bases, incluyendo todos los números enteros entre estos intervalos. Un oligómero antisentido de aproximadamente 14-15 bases es generalmente suficientemente largo para tener una secuencia complementaria única en la dirección del gen de MetRS seleccionado. En ciertas realizaciones, puede requerirse una longitud mínima de bases complementarias para conseguir la Tm de unión requerida, como se analiza en el presente documento.

En ciertas realizaciones, pueden ser adecuados oligómeros antisentido de hasta 40 bases de longitud, en los que al menos un número mínimo de bases, por ejemplo 10-12 bases, son complementarias de la secuencia diana. En general, sin embargo, la captación facilitada o activa en células se optimiza a longitudes de oligómeros menores de aproximadamente 30. Para ciertos oligómeros, descritos adicionalmente posteriormente, se produce un equilibrio óptimo de estabilidad de unión y captación a longitudes de 18-25 bases. Se incluyen oligómeros antisentido (por ejemplo, PNA, LNA, 2'-OMe, MOE) 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 de su secuencia diana de MetRS, o variantes de la misma.

En ciertas realizaciones, los oligómeros antisentido pueden ser 100 % complementarios de una secuencia diana de ácido nucleico de MetRS, o puede incluir desapareamientos, por ejemplo, para acomodar variantes, siempre que un heterodúplex formado entre el oligómero y la secuencia diana de ácido nucleico de MetRS sea suficientemente estable para soportar la acción de nucleasas celulares y otros modos de degradación que pueden producirse *in vivo*. La expresión "secuencia diana" se refiere a una parte del ARN diana contra el que se dirige el oligonucleótido, es decir, la secuencia con la que hibridará el oligonucleótido por formación de pares de bases de Watson-Crick de una secuencia complementaria. En ciertas realizaciones, la secuencia diana puede ser una región contigua de un ARNm de MetRS (por ejemplo, un punto de unión de corte y empalme único de un ARNm de MetRS), o puede estar compuesto de regiones no contiguas del ARNm.

Se analizan posteriormente cadenas principales de oligómeros que son menos susceptibles a escisión por nucleasas. Los desapareamientos, si están presentes, son menos desestabilizantes hacia las regiones finales del dúplex híbrido que en el medio. El número de desapareamientos permitidos dependerá de la longitud del oligómero, el porcentaje de pares de bases G:C en la doble cadena y la posición del desapareamiento o los desapareamientos en la doble cadena, de acuerdo con principios bien entendidos de la estabilidad de dúplex. Aunque dicho oligómero antisentido no es necesariamente 100 % complementario de la secuencia diana de ácido nucleico de MetRS, es eficaz para unirse de forma estable y específicamente con la secuencia diana, de modo que se module una actividad biológica de la diana de ácido nucleico, por ejemplo, expresión de proteína o proteínas de MetRS.

La estabilidad del dúplex formado entre un oligómero y una secuencia diana es una función de la Tm de unión y la susceptibilidad de la doble cadena a escisión enzimática celular. La Tm de un oligonucleótido antisentido con respecto al ARN de secuencia complementaria puede medirse por métodos convencionales, tales como los descritos en Hames *et al*., Nucleic Acid Hybridization, IRL Press, 1985, págs. 107-108 o como se describe en Miyada C.G. y Wallace R.B., 1987, Oligonucleotide hybridization techniques, Methods Enzymol. Vol. 154 págs. 94-107. En ciertas realizaciones, el oligómero antisentido puede tener una Tm de unión, con respecto a un ARN de secuencia complementaria, de temperatura mayor que la corporal y preferentemente mayor de 50 ºC. Se prefieren Tm en el intervalo de 60-80 ºC. De acuerdo con principios bien conocidos, la Tm de un compuesto oligomérico, con respecto a un híbrido de ARN basado en el complementario, puede aumentarse aumentando la relación de bases emparejadas C:G en el dúplex, y/o aumentando la longitud (en pares de bases) del heterodúplex. Al mismo tiempo, para fines de optimización de la captación celular, puede ser ventajoso limitar el tamaño del oligómero antisentido. Por esta razón, se prefieren en general compuestos que muestran alta Tm (50 ºC o más) a una longitud de 25 bases o menos sobre los que requieren más de 25 bases para valores de Tm altos.

Pueden diseñarse oligómeros antisentido para bloquear o inhibir la traducción de ARNm o para inhibir el procesamiento de corte y empalme pre-ARNm natural, o inducir degradación de ARNm diana, y puede decirse que "se dirigen a" o "se dirigen contra" una secuencia diana con la que hibridan. En ciertas realizaciones, la secuencia diana puede incluir cualquier secuencia codificante o no codificante de un transcrito de ARNm de MetRS, y puede por lo tanto estar dentro de un exón o dentro de un intrón. En ciertas realizaciones, la secuencia diana es relativamente única o excepcional entre MetRS (por ejemplo, una MetRS de longitud completa) y es selectiva para

reducir la expresión de un fragmento proteico de MetRS seleccionado, tal como un fragmento proteolítico o variante de corte y empalme. En ciertas realizaciones, el sitio diana incluye un sitio de corte y empalme 3' o 5' de un ARNm preprocesado, o un punto de ramificación. La secuencia diana para un sitio de corte y empalme puede incluir una secuencia de ARNm que tiene su extremo 5' de 1 a aproximadamente 25 a aproximadamente 50 pares de bases cadena abajo de un punto de unión aceptor de corte y empalme o cadena arriba de un punto de unión donador de corte y empalme en un ARNm preprocesado. En ciertas realizaciones, una secuencia diana puede incluir un punto de unión de corte y empalme de un ARNm de MetRS de corte y empalme alternativo, tal como un punto de unión de unión de corte y empalme que no aparece en la MetRS de longitud completa, o es único o excepcional de ese transcrito, porque no aparece o aparece solo escasamente en otras variantes de corte y empalme de MetRS. Se dice más en general que un oligómero está "dirigido contra" una diana biológicamente relevante, tal como polinucleótido de MetRS de referencia, cuando se dirige contra el ácido nucleico de la diana de la manera descrita en el presente documento.

Un oligonucleótido es normalmente complementario de una secuencia diana, tal como ADN o ARN diana. Los términos "complementario" y "complementariedad" se refieren a polinucleótidos (es decir una secuencia de nucleótidos) relacionados con las normas de formación de pares de bases. Por ejemplo, la secuencia "A-G-T" es complementaria a la secuencia "T-C-A". La complementariedad puede ser "parcial", en la que solamente algunas de las bases de ácidos nucleicos coinciden de acuerdo con las normas de formación de pares de bases. O puede ser complementariedad "completa" o "total" (100 %) entre los ácidos nucleicos. El grado de complementariedad entre cadenas de ácido nucleico tiene efectos significativos en la eficacia y la fuerza de hibridación entre cadenas de ácido nucleico. Aunque se desea con frecuencia complementariedad perfecta, algunas realizaciones pueden incluir uno o más pero preferentemente 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 desapareamientos con respecto a la secuencia diana. Se incluyen variaciones en cualquier localización dentro del oligómero. En ciertas realizaciones, se prefieren en general variaciones de secuencia cerca de los extremos de un oligómero frente a variaciones en el interior, y si están presentes son normalmente dentro de aproximadamente 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 nucleótidos del extremo 5' y/o 3' terminal.

La expresión "secuencia de dirección" o en ciertas realizaciones "secuencia de dirección antisentido" se refiere a la secuencia en un oligonucleótido que es complementario (lo que significa, además, sustancialmente complementario) de la secuencia diana en la molécula diana de ADN o ARN. La secuencia completa, o solamente una parte, del compuesto antisentido puede ser complementaria de la secuencia diana. Por ejemplo, en un oligonucleótido que tiene 20-30 bases, aproximadamente 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28 o 29 pueden ser secuencias de dirección que son complementarias de la región diana. normalmente, la secuencia de dirección está formada por bases contiguas, pero puede formarse como alternativa por secuencias no contiguas que cuando se colocan juntas, por ejemplo, de extremos opuestos del oligonucleótido, constituyen una secuencia que abarca la secuencia diana.

Se describen secuencias diana y de dirección como "complementarias" entre sí cuando se produce hibridación en una configuración antiparalela. Una secuencia de dirección puede tener complementariedad "cercana" o "sustancial" a la secuencia diana y aún actuar para el fin de la presente invención, es decir, aún puede ser funcionalmente "complementaria". En ciertas realizaciones, un oligonucleótido puede tener como máximo un desapareamiento con la secuencia diana de 10 nucleótidos y preferentemente como máximo un desapareamiento de 20. Como alternativa, un oligonucleótido puede tener al menos aproximadamente 80 %, 85 %, 90 % de homología de secuencia, y preferentemente al menos 95 % de homología de secuencia, con una secuencia polinucleotídica de referencia de MetRS descrita en el presente documento, o su complemento.

Un oligonucleótido "se hibrida específicamente" con un polinucleótido diana si el oligómero hibrida con una diana (por ejemplo, un polinucleótido de referencia de MetRS o su complemento) en condiciones fisiológicas, con una Tm sustancialmente mayor de 45 ºC, preferentemente al menos 50 ºC, y normalmente 60 ºC-80 ºC o más. Dicha hibridación preferentemente corresponde a condiciones de hibridación rigurosas. A una fuerza iónica y un pH dados, la Tm es la temperatura a la que el 50 % de una secuencia diana hibrida con un polinucleótido complementario. De nuevo, dicha hibridación puede producirse con complementariedad "cercana" o "sustancial" del oligómero antisentido con la secuencia diana, así como con complementariedad exacta.

Las expresiones se unen específicamente o hibridan específicamente se refieren en general a una sonda oligonucleotídica o secuencia polinucleotídica que no solamente se une a su secuencia génica diana pretendida en una muestra en condiciones de hibridación seleccionadas, sino que no se une significativamente con otras secuencias diana en muestra, y por lo tanto diferencia entre su diana pretendida y todas las otras dianas en el grupo diana. Una sonda que hibrida específicamente con su secuencia diana pretendida también puede detectar diferencias de concentración en las condiciones de hibridación seleccionadas, como se describe en el presente documento.

Una molécula oligomérica "resistente a nucleasa" (oligómero) se refiere a una cuya cadena principal es sustancialmente resistente a la escisión por nucleasa, en forma no hibridada o hibridada; por nucleasas extracelulares e intracelulares comunes en el cuerpo; es decir, el oligómero muestra poca o ninguna escisión por nucleasa en condiciones de nucleasa normales en el cuerpo al que se expone el oligómero.

Un "heterodúplex" se refiere a un dúplex entre un oligonucleótido y la parte complementaria de un polinucleótido diana, tal como un ADN o ARN diana. Un "heterodúplex resistente a nucleasa" se refiere a un heterodúplex formado por la unión de un oligómero con su diana complementaria, de modo que el heterodúplex sea sustancialmente resistente a degradación *in vivo* por nucleasas intracelulares y extracelulares, tales como RNasaH, que son capaces de cortar complejos de ARN/ARN o ARN/ADN bicatenarios.

Una "subunidad" de un oligonucleótido se refiere a una unidad de nucleótido (o análogo de nucleótido). El término se puede referir a la unidad de nucleótido con o sin enlace intersubunitario unido, aunque, cuando se hace referencia a una "subunidad con carga", la carga normalmente reside dentro del enlace intersubunitario (por ejemplo, un enlace fosfato o fosforotioato o un enlace catiónico).

Las subunidades cíclicas de un oligonucleótido pueden basarse en ribosa u otro azúcar pentosa o, en ciertas realizaciones, grupos alternos o modificados. Los ejemplos de cadenas principales oligonucleotídicas modificadas incluyen, sin limitación, fosforotioatos, fosforotioatos quirales, fosforoditioatos, fosfotriésteres, aminoalquilfosfotriésteres, metil y otros alquilfosfonatos incluyendo 3'-alquilen fosfonatos y fosfonatos quirales, fosfinatos, fosforamidatos incluyendo 3'-aminofosforamidato y aminoalquilfosforamidatos, tionofosforamidatos, tionoalquilfosfonatos, tionoalquilfosfotriésteres y boranofosfatos que tienen enlaces 3'-5' normales, análogos con enlaces 2'-5' de estos y los que tienen polaridad invertida en los que los pares adyacentes de unidades de nucleósidos tienen enlaces 3'-5' a 5'-3' o 2'-5' a 5'-2'. También se contemplan ácidos nucleicos peptídicos (PNA), ácidos nucleicos bloqueados (LNA), 2'-O-metil oligonucleótidos (2'-OMe), 2'-metoxietoxi oligonucleótidos (MOE), entre otros oligonucleótidos conocidos en la técnica.

El resto de formación de pares de bases de purina o pirimidina es normalmente adenina, citosina, guanina, uracilo, timina o inosina. También se incluyen bases tales como piridin-4-ona, piridin-2-ona, fenilo, pseudouracilo, 2,4,6-trimetoxi benceno, 3-metil uracilo, dihidrouridina, naftilo, aminofenilo, 5-alquilcitidinas (por ejemplo, metilcitidina), 5-alquiluridinas (por ejemplo, ribotimidina), 5-halouridina (por ejemplo, 5-bromouridina) o 6-azapirimidinas o 6-alquilpirimidinas (por ejemplo, 6-metiluridina), propina, quesosina, 2-tiouridina, 4-tiouridina, wibutosina, wibutoxosina, 4-acetiltidina, 5-(carboxihidroximetil) uridina, 5'-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluridina, β-D-galactosilqueosina, 1-metiladenosina, 1-metilinosina, 2,2-dimetilguanosina, 3-metilcitidina, 2-metiladenosina, 2-metilguanosina, N6-metiladenosina, 7-metilguanosina, 5-metoxiaminometil-2-tiouridina, 5-metilaminometiluridina, 5-metilcarbonilmetiluridina, 5-metiloxiuridina, 5-metil-2-tiouridina, 2-metiltio-N6-isopenteniladenosina, β-D-manosilqueosina, ácido uridin-5-oxiacético, 2-tiocitidina, derivados de treonina y otros (Burgin *et al*., 1996, Biochemistry, 35, 14090; Uhlman y Peyman, mencionado anteriormente). Por "bases modificadas" en este aspecto se entienden bases nucleotídicas distintas de adenina (A), guanina (G), citosina (C), timina (T) y uracilo (U), como se ha ilustrado anteriormente; dichas bases pueden usarse en cualquier posición en la molécula antisentido. Los expertos en la materia apreciarán que dependiendo de los usos de los oligómeros, T y U son intercambiables. Por ejemplo, para otras químicas antisentido tales como oligonucleótidos antisentido 2'-O-metilo que son más parecidos a ARN, las bases T pueden mostrarse como U.

Como se ha indicado anteriormente, ciertos oligonucleótidos proporcionados en el presente documento incluyen ácidos nucleicos peptídicos (PNA). Los ácidos nucleicos peptídicos (PNA) son análogos de ADN en los que la cadena principal es estructuralmente homomorfa con una cadena principal de desoxirribosa, que consiste en unidades de N-(2-aminoetil) glicina con los que se unen bases de pirimidina o purina. Los PNA que contienen bases de pirimidina y purina naturales hibridan con oligonucleótidos complementarios que obedecen a las normas de formación de pares de bases de Watson-Crick, e imitan el ADN con respecto a reconocimiento de pares de bases (Egholm, Buchardt *et al*., 1993). La cadena principal de PNA se forma por enlaces peptídicos en lugar de enlaces fosfodiéster, haciéndolos adecuados para aplicaciones antisentido (véase estructura posterior). La cadena principal no tiene carga, dando como resultado dobles cadenas de PNA/ADN o PNA/ARN que muestran estabilidad térmica mayor de lo normal. Los PNA no son reconocidos por nucleasas o proteasas.

Pueden producirse PNA de forma sintética usando cualquier técnica conocida en este campo. El PNA es un análogo de ADN en el que una cadena principal de poliamida reemplaza el anillo de ribosa de fosfato tradicional del ADN. A pesar de un cambio estructural radical de la estructura natural, el PNA es capaz de unirse de forma específica de secuencia en una forma de hélice con ADN o ARN. Las características de PNA incluyen una alta afinidad de unión a ADN o ARN complementario, un efecto desestabilizante provocado por desapareamiento de una única base, resistencia a nucleasas y proteasas, hibridación con ADN o ARN independientemente de la concentración salina y formación de tríplex con ADN de homopurina. Panagene™ ha desarrollado sus monómeros de PNA Bts patentados (Bts; grupo de benzotiazol-2-sulfonilo) y proceso de oligomerización patentado. La oligomerización de PNA usando monómeros de PNA Bts está compuesta de ciclos repetitivos de desprotección, acoplamiento y recubrimiento terminal. Las patentes de Panagene para esta tecnología incluyen los documentos US 6969766, US 7211668, US 7022851, US 7125994, US 7145006 y US 7179896. Las patentes de Estados Unidos representativas que enseñan la preparación de compuestos de PNA incluyen, pero sin limitación, Patentes de Estados Unidos N.º 5.539.082; 5.714.331; y 5.719.262. Pueden encontrarse enseñanzas adicionales de compuestos de PNA en Nielsen *et al*., Science, 1991, 254, 1497.

También se incluyen subunidades de "ácido nucleico bloqueado" (LNA). Las estructuras de LNA se conocen en la

técnica: por ejemplo, Wengel *et al.*, Chemical Communications (1998) 455; Tetrahedron (1998) 54, 3607, y Accounts of Chem. Research (1999) 32, 301); Obika, *et al.*, Tetrahedron Letters (1997) 38, 8735; (1998) 39, 5401, y Bioorganic Medicinal Chemistry (2008) 16, 9230.

Los oligonucleótidos pueden incorporar uno o más LNA; en algunos casos, los compuestos pueden estar compuestos completamente por LNA. Se conocen en la técnica métodos para la síntesis de subunidades de nucleósidos de LNA individuales y su incorporación en oligonucleótidos: Patentes de Estados Unidos 7.572.582; 7.569.575; 7.084.125; 7.060.809; 7.053.207; 7.034.133; 6.794.499; y 6.670.461. Los enlazadores intersubunitarios típicos incluyen restos de fosfodiéster y fosforotioato; como alternativa, pueden emplearse enlazadores que no contienen fósforo. Una realización preferida es un compuesto que contiene LNA, en el que cada subunidad de LNA está separada por una subunidad de ADN (es decir, un nucleótido de desoxirribosa). Compuestos preferidos adicionales están compuestos por subunidades LNA y ADN alternantes en las que el enlazador intersubunitario es fosforotioato.

Ciertos oligonucleótidos pueden comprender subunidades basadas en morfolino que portan restos formadores de pares de bases, unidos por enlaces sin carga o sustancialmente sin carga. Las expresiones "oligómero de morfolino" o "PMO" (oligómero de fosforamidato o fosforodiamidato de morfolino) se refieren a un análogo de oligonucleótido compuesto de estructuras subunitarias de morfolino, en el que (i) las estructuras están unidas entre sí por enlaces que contienen fósforo, de uno a tres átomos de longitud, preferentemente de dos átomos de longitud, y preferentemente sin carga o catiónico, que unen el nitrógeno de morfolino de una subunidad con un carbono exocíclico 5' de una subunidad adyacente, y (ii) cada anillo de morfolino porta una purina o pirimidina o un resto de formación de pares de bases equivalente eficaz para unirse, por formación de enlaces de hidrógeno específicos, con una base en un polinucleótido.

Pueden realizarse variaciones de este enlace, siempre que no interfieran con la unión o actividad. Por ejemplo, el oxígeno unido a fósforo puede sustituirse con azufre (tiofosforodiamidato). El oxígeno 5' puede sustituirse con amino o amino sustituido con alquilo inferior. El nitrógeno colgante unido a fósforo puede estar instituido, monosustituido o disustituido con (opcionalmente sustituido) alquilo inferior. El resto formador de pares de bases de purina o pirimidina es normalmente adenina, citosina, guanina, uracilo, timina o inosina. La síntesis, las estructuras y las características de unión de oligómeros de morfolino se detallan en las Patentes de Estados Unidos N.º 5.698.685, 5.217.866, 5.142.047, 5.034.506, 5.166.315, 5.521.063 y 5.506.337, y solicitudes de PCT N.º. PCT/US07/11435 (enlaces catiónicos) y US08/012804 (síntesis mejorada).

Las subunidades de morfolino pueden también unirse por enlaces intersubunitarios no basados en fósforo, como se describe adicionalmente posteriormente, en los que al menos un enlace está modificado con un grupo catiónico colgante como se ha descrito anteriormente. Podrían usarse otros enlaces análogos de oligonucleótidos que no tienen carga en su estado no modificado pero que podrían también portar un sustituyente de amina colgante. Por ejemplo, podría emplearse un átomo de nitrógeno 5' en un anillo de morfolino en un enlace de sulfamida o un enlace de urea (en los que se reemplaza fósforo con carbono o azufre, respectivamente) y modificarse de una manera análoga al átomo de nitrógeno 5' en la estructura (b3) anterior.

Ciertas realizaciones incluyen oligómeros de morfolino sustancialmente sin carga, tales como un oligómero de morfolino con enlaces fosforodiamidato sustancialmente sin carga. Una cadena principal que contiene fósforo, sustancialmente sin carga, en un análogo de oligonucleótido es una en la que una mayoría de los enlaces subunitarios, por ejemplo, entre 50 y 100 %, normalmente al menos 60 % a 100 % o 75 % u 80 % de sus enlaces, no tienen carga a pH fisiológico, y contienen un único átomo de fósforo. Los ejemplos de oligonucleótidos de morfolino que tienen enlaces de cadena principal que contienen fósforo incluyen oligonucleótidos de morfolino con enlaces de fosforamidato y fosforodiamidato. Ciertas realizaciones pueden contener grupos con carga positiva preferentemente a aproximadamente 10 %-50 % de sus enlaces de cadena principal.

Las propiedades de las subunidades basadas en morfolino incluyen, por ejemplo, la capacidad para unirse en una forma oligomérica por enlaces de cadena principal estables sin carga o con carga positiva, la capacidad para soportar una base de nucleótido (por ejemplo, adenina, citosina, guanina, timidina, uracilo e hipoxantina) de modo que el polímero formado pueda hibridar con un ácido nucleico diana de bases complementarias, incluyendo ARN diana, valores de Tm por encima de aproximadamente 45 ºC en oligonucleótidos relativamente cortos (por ejemplo, 10-15 bases), la capacidad del oligonucleótido para transportarse de forma activa o pasiva a células de mamífero, y la capacidad del heterodúplex de oligonucleótido antisentido:ARN para resistir la degradación por RNasa o RNasaH, respectivamente.

En ciertas realizaciones, un oligonucleótido sustancialmente sin carga puede modificarse para incluir enlaces con carga, por ejemplo, hasta aproximadamente 1 por cada 2-5 enlaces sin carga, tal como aproximadamente 4-5 por cada 10 enlaces sin carga. En ciertas realizaciones, puede verse mejora óptima en la actividad antisentido cuando aproximadamente el 25 % de los enlaces de cadena principal son catiónicos. En ciertas realizaciones, puede verse potenciación con un número pequeño, por ejemplo, 10-20 % de enlaces catiónicos, o cuando el número de enlaces catiónicos está en el intervalo de 50-80 %, tal como aproximadamente 60 %. En ciertas realizaciones las cargas de cadena principal catiónicas pueden potenciarse adicionalmente distribuyendo el volumen de las cargas cerca de los

enlaces de cadena principal “de región central” del oligonucleótido antisentido, por ejemplo, en un oligonucleótido de 20 unidades con 8 enlaces de cadena principal catiónicos, que tienen al menos 70 % de estos enlaces con carga localizados en los 10 enlaces más centrales.

Pueden usarse oligonucleótidos que se dirigen a una o más partes de una secuencia de referencia polinucleotídica de MetRS o su complemento en cualquiera de los métodos terapéuticos, de diagnóstico o de cribado de fármacos descritos en el presente documento y evidentes para los expertos en la materia.

B. Agentes de Interferencia de ARN

Ciertas realizaciones se refieren a agentes de interferencia de ARN (ARNi) que se dirigen a uno o más transcritos de ARNm de un polinucleótido de referencia de metionil-ARNt sintetasa (MetRS), incluyendo fragmentos y variantes de corte y empalme de los mismos. También se incluyen métodos de uso de los mismos para modular los niveles de un transcrito de MetRS seleccionado, tal como una variante de corte y empalme o fragmento proteolítico endógeno de MetRS.

El término "bicatenario" significa dos cadenas de ácido nucleico separadas que comprenden una región en la que al menos una parte de las cadenas son suficientemente complementarias para formar enlace de hidrógeno y formar una estructura de dúplex. El término "dúplex" o "estructura de dúplex" se refiere a la región de una molécula bicatenaria en la que dos cadenas separadas son sustancialmente complementarias, y por lo tanto hibridan entre sí. "ARNbc" se refiere a una molécula de ácido ribonucleico que tiene una estructura de dúplex que comprende dos cadenas de ácido nucleico complementarias y antiparalelas (es decir, las cadenas con sentido y antisentido). No todos los nucleótidos de ARNbc deben mostrar pares de bases de Watson-Crick; las dos cadenas de ARN pueden ser sustancialmente complementarias. Las cadenas de ARN pueden tener el mismo número de nucleótidos o uno diferente.

En ciertas realizaciones, un ARNbc es o incluye una región que es al menos parcialmente complementaria del ARN diana. En ciertas realizaciones, el ARNbc es completamente complementario del ARN diana. No es necesario que haya complementariedad perfecta entre el ARNbc y la diana, pero la correspondencia debe ser suficiente para permitir que el ARNbc, o un producto de escisión del mismo, dirija el silenciamiento específico de secuencia, tal como por escisión por ARNi del ARN diana. La complementariedad, o el grado de homología con la cadena diana, es normalmente más crítico en la cadena antisentido. Aunque se desea con frecuencia complementariedad perfecta, particularmente en la cadena antisentido, algunas realizaciones pueden incluir uno o más pero preferentemente 6, 5, 4, 3, 2 o menos desapareamientos con respecto al ARN diana. Los desapareamientos se toleran más en las regiones terminales, y si están presentes están preferentemente en una región o regiones terminales, por ejemplo, a una distancia de 6, 5, 4 o 3 nucleótidos del extremo 5’ y/o 3’ terminal. Solo es necesario que la cadena con sentido sea sustancialmente complementaria con la cadena antisentido para mantener el carácter general bicatenario de la molécula.

Como se usa en el presente documento, "ARNbc modificado" se refiere a una molécula de ARNbc que comprende al menos una alteración que la hace más resistente a nucleasas (por ejemplo, proteína quinasa) que una molécula de ARNbc idéntica que reconoce el mismo ARN diana. Los ARNbc modificados pueden incluir un saliente de nucleótidos monocatenario y/o al menos un nucleótido sustituido.

Como se usa en el presente documento, un "saliente de nucleótidos" se refiere al nucleótido o los nucleótidos no emparejados que protruyen de la estructura de dúplex cuando un extremo 3’ de la cadena de ARN se extiende más allá del extremo 5’ de la otra cadena complementaria o viceversa. "Romo" o "extremo romo" significa que no hay nucleótidos desapareados en ese extremo del ARNbc, es decir, no hay saliente de nucleótidos. Un ARNbc de "extremos romos" es un ARNbc que es bicatenario sobre su longitud completa, es decir, no hay ningún saliente de nucleótidos en uno de los extremos de la molécula.

La expresión "par de bases terminal", como se usa en el presente documento, se refiere al último par de bases de nucleótidos en un extremo de la región de dúplex de una molécula bicatenaria. Por ejemplo, si un ARNbc u otra molécula tiene extremos romos (es decir, no tiene ningún saliente de nucleótidos), los últimos pares de bases de nucleótidos en ambos extremos de la molécula son pares de bases terminales. Cuando un ARNbc u otra molécula tiene un saliente de nucleótidos en uno o ambos extremos de la estructura de dúplex, el último o los últimos pares de bases de nucleótidos inmediatamente adyacentes al saliente o los salientes de nucleótidos es el par de bases terminal en ese extremo o esos extremos de la molécula.

En ciertas realizaciones, los métodos proporcionados en el presente documento pueden utilizar moléculas de ácido ribonucleico bicatenario (ARNbc) como agentes moduladores, para reducir la expresión de un transcrito de MetRS tal como un fragmento seleccionado o variante de corte y empalme. Los ARNbc generalmente comprenden dos cadenas individuales. Una cadena del ARNbc comprende una secuencia de nucleótidos que es sustancialmente idéntica a una parte del gen diana o región la región diana (la cadena "con sentido"), y la otra cadena (la cadena "complementaria" o "antisentido") comprende una secuencia que es sustancialmente complementaria de una parte de la región diana. Las cadenas son suficientemente complementarias para hibridar para formar una estructura de

dúplex. En ciertas realizaciones, la cadena de ARN complementaria puede ser menor de 30 nucleótidos, menor de 25 nucleótidos de longitud, o incluso de 19 a 24 nucleótidos de longitud. En ciertos aspectos, la secuencia de nucleótidos complementaria puede ser de 20 a 23 nucleótidos de longitud, o 22 nucleótidos de longitud.

En ciertas realizaciones, al menos una de las cadenas de ARN comprende un saliente de nucleótidos de 1 a 4 nucleótidos de longitud. En otras realizaciones, el ARNbc puede comprender además al menos un nucleótido modificado químicamente. En ciertos aspectos, un ARNbc que comprende un saliente monocatenario de 1 a 4 nucleótidos puede comprender una molécula en la que el nucleótido desapareado del saliente monocatenario que está directamente adyacente al par de nucleótidos terminal contiene una base de purina. En otros aspectos, los últimos pares de nucleótidos complementarios en ambos extremos de un ARNbc son un par de G-C o al menos dos de los últimos cuatro pares de nucleótidos terminales son pares G-C.

Ciertas realizaciones de la presente invención pueden comprender micro-ARN. Los micro-ARN representan un gran grupo de ARN pequeños producidos de forma natural en organismos, algunos de los cuales regulan la expresión de genes diana. Se forman micro-ARN de un trascrito precursor en horquilla monocatenario de aproximadamente 70 nucleótidos por Dicer. (V. Ambros *et al.*, Current Biology 13:807, 2003). Ciertos micro-ARN pueden transcribirse como precursores de ARN en horquilla, que después se procesan a sus formas maduras por enzima Dicer.

Ciertas realizaciones también pueden emplear ARN de interferencia cortos (ARNip). En ciertas realizaciones, la primera cadena del oligonucleótido bicatenario contiene dos restos de nucleósidos más que la segunda cadena. En otras realizaciones, la primera cadena y la segunda cadena tienen el mismo número de nucleósidos; sin embargo, la primera y segunda cadenas pueden compensarse de modo que los dos nucleósidos terminales de la primera y la segunda cadenas no estén emparejados con un resto en la cadena complementaria. En ciertos casos, los dos nucleósidos que no están emparejados son restos de timidina.

También se incluyen ARN en horquilla cortos (ARNhp) y micro ARN (miARN). Una estructura bicatenaria de un ARNhp se forma por una cadena de ARN autocomplementaria sencilla, y la formación de dúplex de ARN puede iniciarse bien dentro o bien fuera de la célula. Los microARN (miARN) son ARN no codificantes pequeños de 20-22 nucleótidos, normalmente escindidos de estructuras precursoras de ARN regresivas de ~70 nucleótidos conocidas como pre-miARN.

En casos en los que el agente de modulación comprende ARNip, el agente debería incluir una región de homología suficiente con la región diana, y ser de suficiente longitud con respecto a nucleótidos, de modo que el ARNip, o un fragmento del mismo, pueda mediar en la regulación negativa del ARN diana. Se entenderá que el término "ribonucleótido" o "nucleótido" puede, en el caso de un ARN modificado o un sustituto de nucleótido, también referirse a un nucleótido modificado, o resto de remplazo sustituto en una o más posiciones. Por lo tanto, un agente de ARNip es o incluye una región que es al menos parcialmente complementaria del ARN diana, como se describe en el presente documento.

Además, un agente modulador de ARNip puede modificarse o incluir sustitutos de nucleósidos. Las regiones monocatenarias de un agente de ARNip pueden modificarse o incluir sustitutos de nucleósidos, por ejemplo, la región o las regiones no emparejadas de una estructura en horquilla, por ejemplo, una región que une dos regiones complementarias, pueden tener modificaciones o sustitutos de nucleósidos. La modificación para estabilizar uno o más extremos 3' o 5' de un agente de ARNip, por ejemplo, contra exonucleasas, o para favorecer que el agente de ARNip antisentido entre en RISC también es útil. Las modificaciones pueden incluir enlazadores de aminoácidos C3 (o C6, C7, C12), enlazadores de tiol, enlazadores de carboxilo, espaciadores no nucleotídicos (C3, C6, C9, C12, abásicos, trietilenglicol, hexaetilenglicol), reactivos de biotina o fluoresceína especiales que aparecen como fosforamiditas y que tienen otro grupo hidroxilo protegido por DMT, permitiendo múltiples acoplamientos durante la síntesis de ARN.

Los agentes de ARNip pueden incluir, por ejemplo, moléculas que son suficientemente largas para desencadenar la respuesta de interferón (que puede escindirse por Dicer (Bernstein *et al.*, 2001. Nature, 409:363-366) y entrar en un RISC (complejo de silenciamiento inducido por ARNi)), además de moléculas que son suficientemente cortas para que no desencadenen la respuesta de interferón (moléculas que también pueden escindirse por Dicer y/o entrar en un RISC), por ejemplo, moléculas que son de un tamaño que permite la entrada en un RISC, por ejemplo, moléculas que se asemejan a productos de escisión de Dicer. Un agente modulador de ARNip, o un producto de escisión del mismo, puede regular negativamente un gen diana, por ejemplo, induciendo ARNi con respecto a un ARN diana, preferentemente una diana de MetRS tal como una variante de corte y empalme seleccionada.

Cada cadena de un agente de ARNip puede ser igual a o menor de 35, 30, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16 o 15 nucleótidos de longitud. La cadena es preferentemente de al menos 19 nucleótidos de longitud. Por ejemplo, cada cadena puede ser de entre 21 y 25 nucleótidos de longitud. Los agentes de ARNip preferidos tienen una región dúplex de 17, 18, 19, 29, 21, 22, 23, 24 o 25 pares de nucleótidos y uno o más salientes, preferentemente uno o dos salientes 3', de 2-3 nucleótidos.

Además de la homología con el ARN diana y la capacidad para regular negativamente un gen diana, un agente de

ARNip puede tener una o más de las siguientes propiedades: puede, a pesar de las modificaciones, incluso en un número muy grande, o todos los nucleósidos, tener una cadena antisentido que puede presentar bases (o bases modificadas) en el armazón tridimensional apropiado para poder formar pares de bases correctos y formar una estructura de dúplex con un ARN diana homólogo que es suficiente para permitir la regulación negativa de la diana, por ejemplo, por escisión del ARN diana; puede, a pesar de las modificaciones, incluso en un número muy grande, o todos los nucleósidos, aún tener propiedades "de tipo ARN", es decir, puede poseer la estructura general, las propiedades químicas y físicas de una molécula de ARN, incluso aunque no exclusivamente, o incluso parcialmente, del contenido basado en ribonucleótidos. Por ejemplo, un agente de ARNip puede contener, por ejemplo, una cadena con sentido y/o una antisentido en la que todos los azúcares de nucleótidos contienen por ejemplo 2' flúor en lugar de 2' hidroxilo. Aún puede esperarse que este agente que contiene desoxirribonucleótido muestre propiedades de tipo a ARN. Aunque sin desear quedar ligado a la teoría, el flúor electronegativo prefiere una orientación axial cuando se une a la posición C2' de ribosa. Esta referencia espacial de flúor puede, a su vez, obligar a los azúcares a adoptar un pliegue C_3'-endo. Este es el mismo modo de pliegue que se observa en moléculas de ARN y da lugar a la hélice del tipo familia A característica de ARN. Además, ya que el flúor es un buen aceptor de enlaces de hidrógeno, puede participar en las mismas interacciones de enlaces de hidrógeno con moléculas de agua que se sabe que estabilizan estructuras de ARN. En general, se prefiere que un resto modificado en la posición de azúcar 2' sea capaz de entrar en el enlace de H que es más característico del resto OH de un ribonucleótido que el resto H de un desoxirribonucleótido.

Un "agente de ARNi monocatenario" como se usa en el presente documento, es un agente de ARNi que está compuesto de una única molécula. Puede incluir una región en dúplex, formada por emparejamiento intracatenario, por ejemplo, puede ser, o incluir, una estructura en horquilla o en mango de sartén. Los agentes moduladores de ARNi monocatenarios son preferentemente antisentido con respecto a la molécula diana. Un agente de ARNi monocatenario debería ser suficientemente largo para que pueda entrar en el RISC y participar en escisión mediada por RISC de un ARNm diana. Un agente de ARNi monocatenario es de al menos 14, y más preferentemente al menos 15, 20, 25, 29, 35, 40 o 50 nucleótidos de longitud. Es preferentemente de menos de 200, 100 o 60 nucleótidos de longitud.

Los agentes moduladores de ARNi en horquilla pueden tener una región de dúplex igual a o al menos 17, 18, 19, 29, 21, 22, 23, 24 o 25 pares de nucleótidos. La región de dúplex puede ser preferentemente igual a o menor de 200, 100 o 50, de longitud. Ciertos intervalos para la región de dúplex son 15-30, 17 a 23, 19 a 23 y 19 a 21 pares de nucleótidos de longitud. La horquilla puede tener un saliente monocatenario o región despareada terminal, preferentemente la 3', y preferentemente del lado antisentido de la horquilla. En ciertas realizaciones, los salientes son de 2-3 nucleótidos de longitud.

Ciertos agentes moduladores utilizados de acuerdo con los métodos proporcionados en el presente documento pueden comprender oligonucleótidos de ARNi tales como oligonucleótidos quiméricos o "quimeras" que contienen dos o más regiones químicamente definidas, cada una compuesta de al menos una unidad monomérica, es decir, un nucleótido en el caso de un compuesto oligonucleotídico. Estos oligonucleótidos normalmente contienen al menos una región en la que el oligonucleótido está modificado para conferir al oligonucleótido una resistencia aumentada a la degradación por nucleasa, captación celular aumentada y/o afinidad de unión aumentada por el ácido nucleico diana. En consecuencia, pueden obtenerse con frecuencia resultados comparables con oligonucleótidos más cortos cuando se usan oligonucleótidos quiméricos, en comparación con oligodesoxinucleótidos de fosforotioato. Pueden formarse oligonucleótidos quiméricos como estructuras compuestas de dos o más oligonucleótidos, oligonucleótidos modificados, oligonucleótidos y/o miméticos oligonucleotídicos como se ha descrito anteriormente. Dichos oligonucleótidos también se han denominado en la técnica híbridos u oligómeros con huecos. Las patentes de Estados Unidos representativas que enseñan la preparación de dichas estructuras híbridas incluyen, pero sin limitación, Patentes de Estados Unidos N.º 5.013.830; 5.149.797; 5.220.007; 5.256.775; 5.366.878; 5.403.711; 5.491.133; 5.565.350; 5.623.065; 5,652,355; 5.652.356; 5.700.922; y 5.955.589. En ciertas realizaciones, el oligonucleótido quimérico es ARN-ADN, ADN-ARN, ARN-ADN-ARN, ADN-ARN-ADN o ARN-ADN-ARN-ADN, en el que el oligonucleótido es de entre 5 y 60 nucleótidos de longitud.

En un aspecto de la invención, los agentes de ARNi están relacionados con un oligonucleótido que comprende al menos un ligando unido con una nucleobase alterada o no natural. Un gran número de compuestos pueden actuar como la base alterada. La estructura de la base alterada es importante en la medida en que la base alterada no debería evitar sustancialmente la unión del oligonucleótido con su diana, por ejemplo, ARNm. En ciertas realizaciones, la base alterada es difluorotolilo, nitropirrolilo, nitroimidazolilo, nitroindolilo, naftalenilo, antrancenilo, piridinilo, quinolinilo, pirenilo o el radical divalente de una cualquiera de las nucleobases no naturales descritas en el presente documento. En ciertas realizaciones, la nucleobase no natural es difluorotolilo, nitropirrolilo o nitroimidazolilo. En ciertas realizaciones, la nucleobase no natural es difluorotolilo. Una amplia diversidad de ligandos se conocen en la técnica y son susceptibles a la presente invención. Por ejemplo, el ligando puede ser un esteroide, ácido biliar, lípido, ácido fólico, piridoxal, B12, riboflavina, biotina, compuesto aromático, compuesto policíclico, éter de corona, intercalador, molécula de escisión, agente de unión de proteínas o carbohidrato. En ciertas realizaciones, el ligando es un esteroide o compuesto aromático. En ciertos casos, el ligando es colesterilo.

En otras realizaciones, el agente de ARNi es un oligonucleótido unido a un ligando para los fines de mejorar la

dirección y captación celular. Por ejemplo, un agente de ARNi puede estar unido a un anticuerpo, o fragmento de unión al antígeno del mismo. Como ejemplo adicional, un agente de ARNi puede estar unido con una molécula de unión a ligando específica, tal como un polipéptido o fragmento polipeptídico que se une específicamente con un receptor de superficie celular particular.

En otras realizaciones, el agente modulador comprende una nucleobase no natural, como se describe en el presente documento. En ciertos casos, el resto de azúcar ribosa que aparece de forma natural en nucleósidos se reemplaza con un azúcar hexosa. En ciertos aspectos, el azúcar hexosa es una alosa, altrosa, glucosa, manosa, gulosa, idosa, galactosa, talosa o un derivado de las mismas. En una realización preferida, la hexosa es una D-hexosa. En ciertos casos, el resto de azúcar ribosa que aparece de forma natural en nucleósidos se reemplaza con un anillo de heteroalquilo policíclico o un grupo ciclohexenilo. En ciertos casos, el grupo de heteroalquilo policíclico es un anillo bicíclico que contiene un átomo de oxígeno en el anillo. En ciertos casos, el grupo de heteroalquilo policíclico es un biciclo[2.2.1]heptano, un biciclo[3.2.1]octano o un biciclo[3.3.1]nonano. Los ejemplos de agentes de ARNi modificados también incluyen oligonucleótidos que contienen cadenas principales modificadas o enlaces internucleosídicos no naturales, como se describe en el presente documento.

La presente invención abarca además oligonucleótidos que emplean ribozimas. Las moléculas de ARN sintéticas y derivados de las mismas que catalizan actividades endorribonucleasa altamente específicas se conocen como ribozimas. (Véase, por ejemplo, Patente de Estados Unidos N.º 5.543.508 de Haseloff *et al.*, y Patente de Estados Unidos N.º 5.545.729 de Goodchild *et al.*) Las reacciones de escisión se catalizan por las moléculas de ARN en sí mismas. En moléculas de ARN de origen natural, los sitios de escisión autocatalizados se localizan dentro de regiones altamente conservadas de estructura secundaria de ARN (Buzayan *et al.*, Proc. Nat. Acad. Sci. USA, 1986, 83, 8859; Forster *et al.* Cell, 1987, 50, 9). Se han modificado moléculas de ARN autocatalíticas de origen natural para generar ribozimas que pueden dirigirse a una molécula de ARN patógena o celular particular con un alto grado de especificidad. Por lo tanto, las ribozimas cumplen el mismo fin general como oligonucleótidos antisentido (es decir, modulación de la expresión de un gen específico) y, como los oligonucleótidos, son ácidos nucleicos que poseen partes significativas monocatenarias.

En ciertos casos, los agentes de ARNi u oligonucleótidos antisentido para uso con los métodos proporcionados en el presente documento pueden modificarse por un grupo no ligando. Varias moléculas no ligando se han conjugado con oligonucleótidos para potenciar la actividad, distribución celular, dirección celular o captación celular del oligonucleótido, y están disponibles en la bibliografía científica procedimientos para realizar dichas conjugaciones. Dichos restos no ligando han incluido restos lipídicos, tales como colesterol (Letsinger *et al.*, Proc. Nat. Acad. Sci. USA, 1989, 86:6553), péptidos ricos en arginina, ácido cólico (Manoharan *et al.*, Bioorg, Med. Chem. Lett., 1994, 4:1053), un tioéter, por ejemplo, hexil-5-tritiltiol (Manoharan *et al.*, Ann, NY Acad. Sci., 1992, 660: Bioorg, Med. Chem. Let., 1993, 3:2765), un tiocolesterol (Oberhauser *et al.*, Nucl. Acids Res., 1992, 20:533), una cadena alifática, por ejemplo, restos de dodecandiol o undecilo (Saison-Behmoaras *et al.*, EMBO J., 1991, 10:111, Kabanov *et al.*, FEBS Lett., 1990, 259:327, Svinarchuk *et al.*, Biochimie, 1993, 75:49), un fosfolípido, por ejemplo di-hexadecil-rac-glicerol o 1,2-di-O-hexadecil-rac-glicero-3-H-fosfonato de trietilamonio (Manoharan *et al.*, Tetrahedron Lett., 1995, 36:3651, Shea *et al.*, Nucl. Acids Res., 1990, 18:3777), una poliamina o una cadena de polietilenglicol (Manoharan *et al.*, Nucleosides & Nucleotides, 1995, 14:969), o ácido adamantano acético (Manoharan *et al.*, Tetrahedron Lett., 1995, 36:3651), un resto de palmitilo (Mishra *et al.*, Biochim. Biophys. Acta, 1995, 1264:229), o un resto de octadecilamina o hexilamino-carbonil-oxicolesterol (Crooke *et al.*, J. Pharmacol. Exp. Ther., 1996, 277:923). Se han enumerado anteriormente patentes de Estados Unidos representativas que enseñan la preparación de dichos conjugados oligonucleotídicos. Los protocolos de conjugación típicos implican la síntesis de oligonucleótidos que portan un enlazador amino en una o más posiciones de la secuencia. El grupo amino se hace reaccionar después con la molécula que se conjuga usando reactivos de activación o acoplamiento apropiados. La reacción de conjugación puede realizarse bien con el oligonucleótido aún unido al soporte sólido o después de la escisión del oligonucleótido en fase de solución. La purificación del conjugado oligonucleotídico por HPLC normalmente proporciona el conjugado puro.

Pueden encontrarse ejemplos adicionales de agentes de ARNi en las Publicaciones de Solicitud de Estados Unidos N.º 2007/0275465, 2007/0054279, 2006/0287260, 2006/0035254, 2006/0008822. También se incluyen sistemas de suministro de vectores que son capaces de expresar las secuencias de dirección a MetRS descritas en el presente documento. Se incluyen vectores que expresan ARNip u otras moléculas de interferencia de ARN formadoras de dúplex.

Un sistema de vector o construcción de ácido nucleico puede comprender un único vector o plásmido, dos o más vectores o plásmidos, que juntos contienen el ADN total para introducir en el genoma de la célula hospedadora, o un transposón. La elección del vector dependerá normalmente de la compatibilidad del vector con la célula hospedadora en la que va a introducirse el vector. En el presente caso, el vector o la construcción de ácido nucleico es preferentemente uno que es operativamente funcional en una célula de mamífero, tal como una célula muscular. El vector también puede incluir también un marcador de selección tal como un gen de resistencia a antibiótico o fármaco, o un gen indicador (es decir, proteína verde fluorescente, luciferasa), que puede usarse para selección o identificación de transformantes o transfectantes adecuados. Los sistemas de suministro a modo de ejemplo pueden incluir sistemas de vectores víricos (es decir, transducción mediada por virus) incluyendo, pero sin limitación,

vectores retrovíricos (por ejemplo, lentivíricos), vectores adenovíricos, vectores víricos adenoasociados y vectores víricos del herpes, entre otros conocidos en la técnica.

XI. DESCUBRIMIENTO DE FÁRMACOS

Ciertas realizaciones se relacionan con el uso de polipéptidos de MetRS, anticuerpos o polinucleótidos en el descubrimiento de fármacos, normalmente para identificar agentes que modulan una o más de las actividades no canónicas del polipéptido de MetRS de referencia, por ejemplo, el fragmento proteico de MetRS. Por ejemplo, ciertas realizaciones incluyen métodos para identificación de uno o más "compañeros de unión celular" de un polipéptido de referencia de MetRS, tal como una proteína celular, un lípido, un ácido nucleico u otra molécula hospedadora que interacciona directamente o físicamente con el polipéptido de MetRS. Los ejemplos particulares incluyen por ejemplo receptores de superficie celular, tales como GPCR, dominios de interacción proteína-proteína, y dominios extracelulares o intracelulares de los mismos.

También se incluyen métodos para identificar moléculas hospedadoras que participan en una o más actividades no canónicas del polipéptido de MetRS, incluyendo moléculas que interaccionan directa o indirectamente con el compañero de unión celular y regulan su papel en una actividad no canónica, o están reguladas por el compañero de unión. Dichas moléculas hospedadoras incluyen componentes tanto corriente arriba como corriente abajo de la ruta no canónica, normalmente relacionados con aproximadamente 1, 2, 3, 4, 5 o más etapas identificables en la ruta, en relación con la interacción de compañero de unión celular/proteína MetRS.

Ciertos aspectos incluyen métodos de identificación de un compuesto (por ejemplo, polipéptido) u otro agente que es agonista de o antagoniza la actividad no canónica de un polipéptido de referencia de MetRS o variante activa del mismo, tal como interaccionando con el polipéptido de MetRS y/o uno o más de sus compañeros de unión celular. También se incluyen métodos para identificar agentes que modulan la expresión (por ejemplo, corte y empalme) de variantes de corte y empalme de MetRS, o modulan la actividad de proteasas que de otro modo regulan la producción de fragmentos proteicos de MetRS endógenos (resectinas) en las proteínas.

Ciertas realizaciones incluyen, por lo tanto, métodos para identificar un compañero de unión de un polipéptido de referencia de MetRS, que comprenden: a) combinar el polipéptido de MetRS con una muestra biológica en condiciones adecuadas, y b) detectar la unión específica del polipéptido de MetRS con un compañero de unión, identificando de este modo un compañero de unión que se une específicamente con el polipéptido de referencia de MetRS. También se incluyen métodos para explorar un compuesto que se une específicamente con un polipéptido de referencia de MetRS o un compañero de unión del polipéptido de MetRS, que comprenden a) combinar el polipéptido o el compañero de unión a 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 este modo un compuesto que se une específicamente con el polipéptido o su compañero de unión. En ciertas realizaciones, el compuesto es un polipéptido o péptido. En ciertas realizaciones, el compuesto es una molécula pequeña u otro compuesto químico (por ejemplo, no biológico). En ciertas realizaciones, el compuesto es un peptidomimético.

Puede emplearse cualquier método adecuado para detectar interacciones proteína-proteína para identificar proteínas celulares que interaccionan con un polipéptido de referencia de MetRS, interaccionan con uno o más de sus compañeros de unión celular, o ambos. Los ejemplos de métodos tradicionales que pueden emplearse incluyen co-inmunoprecipitación, reticulación y copurificación mediante gradientes o columnas cromatográficas de lisados celulares o proteínas obtenidas de lisados celulares, principalmente para identificar proteínas en el lisado que interaccionan con el polipéptido de MetRS.

En estas realizaciones y otras relacionadas, al menos una parte de la secuencia de aminoácidos de una proteína que interacciona con un polipéptido de MetRS o su compañero de unión puede determinarse usando técnicas bien conocidas por los expertos en la materia, tal como mediante la técnica de degradación de Edman. Véase, por ejemplo, Creighton Proteins: Structures and Molecular Principles, W. H. Freeman & Co., N.Y., págs. 34 49, 1983. La secuencia de aminoácidos obtenida puede usarse como una guía para la generación de mezclas de oligonucleótidos que pueden usarse para explorar con respecto a secuencias génicas que codifican dichas proteínas. El cribado puede conseguirse, por ejemplo, mediante hibridación convencional o técnicas de PCR, como se describe en el presente documento y se conoce en la técnica. Se conocen bien técnicas para la generación de mezclas de oligonucleótidos y el cribado. Véase, por ejemplo, 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, pueden emplearse métodos en la identificación simultánea de genes que codifican el compañero de unión u otro polipéptido. Estos métodos incluyen, por ejemplo, cribado de bibliotecas de expresión, de una manera similar a la técnica bien conocida de cribado de anticuerpos de bibliotecas lambda-gt11, usando proteína de MetRS marcada, u otro polipéptido, péptido o proteína de fusión, por ejemplo, un polipéptido de MetRS variante o dominio de MetRS fusionado con un marcador (por ejemplo, una enzima, flúor, proteína luminiscente o colorante), o un dominio Ig-Fc.

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

Brevemente, utilizando dicho sistema, pueden construirse 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 con una secuencia de nucleótidos de referencia de MetRS (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 con un ADNc (o colección de ADNc) que codifica una proteína o proteínas desconocidas que se han recombinado en el plásmido como parte de una biblioteca de ADNc. El plásmido de fusión de dominio de unión a ADN y la biblioteca de ADNc activador pueden transformarse en una cepa de la levadura *Saccharomyces cerevisiae* que contiene un gen indicador (por ejemplo, HBS o lacZ) cuya región reguladora contiene el sitio de unión del activador de la transcripción. Una de las proteínas híbridas sola no puede activar la transcripción del gen indicador: el híbrido de dominio de unión ADN no puede porque no proporciona función de activación y el híbrido del dominio de activación no puede porque no puede localizarse en 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 indicador, que se detecta por un ensayo para el producto de gen indicador.

El sistema de dos híbridos u otra metodología tal pueden usarse para explorar bibliotecas de dominio de activación con respecto a proteínas que interaccionan con el producto génico "cebo". Como ejemplo, y no como limitación, un polipéptido de referencia de MetRS o variante puede usarse como el producto génico cebo. Un compañero de unión de MetRS también puede usarse como un producto génico "cebo". Se fusionan secuencias de ADNc o genómicas totales con el ADN que codifica un dominio de activación. Esta biblioteca y un plásmido que codifica un híbrido de un producto génico de MetRS cebo fusionado con el dominio de unión a ADN se cotransforman en una cepa indicadora de levadura, y los transformantes resultantes se exploran con respecto a los que expresan el gen indicador.

Una biblioteca de ADNc de la línea celular de la que van a detectarse proteínas que interaccionan con productos génicos de MetRS cebo puede prepararse usando métodos practicados rutinariamente en la técnica. Por ejemplo, los fragmentos de ADNc pueden insertarse en un vector de modo que se fusionen de forma traduccional con el dominio de activación transcripcional de GAL4. Esta biblioteca puede co-transformarse junto con el plásmido de fusión de GAL4-gen cebo en una cepa de levadura, que contiene un gen de lacZ conducido por un promotor que contiene secuencia de activación de GAL4. Una proteína codificada por ADNc, fusionada con un dominio de activación transcripcional de GAL4, que interacciona con el producto génico cebo, reconstituirá una proteína GAL4 activa y de este modo conducirá la expresión del gen de HIS3. Pueden detectarse colonias, que expresan HIS3, por su cultivo en placas de Petri que contienen medios basados en agar semisólidos que carecen de histidina. El ADNc puede después purificarse a partir de estas cepas, y usarse para producir y aislar la proteína que interacciona con el gen de MetRS cebo usando técnicas practicadas de forma rutinaria en este campo.

También se incluyen sistemas de tres híbridos, que permiten la detección de interacciones de ARN-proteína en levadura. Véase, por ejemplo, Hook *et al.*, RNA. 11:227-233, 2005. En consecuencia, estos métodos y otros relacionados pueden usarse para identificar un compañero de unión celular de un polipéptido de MetRS, y para identificar otras proteínas o ácidos nucleicos que interaccionan con el polipéptido de MetRS, el compañero de unión celular o ambos.

Ciertas realizaciones se refieren al uso de enfoques de cribado de interactoma. Los ejemplos particulares incluyen cribado basada en dominio proteico (véase, por ejemplo, Boxem *et al.*, Cell. 134:534-545, 2008 y Yu *et al.*, Science 322:10-110, 2008).

Como se ha indicado anteriormente, una vez aislados, pueden identificarse compañeros de unión y pueden, a su vez, usarse junto con técnicas convencionales para identificar proteínas u otros compuestos con los que interacciona. Ciertas realizaciones se refieren por lo tanto a métodos de cribado de un compuesto que se une específicamente con el compañero de unión de un polipéptido de referencia de MetRS, que comprenden a) combinar el compañero de unión a al menos un compuesto de ensayo en condiciones adecuadas y b) detectar la unión del compañero de unión al compuesto de ensayo, identificando de este modo un compuesto que se une específicamente con el compañero de unión. En ciertas realizaciones, el compuesto de ensayo es un polipéptido. En ciertas realizaciones, el compuesto de ensayo es un compuesto químico, tal como un compuesto de molécula pequeña o peptidomimético.

Ciertas realizaciones incluyen métodos de cribado para un compuesto que modula la actividad de un polipéptido de referencia de MetRS, que comprende 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 el que un cambio en la actividad del polipéptido en presencia del compuesto de ensayo es indicativo de un compuesto que modula la actividad del polipéptido. Ciertas realizaciones incluyen métodos de cribado para un compuesto que modula la actividad de un compañero de

unión de un polipéptido de referencia de MetRS, que comprende: 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 el que un cambio en la actividad del compañero de unión en presencia del compuesto de ensayo es indicativo de un compuesto que modula la actividad del compañero de unión. normalmente, estas realizaciones y otras relacionadas incluyen evaluar una actividad no canónica seleccionada que se asocia con el polipéptido de MetRS o su compañero de unión. Se incluyen condiciones *in vitro* e *in vivo*, tales como condiciones de cultivo celular.

Ciertas realizaciones incluyen métodos de cribado de un compuesto con respecto a eficacia como un agonista completo o parcial de un polipéptido de referencia de MetRS o un fragmento activo o variante del mismo, que comprende: a) exponer una muestra que comprende el polipéptido a un compuesto y b) detectar la actividad agonista en la muestra, normalmente midiendo un aumento en la actividad no canónica del polipéptido de MetRS. Ciertos métodos incluyen a) exponer una muestra que comprende un compañero de unión del polipéptido de MetRS a un compuesto y b) detectar la actividad agonista en la muestra, normalmente midiendo un aumento en la actividad no canónica seleccionada del polipéptido de MetRS. Ciertas realizaciones incluyen composiciones que comprenden un compuesto agonista identificado por el método y un vehículo o excipiente farmacéuticamente aceptable.

También se incluyen métodos de cribado de un compuesto con respecto a eficacia como un antagonista completo o parcial de un polipéptido de referencia de MetRS, que comprenden a) exponer una muestra que comprende el polipéptido a un compuesto y b) detectar la actividad antagonista en la muestra, normalmente midiendo una reducción en la actividad no canónica del polipéptido de MetRS. Ciertos métodos incluyen a) exponer una muestra que comprende un compañero de unión del polipéptido de MetRS a un compuesto y b) detectar la actividad antagonista en la muestra, normalmente midiendo una reducción en la actividad no canónica seleccionada del polipéptido de MetRS. Ciertas realizaciones incluyen composiciones que comprenden un compuesto antagonista identificado por el método y un vehículo o excipiente farmacéuticamente aceptable.

En ciertas realizaciones, pueden diseñarse sistemas *in vitro* para identificar compuestos capaces de interaccionar con o modular una secuencia de referencia de MetRS o su compañero de unión. Algunos de los compuestos identificados por dichos sistemas pueden ser útiles, por ejemplo, en la modulación de la actividad de la ruta y en la elaboración de componentes de la ruta de sí misma. También pueden usarse en exploraciones para identificar compuestos que alteran las interacciones entre componentes de la ruta; o pueden alterar dichas interacciones directamente. Un enfoque ejemplar implica preparar una mezcla de reacción del polipéptido de MetRS y un compuesto de ensayo en condiciones y durante un tiempo suficiente para permitir que los dos interaccionen y se unan, formando de este modo un complejo que puede retirarse de y/o detectarse en la mezcla de reacción.

Pueden realizarse ensayos de cribado *in vitro* de diversas maneras. Por ejemplo, un polipéptido de MetRS, un compañero de unión celular, o compuesto o compuestos de ensayo puede anclarse en una fase sólida. En estas realizaciones y otras relacionadas, los complejos resultantes pueden capturarse y detectarse en la fase sólida al final de la reacción. En un ejemplo de dicho método, el polipéptido de MetRS y/o su compañero de unión se anclan en una superficie sólida, y el compuesto o los compuestos de ensayo, que no están anclados, pueden marcarse, directa o indirectamente, de modo que su captura por el componente en la superficie sólida pueda detectarse. En otros ejemplos, el compuesto o los compuestos de ensayo se anclan en la superficie sólida, y el polipéptido de MetRS y/o su compañero de unión, que no están anclados, se marcan o son de algún modo detectables. En ciertas realizaciones, pueden utilizar convenientemente placas de microtitulación como la fase sólida. El componente anclado (o compuesto de ensayo) puede inmovilizarse por uniones no covalentes o covalentes. Puede conseguirse unión no covalente simplemente recubriendo la superficie sólida con una solución de la proteína y secando. Como alternativa, un anticuerpo inmovilizado, preferentemente un anticuerpo monoclonal, específico para la proteína para inmovilizar puede usarse para anclar la proteína a la superficie sólida. Las superficies pueden prepararse previamente y almacenarse.

Para realizar un ensayo ejemplar, el componente no inmovilizado se añade normalmente a la superficie recubierta que contiene el componente anclado. Después de completarse la reacción, se retiran los componentes que no han reaccionado (por ejemplo, lavando) en condiciones tales que cualquier complejo específico formado permanezca inmovilizado en la superficie sólida. La detección de complejos anclados a la superficie sólida puede conseguirse de diversas maneras. Por ejemplo, cuando el componente no inmovilizado previamente está premarcado, la detección de marcador inmovilizado en la superficie indica que se formaron complejos. Cuando el componente no inmovilizado previamente no está premarcado, puede usarse un marcador indirecto para detectar complejos anclados en la superficie; por ejemplo, usando un anticuerpo marcado específico para el componente no inmovilizado previamente (el anticuerpo, a su vez, puede marcarse directamente o marcarse indirectamente con un anticuerpo anti-Ig marcado).

Como alternativa, la presencia o ausencia de unión de un compuesto de ensayo puede determinarse, por ejemplo, usando resonancia de plasmón superficial (RPS) y el cambio en el ángulo de resonancia como un índice, en el que un polipéptido de MetRS o un compañero de unión celular está inmovilizado en la superficie de una microplaca

sensora disponible en el mercado (por ejemplo, fabricada por BIACORE™) de acuerdo con un método convencional, el compuesto de ensayo está en contacto con la misma y la microplaca sensora se ilumina con una luz de una longitud de onda particular desde un ángulo particular. La unión de un compuesto de ensayo también puede medirse detectando la aparición de un pico correspondiente al compuesto de ensayo por un método en el que un polipéptido de MetRS o un compañero de unión celular se inmoviliza en la superficie de una microplaca proteica adaptable a un espectrómetro de masas, un compuesto de ensayo se pone en contacto con la misma y se combina un método de ionización tal como MALDI-EM, ISE-EM, FAB-EM y similares con un espectrómetro de masas (por ejemplo, espectrómetro de masas de doble enfoque, espectrómetro de masas cuadripolo, espectrómetro de masas de tiempo de vuelo, espectrómetro de masas de transformación de Fourier, espectrómetro de masas de ciclotrón de iones y similares).

En ciertas realizaciones, pueden usarse ensayos basados en células, ensayos basados en vesículas de membrana o ensayos basados en fracción de membrana para identificar compuestos que modulan interacciones en la ruta no canónica del polipéptido de MetRS seleccionado. Para este fin, pueden usarse todas las líneas celulares que expresan un polipéptido de MetRS y/o un compañero de unión, o una proteína de fusión que contiene un dominio o fragmento de dichas proteínas (o una combinación de las mismas), o líneas celulares (por ejemplo, células COS, células CHO, células HEK293, células Hela, etc.) que se han modificado por ingeniería genética para expresar dichas proteína o proteínas o proteína de fusión o proteínas de fusión. Pueden identificarse compuesto o compuestos de ensayo que influyen en la actividad no canónica supervisando un cambio (por ejemplo, un cambio estadísticamente significativo) en esa actividad en comparación con un control o una cantidad predeterminada.

Para las realizaciones relacionadas con agentes antisentido y de ARNi, por ejemplo, también se incluyen métodos de cribado de un compuesto con respecto a eficacia en la alteración de la expresión de un polinucleótido de referencia de MetRS, que comprenden: a) exponer una muestra que comprende el polinucleótido de referencia de MetRS a un compuesto tal como un oligonucleótido antisentido potencial, y b) detectar la expresión alterada del polinucleótido de MetRS. En ciertos ejemplos no limitantes, estas realizaciones y otras relacionadas pueden emplearse en ensayos basados en células o en ensayos de traducción sin células, de acuerdo con técnicas rutinarias en este campo. También se incluyen los agentes antisentido y de ARNi identificados por dichos métodos.

Pueden usarse también anticuerpos para fragmentos proteicos de MetRS en ensayos de cribado, tal como para identificar un agente que se une específicamente con una MetRS, confirmar la especificidad o afinidad de un agente que se une a un fragmento proteico de MetRS, o identificar el sitio de interacción entre el agente y el fragmento proteico de MetRS. Se incluyen ensayos en los que el anticuerpo se usa como un inhibidor competitivo del agente. Por ejemplo, un anticuerpo que se une específicamente con un fragmento proteico de MetRS con una afinidad conocida puede actuar como un inhibidor competitivo de un agente seleccionado y puede usarse para calcular la afinidad del agente por el fragmento proteico de MetRS. Además, uno o más anticuerpos que se unen específicamente con epítopos o sitios conocidos de un fragmento proteico de MetRS pueden usarse como un inhibidor competitivo para confirmar si el agente se une o no en ese mismo sitio. Otras variaciones resultarán evidentes para los expertos en la materia.

También se incluyen cualquiera de los métodos anteriores, u otros métodos de cribado conocidos en la técnica, que se adaptan para el cribado de alto rendimiento (HTS). La HTS usa normalmente automatización para procesar un cribado del ensayo frente a una biblioteca de compuestos candidatos, por ejemplo, un ensayo que mide un aumento o una reducción en una actividad no canónica, como se describe en el presente documento.

Cualquiera de los métodos de cribado proporcionados en el presente documento puede utilizar bibliotecas de moléculas pequeñas o bibliotecas generadas por química combinatoria. Se conocen en la técnica bibliotecas de mezclas químicas y/o biológicas, tales como extractos fúngicos, bacterianos o de algas y pueden explorarse con cualquiera de los ensayos de la invención. Pueden encontrarse ejemplos para la síntesis de bibliotecas moleculares en: (Carell *et al.*, 1994a, Carell *et al.*, 1994b, Cho *et al.*, 1993, DeWitt *et al.*, 1993, Gallop *et al.*, 1994; Zuckermann *et al.*, 1994).

Pueden presentarse bibliotecas de compuestos en solución (Houghten *et al.*, 1992) o en perlas (Lam *et al.*, 1991), en microplacas (Fodor *et al.*, 1993), bacterias, esporas (Ladner *et al.,* Patente de Estados Unidos N.º 5.223.409, 1993), plásmidos (Cull *et al.*, 1992) o en fagos (Cwirla *et al.*, 1990, Devlin *et al.*, 1990, Felici *et al.*, 1991, Ladner *et al.* Patente de Estados Unidos N.º 5.223.409, 1993, Scott y Smith, 1990). Las realizaciones de la presente invención abarcan el uso de diferentes bibliotecas para la identificación de moduladores de molécula pequeña de uno o más fragmentos proteicos de MetRS, sus compañeros de unión celular y/o sus actividades no canónicas relacionadas. Las bibliotecas útiles para los fines de la invención incluyen, pero sin limitación, (1) bibliotecas químicas, (2) bibliotecas de productos naturales y (3) bibliotecas combinatorias comprendidas por péptidos aleatorios, oligonucleótidos y/o moléculas orgánicas.

Las bibliotecas químicas consisten en análogos estructurales de compuestos conocidos o compuestos que se identifican como “aciertos” o “candidatos” mediante cribado de productos naturales. Se derivan bibliotecas de productos naturales a partir de colecciones de microorganismos, animales, plantas u organismos marinos que se usan para crear mezclas para explorar mediante: (1) fermentación y extracción de caldos de microorganismos del

suelo, vegetales o marinos o (2) extracción de plantas u organismos marinos. Las bibliotecas de productos naturales incluyen poliquétidos, péptidos no ribosómicos y variantes (de origen no natural) de los mismos. Véase, por ejemplo, Cane *et al.*, Science 282:63-68, 1998. Las bibliotecas combinatorias pueden estar compuestas de grandes números de péptidos, oligonucleótidos o compuestos orgánicos como una mezcla. Son relativamente fáciles de preparar mediante métodos de síntesis automática tradicionales, PCR, clonación o métodos de síntesis patentados.

Más específicamente, una biblioteca química combinatoria es una colección de diversos compuestos químicos generados mediante síntesis química o síntesis biológica, combinando varios "componentes básicos" químicos tales como reactivos. Por ejemplo, se forma una biblioteca química combinatoria lineal tal como una biblioteca de polipéptidos combinando un conjunto de componentes básicos químicos (aminoácidos) de todas las maneras posibles para una longitud de compuesto dada (es decir, el número de aminoácidos en un compuesto polipeptídico). Pueden sintetizarse millones de compuestos químicos mediante dicha mezcla combinatoria de componentes básicos químicos.

Para una revisión de química combinatoria y bibliotecas creadas a partir de la misma, véase, por ejemplo, Huc, I. y Nguyen, R. (2001) Comb. Chem. High Throughput Screen 4:53-74; Lepre, C A. (2001) Drug Discov. Hoy 6:133-140; Peng, S. X. (2000) Biomed. Chromatogr. 14:430-441; Bohm, H. J. y Stahl, M. (2000) Curr. Opin. Chem. Biol. 4:283-286; Barnes, C y Balasubramanian, S. (2000) Curr. Opin. Chem. Biol. 4:346-350; Lepre, Enjalbal, C, *et al.*, (2000) Mass Septrom Rev. 19:139-161; Hall, D. G., (2000) Nat. Biotechnol. 18:262-262; Lazo, J. S., y Wipf, P. (2000) J. Pharmacol. Exp. Ther. 293:705-709; Houghten, R. A., (2000) Ann. Rev. Pharmacol. Toxicol. 40:273-282; Kobayashi, S. (2000) Curr. Opin. Chem. Biol. (2000) 4:338-345; Kopylov, A. M. y Spiridonova, V. A. (2000) Mol. Biol. (Mosk) 34:1097-1113; Weber, L. (2000) Curr. Opin. Chem. Biol. 4:295-302; Dolle, R. E. (2000) J. Comb. Chem. 2:383-433; Floyd, C D., *et al.*, (1999) Prog. Medicina. Chem. 36:91-168; Kundu, B., *et al.*, (1999) Prog. Drug Res. 53:89-156; Cabilly, S. (1999) Mol. Biotechnol. 12:143-148; Lowe, G. (1999) Nat. Pinchar. Rep. 16:641-651; Dolle, R.E. y Nelson, K. H. (1999) J. Comb. Chem. 1:235-282; Czarnick, A. W. y Keene, J. D. (1998) Curr. Biol. 8:R705-R707; Dolle, R. E. (1998) Mol. Divers. 4:233-256; Myers, P.L., (1997) Curr. Opin. Biotechnol. 8:701-707; y Pluckthun, A. y Cortese, R. (1997) Biol. Chem. 378:443.

Están disponibles en el mercado dispositivos para la preparación de bibliotecas combinatorias (véase, por ejemplo, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville Ky., Symphony, Rainin, Woburn, Massachusetts, 433A Applied Biosystems, Foster City, Calif., 9050 Plus, Millipore, Bedford, Mass.). Además, numerosas bibliotecas combinatorias están disponibles en sí mismas en el mercado (véase, por ejemplo, ComGenex, Princeton, NJ, Asinex, Moscú, Ru, Tripos, Inc., St. Louis, Mo. ChemStar, Ltd., Moscú, RU, 3D Pharmaceuticals, Exton, Pa., Martek Biosciences, Columbia, MD, etc.).

XII. MÉTODOS DE USO

Las realizaciones de la presente invención incluyen métodos terapéuticos de tratamiento. En consecuencia, los agentes de MetRS descritos en el presente documento, incluyendo polipéptidos de MetRS, polinucleótidos de MetRS, vectores basados en polinucleótidos de MetRS, células hospedadoras que expresan MetRS, oligonucleótidos antisentido, agentes de ARNi, así como agentes de unión tales como péptidos, anticuerpos y fragmentos de unión a antígeno, peptidomiméticos y otras moléculas pequeñas, pueden usarse para tratar diversas enfermedades o afecciones no limitantes asociadas con las actividades no canónicas de una MetRS de referencia. Los ejemplos de dichas actividades no canónicas incluyen modulación de la señalización extracelular, modulación de la proliferación celular, modulación de la migración celular, modulación de la diferenciación celular (por ejemplo, hematopoyesis, neurogénesis, miogénesis, osteogénesis y adipogénesis), modulación de la apoptosis u otras formas de muerte celular, modulación de la angiogénesis, modulación de la unión celular, modulación del metabolismo celular, modulación de la producción o actividad de citocinas, modulación de la actividad de receptores de citocinas, modulación de la captación celular, o secreción, inmunomodulación, modulación de la inflamación, modulación de los procesos metabólicos tales como control de glucosa y similares.

Se incluyen terapias basadas en polinucleótidos, tales como terapias antisentido y terapias de interferencia de ARNi, que normalmente están relacionadas con la reducción de la expresión de una molécula diana, tal como un fragmento endógeno de una MetRS o un compañero de unión celular de un polipéptido de MetRS, que contribuye de otro modo a su actividad no canónica. Las terapias antisentido o de ARNi normalmente antagonizan la actividad no canónica, tal como reduciendo la expresión del polipéptido de referencia de MetRS. También se incluyen polipéptidos o péptidos, anticuerpos o fragmento de unión a antígeno, peptidomiméticos u otras terapias basadas en moléculas pequeñas, que son agonistas de o antagonizan la actividad canónica de un polipéptido de referencia de MetRS, tal como interaccionando directamente con el polipéptido de MetRS, su compañero o sus compañeros de unión celular, o ambos.

Estas realizaciones y otras relacionadas incluyen métodos para usar los agentes de MetRS o composiciones de la presente invención para tratar una célula, un tejido o un sujeto. Las células o los tejidos que pueden tratarse o modularse por la presente invención son preferentemente células o tejidos de mamíferos, o más preferentemente células o tejidos humanos. Dichas células o tejidos pueden ser de un estado sano o de un estado enfermo.

En ciertas realizaciones, por ejemplo, se proporcionan métodos para modular actividades celulares terapéuticamente relevantes incluyendo, pero sin limitación, metabolismo celular, diferenciación celular, proliferación celular, captación celular, secreción celular, muerte celular, movilización celular, migración celular, transcripción génica, traducción de ARNm, impedancia celular, respuestas inmunitarias, respuestas inflamatorias, y similares, que comprenden poner en contacto una célula con un agente de MetRS o composición como se describe en el presente documento. En ciertas realizaciones, la célula está en un sujeto. En consecuencia, las composiciones de MetRS pueden emplearse en tratamiento esencialmente de cualquier célula o tejido o sujeto que se beneficiaría de la modulación de una o más de dichas actividades.

Los agentes de MetRS y composiciones también pueden usarse en cualquiera de varios contextos terapéuticos incluyendo, por ejemplo, los relacionados con el tratamiento o prevención de enfermedades neoplásicas, enfermedades o afecciones del sistema inmunitario (por ejemplo, enfermedades autoinmunitarias e inflamación), enfermedades infecciosas, enfermedades metabólicas, enfermedades neuronales/neurológicas, enfermedades musculares/cardiovasculares, enfermedades asociadas con hematopoyesis aberrante, enfermedades asociadas con miogénesis aberrante, enfermedades asociadas con neurogénesis aberrante, enfermedades asociadas con adipogénesis aberrante, enfermedades asociadas con osteogénesis aberrante, enfermedades asociadas con angiogénesis aberrante, enfermedades asociadas con supervivencia celular aberrante, enfermedades asociadas con captación de lípidos aberrante, enfermedades asociadas con envejecimiento (por ejemplo, pérdida de audición, neuropatías periféricas o autónomas, demencia senil, retinopatía) y otras.

Por ejemplo, en ciertas realizaciones ilustrativas, las composiciones de MetRS de la invención pueden usarse para modular la angiogénesis, por ejemplo, mediante modulación de la proliferación y/o señalización de células endoteliales. La proliferación y/o señalización de células endoteliales puede supervisarse usando una línea celular apropiada (por ejemplo, células de pulmón endoteliales microvasculares humanas (HMVEC-L) y células endoteliales de vena umbilical humana (HUVEC)) y usando un ensayo apropiado (por ejemplo, ensayos de migración de células endoteliales, ensayos de proliferación de células endoteliales, ensayos de formación de tubos, ensayos de tapón de matrigel, etc.), muchos de los cuales se conocen y están disponibles en la técnica.

Por lo tanto, en realizaciones relacionadas, las composiciones de la invención pueden emplearse en el tratamiento de esencialmente cualquier célula o tejido o sujeto que se beneficiaría de la modulación de la angiogénesis. Por ejemplo, en algunas realizaciones, una célula o tejido o sujeto que experimenta o es susceptible a angiogénesis (por ejemplo, una afección angiogénica) puede ponerse en contacto con una composición adecuada de la invención para inhibir una afección angiogénica. En otras realizaciones, una célula o un tejido que experimenta o es susceptible a angiogénesis insuficiente (por ejemplo, una afección angiostática) puede ponerse en contacto con una composición apropiada de la invención para interferir con la actividad angiostática y/o promover la angiogénesis.

También se incluyen métodos para modular la hematopoyesis y afecciones relacionadas. Los ejemplos de procesos hematopoyéticos que pueden modularse por los polipéptidos de MetRS de la invención incluyen, sin limitación, la formación de células mieloides (por ejemplo, células eritroides, mastocitos, monocitos/macrófagos, células dendríticas mieloides, granulocitos tales como basófilos, neutrófilos y eosinófilos, megacariocitos, plaquetas) y células linfoides (por ejemplo, linfocitos citolíticos naturales, células dendríticas linfoides, linfocitos B y linfocitos T). Ciertos procesos hematopoyéticos específicos incluyen eritropoyesis, granulopoyesis, linfopoyesis, megacariopoyesis, trombopoyesis y otros. También se incluyen métodos de modulación del tráfico o la movilización de células hematopoyéticas, incluyendo células madre hematopoyéticas, células progenitoras, eritrocitos, granulocitos, linfocitos, megacariocitos y trombocitos.

Los métodos de modulación de la hematopoyesis pueden practicarse *in vivo*, *in vitro*, *ex vivo* o en cualquier combinación de los mismos. Estos métodos pueden practicarse en cualquier muestra biológica, cultivo celular o tejido que contenga células madre hematopoyéticas, células progenitoras hematopoyéticas u otras células madre o progenitoras que sean capaces de diferenciar a lo largo del linaje hematopoyético (por ejemplo, células madre derivadas de tejido adiposo). Para métodos *in vitro* y *ex vivo*, pueden aislarse células madre y células progenitoras, bien de origen hematopoyético o de otro modo, y/o identificarse de acuerdo con las técnicas y características descritas en el presente documento y conocidas en este campo.

Las composiciones de la invención también pueden ser útiles como inmunomoduladores para tratar indicaciones anti o proinflamatorias modulando las células que median, directa o indirectamente, en enfermedades, afecciones y trastornos autoinmunitarios y/o inflamatorios. La utilidad de las composiciones de la invención como inmunomoduladores o moduladores de inflamación puede supervisarse usando cualquiera de varias técnicas conocidas y disponibles en este campo incluyendo, por ejemplo, ensayos de migración (por ejemplo, usando leucocitos o linfocitos) o ensayos de viabilidad celular (por ejemplo, usando linfocitos B, linfocitos T, monocitos o linfocitos NK).

La "inflamación" se refiere en general a la respuesta biológica de tejidos a estímulos perjudiciales, tales como patógenos, células dañadas (por ejemplo, heridas) e irritantes. La expresión "respuesta inflamatoria" se refiere a los mecanismos específicos por los que se consigue y se regula inflamación, incluyendo, únicamente como ilustración, activación o migración de células inmunitarias, producción de citocinas, vasodilatación, incluyendo liberación de

quinasa, fibrinolisis y coagulación, entre otros descritos en el presente documento y conocidos en la técnica.

Las señales clínicas de 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-8ª Ed., 2009 Elsevier, Londres, Miller, LM, Pathology Lecture Notes, Atlantic Veterinary College, Charlottetown, PEI, Canadá). La inflamación crónica se asocia con diversas 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, insuficiencia cardiaca congestiva, fibromialgia, fibrosis, ataque cardiaco, insuficiencia renal, lupus, pancreatitis, ictus, complicaciones quirúrgicas, enfermedad inflamatoria del pulmón, enfermedad inflamatoria del intestino, aterosclerosis, trastornos neurológicos, diabetes, trastornos metabólicos, obesidad y psoriasis, entre otras descritas en el presente documento y conocidas en la técnica . Por lo tanto, las composiciones de MetRS pueden usarse para tratar o controlar inflamación crónica, modular cualquiera de una o más de las respuestas inflamatorias crónicas individuales, o tratar una cualquiera o más enfermedades o afecciones asociadas con inflamación crónica.

Los criterios para evaluar las señales y síntomas de afecciones inflamatorias y otras, incluyendo para fines de realización de diagnóstico diferencial y también para supervisar los tratamientos tales como determinar si se ha administrado en el transcurso del tratamiento una dosis terapéuticamente eficaz, por ejemplo, determinando la mejora de acuerdo con criterios clínicos aceptados, resultarán evidentes para los expertos en la materia y se ejemplifican por las enseñanzas de, por ejemplo, Berkow *et al*., Eds., The Merck Manual, 16ª edición, Merck and Co., Rahway, NJ, 1992; Goodman *et al.*, Eds., Goodman y Gilman, The Pharmacological Basis of Therapeutics, 10ª edición, Pergamon Press, Inc., Elmsford, N.Y., (2001); Avery's Drug Treatment: Principios y Práctica de Farmacología Clínica y Terapéutica, 3ª edición, ADIS Press, Ltd., Williams y Wilkins, Baltimore, MD. (1987); Ebadi, Pharmacology, Little, Brown and 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).

En otras realizaciones, las composiciones de MetRS de la invención pueden usarse para modular la proliferación y/o supervivencia celular y, en consecuencia, para tratar o prevenir enfermedades, trastornos o afecciones caracterizados por anomalías en la proliferación y/o supervivencia celular. Por ejemplo, en ciertas realizaciones, las composiciones de MetRS pueden usarse para modular la apoptosis y/o para tratar enfermedades o afecciones asociadas con apoptosis anómala. La apoptosis puede supervisarse por cualquiera de varias técnicas disponibles conocidas y disponibles en este campo incluyendo, por ejemplo, ensayos que miden la fragmentación de ADN, alteraciones en la asimetría de membrana, activación de caspasas apoptóticas y/o liberación de citocromo C y AIF.

El progreso de estas y otras terapias (por ejemplo, terapias *ex vivo*) puede supervisarse fácilmente por métodos y ensayos convencionales y basándose en criterios conocidos por los médicos u otros expertos en la materia.

XIII. FORMULACIONES FARMACÉUTICAS, ADMINISTRACIÓN Y KITS

Las realizaciones de la presente invención incluyen polinucleótidos de MetRS, polipéptidos de MetRS, células hospedadoras que expresan polipéptidos de MetRS, agentes de unión, agentes moduladores u otros compuestos descritos en el presente documento, formulados en soluciones farmacéuticamente aceptables o fisiológicamente aceptables para administración a una célula o un animal, bien solos, o bien en combinación con una o más modalidades adicionales de terapia. También se entenderá que, si se desea, las composiciones de la invención pueden administrarse también en combinación con otros agentes, tales como, por ejemplo, otras proteínas u otros polipéptidos o diversos agentes farmacéuticamente activos. No hay prácticamente ningún límite a los otros componentes que también pueden incluirse en las composiciones, siempre que los agentes adicionales no afecten de forma adversa a los efectos moduladores u otros que se deseen conseguir.

En las composiciones farmacéuticas de la invención, los expertos en la materia conocen bien la formulación de excipientes y soluciones de vehículo farmacéuticamente aceptables, así como el desarrollo de regímenes de dosificación y tratamiento adecuados para usar las composiciones particulares descritas en el presente documento en diversos regímenes de tratamiento, incluyendo, por ejemplo, administración oral, parenteral, intravenosa, intranasal, subcutánea e intramuscular y formulación.

En ciertas aplicaciones, las composiciones farmacéuticas o terapéuticas de la invención no estimulan una reacción inmunitaria. En otras realizaciones, las composiciones farmacéuticas o terapéuticas de la invención, que comprenden normalmente uno o más polipéptidos o polinucleótidos de MetRS, estimulan una reacción inmunitaria, tal como actuando como un adyuvante en una vacuna o composición relacionada, o estando presente en una composición junto con un adyuvante o agente separado estimula una respuesta inmunitaria.

En ciertas realizaciones, los agentes de MetRS tales como polipéptidos de MetRS, polinucleótidos de MetRS y anticuerpos tienen una solubilidad que es deseable para el modo particular de administración, tal como administración intravenosa. Los ejemplos de solubilidades deseables incluyen al menos aproximadamente 1 mg/ml, al menos aproximadamente 10 mg/ml, al menos aproximadamente 25 mg/ml y al menos aproximadamente 50

mg/ml.

En ciertas aplicaciones, las composiciones farmacéuticas desveladas en el presente documento pueden suministrarse 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.

En ciertas circunstancias será deseable suministrar las composiciones farmacéuticas desveladas en el presente documento por vía parenteral, por vía subcutánea, por vía intravenosa, por vía intramuscular, por vía intraarterial, por vía intrarrectal, por vía intraparenquimal, por vía intraventricular, por vía intrauretral, por vía intraesternal, por vía intracraneal, por vía intrasinovial, o incluso por vía intraperitoneal como se describe, por ejemplo, en la Patente de Estados Unidos N.º 5.543.158; Patente de Estados Unidos N.º 5.641.515 y la Patente de Estados Unidos 5.399.363. Los dispositivos adecuados para administración parenteral incluyen inyectores de aguja (incluyendo microagujas), inyectores sin agujas y técnicas de infusión.

Pueden prepararse soluciones de los compuestos activos como base libre o sales farmacológicamente aceptables en agua convenientemente mezclados con un tensioactivo, tal como hidroxipropilcelulosa. También pueden prepararse dispersiones en glicerol, polietilenglicoles líquidos, y mezclas de los mismos y en aceites. En condiciones habituales de almacenamiento y uso, estas preparaciones contienen un conservante para evitar el crecimiento de microorganismos.

Las formas farmacéuticas adecuadas para uso inyectable incluyen soluciones o dispersiones acuosas estériles y polvos estériles para la preparación extemporánea de soluciones o dispersiones inyectables estériles (Patente de Estados Unidos N.º 5.466.468). En todos los casos la forma debería ser estéril y debería ser fluida en la medida en que exista fácil inyectabilidad. Debería ser estable en las condiciones de fabricación y almacenamiento y debería conservarse frente a la acción contaminante de microorganismos, tales como bacterias y hongos. El vehículo puede ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol y polietilenglicol líquido, y similares), mezclas adecuadas de los mismos y/o aceites vegetales. Puede mantenerse la fluidez apropiada, por ejemplo, mediante el uso de un recubrimiento, tal como lecitina, mediante el mantenimiento del tamaño de partícula requerido en caso de dispersión y mediante el uso de tensioactivos. La prevención de la acción de microorganismos puede facilitarse por diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido sórbico, timerosal y similares. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares o cloruro sódico. Puede proporcionarse absorción prolongada de las composiciones inyectables mediante el uso en las composiciones de agentes que retardarán la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Para administración parenteral en una solución acuosa, por ejemplo, la solución debería tamponarse convenientemente si es necesario y el diluyente líquido hacerse isotónico primero con suficiente solución salina o glucosa. Estas soluciones acuosas particulares son especialmente adecuadas para administración intravenosa, intramuscular, subcutánea e intraperitoneal. En relación con esto, a la luz de la presente invención, los expertos en la materia conocerán un medio acuoso estéril que pueda emplearse. Por ejemplo, una dosificación puede disolverse en 1 ml de solución de NaCl isotónica y añadirse a 1.000 ml de líquido de hipodermoclisis o inyectarse en el sitio propuesto de infusión (véase, por ejemplo, Remington's Pharmaceutical Sciences, 15ª edición, págs. 1035-1038 y 1570-1580). Se producirá necesariamente algo de variación en la dosificación dependiendo de la condición del sujeto que se trata. La persona responsable de la administración determinará, en cualquier caso, la dosis apropiada para el sujeto individual. Además, para administración humana, las preparaciones deberían cumplir los criterios de esterilidad, pirogenicidad y seguridad y pureza generales que requiere la Oficina de Patrones Biológicos de la FDA.

Pueden prepararse soluciones inyectables estériles incorporando los compuestos activos en la cantidad requerida en el disolvente apropiado con los diversos otros ingredientes enumerados anteriormente, según se requiera, seguido de esterilización filtrada. En general, se preparan dispersiones incorporando los diversos principios activos esterilizados en un vehículo estéril que contiene el medio de dispersión básico y los otros ingredientes requeridos de los enumerados anteriormente. En el caso de polvos estériles para la preparación de soluciones inyectables estériles, los métodos preferidos de preparación son técnicas de secado al vacío y liofilización que producen un polvo del principio activo más cualquier ingrediente deseado adicional a partir de una solución esterilizada por filtración previamente del mismo.

Las composiciones desveladas en el presente documento pueden formularse en una forma neutra o salina. Las sales farmacéuticamente aceptables incluyen las sales de adición de ácidos (formadas con los tres 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 derivar de bases inorgánicas tales como, por ejemplo, sodio, potasio, amonio, calcio o hidróxidos férricos, y bases orgánicas tales como isopropilamina, trimetilamina, histidina, procaína y similares. Tras la formulación, las soluciones se administrarán de una manera compatible con la formulación de dosificación y en una cantidad tal que sea terapéuticamente eficaz. Las formulaciones se administran fácilmente en diversas formas de dosificación tales como soluciones inyectables, cápsulas de liberación de fármacos y similares.

Como se usa en el presente documento, "vehículo" incluye todos y cada uno de los disolventes, medios de dispersión, vehículos, recubrimientos, diluyentes, agentes antibacterianos y antifúngicos, agentes retardantes de la absorción e isotónicos, tampones, soluciones transportadoras, suspensiones, coloides y similares. El uso de dichos medios y agentes para sustancias farmacéuticas activas se conoce bien en la técnica. Excepto en la medida en que cualquier medio o agente convencional sea incompatible con el principio activo, se contempla su uso en las composiciones terapéuticas. También pueden incorporarse en las composiciones principios activos complementarios.

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

En ciertas realizaciones, las composiciones farmacéuticas pueden suministrarse por pulverizaciones intranasales, inhalación y/u otros vehículos de suministro de aerosol. Se han descrito métodos para suministrar genes, polinucleótidos y composiciones peptídicas directamente a los pulmones mediante pulverizaciones de aerosol nasal, por ejemplo, en la Patente de Estados Unidos N.º 5.756.353 y la Patente de Estados Unidos N.º 5.804.212. De forma similar, el suministro de fármacos usando resinas de micropartículas intranasales (Takenaga *et al*., 1998) y compuestos de lisofosfatidil-glicerol (Patente de Estados Unidos N.º 5.725.871) también se conocen bien en la técnica farmacéutica. De forma similar, se describe suministro farmacológico transmucoso en forma de una matriz de soporte de politetrafluoroetileno en la Patente de Estados Unidos N.º 5.780.045.

Las composiciones farmacéuticas pueden formularse para ser de liberación inmediata y/o sostenida. Las composiciones de liberación sostenida incluyen liberación retardada, modificada, por pulsos, controlada, dirigida y programada. Por lo tanto, las composiciones pueden formularse como una suspensión o como un sólido, semisólido o líquido tixotrópico para administración como un depósito implantado que proporciona liberación sostenida de los polinucleótidos de MetRS, polipéptidos de MetRS, agentes de unión, agentes moduladores y otros agentes activos. Los ejemplos de dichas formulaciones incluyen sin limitación estents recubiertos con fármaco y semisólidos de suspensiones que comprenden vesículas lamelares o micropartículas cargadas con fármaco de ácido poli(DL-láctico-co-glicólico) (PGLA), poli(DL-lactida-coglicólido) (PLG) o poli(lactida) (PLA), hidrogeles (Hoffman AS: Ann. N.Y. Acad. Sci.), sistemas de nanopartículas de poliaminoácido, comercializados con la marca comercial MEDUSA® desarrollada por Flamel Technologies Inc., sistemas de geles no acuosos comercializados con la marca comercial ATRIGEL® desarrollada por Atrix, Inc. y formulaciones de liberación extendida de acetato isoburitato de sacarosa comercializadas con el nombre comercial SABER® desarrollado por Durect Corporation y sistemas basados en lípidos desarrollados por SkyePharma y comercializados con la marca comercial DEPOFOAM®.

Se conocen en la técnica dispositivos de liberación sostenida capaces de suministrar dosis deseadas de las composiciones farmacéuticas durante periodos de tiempo prolongados. Por ejemplo, las Patentes de Estados Unidos N.º 5.034.229; 5.557.318; 5.110.596; 5.728.396; 5.985.305; 6.113.938; 6.156.331; 6.375,978; y 6.395.292; enseñan dispositivos conducidos de forma osmótica capaces de suministrar una formulación de agente activo, tal como una solución o una suspensión, a una velocidad deseada durante un periodo de tiempo prolongado (es decir, un periodo que varía de más de una semana hasta un año o más). Otros dispositivos de liberación sostenida a modo de ejemplo incluyen bombas de tipo regulador que proporcionan flujo constante, flujo ajustable o flujo programable de formulaciones de agentes beneficiosos, que están disponibles en Medtronic, incluyendo las bombas intratecales comercializadas con el nombre comercial SYNCHROMED INFUSION SYSTEM®, los sistemas de Johnson and Johnson comercializados con el nombre comercial bombas de división CODMAN®, y bombas de tecnologías INSET®. Se describen ejemplos adicionales de dispositivos en las Patentes de Estados Unidos N.º 6.283.949; 5.976.109; 5.836.935; y 5.511.355.

En ciertas realizaciones, el suministro puede producirse mediante el uso de liposomas, nanocápsulas, micropartículas, microesferas, partículas lipídicas, vesículas y similares, para la introducción de las composiciones de la presente invención en células hospedadoras adecuadas. En particular, las composiciones de la presente invención pueden formularse para el suministro, bien encapsuladas en una partícula lipídica, un liposoma, una vesícula, una nanoesfera, una nanopartícula o similares. La formulación y el uso de dichos vehículos de suministro pueden llevarse a cabo usando técnicas conocidas y convencionales.

En ciertas realizaciones, los agentes proporcionados en el presente documento pueden unirse a un sustrato sólido farmacéuticamente aceptable, incluyendo sustratos biocompatibles y biodegradables tales como polímeros y matrices. Los ejemplos de dichos sustratos sólidos incluyen, sin limitación, poliésteres, hidrogeles (por ejemplo, poli(2-hidroxietilmetacrilato) o poli(alcohol vinílico)), polilactidas (Patente de Estados Unidos N.º 3.773.919), copolímeros de ácido L-glutámico y γ-etil-L-glutamato, etilen-vinilacetato no degradable, copolímeros de ácido láctico-ácido glicólico degradables tales como poli(ácido láctico-co-glicólico) (PLGA) y el LUPRON DEPOT™ (microesferas inyectables compuestas de copolímero de ácido láctico-ácido glicólico y acetato de leuprolide), ácido poli-D-(-)-3-hidroxibutírico, colágeno, metal, hidroxiapatita, biovidrio, aluminado, materiales biocerámicos y proteínas

purificadas.

En una realización particular, el sustrato sólido comprende polímeros biodegradables comercializados con la marca comercial ATRIGEL™ (QLT, Inc., Vancouver, BC). El sistema de suministros de fármacos ATRIGEL® consiste en polímeros biodegradables disueltos en vehículos biocompatibles. Los productos farmacéuticos pueden mezclarse en este sistema de suministro líquido en el momento de fabricación o, dependiendo del producto, pueden añadirse posteriormente por el médico en el momento de uso. Cuando el producto líquido se inyecta en el espacio subcutáneo a través de una aguja de calibre pequeño o se coloca en sitios de tejido accesible a través de una cánula, el agua en los fluidos tisulares provoca que el polímero se precipite y atrape el fármaco en un implante sólido. El fármaco encapsulado dentro del implante se libera después de una manera controlada a medida que la matriz polimérica se biodegrada con el tiempo.

También pueden administrarse composiciones farmacéuticas para uso de la presente invención por vía tópica, por vía (intra) dérmica o por vía transdérmica a la piel o mucosa. Las formulaciones típicas para este fin incluyen geles, hidrogeles, lociones, soluciones, cremas, pomadas, polvos de uso externo, apósitos, espumas, películas, parches cutáneos, obleas, implantes, esponjas, fibras, vendas y microemulsiones. También pueden usarse liposomas. Los vehículos típicos incluyen alcohol, agua, aceite mineral, vaselina líquida, vaselina blanca, glicerina, polietilenglicol y propilenglicol. Pueden incorporarse potenciadores de la penetración, véase, por ejemplo, Finnin y Morgan: J. Pharm. Sci. 88 (10):955-958, (1999). Otros medios de administración tópica incluyen suministro por electroporación, iontoforesis, fonoforesis, sonoforesis e inyección por microaguja o sin agujas, por ejemplo usando los sistemas comercializados con las marcas comerciales POWDERJECT™ y BIOJECT™.

Se conocen bien en la técnica métodos de formulación y se desvelan, por ejemplo, en Remington: The Science and Practice of Pharmacy, Mack Publishing Company, Easton, PA., 20ª edición, ISBN: 0683306472 (2000). Las composiciones y agentes proporcionados en el presente documento pueden administrarse de acuerdo con los métodos de la presente invención en cualquier régimen de dosificación terapéuticamente eficaz. La cantidad y la frecuencia de dosificación se seleccionan para crear un nivel eficaz del agente sin efectos perjudiciales. La cantidad eficaz de un compuesto de la presente invención dependerá de la vía de administración, el tipo de animal de sangre caliente que se trate, y las características físicas del animal de sangre caliente específico que se considere. Los practicantes expertos en la técnica médica conocen bien estos factores y su relación para determinar esta cantidad. Esta cantidad y el método de administración pueden adaptarse para conseguir eficacia óptima pero dependerán de factores tales como el peso, la dieta, la medicación simultánea y otros factores que reconocerán los expertos en la técnica médica.

En realizaciones particulares, la cantidad de una composición o un agente administrado variará en general de una dosificación de aproximadamente 0,1 a aproximadamente 100 mg/kg/día, y normalmente de aproximadamente 0,1 a 10 mg/kg cuando se administra por vía oral o por vía intravenosa. En realizaciones particulares, una dosificación es de 5 mg/kg o 7,5 mg/kg. En diversas realizaciones, la dosificación es de aproximadamente 50-2500 mg por día, 100-2500 mg/día, 300-1800 mg/día o 500-1800 mg/día. En una realización, la dosificación es de entre aproximadamente 100 y 600 mg/día. En otra realización, la dosificación es de entre aproximadamente 300 y 1.200 mg/día. En realización particulares, la composición o el agente se administra a una dosificación de 100 mg/día, 240 mg/día, 300 mg/día, 600 mg/día, 1000 mg/día, 1200 mg/día o 1800 mg/día, en una o más dosis por día (es decir, cuando las dosis combinadas consiguen la dosificación diaria deseada). En realizaciones relacionadas, una dosis es de 100 mg bid, 150 mg bid, 240 mg bid, 300 mg bid, 500 mg bid, o 600 mg bid. En diversas realizaciones, la composición o el agente se administra en dosificaciones individuales o repetidas. La dosificación inicial y dosificaciones posteriores pueden ser iguales o diferentes.

En ciertas realizaciones, una composición o un agente se administra en una única dosificación de 0,1 a 10 mg/kg o 0,5 a 5 mg/kg. En otras realizaciones, una composición o un agente se administra en una dosificación de 0,1 a 50 mg/kg/día, de 0,5 a 20 mg/kg/día o de 5 a 20 mg/kg/día.

En ciertas realizaciones, una composición o un agente se administra por vía oral o por vía intravenosa, por ejemplo, mediante infusión durante un periodo de tiempo de aproximadamente, por ejemplo, 10 minutos a 90 minutos. En otras realizaciones relacionadas, se administra una composición o un agente por infusión continua, por ejemplo, a una dosificación de entre aproximadamente 0,1 y aproximadamente 10 mg/kg/h durante un periodo de tiempo. Aunque el periodo de tiempo puede variar, en ciertas realizaciones el periodo de tiempo puede ser de entre aproximadamente 10 minutos a aproximadamente 24 horas o entre aproximadamente 10 minutos y aproximadamente tres días

En realizaciones particulares, una cantidad eficaz o cantidad terapéuticamente eficaz es una cantidad suficiente para conseguir una concentración total de la composición o el agente en el plasma sanguíneo de un sujeto con una $C_{máx}$ de entre aproximadamente 0,1 µg/ml y aproximadamente 20 µg/ml o entre aproximadamente 0,3 µg/ml y aproximadamente 20 µg/ml. En ciertas realizaciones, una dosificación oral es una cantidad suficiente para conseguir una concentración en plasma sanguíneo ($C_{máx}$) de entre aproximadamente 0,1 µg/ml y aproximadamente 5 µg/ml o entre aproximadamente 0,3 µg/ml y aproximadamente 3 µg/ml. En ciertas realizaciones, una dosificación intravenosa es una cantidad suficiente para conseguir una concentración en plasma sanguíneo ($C_{máx}$) de entre aproximadamente

1 µg/ml y aproximadamente 10 µg/ml o entre aproximadamente 2 µg/ml y aproximadamente 6 µg/ml. En una realización relacionada, la concentración total de un agente en el plasma sanguíneo del sujeto tiene una concentración valle media de menos de aproximadamente 20 µg/ml y/o una concentración de estado estacionario de menos de aproximadamente 20 µg/ml. En una realización adicional, la concentración total de un agente en el plasma sanguíneo del sujeto tiene una concentración valle media de menos de aproximadamente 10 µg/ml y/o una concentración de estado estacionario de menos de aproximadamente 10 µg/ml.

En otra realización más, la concentración total de un agente en el plasma sanguíneo del sujeto tiene una concentración valle media de entre aproximadamente 1 ng/ml y aproximadamente 10 µg/ml y/o una concentración de estado estacionario de entre aproximadamente 1 ng/ml y aproximadamente 10 µg/ml. En una realización, la concentración total de un agente en el plasma sanguíneo del sujeto tiene una concentración valle media de entre aproximadamente 0,3 µg/ml y aproximadamente 3 µg/ml y/o una concentración de estado estacionario de entre aproximadamente 0,3 µg/ml y aproximadamente 3 µg/ml.

En realizaciones particulares, una composición o un agente se administra en una cantidad suficiente para conseguir en el mamífero una concentración en plasma sanguíneo que tiene una concentración valle media de entre aproximadamente 1 ng/ml y aproximadamente 10 µg/ml y/o una concentración de estado estacionario de entre aproximadamente 1 ng/ml y aproximadamente 10 µg/ml. En realizaciones relacionadas, la concentración total del agente en el plasma sanguíneo del mamífero tiene una concentración valle media de entre aproximadamente 0,3 µg/ml y aproximadamente 3 µg/ml y/o una concentración de estado estacionario de entre aproximadamente 0,3 µg/ml y aproximadamente 3 µg/ml.

En realizaciones particulares de la presente invención, la cantidad eficaz de una composición o un agente, o la concentración en plasma sanguíneo de composición o agente se consigue o mantiene, por ejemplo, durante al menos 15 minutos, al menos 30 minutos, al menos 45 minutos, al menos 60 minutos, al menos 90 minutos, al menos 2 horas, al menos 3 horas, al menos 4 horas, al menos 8 horas, al menos 12 horas, al menos 24 horas, al menos 48 horas, al menos 3 días, al menos 4 días, al menos 5 días, al menos 6 días, al menos una semana, al menos 2 semanas, al menos un mes, al menos 2 meses, al menos 4 meses, al menos 6 meses, al menos un año, al menos 2 años o más de 2 años.

En ciertas realizaciones basadas en polipéptidos, la cantidad de polipéptido administrado estará normalmente en el intervalo de aproximadamente 0,1 µg/kg a aproximadamente 0,1 mg/kg a aproximadamente 50 mg/kg de peso corporal de paciente. Dependiendo del tipo y gravedad de la enfermedad, de aproximadamente 0,1 µg/kg a aproximadamente 0,1 mg/kg a aproximadamente 50 mg/kg de peso corporal (por ejemplo, aproximadamente 0,1-15 mg/kg/dosis) de polipéptido puede ser una dosificación candidata inicial para administración al paciente, bien, por ejemplo, mediante una o más administraciones separadas, o bien mediante infusión continua. Por ejemplo, un régimen de dosificación puede comprender administrar una dosis de carga inicial de aproximadamente 4 mg/kg, seguido de una dosis de mantenimiento semanal de aproximadamente 2 mg/kg de polipéptido, o aproximadamente la mitad de la dosis de carga. Sin embargo, pueden ser útiles otros regímenes de dosificación. Una dosificación diaria típica podría variar de aproximadamente 0,1 µg/kg a aproximadamente 1 µg/kg hasta 100 mg/kg o más, dependiendo de los factores mencionados anteriormente. Para administraciones repetidas durante varios días o más, dependiendo de la afección, el tratamiento se mantiene hasta que se produce una deleción deseada de los síntomas de enfermedad.

En realizaciones particulares, la dosificación eficaz consigue los niveles en plasma sanguíneo o concentración de valle media de una composición o un agente descrito en el presente documento. Estos pueden determinarse fácilmente usando procedimientos rutinarios.

Las realizaciones de la presente invención, en otros aspectos, proporcionan 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 invención, como se describe en el presente documento. Los kits pueden incluir instrucciones escritas sobre cómo usar dichas composiciones (por ejemplo, para modular la señalización celular, angiogénesis, cáncer, afecciones inflamatorias, diagnóstico, etc.).

Los kits del presente documento también pueden incluir uno o más agentes terapéuticos adicionales u otros componentes adecuados o deseados para la indicación que se trate, o para la aplicación de diagnóstico deseada. Un agente terapéutico adicional puede estar contenido en un segundo recipiente, si se desea. Los ejemplos de agentes terapéuticos adicionales incluyen, pero sin limitación, agentes antineoplásicos, agentes antiinflamatorios, agentes antibacterianos, agentes antivíricos, agentes angiogénicos, etc.

Los kits del presente documento también pueden incluir una o más jeringas u otros componentes necesarios o deseados para facilitar un modo pretendido de suministro (por ejemplo, estents, depósitos implantables, etc.).

Los siguientes ejemplos se proporcionan solamente como ilustración y no como limitación. Los ejemplos que no quedan dentro del alcance de las reivindicaciones son solamente para fines ilustrativos. Los expertos en la materia reconocerán fácilmente diversos parámetros no críticos que podrían cambiarse o modificarse para producir

resultados esencialmente similares.

XIV. EJEMPLOS

MÉTODOS GENERALES. A no ser que se indique de otro modo, en los ejemplos posteriores, se usaron los siguientes métodos generales para optimización génica, expresión de proteínas a pequeña y gran escala, purificación de proteínas, perfiles transcripcionales y cribado para preparar y caracterizar los polipéptidos de MetRS descritos en los Ejemplos posteriores.

SÍNTESIS Y CLONACIÓN DE GENES EN VECTORES DE EXPRESIÓN

Se optimizaron los codones de secuencias polinucleotídicas que codificaban versiones marcadas en epítopos de los polipéptidos de MetRS y se clonaron en vectores de expresión bacterianos usando los métodos enumerados posteriormente.

En el método (1), se sintetiza ADN de codones optimizados para *E. coli* (Welch *et al*., PLoS ONE 4 (9): e7007 doi: 10.1371/journal.pone.0007002) que codifica cada polipéptido de MetRS por DNA 2.0 (Menlo Park, CA), y se sintetizan dos versiones de cada polipéptido de MetRS, que contienen un marcador epitópico combinado N-terminal o C-terminal que comprende tanto un marcador de seis histidinas como un marcador epitópico de V5.

Se sintetiza ADN que codifica los polipéptidos de MetRS marcados en el extremo N-terminal con una extensión 5’ que codifica en orientación de 5' a 3', un sitio de unión a ribosoma (rbs (subrayado posteriormente)), sitio de restricción NdeI, marcador de seis histidinas y un marcador epitópico de V5, (AGGAGGTAAAACATATGCATCATCATCATCATCACGGTAAGCCTATCCCTA ACCCTTTGCTCGGTCTCGAT-TCTACG) (**SEQ ID NO 1**), que se fusiona en fase con la fase abierta de lectura del polipéptido de MetRS predicha. En casos en los que el polipéptido de MetRS comprende un resto de metionina de inicio nativo predicho (ATG), o el primer resto de aminoácido del polipéptido de MetRS predicho es Met, este se suprime. Al final de la fase abierta de lectura de polipéptido de MetRS predicha, se añaden dos codones de terminación y un sitio XhoI (TAATGACTCGAG) (**SEQ ID NO 2**).

Se sintetiza ADN que codifica los polipéptidos de MetRS marcados en el extremo C terminal con una extensión 5’ que codifica un sitio rbs (subrayado posteriormente) y un sitio de restricción NdeI que recapitula el codón de inicio nativo predicho para el polipéptido de MetRS, o inserta un ATG en fase con la fase de lectura de polipéptido de MetRS predicha, (AGGAGATAAAA-CATATG) (**SEQ ID NO 3**). En diferentes realizaciones, el sitio de unión a ribosoma puede comprender las secuencias "AGGAG-GTAAAACAT" (**SEQ ID NO 4**), "AGGAGATAAAACAT" (SEQ ID NO 5) o GAAGGAGATATACAT (**SEQ ID NO 6**). En el extremo 3’ de la fase abierta de lectura del polipéptido de MetRS predicha, se sintetiza una extensión 3' que codifica en orden de 5’ a 3', un marcador epitópico V5, un marcador de seis histidinas, dos codones de terminación y un sitio XhoI (GGTAAGCCTATCCCTAACCCTCTC-CTCGGTCTCGATTCTACGCACCACCATC ATCACCATTAATGACTCGAG) (**SEQ ID NO 7**), que se fusiona en fase con la fase abierta de lectura del polipéptido de MetRS predicha. Si el polipéptido de MetRS incluía un codón de terminación nativo predicho, este se suprimió.

Se subclonan secuencias de ADN sintetizadas que codifican los polipéptidos de MetRS en vector pJExpress411 (DNA 2.0). Después de secuenciar para confirmar la síntesis del producto correcto, se transforman vectores de expresión en bacterias para expresión de proteínas como se describe más completamente posteriormente.

En el método (2), se sintetiza ADN con codones optimizados para *E. coli* (Ermolaeva MD (2001) Curr. Iss. Mol. Biol. 3 (4) 91-7) que codifica cada polipéptido de MetRS por GENEWIZ (South Plainfield, NJ). Cada secuencia polinucleotídica que codifica el polipéptido de MetRS se sintetizó con extensiones 5’ y 3' cortas que comprendían sitios de restricción únicos para clonación posterior.

Específicamente se insertó un sitio de restricción de BamHI en el extremo 5’ de la fase abierta de lectura predicha. En casos en los que el polipéptido de MetRS comprende un resto de metionina de inicio nativo predicho (ATG), o el primer resto de aminoácido del polipéptido de MetRS predicho es Met, este se suprimió. Adicionalmente se insertó un sitio de restricción XhoI en el extremo 3’ de la fase abierta de lectura predicha. En casos en los que el polipéptido de MetRS comprende un codón de terminación nativo predicho, este se suprimió.

Después de digestión de restricción, las secuencias de ADN resultantes se subclonan en vectores pET-24b modificados (EMD, Gibbstown, NJ) que contienen un marcador epitópico combinado N-terminal (pET24b_N-6XHis/V5) o C-terminal (pET24b_C-V5/6XHis) que comprende tanto un marcador de seis histidinas como un marcador epitópico V5 (modificación de vector por GENEWIZ, (South Plainfield, NJ).

Después de digestión de restricción, y clonación, el ADN que codifica el polipéptido de MetRS marcado en N se clona en el vector marcado en N (pET24b_N-6XHis/V5), que comprende una secuencia de ADN 5’ que codifica seis histidinas y un marcador epitópico V5, (CATATGCATCATCATCATCATCACGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTC TACG-GGATCC),

(**SEQ ID NO 8**), en fase con un codón de inicio (ATG) incluido dentro del sitio de restricción NdeI. Esta extensión 5' se fusiona con la fase abierta de lectura del polipéptido de MetRS predicha a través de un enlazador de 2 aminoácidos corto (GS).

En el extremo 3' de la fase abierta de lectura predicha, el ADN que codifica el polipéptido de MetRS marcado en N comprende una secuencia de ADN que codifica una extensión de 2 aminoácidos (LE) seguida de dos codones de terminación (CTCGAG-TAATGA) **(SEQ ID NO 9).**

Después de digestión de restricción, y clonación, el ADN que codifica el polipéptido de MetRS marcado en C clonado en el vector marcado en C (pET24b_C-V5/6XHis), comprende una secuencia 5' que codifica un codón de inicio (ATG) incluido dentro del sitio de restricción NdeI que se fusiona con la fase de lectura del polipéptido de MetRS predicho a través de un enlazador de 2 aminoácidos corto (GS), (CATATGGGATCC) **(SEQ ID NO 10).**

En el extremo 3' de la fase abierta de lectura predicha, el ADN que codifica el polipéptido de MetRS marcado en C comprende una secuencia de ADN 3' que codifica un enlazador de 2 aminoácidos corto (LE) seguido de un marcador epitópico V5 seguido de seis histidinas y dos codones de terminación,

CTCGAGGGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGCACCACCACCACCACCACTAATGA **(SEQ ID NO 11).**

EXPRESIÓN, PURIFICACIÓN Y CARACTERIZACIÓN BIOFÍSICA DE POLIPÉPTIDO DE METRS

Se expresan polipéptidos de MetRS marcados con His6x en bacterias en un formato de rendimiento medio y/o en cultivos de matraces de mayor escala dependiendo de la cantidad de proteína requerida. Se purifican polipéptidos de MetRS usando cromatografía de intercambio iónico y de afinidad como se describe posteriormente, y como se especifica para experimentos específicos.

Cultivos bacterianos: Se transforman 100 ng de vector de expresión que comprende ADN de codones optimizados que codifica cada polipéptido de MetRS (como se ha descrito anteriormente) en bacterias *E. coli* competentes BL21(DE3) (EMD chemicals, cat. n.º 69450) a 42 ºC Durante 30 segundos en placas de PCR. También se evalúan las cepas C41(DE3) (Lucigen, cat. n.º 60442), HMS174(DE3) (EMD chemicals, cat. n.º 69453) y Origami2(DE3) (EMD chemicals, cat. n.º 71345). Las placas se colocan en hielo durante 2 minutos y se añaden 100 µl de medio SOC, seguido de una incubación de 1 hora a 37 ºC. Se añaden 5 ml de medio de autoinducción (EMD chemicals, cat. n.º 71491) complementado con kanamicina (100 µg/ml) a cada pocillo de un bloque de 24 pocillos (Qiagen, cat. n.º 19583). Las reacciones de transformación se añaden a los pocillos individuales, el bloque se sella con película adhesiva (VWR, cat. n.º 60941-078) y se incuban durante una noche a 250 rpm en un agitador a 37 ºC. Cuando se usan condiciones de baja temperatura (25 ºC), se lleva a cabo en su lugar incubación durante 48 horas.

Para expresión a mayor escala, se añaden 200 ml de medio de autoinducción complementado con kanamicina (100 µg/ml) en matraces de Erlenmeyer de 500 ml con tapas de ventilación (Corning, cat. n.º 431401). Las reacciones de transformación se añaden a los matraces individuales y se incuban durante 30 horas a 250 rpm en un agitador a 37 ºC.

***Aislamiento de proteínas*:** Después de que el cultivo alcance la fase estacionaria (normalmente DO_{600} de 3-6), los bloques se centrifugan a 3600 x g durante 10 minutos. El medio se aspira cuidadosamente y los bloques se congelan a -80 ºC o -20 ºC durante 10 minutos. Después se permite que los bloques se descongelen a temperatura ambiente y se añade 1 ml de tampón de lisis (100 ml de Bugbuster complementado con 200 µl de lisonasa (EMD chemicals, cat. no 71370) y se añaden inhibidores de proteasa "completos mini sin EDTA" (Roche, cat. No. 11 836 170 001)) a cada pocillo. Los sedimentos se resuspenden pipeteando repetidas veces hasta que no hay ningún grumo visible y se transfieren a tubos eppendorf, seguido de una incubación de 10-20 minutos en un agitador a temperatura ambiente. Después de centrifugar a 16.000 g durante 10 minutos a 4 ºC, los lisados se cargan en una placa TurboFilter 96 incluida en el kit Ni-NTA Superflow 96 BioRobot (Qiagen, cat. n.º 969261) y se centrifugan a 500 g durante 5-10 minutos.

Para expresión a mayor escala, el cultivo de fase estacionaria se transfiere a frascos de 500 ml y se centrifuga a 6.000 g durante 10 minutos. El medio se decanta y el sedimento se almacena a -80 ºC o -20 ºC antes de procesamiento posterior. Después se permite que el sedimento se descongele a temperatura ambiente y se añaden 20 ml de tampón de lisis a cada frasco. Los sedimentos se resuspenden por pipeteo repetido hasta que no hay ningún grumo visible, seguido de incubación durante 20 minutos en un agitador a temperatura ambiente. Después de centrifugación a 10.000 g durante 30 minutos a 4 ºC, los lisados se transfieren a tubos o frascos limpios. Si se pasan cantidades traza de residuos durante la transferencia, la muestra se centrifuga de nuevo o se pasa a través de una membrana de acetato de celulosa de 0,45 µm (Corning, cat. n.º 430314) para clarificación adicional.

Purificación de afinidad: Se carga una placa QIAFilter 96 con 200 µl de suspensión de Ni-NTA Superflow incluida en el kit Ni-NTA Superflow 96 BioRobot y la resina se equilibra añadiendo 600 µl de tampón de unión (fosfato sódico 20 mM, cloruro sódico 500 mM e imidazol 10 mM, pH 7,5). Se aplica un vacío de -50,67 kPa hasta que todo el

tampón ha pasado a través de la resina. Los lisados celulares clarificados de la etapa previa se cargan después en la placa de QIAFilter® y se permite que se unan durante 5 minutos. Se aplica un vacío de -10,13 kPa durante aproximadamente 5 minutos hasta que todas las muestras han pasado a través de la resina. La resina se lava después con 1 ml de tampón de unión, seguido de dos lavados con 1 ml de tampón de unión que contiene Triton X-100 0,1 %. La resina se lava después 10 minutos con 1 ml de tampón de unión sin Triton X-100. Los polipéptidos de MetRS marcados con His6x unidos se eluyen con 450 µl de tampón de elución (fosfato sódico 20 mM, cloruro sódico 500 mM e imidazol 500 mM, pH 7,5) y se almacenan a 4 ºC.

Para expresión a mayor escala, se carga una columna desechable vacía "Poly-Prep" (Bio-Rad, cat. n.º. 731-1550) con 1 ml de suspensión Ni-NTA Superflow (Qiagen, n.º cat. 30450) y la resina de 0,5 ml se equilibra añadiendo 5 ml de tampón de unión. El lisado celular clarificado de la etapa previa se carga después en la columna y se permite que la atraviese por gravedad. La resina se lava primero con 50 ml de tampón de unión más Triton X-100 0,1 %, después se lava con 50 ml de tampón de unión sin Triton X-100. Se eluyen los polipéptidos de MetRS marcados con His6x unidos con 2 ml de tampón de elución y se almacenan a 4 ºC.

Etapas de desalación y pulido: Para polipéptidos de MetRS con una masa molecular de >10 kDa, se aclara la membrana Omega 10K de un papel de filtro AcroPrep 96 (Pall, n.º cat. 5034) con 20 µl de PBS 1X y la placa se coloca en un colector de vacío (>33,78 kPa) hasta que todo el líquido atraviesa. Los eluatos de la etapa previa (Ni-NTA) se distribuyen en cada pocillo y el vacío se aplica hasta que todos los líquidos atraviesan. Estas etapas se repiten hasta que se ha procesado el volumen de eluato total (450 µl). Se recuperan polipéptidos de MetRS añadiendo 180 µl de PBS 1X pH 7,4 a cada pocillo, pipeteando arriba y abajo 10 veces cuidadosamente y después se transfieren a un bloque limpio. Esta etapa se repite para producir un volumen total de 360 µl por pocillo y el bloque se almacena a 4 ºC. Para polipéptidos de MetRS con una masa molecular de <10 kDa, los eluatos de Ni-NTA se cargan en una unidad de filtro de centrífuga Amicon Ultra-15 con membrana Ultracel-3 (Millipore, cat. n.º UFC900308), seguido de la adición de 10 ml de PBS 1X y una centrifugación a 3.600 g durante 10-30 minutos hasta que el volumen es menor de 360 µl. Las muestras se recuperan y se añade PBS 1X a un volumen final de 360 µl.

Para retirar endotoxinas, se aclara un papel de filtro AcroPrep Advance con membrana Mustang Q (Pall, n.º cat. 8171) con 300 µl de PBS 1X y se centrifuga a 1.000 g durante 5 minutos para eliminar el tampón. Los polipéptidos desalados de MetRS (360 µl/pocillo) se añaden a la placa de filtro y se incuban en un agitador durante 5-10 minutos. La placa se centrifuga a continuación a 1.000 g durante 5-10 minutos y las fracciones de flujo continuo que contienen los polipéptidos de MetRS se recogen y se almacenan a 4 ºC.

Para la expresión a gran escala, los eluatos de Ni-NTA se cargan en una unidad de filtro de centrífuga Ultra-15 de Amicon con membrana Ultracel-3 o Ultracel-10 (Millipore, cat n.º UFC900308 o UFC901008) dependiendo del peso molecular del polipéptido de MetRS y luego se centrifugan a 3.600 g durante 10-30 minutos hasta que el volumen se reduce a 250 µl. Las muestras se mezclan en 10 ml de PBS 1X, pH 7,4 y se centrifugan de nuevo a 3.600 g durante 10-30 minutos hasta que el volumen es de aproximadamente 250 µl. Este paso se repite una vez más, los sobrenadantes se recuperan y se añade PBS 1X a un volumen final de 1,5 ml.

Para eliminar endotoxinas, se enjuaga una membrana intercambiadora de aniones fuerte Sartobind Q5 (Sartorius, cat n.º Q5F) con 1 ml de PBS 1X y los polipéptidos de MetRS se pasan lentamente a través de la membrana usando una jeringa de plástico. La fracción de flujo continuo que contiene los polipéptidos de MetRS se recoge en un bloque de 96 pocillos profundos que se sella y se almacena a 4 ºC.

Se purifican polipéptidos de MetRS marcados con His6x expresados en bacterias y hallados en cuerpos de inclusión usando cromatografía de afinidad y una serie de etapas de replegamiento, como se describe posteriormente.

Cultivos bacterianos: se transforman 100 ng de plásmido que codifica cada polipéptido de MetRS en bacterias *E. coli* competentes BL21(DE3) ((EMD chemicals, cat. n.º 69450) o C41(DE3) (Lucigen, cat. n.º 60442) a 42 ºC durante 30 segundos en placas de PCR. Las placas se colocan en hielo durante 2 minutos y se añaden 100 µl de medio SOC, seguido de una incubación de 1 hora a 37 ºC. Se añaden 5 ml de medio de autoinducción (EMD chemicals, cat. n.º 71491) complementado con kanamicina (100 µg/ml) a cada pocillo de un bloque de 24 pocillos (Qiagen, cat. nº. 19583). Las reacciones en transformación se añaden a los pocillos individuales, el bloque se sella con película adhesiva (VWR, cat. n.º 60941-078) y se incuba durante una noche a 250 rpm en un agitador a 37 ºC.

Para expresión a mayor escala, se añaden 200 ml de medio de autoinducción complementado con kanamicina (100 µg/ml) en matraces de Erlenmeyer de 500 ml con tapas de ventilación (Corning, cat. n.º. 431401). Las reacciones de esta transformación se añaden a los matraces individuales y se incuban durante 30 horas a 250 rpm en un agitador a 37 ºC.

Aislamiento: Después de que los cultivos alcancen fase estacionaria (DO_{600} típica de 3-6), los bloques se centrifugan a 3.600 x g durante 10 minutos. El medio se aspira cuidadosamente y los bloques se congelan a -80 ºC o -20 ºC durante 10 minutos. Después se permite que los bloques se descongelen a temperatura ambiente y se añade 1 ml de tampón de lisis (100 ml de Bugbuster complementado con 200 µl de lisonasa (EMD chemicals, cat. n.º 71370) e inhibidor de proteasa "completo mini sin EDTA" (Roche, cat. n.º 11 836 170 001)) a cada pocillo. Los

sedimentos se resuspenden por pipeteo repetido hasta que no hay ningún grumo visible y se transfieren a tubos eppendorf, seguido de una incubación de 10-20 minutos en un agitador a temperatura ambiente. Después de centrifugación a 16.000 x g durante 10 minutos a 4 ºC, los lisados solubles se descartan y los cuerpos de inclusión se resuspenden exhaustivamente en tampón de unión desnaturalizante (fosfato sódico 20 mM, cloruro sódico 500 mM, clorhidrato de guanidina 6 M, imidazol 10 mM, pH 7,5). Las muestras se centrifugaron a 16.000 x g durante 10 minutos y los sobrenadantes se cargan en una placa TurboFilter 96 incluida en el kit Ni-NTA Superflow 96 BioRobot (Qiagen, cat. n.º 969261) seguido de centrifugación a 500 g durante 5-10 minutos. Los filtrados se recogen en un bloque de 96 pocillos limpio (Greiner, cat. n.º 780286).

Para la expresión a mayor escala, el cultivo de fase estacionaria se transfiere a frascos de 500 ml y se centrifuga a 6.000 g durante 10 minutos. El medio se decanta y el sedimento se almacena a -80 ºC o -20 ºC antes de su procesamiento posterior. Después se permite que el sedimento se descongele a temperatura ambiente y se añaden 20 ml de tampón de lisis a cada frasco. Los sedimentos se resuspenden por pipeteo repetido hasta que no hay ningún grumo visible, seguido de incubación durante 20 minutos en un agitador a temperatura ambiente. Después de centrifugación a 10.000 g durante 30 minutos a 4 ºC, los lisados solubles se descartan y los cuerpos de inclusión insolubles se resuspenden exhaustivamente en tampón de unión desnaturalizante.

Purificación de afinidad: se carga una placa QIAFilter 96 con 200 µl de suspensión de Ni-NTA Superflow incluida en el Kit de Ni-NTA Superflow 96 BioRobot y la resina se equilibra añadiendo 600 µl de tampón de unión desnaturalizante (véase anteriormente). Se aplica un vacío de -50,67 kPa hasta que todo el tampón pase a través de la resina. Las muestras desnaturalizadas clarificadas de la etapa previa se cargan después en la placa QIAFilter® 96 y se permite que se unan durante 5 minutos. Se aplica un vacío de aproximadamente 10,13 kPa durante aproximadamente 5 minutos hasta que todas las muestras pasan a través de la resina. La resina se lava después con 1 ml de tampón de unión desnaturalizante, seguido de cinco lavados con 1 ml de tampón de unión desnaturalizante que contiene Triton X-100 0,1 %. La resina se lava después 15 veces con 1 ml de tampón de unión desnaturalizante sin Triton X-100. Los polipéptidos de MetRS marcados con His 6x unidos se eluyen después con 450 µl de tampón de elución desnaturalizante (fosfato sódico 20 mM, cloruro sódico 500 mM, clorhidrato de guanidina 6 M e imidazol 500 mM, pH 7,5) y se almacenan a 4 ºC.

Para expresión a mayor escala, se carga una columna desechable vacía “Poly-Prep” (Bio-Rad, n.º de catálogo 731-1550) con 1 ml de suspensión de Ni-NTA Superflow (Qiagen, cat n.º 30450) y los 0,5 ml de resina se equilibran añadiendo 5 ml de tampón de unión desnaturalizante (véase anteriormente). Los cuerpos de inclusión desnaturalizados de la etapa anterior se cargan después en la columna y se permite que atraviesen por gravedad. La resina se lava en primer lugar con 50 ml de tampón de unión desnaturalizante más Triton X-100 0,1 %, después se lava con 50 ml de tampón de unión desnaturalizante sin Triton X-100. Los polipéptidos de MetRS marcados con His 6x unidos se eluyen con 2 ml de tampón de elución desnaturalizante y se almacenan a 4 ºC.

Replegamiento: para polipéptidos de MetRS >10 kDa, la membrana de 10 K Omega de una placa de filtro AcroPrep 96 (Pall, cat. n.º 5034) se aclara con 20 µl de PBS 1X y la placa se coloca en un colector de vacío (>33,78 kPa) hasta que todo el líquido atraviesa. Los eluatos de la etapa previa (Ni-NTA) se distribuyen en cada pocillo y se aplica al vacío hasta que todo el líquido atraviesa. Estas etapas se repiten hasta que se ha procesado el volumen de eluato total (450 µl). Se recuperan polipéptidos de MetRS añadiendo 200 µl de tampón de replegamiento que contiene Tris 50 mM, cloruro sódico 250 mM, cloruro potásico 10 mM, cloruro magnésico 2 mM, cloruro cálcico 2 mM, sacarosa 400 mM, arginina 500 mM, DTT 1 mM y Polisorbato 80 0,01 %, pH 7,4) a cada pocillo, pipeteando arriba y abajo 10 veces cuidadosamente, y después se transfieren a un bloque limpio. Esta etapa se repite para producir un volumen total de 400 µl por pocillo y el bloque se coloca en el agitador durante una noche a 4 ºC. Para polipéptidos de MetRS <10 kDa, los eluatos de Ni-NTA se cargan en una unidad de filtro de centrífuga Amicon Ultra-15 con membrana Ultracel-3 (Millipore, cat. n.º UFC900308), seguido de la adición de 10 ml de tampón de replegamiento y una centrifugación a 3.600 g durante 10-30 minutos hasta que el volumen es menor de 400 µl. Las muestras se recuperan y se añade tampón de replegamiento extra a un volumen final de 400 µl. Las muestras se transfieren a un bloque de 96 pocillos, se sella con película y se coloca en un agitador durante una noche a 4 ºC.

Para cultivos a mayor escala, los eluatos de Ni-NTA se cargan en una unidad de filtro de centrífuga Amicon Ultra-15 con membrana Ultracel-3 o Ultracel-10 (Millipore, cat. n.º UFC900308 o UFC901008 dependiendo del peso molecular del polipéptido de MetRS) y después se centrifugan a 3.600 g durante 10-30 minutos hasta que el volumen se reduce a aproximadamente 500 µl. Para polipéptidos d MetRS con pI > 7, las muestras se diluyen 20 veces en el siguiente tampón: acetato sódico 50 mM, cloruro sódico 10 mM, cloruro potásico 0,4 mM, EDTA 1 mM, sacarosa 400 mM, arginina 500 mM, DTT 1 mM y polisorbato 80 0,01 %, pH 6,0. Para polipéptidos de MetRS con pI < 7, las muestras se diluyen 20 veces en el siguiente tampón: Tris 50 mM, cloruro sódico 250 mM, cloruro potásico 10 mM, cloruro de magnesio 2 mM, cloruro cálcico 2 mM, sacarosa 400 mM, arginina 500 mM, DTT 1 mM y polisorbato 80 0,01 %, pH 8,0. Las muestras se incuban en un agitador a 4 ºC durante una noche.

Etapas de desalación y pulido: después de incubación durante una noche, el bloque de 96 pocillos se centrifuga a 3.600 g para retirar cualquier agregado potencial. Los sobrenadantes se someten después a intercambio de tampón con PBS 1X (Invitrogen, cat. n.º 10010). Para polipéptidos de MetRS > 10 kDa, la membrana Omega 10 K de una placa de filtro AcroPrep 96 se aclara con 20 µl de PBS 1X y la placa se coloca en un colector de vacío (> 33,78 kPa)

hasta que todo el líquido atraviesa. Las muestras en el tampón de replegamiento se distribuyen en cada pocillo y se aplica el vacío hasta que atraviesa todo el líquido. Estas etapas se repiten hasta que se ha procesado el volumen de muestra total (400 µl). Se recuperan polipéptidos de MetRS añadiendo 180 µl de PBS 1X pH 7,4 a cada pocillo, pipeteando arriba y abajo 10 veces cuidadosamente, y se transfiere después a un bloque limpio. Esta etapa se repite hasta producir un volumen total de 360 µl por pocillo y el bloque se almacena a 4 ºC. Para polipéptidos de MetRS <10 kDa, las muestras replegadas se cargan en una unidad de filtro de centrífuga Amicon Ultra-15 con membrana Ultracel-3 (Millipore, cat. n.º UFC900308) seguido de la adición de 10 ml de PBS 1X y centrifugación a 3.600 g durante 10-30 minutos hasta que el volumen es menor de 360 µl. Las muestras se recuperan y se añade PBS 1X a un volumen final de 360 µl.

Para retirar endotoxinas, se aclara una placa de filtro AcroPrep Advance con membrana Mustang Q (Pall, cat. n.º 8171) con 300 µl de PBS 1X y se centrifuga a 1.000 g durante 5 minutos para retirar el tampón. Los polipéptidos de MetRS (360 µl/pocillo) se añaden a la placa de filtro y se incuban en un agitador durante 5-10 minutos. La placa se centrifuga después a 1.000 g durante 5-10 minutos y las fracciones de flujo continuo que contienen los polipéptidos de MetRS se recogen y se almacenan a 4 ºC.

Para cultivos a mayor escala, después de la incubación de una noche, las muestras replegadas se centrifugan a 10.000 g durante 10 minutos para retirar cualquier agregado insoluble. El sobrenadante se carga en una unidad de filtro de centrífuga Amicon Ultra-15 y se centrifuga a 3.600 g hasta que el volumen se reduce a 250 µl. Las muestras se mezclan en 10 ml de PBS 1X y se centrifugan de nuevo a 3.600 g durante 10-30 minutos hasta que el volumen es de aproximadamente 250 µl. Obsérvese que el pH de PBS 1X se ajusta para coincidir con el pH del tampón de replegamiento, bien pH 6,0 o pH 8,0. Esta etapa se repite una vez más, los sobrenadantes se recuperan y se añade PBS 1X a un volumen final de 1,5 ml.

Para retirar endotoxinas, se lava abundantemente una membrana de intercambio aniónico fuerte Sartobind Q 5 (Sartorius, cat. n.º Q5F) con 1 ml de PBS 1X y se hace que los polipéptidos de MetRS atraviesen lentamente la membrana usando una jeringa de plástico. La fracción de flujo continuo que contiene los polipéptidos de MetRS se recoge en un bloque 96 pocillos profundos que se sella y se almacena a 4 ºC.

CARACTERIZACIÓN BIOFÍSICA: Todos los polipéptidos de MetRS purificados se analizan por SDS-PAGE, su concentración se determina basándose en A_{280} y se calcula su coeficiente de extinción (ProtParam en servidor ExPASy). Los niveles de endotoxina se miden por el ensayo LAL Cromogénico de criterios de valoración QCL-1000 (Lonza, cat. n.º 50-648U) de acuerdo con las instrucciones del fabricante.

Dispersión de luz dinámica: un instrumento Wyatt Technology DynaPro 99 y el controlador de temperatura (20 ºC) se calientan durante 15 minutos antes del experimento seguido de conexión del software de Dynamics con el instrumento. El tiempo de adquisición se establece en 10 segundos para múltiples adquisiciones y la potencia del láser se ajusta al 100 %. La cubeta de cuarzo se lava exhaustivamente con agua desionizada y metanol antes de la adición de la muestra de proteínas (15 µl a una concentración de aproximadamente 1 mg/ml en PBS). Se retiran las burbujas de aire golpeando la cubeta suavemente antes de insertarla en el soporte con el lado esmerilado a la izquierda. Si la intensidad es demasiado alta (se muestra un mensaje de advertencia en la pantalla), la muestra se diluye adicionalmente con PBS hasta que la intensidad se reduce hasta un intervalo normal. Los datos recogidos incluyen radio hidrodinámico, polidispersión, peso molecular promedio predicho, porcentaje de intensidad y porcentaje de masa.

Cromatografía de exclusión por tamaño: la muestra de proteína se diluye hasta una concentración de aproximadamente 5-10 mg/ml en PBS antes de cargarse en un asa de muestras de 100 µl en el General Electric AKTA FPLC. Se usa la columna de exclusión por tamaño Superdex 200 10/300 GL (General Electric, cat. n.º 17-5175-01) para separación. La columna se equilibra en primer lugar con 1,5 volúmenes de columna (VC) de tampón de PBS 1X, seguido de inyección de muestras. La columna se procesa en 1 VC de tampón de PBS 1x (flujo isocrático) con control de la absorbancia a 280 nm. El área del pico se integra y el porcentaje se calcula con el software Unicorn. El volumen de elución se usa para estimar el peso molecular basándose en la comparación con kits de calibración de filtración en gel (General Electric, cat. n.º 28-4038-41 y 28-4038-42).

Recuperación de proteínas tras almacenamiento a alta concentración: se transfieren 10 µl de los polipéptidos de MetRS concentrados hasta ≥ 10 mg/ml usando una unidad de filtro Amicon Ultra-15 (Millipore, cat. n.º UFC901024 o UFC900324, dependiendo del peso molecular) a un tubo de microcentrífuga limpio. La muestra se almacena a temperatura ambiente durante una semana seguido de centrifugación a 16.000 g durante 10 minutos para sedimentar cualquier precipitado. La concentración del sobrenadante se determina por un ensayo de proteínas de Bradford y se compara con la concentración medida antes de la exposición de una semana a temperatura ambiente. La recuperación se expresa como porcentaje de la concentración de partida.

Caracterización de polipéptidos de MetRS por CL-EM: se diluyen polipéptidos de MetRS purificados (1 mg/ml) 1:10 en ácido fórmico 0,1 % y se cargan 0,6 µg de proteína con un automuestreador Dionex en una columna capilar C4. La columna capilar se prepara cortando 150 mm de tubo de sílice fusionado (DO de 0,36 mm por DI de 0,1 mm, Polymicro Technologies, cat. n.º 2000023). El capilar se extrae en un extremo con un Extractor de Fibra de Láser

Suter Instrument y se corta con un cortador de sílice fusionado para generar una punta de 5 µm. El capilar se envasa a la longitud de 75 mm con resina C4 (5 µm, 300 Å, Michrom, cat. n.º PM5/64300/00) usando una bomba de presión. El análisis de CL-EM se realiza en un espectrómetro de masas de trampa iónica ThermoFisher LTQ acoplado a un sistema de HPLC Dionex Ultimate3000. El analito se eluye de la columna usando un gradiente de 35 minutos de acetonitrilo 5-70 % en ácido fórmico 0,1 % a un caudal de 0,9 µl/min. El LTQ se opera en un modo de cribado de EM completo (300-2.000 m/z) con una tensión de pulverización de 2,5 kV.

Recogida y análisis de datos: se almacenan datos de espectrometría de masas en bruto en archivos RAW generados por XCalibur que se ejecuta en el espectrómetro de masas LTQ XL. Los espectros de EM de los picos principales en el cromatógrafo se analizan adicionalmente con algoritmo de desconvolución de ThermoFisher ProMass para obtener los pesos moleculares del polipéptido de MetRS.

ANÁLISIS FUNCIONAL DE POLIPÉPTIDOS DE METRS

PERFILES TRANSCRIPCIONALES

Antecedentes y relevancia terapéutica: además de técnicas de identificación de dianas tradicionales, han surgido recientemente herramientas genómicas como enfoques importantes para ayudar a dilucidar el mecanismo de acción de polipéptidos de MetRS y pueden proporcionar información directa sobre la relevancia terapéutica temprano en el proceso de descubrimiento de fármacos. Para facilitar un entendimiento de la utilidad terapéutica potencial, se cultivan tipos celulares humanos primarios con polipéptidos de MetRS y se evalúa el perfil transcripcional en dos puntos temporales separados después de incubación con polipéptidos de MetRS.

Los tipos celulares elegidos para perfiles de transcripción se basan en las capacidades pluripotenciales de las células en cuestión y el potencial para identificar polipéptidos de MetRS de valor terapéutico directo. Por ejemplo, las células madre mesenquimales (MSC) pueden diferenciarse en linajes osteogénico, adipogénico, condrogénico, miocárdico o neuronal cuando se exponen a estímulos específicos, haciéndolas atractivas para el entendimiento de la relevancia potencial de los polipéptidos de MetRS para una amplia serie de tipos celulares y enfermedades.

Además de soportar células hematopoyéticas, también puede inducirse que células del estroma de la médula ósea se diferencien en células de diferente linaje de tejido conectivo, tal como hueso, cartílago y grasa. El potencial de células madre mesenquimales humanas (hMSC) para mantener la multipotencialidad y proliferar extensivamente *in vitro* proporciona nuevas vías para la terapia basada en células en la restauración de tejido dañado o enfermo. Informes recientes también indican que las HMSC tienen capacidad de cruce de destino celular a través de límites de capas germinales. Además de diferenciarse en múltiples linajes del mesodermo, estas células también pueden diferenciarse en neuronas de origen ectodérmico y células de tipo hepatocitos de origen endodérmico. Durante el proceso de diferenciación, estas células pueden modificar los patrones de expresión de ciertos transcritos específicos de linaje.

En consecuencia la capacidad de polipéptidos de MetRS específicos para modular patrones específicos de genes en HMSC de una manera dependiente del tiempo demuestra que estas proteínas desempeñan funciones potencialmente significativas en una amplia serie de rutas de diferenciación, así como enfermedades y trastornos resultantes de la disfunción, o el deterioro de estos procesos, o los tipos celulares correspondientes. Además los polipéptidos de MetRS con la capacidad de modular la transcripción génica en MSC tienen una utilidad terapéutica significativa para permitir la modulación *in vitro* o *in vivo* de hematopoyesis, neurogénesis, miogénesis, osteogénesis y adipogénesis, así como en una amplia serie de trastornos y enfermedades, incluyendo por ejemplo respuestas inflamatorias, autoinmunidad, cáncer, degeneración neuronal, distrofia muscular, osteoporosis y lipodistrofia.

Las células musculares esqueléticas humanas (HSkMC) pueden experimentar diferenciación para mostrar miofilamentos de actina y miosina, y se han usado en el estudio de enfermedades musculares genéticas tales como hipertermia maligna. Las HSkMC también tienen el potencial de actuar como un injerto cardíaco, reparando el daño al corazón. Recientemente, se han usado células del músculo esquelético humano cultivadas en experimentos de microgravedad para estudiar los efectos de ambientes de gravedad baja en el músculo esquelético humano.

En consecuencia la capacidad de polipéptidos de MetRS específicos para modular patrones específicos de genes en HSkMC de una manera dependiente del tiempo demuestra que estas proteínas desempeñan funciones potencialmente significativas en los procesos de miogénesis, así como enfermedades y trastornos resultantes de la disfunción, o el deterioro de estos procesos así como desarrollo de células musculares o metabolismo. En consecuencia los polipéptidos de MetRS con la capacidad de modular la transcripción génica en células musculares tienen utilidad terapéutica en una amplia serie de enfermedades incluyendo por ejemplo el tratamiento de enfermedad metabólica, caquexia, diversas afecciones de debilitamiento muscular, así como enfermedades musculoesqueléticas.

Métodos: la capacidad de polipéptidos de MetRS para modular la expresión génica se evalúa usando un enfoque de PCR cuantitativa en tiempo real (RT-qPCR) microfluídico de alto rendimiento (Fluidigm Corporation). (Véase Petriv *et al*., (2010) PNAS (doi/10.1073/pnas.1009320107) en Células del Estroma de Médula Ósea Humana (HMSC) y

Células del Músculo Esquelético Humano (HSkMC)). En los experimentos indicados en el presente documento, se obtuvieron HSkMC humanas (Cat n.º 150-05f) y HMSC (Cat n.º 492-05f) de Cell Applications. Se crioconservaron células HMSC en un segundo pase y pueden cultivarse y propagarse a 10 duplicaciones de población. Aquí se usan HMSC en el 6º Pase. Se crioconservan células de Músculo Esquelético Humano (HSkMC) en el segundo pase y pueden cultivarse y propagarse durante al menos 15 duplicaciones de población. En los experimentos presentados en el presente documento se usan HSkMC en el pase 6 después de recogida de donante humano normal.

En ambos casos, las células se siembran en placas a 50000 células/ml en 100 µl de volumen de medio de cultivo y se exponen a polipéptidos de MetRS a una concentración de 250 nM, o como se indica de otro modo posteriormente, durante 24 horas y 72 horas. Los controles incluyen medio de diferenciación con un cóctel convencional para promover (1) Adipogénesis, (2) Osteogénesis, (3) Condrogénesis y (4) Formación de miotubo del músculo esquelético. Los controles adicionales incluyen pocillos no tratados que contienen solamente medio de cultivo. Se procesaron dos pocillos para cada control de Diferenciación. Controles: todos los medios se prepararon utilizando DMEM como el medio basal. Se siguió la bibliografía convencional y se obtuvo medio de diferenciación de Cell Applications. Según el proveedor, los medios de diferenciación contenían los siguientes aditivos: cóctel de diferenciación de músculo esquelético: FBS, insulina, glutamina, FGF, EGF; Cóctel de adipogénesis: insulina, dexametasona e IBMX; Cóctel de osteogénesis: FBS, dexametasona, ascorbato 2 fosfato, beta-glicerofosfato; Cóctel de condrogénesis: insulina, ascorbato-2-fosfato y TGIF-β1.

Se utilizan protocolos convencionales para usar un Kit de Expresión Génica Cells-to-CT™ de ABI (Applied Biosystems, n.º de Artículo AM1728) TAQMAN® para lisar células y recoger material genómico. Se usa una mezcla Pre-Amp ABI (Applied Biosystems, n.º de Artículo 4391128) para iniciar la preamplificación. Se crean cebadores específicos de genes usando un programa Primer 3 y se obtiene de IDT technologies. Se usaron matrices de perfiles Fluidigm (n.º de Artículo BMK-M-96.96) para PCR cuantitativa real con reactivos de carga Fluidigm convencionales y dispositivos de pipeteo. La **Tabla E1** a continuación enumera los genes perfilados.

Tabla E1
Lista de genes evaluados en el perfil transcripcional

Lista Única Compilada	refseq_nt	Número_completo_	Sinónimos
ABCA1	NM_005502	casete de unión a ATP, subfamilia A (ABC1), miembro 1	ABC-1|ABC1|CERP|FLJ14958|HDLDT1 |MGC164864|MGC 165011|TGD
ACTB	NM_001101	actina, beta	PS1TP5BP1
ACTG1	NM_001614	actina, gamma 1	ACT|ACTG|DFNA20|DFNA26|
ACVR2B	NM_001106	receptor de activina A, tipo IIB	ACTRIIB|ActR-IIB|MGC116908
APOA1	NM_000039	apolipoproteína A-I	MGC117399
ARNT	NM_178427	translocador nuclear de receptor de aril hidrocarburo	HIF-1beta|HIF1B|HIF1BETA|TANGO|bHLHe2
BAD	NM_032989	agonista de muerte celular asociado a BCL2	BBC2|BCL2L8
BCL2	NM_000657	CLL de linfocitos B/linfoma 2	Bcl-2
BMP2	NM 001200	proteína morfogenética del hueso 2	BMP2A
BMP4	NM_130851	proteína morfogenética del hueso 4	BMP2B|BMP2B1|MCOPS6|OFC11| ZYME
C3AR1	NM_004054	receptor 1 de componente de complemento 3a	AZ3B|C3AR|HNFAG09
CASP3	NM_032991	caspasa 3, cisteína peptidasa relacionada con apoptosis	CPP32|CPP32B|SCA-1
CAV1	NM 001753	caveolina 1, proteína de caveolas, 22 kDa	BSCL3|CGL3|MSTP085|VIP21
CDH5	NM_001795	cadherina 5, tipo 2 (endotelio vascular)	7B4|CD144|FLJ17376
CFLAR	NM_003879	regulador de la apoptosis de tipo FADD y CASP8	CASH|CASP8AP1|CLARP|Casper| FLAME|FLAME-1|FLAME1|FLIP| I-FLICE |MRIT|cFLIP|c-FLIPL|c-FLIPR|c-FLIPS

COMP	NM_000095	proteína de matriz oligomérica de cartílago	EDM1|EPD1|MED|MGC131819| MGC149768| PSACH|THBS5
CSF1	NM_172212	factor estimulante de colonias 1 (macrófagos)	MCSF|MGC31930
CTGF	NM_001901	factor de crecimiento tisular conectivo	CCN2|HCS24|IGFBP8|MGC102839|NOV2
CTNNB1	NM_001904	catenina (proteína asociada a cadherina), beta 1, 88 kDa	CTNNB|DKFZp686D02253|FLJ25606| FLJ37923
DAAM1	NM_014992	activador asociado a dishevelled de morfogénesis 1	FLJ41657|KIAA0666
ELN	NM_001081755	elastina	FLJ38671 |FLJ43523|SVAS|WBS|WS
EN01	NM_001428	enolasa 1, (alfa)	EN01L1|MPB1|NNE|PPH
FABP3	NM_004102	proteína de unión a ácido graso 3, músculo y corazón (inhibidor del crecimiento derivado de mama)	FABP11 |H-FABP|MDGI|0-FABP
FAK	NM_001199649	quinasa de adhesión focal	fakl
FGF4	NM_002007	factor de crecimiento de fibroblastos 4	HBGF-4|HST|HST-1|HSTF1|K-FGF| KFGF
FIGF	NM_004469	factor de crecimiento inducido por c-fos (factor de crecimiento endotelial vascular D)	VEGF-D|VEGFD
FLT1	NM_002019	tirosina quinasa relacionada con fms 1 (receptor de factor de permeabilidad vascular/factor de crecimiento endotelial vascular)	FLT|VEGFR1
FOXA1	NM_004496	caja forkhead A1	HNF3A|MGC33105|TCF3A
GAPDH	NM_002046	gliceraldehído-3-fosfato deshidrogenasa	G3PD|GAPD|MGC88685
GFAP	NM_002055	proteína ácida fibrilar glial	FLJ45472
SLC2A4	NM_001042	familia de vehículo de soluto 2 (transportador de glucosa facilitado), miembro 4	GLUT4
HAND1	NM_004821	expresado por derivados de corazón y cresta neural 1	Hxt|Thing1|bHLHa27|eHand
HIF1A	NM_181054	factor inducible por hipoxia 1, subunidad alfa (factor de transcripción básico de hélice-bucle-hélice)	HIF-1alpha|HIF1|HIF1-ALFA|MOP1|PASD8|bHLHe78
HK2	NM_000189	hexoquinasa 2	DKFZp686M1669|HKII|HXK2
HMGB1	NM_002128	caja de grupo de alta movilidad 1	DKFZp686A04236|HMG1|HMG3|SBP-1
HNF4A	NM_178850	factor nuclear de hepatocitos 4, alfa	FLJ39654|HNF4|HNF4a7|HNF4a8|HNF4a9| HNF4alfa|MODY|MODY1|NR2A1| NR2A21|TCF|TCF14
HPRT1	NM_000194	hipoxantina fosforribosiltransferasa 1	HGPRT|HPRT
HSPB 1	NM_001540	proteína de choque térmico de 27 kDa 1	CMT2F|DKFZp586P1322|HMN2B| HS.76067| HSP27|HSP28|Hsp25|SRP27

ICAM1	NM_000201	molécula de adhesión intracelular 1	BB2|CD54|P3.58
IFNG	NM_000619	interferón, gamma	IFG|IFI
IGF1	NM_001111285	factor de crecimiento de tipo insulina 1 (somatomedina C)	IGF-I|IGF1A|IGFI
IGF2	NM_001127598	factor de crecimiento de tipo insulina 2 (somatomedina A)	C11orf43|FLJ22066|FLJ44734|INSIGF| pp9974
IGFBP3	NM_001013398	proteína de unión a factor de crecimiento de tipo insulina 3	BP-531|BP3
IGFBP5	NM_000599	proteína de unión a factor de crecimiento de tipo insulina 5	IBP5
IKBKB	NM_001556	inhibidor de potenciador de gen de polipéptido ligero kappa en linfocitos B, quinasa beta	FLJ33771|FLJ36218|FLJ38368|FLJ40509| IKK-beta|IKK2|IKKB|MGC131801| NFKBIKB
IL10	NM_000572	interleucina 10	CSIF|IL-10|IL10A|MGC126450|MGC126451|TGIF
IL1B	NM_000576	interleucina 1, beta	IL-1|IL1-BETA|IL1F2
IL3	NM_000588	interleucina 3 (factor estimulante de colonias, múltiple)	IL-3|MCGF|MGC79398|MGC79399|MULTI-CSF
IL4	NM_172348	interleucina 4	BCGF-1|BCGF1|BSF1|IL-4|MGC79402
IL5	NM_000879	interleucina 5 (factor estimulante de colonias, eosinófilos)	EDF|IL-5|TRF
IL6R	NM_181359	receptor de interleucina 6	CD126|IL-6R-1|IL-6R-alfa|IL6RA| MGC104991
IL8	NM_000584	interleucina 8	CXCL8|GCP-1|GCP1|LECT|LUCT| LYNAP| MDNCF |MONAP|NAF|NAP-1|NAP1
ITGA5	NM_002205	integrina, alfa 5 (receptor de fibronectina, polipéptido alfa)	CD49e|FNRA|VLA5A
KDR	NM_002253	receptor de dominio de inserto de quinasa (una tirosina quinasa receptora de tipo III)	CD309|FLK1|VEGFR|VEGFR2
LEP	NM_000230	Leptina	FLJ94114|OB|OBS
LPL	NM_000237	lipoproteína lipasa	HDLCQ11|LIPD
MAPK11	NM_002751	proteína quinasa activada por mitógeno 11	P38B|P38BETA2|PRKM11|SAPK2|SAPK2B|p38-2|p38Beta
MMP1	NM_002421	metalopeptidasa de matriz 1 (colagenasa intersticial)	CLG|CLGN
MMP3	NM_002422	metalopeptidasa de matriz 3 (estromelisina 1, progelatinasa)	CHDS6|MGC126102|MGC126103| MGC126104| MMP-3|SL-1|STMY|STMY1|STR1
MYH1	NM_005963	miosina, cadena pesada 1, músculo esquelético, adulto	MGC133384|MYHSA1|MYHa |MyHC-2X/D| MyHC-2x
MYH11	NM_022844	miosina, cadena pesada 11, músculo liso	AAT4|DKFZp686D10126| DKFZp686D19237| FAA4|FLJ35232|MGC126726|MGC32963|SMHC|SMMHC
MYH7	NM_000257	miosina, cadena pesada 7, músculo cardiaco, beta	CMD1S|CMH1|DKFZp451F047| MGC138376| MGC138378|MPD1|MYHCB|SPMD| SPMM
MYOD1	NM_002478	diferenciación miogénica 1	MYF3|MYOD|PUM|bHLHc1

NFATC1	NM_172390	factor nuclear de linfocitos T activados, citoplasmático, dependiente de calcineurina 1	MGC138448|NF-ATC|NFAT2|NFATc
NFATC2	NM_173091	factor nuclear de linfocitos T activados, citoplasmático, dependiente de calcineurina 2	NFAT1|NFATP
NFKB1	NM_003998	factor nuclear de potenciador de gen de polipéptido ligero kappa en linfocitos B 1	DKFZp686C01211|EBP-1|KBF1| MGC54151| NF-kappa-B|NF p105| NFKB-p50|p105|p50-kappaB|NFKB-
NOS2	NM_000625	óxido nítrico sintasa 2, inducible	HEP-NOS|INOS|NOS|NOS2A
NOTCH1	NM_017617	notch 1	TAN1|hN1
NR3C1	NM_001024094	subfamilia del receptor nuclear 3, grupo C, miembro 1 (receptor de glucocorticoides)	GCCR|GCR|GR|GRL
NRP2	NM_201279	neuropilina 2	MGC126574|NP2|NPN2|PR02714|VEGF165R2
PAX7	NM_013945	caja emparejada 7	FLJ37460|HUP1|PAX7B|RMS2
PDGFB	NM_033016	polipéptido beta de factor de crecimiento derivado de plaquetas (homólogo de oncogén vírico de sarcoma de simio (v-sis))	FLJ12858|PDGF2|SIS|SSV|c-sis
PDK4	NM_002612	piruvato deshidrogenasa quinasa, isozima 4	FLJ40832
PLA2G1B	NM_000928	fosfolipasa A2, grupo IB (páncreas)	MGC119834|MGC119835|PLA2|PLA2A| PPLA2
PLIN1	NM_002666	proteína asociada a gota lipídica	perilipina
PPARG	NM_138712	receptor gamma activado por proliferador de peroxisoma	CIMT1|GLM1|NR1C3|PPARG1|PPARG2| PPARgamma
QARS	NM_005051	glutaminil-ARNt sintetasa	GLNRS|PR02195
RHOA	NM_001664	familia génica del homólogo de ras, miembro A	ARH12|ARHA|RH012|RHOH12
RUNX1	NM_001754	factor de transcripción relacionado con runt 1	AML1|AML1-EVI-1|AMLCR1|CBFA2|EVI-1|PEBP2aB
RXRA	NM_002957	receptor retinoide X, alfa	FLJ00280|FLJ00318|FLJ16020|FLJ16733|MGC102720|NR2B1
SERPINE1	NM 001165413	inhibidor de serpina peptidasa, clado E (nexina, inhibidor de activador de plasminógeno tipo 1), miembro 1	PAI|PAI-1|PAI1|PLANH1
SMAD2	NM_005901	miembro de la familia de SMAD 2	JV18|JV18-1|MADH2|MADR2|MGC22139| MGC34440|hMAD-2|hSMAD2
SMAD4	NM_005359	miembro de la familia de SMAD 4	DPC4|JIP|MADH4
TERT	NM_198255	transcriptasa inversa de telomerasa	EST2|TCS1 |TP2|TRT|hEST2
TGFB1	NM_000660	factor de crecimiento transformante, beta 1	CED|DPD1|LAP|TGFB|TGFbeta

TGFB3	NM_003239	factor de crecimiento transformante, beta 3	ARVD|FLJ16571|TGF-beta3
THBS4	NM_003248	trombospondina 4	TSP4
TNF	NM_000594	factor de necrosis tumoral	DIF|TNF-alfa|TNFA|TNFSF2
TUBB	NM_178014	tubulina, beta	M40|MGC117247|MGC16435|OK/SW-c1.56|TUBB1|TUBB5
TUBB1	NM_030773	tubulina, beta 1	isoforma de tubulina beta (1)
TUBG1	NM_001070	tubulina, gamma 1	GCP-1|TUBG|TUBGCP1
VCAM1	NM_080682	molécula de adhesión de células vasculares 1	CD106|DKTZp779G2333|INCAM-100| MGC99561
VEGFA	NM_003376	factor de crecimiento endotelial vascular A	MGC70609|MVCD1|VEGF|VPF
VIM	NM_003380	vimentina	FLJ36605
WISP1	NM_080838	proteína 1 de ruta de señalización inducible por WNT1	CCN4|WISP1c|WISP1i|WISP1tc
WNT1	NM_005430	familia de sitio de integración de MMTV tipo wingless, miembro 1	INT1

Análisis bioinformático: se convierten datos recuperados en formato .csv de la máquina Biomark de Fluidigm a un formato tabular que incluye información de muestra, ARNm y repetición junto con el valor de fluorescencia en bruto. Las reacciones de PCR que fracasan se marcan como ausentes. Se combinaron experimentos múltiples después de normalizar con respecto a expresión total de especies de ARNm. Toda la expresión de ARNm medida se filtra basándose en el requisito de detección en al menos 2 de todos los replicados biológicos ensayados. Se evaluó la media de desviación técnica, biológica y de conjunto en conjuntos de datos completos.

Para análisis de datos se normalizan en primer lugar los valores de Ct para todos los genes de interés con respecto a los valores de Ct promedio para genes constitutivos a partir de la muestra correspondiente para obtener valores de ΔCt (ΔCt = Ct del gen – Ct de genes constitutivos promedio). Después se normalizan los genes de cada muestra con respecto al mismo gen en control no tratado para obtener valores de ΔΔCt (ΔΔCt = ΔCt de muestra de control - ΔCt de muestra tratada).

Para obtener valores de factor de cambio se someten genes regulados positivamente (es decir ΔΔCt mayores de 0) al siguiente cálculo: factor de cambio = 2^ΔΔCt. Para genes regulados negativamente (es decir ΔΔCt menores de 0): factor de Cambio = -(2^|ΔΔCt|).

ENSAYOS DE PROLIFERACIÓN CELULAR (ENSAYOS A1-A11 EN LAS TABLAS DE DATOS POSTERIORES)

Antecedentes y relevancia terapéutica: la capacidad de modular la tasa de proliferación celular y apoptosis de diferentes tipos celulares representa una propiedad fundamental de muchos compuestos terapéuticos, y es directamente relevante para el tratamiento y prevención de una amplia serie de enfermedades y trastornos.

En consecuencia los polipéptidos de MetRS con la capacidad de modular la tasa de proliferación celular y/o apoptosis tienen utilidad terapéutica significativa en una amplia serie de enfermedades incluyendo, como factores de crecimiento, y factores de diferenciación para células madre, y en regímenes de tratamiento para potenciar o suprimir la proliferación de tipos celulares específicos de interés *in vivo* o *in vitro*, incluyendo por ejemplo células hematopoyéticas, células inmunomoduladoras, cáncer y para el tratamiento y prevención de enfermedades asociadas con envejecimiento, incluyendo por ejemplo neurodegeneración, neuropatía periférica y pérdida de tono muscular y de tejido blando.

Métodos: se evalúan los efectos de los polipéptidos de MetRS en la proliferación celular usando uno o más de los métodos enumerados posteriormente, y como se elabora más específicamente en los métodos posteriores.

Hoechst 33432. Se realizan recuentos celulares convencionales para evaluar la proliferación usando Hoechst 33432, que es una contratinción nuclear permeante celular que emite fluorescencia azul cuando se une a ADNbc. Está disponible como una solución (Invitrogen Cat n.º H-3570) que se usa a una concentración final de 1 µg/ml en medio o PBS. Las células se cultivan en placas de 96 pocillos en presencia de polipéptidos de MetRS durante un tiempo de crecimiento convencional de 48 horas, o más tiempo dependiendo del tipo celular y como se describe en los ejemplos posteriores.

ATP-lite. Los niveles de ATP celulares se correlacionan con salud celular y pueden determinarse fácilmente usando diversos kits disponibles en el mercado. ATP-lite (Perkin-Elmer, Cat n.º 6016947 Boston, MA 02481) que es una mezcla homogénea de solución de lisis y reactivo de detección de ATP se premezcla antes de su uso y se usa a una

relación v:v 1:1 con células cultivadas. Las placas se incuban durante 5 minutos para promover la lisis y las placas se miden usando un lector de placas luminiscente. Las células se cultivan en placas de 96 pocillos en presencia de polipéptidos de MetRS durante un tiempo de crecimiento convencional de 48 horas, o más dependiendo del tipo celular y como se describe en los ejemplos posteriores.

***ALAMARBLUE*®** (Resazurina) es un indicador de viabilidad celular que se basa en el estado redox de las células. Resazurina, el principio activo, es un compuesto permeable a células, no tóxico, que es de color azul y prácticamente no fluorescente cuando está presente en su forma oxidada. Sin embargo tras entrar en células viables normales, la resazurina se reduce rápidamente a resorufina, que produce una señal de fluorescencia roja. Las células viables convierten continuamente resazurina en resorufina, generando de este modo una medida cuantitativa de viabilidad y citotoxicidad. La falta de toxicidad permite exposición a largo plazo de células a resazurina sin impacto negativo; se ha descubierto que las células que han crecido en presencia de resazurina producen números similares de células viables a las células de control, como se determina por análisis citométrico de flujo.

Se realizan mediciones añadiendo una solución de Resazurina/ALAMARBLUE® a células, incubándolas durante 1-4 horas y leyendo la fluorescencia o absorbancia. La cantidad de fluorescencia o absorbancia es proporcional al número de células vivas y corresponde a la actividad metabólica de las células. Las células dañadas y no viables tienen actividad metabólica innata menor y generan por lo tanto una señal proporcionalmente menor que las células sanas. Después de incubación con ALAMARBLUE®, las muestras pueden medirse fácilmente en instrumentación de fluorescencia y absorbancia. Para lecturas de fluorescencia se usan ajustes de filtro de excitación a 530 nm y emisión a 590 nm.

Las células se cultivan en placas de 96 pocillos en presencia de polipéptidos de MetRS durante un tiempo de cultivo convencional de 48 horas o más dependiendo del tipo celular y como se describe en los ejemplos posteriores.

CAPTACIÓN DE LDL ACETILADO EN CÉLULAS HEPATOCÍTICAS HUMANAS HEPG2C3A. (ENSAYO B1 EN LAS TABLAS DE DATOS POSTERIORES)

Antecedentes y relevancia terapéutica: LDL es el principal vehículo de colesterol en la sangre, representando más del 60 % del colesterol en plasma. En seres humanos, el receptor de LDL hepático es responsable de eliminar aproximadamente el 70 % de la LDL en plasma de la circulación. La LDL internalizada se degrada a colesterol libre y aminoácidos en el lisosoma. El hígado es el órgano más importante para catabolismo de LDL y actividad receptora de LDL en seres humanos. La LDL que no está internalizada y permanece en circulación puede transportarse por células endoteliales a la pared vascular, dando como resultado la formación de placas ateroscleróticas. La LDL en circulación también puede captarse por macrófagos y esto también puede contribuir a la formación de placas. Se cree que el aumento de la captación de LDL en tejido hepático es beneficioso para la salud humana y el hallazgo de productos terapéuticos seguros y eficaces que puedan regular positivamente este proceso puede proporcionar nuevas terapias para enfermedades cardiovasculares y metabólicas. Para investigar si las propiedades únicas de polipéptidos de MetRS pueden regular la captación de LDL acetilada, se emplea un ensayo convencional para medir la captación de LDL acetilada en células HepG2C3a.

En consecuencia los polipéptidos de MetRS con la capacidad de modular la captación de LDL tienen utilidad terapéutica significativa en una amplia serie de enfermedades incluyendo por ejemplo el tratamiento de hipercolesteremia, hiperlipidemia, diabetes de tipo 1 y 2, síndrome metabólico y enfermedades vasculares incluyendo aterosclerosis.

Métodos: se mantienen células HEPG2C3a (ATCC n.º CRL-10741) en medio Esencial Mínimo de Eagle (EMEM) complementado con FBS 10 % (HyClone Cat n.º SH30910.03), penicilina 50 u/ml/estreptomicina 50 µg/ml (Invitrogen) en 15 ml de medio en matraces de 75 ml. Las células se cultivan a 37 ºC, CO_2 5 %, en un ambiente humidificado y se utilizan en campanas de cultivo tisular certificadas para BSL2 usando técnica estéril y equipo de protección individual apropiado incluyendo gafas, guantes y batas de laboratorio. HEPG2C3a expresa el receptor de LDL y son competentes para captación de LDL acetilada cuando se cultivan en placas recubiertas con colágeno de fondo transparente. Se siembra un volumen de 100 µl de células en placas recubiertas con colágeno (Invitrogen Cat n.º A11428) durante una noche en medio completo (véase anteriormente) a una densidad celular de 50.000 células/ml. Las células se lavan una vez con PBS (Invitrogen Cat n.º 10010) y se añaden 80 µl de EMEM sin suero a cada pocillo. Se añaden polipéptidos de MetRS a una concentración final de 250 nM por pocillo en un volumen uniforme en PBS estéril a cada pocillo. Se coloca un polipéptido de MetRS único en cada pocillo. Las células se privan de suero y se exponen a los polipéptidos de MetRS durante 16 horas. Después de la incubación de 16 horas, el sobrenadante se recoge y se mide ICAM soluble usando un kit de ELISA convencional de RND Systems (Cat n.º DY643), y se añade medio sin suero complementado con ac-LDL 5 µg/ml (Alexa Fluor 488 marcado Cat n.º L23380, Invitrogen) a cada pocillo. Después de una incubación de 2 horas a 37 ºC con CO_2 5 %, las células se lavan dos veces con PBS estéril antes de añadirse 100 µl de PBS a cada pocillo para cuantificación. Las placas se analizaron con respecto a intensidad de fluorescencia total usando una lectura inferior en un lector de placas fluorescente Victor X5 (Perkin Elmer) a una longitud de onda de excitación centrada alrededor de 485 nm, y una longitud de onda de emisión centrada alrededor de 535 nm. Las células se tiñen con colorante Hoechst y se lee la intensidad fluorescente con Excitación a 405 nm/Emisión a 450 nm para confirmar que el número de células total es uniforme

en toda la placa.

REGULACIÓN DE ESTALLIDO OXIDATIVO DE NEUTRÓFILOS HUMANOS Y PRODUCCIÓN DE ELASTASA (ENSAYOS C1-C3 EN LAS TABLAS DE DATOS POSTERIORES)

ESTALLIDO OXIDATIVO DE NEUTRÓFILOS

Antecedentes y relevancia terapéutica: la fagocitosis por neutrófilos polimorfonucleares y monocitos constituye una rama esencial de la defensa del hospedador contra infecciones por microorganismos incluyendo bacterias y hongos. El proceso fagocítico puede separarse en varios estadios principales: quimiotaxis (migración de fagocitos a sitios inflamatorios), unión de partículas con la superficie celular de fagocitos, ingesta (fagocitosis) y destrucción intracelular por mecanismos dependientes de oxígeno (estallido oxidativo) e independientes de oxígeno. Se observa actividad de estallido reducida o ausente en defectos innatos como la enfermedad granulomatosa crónica (CGD). CGD es un grupo heterogéneo de trastornos heredados que habitualmente se manifiesta durante los primeros dos años de vida. La enfermedad se caracteriza por infecciones repetidas y con peligro para la vida provocadas por organismos bacterianos y fúngicos. Estas infecciones normalmente consisten en neumonía, linfadenitis o abscesos que implican ganglios linfáticos, pulmones e hígado. La NADPH oxidasa es el sistema enzimático responsable de producir anión superóxido, que se convierte rápidamente en peróxido de hidrógeno y radicales de hidroxilo. Las anomalías en los péptidos constituyentes del sistema de enzima NADPH oxidasa conducen a las disfunciones características de CGD. Los neutrófilos de pacientes con CGD no consiguen producir un estallido oxidativo significativo después de la estimulación. Se describen diferentes formas de CGD (CGD ligada a X clásica y patrones recesivos autosómicos). El estallido oxidativo de granulocitos está alterado en trasplante, estadios tardíos de infección por VIH y en los ancianos, haciendo a estas poblaciones más susceptibles de infección secundaria y empeoramientos de enfermedad inflamatoria. Diversos inmunomoduladores (por ejemplo, citocinas (GM-CSF, G-CSF, TNF) o fármacos) también parecen tener efectos en el estallido oxidativo. Existe potencial para proteínas con capacidad de regular positivamente o regular negativamente el estallido oxidativo de una manera terapéutica para que sea útil para diversas patologías diferentes.

Métodos: El ligando de proteína quinasa C forbol 12-miristato-13-acetato (PMA) puede utilizarse en este ensayo como un agonista del proceso de estallido oxidativo. La sangre completa heparinizada se mezcla con dextrano estéril (concentración final de 0,6 %) durante 1 hora y se permite que se separen en capas. La capa inferior contiene neutrófilos, monocitos y glóbulos rojos. Se utiliza una etapa de lisis por cloruro de amonio para retirar todos los RBC y permanece una población pura al 97 % de neutrófilos con aproximadamente 3 % de contaminación de monocitos después de la etapa de lisis. Tras estimulación, los granulocitos y monocitos producen metabolitos de oxígeno reactivos (anión superóxido, peróxido de hidrógeno, ácido hipocloroso) que destruyen bacterias dentro del fagosoma. La formación de los oxidantes reactivos durante el estallido oxidativo puede supervisarse por la adición y oxidación de Amplex Red. Se analiza después el porcentaje de células que tienen radicales de oxígeno reactivos producidos así como su intensidad de fluorescencia media usando un lector de placas fluorescente. El ciclo temporal típico para esta reacción es de 10 minutos, viéndose un estallido evidente a los 2 minutos y viéndose un descenso de la señal a los 20 minutos. Este ensayo puede procesarse en modo agonista en ausencia de PMA o en modo antagonista, con administración simultánea de polipéptidos de MetRS y PMA a una concentración que está por debajo de la CE50 para este compuesto.

REGULACIÓN DE PRODUCCIÓN DE ELASTASA POR NEUTRÓFILOS HUMANOS

Antecedentes y relevancia terapéutica: la elastasa de neutrófilos es una serina proteasa que se ha implicado como poseedora de un papel específico en el desarrollo de una amplia serie de enfermedades humanas, incluyendo trastornos inflamatorios del pulmón y el sistema cardiovascular. Aunque su papel fisiológico clave está en la defensa innata del hospedador, también puede participar en la remodelación tisular y poseer acciones secretagogas que se reconocen ahora como importantes para señales inflamatorias locales. La actividad elastasa de neutrófilos se ha implicado en el desarrollo de enfisema durante varias décadas, sin embargo solamente en momentos relativamente recientes se ha atribuido una fracción patógena a esta serina proteinasa en situaciones en las que se produce deposición de matriz extracelular excesiva. El uso de modelos animales manipulados genéticamente está comenzando a descubrir los modos potenciales en los que sus acciones podrían influir en la reparación de pulmón fibrótico. Las pruebas emergentes sugieren que la implicación de rutas celulares con efectos más directos en la generación de mediadores fibrogenéticos y síntesis de colágeno parece sustentar las acciones de la elastasa de neutrófilos en la promoción de la acumulación de matriz en el pulmón. La elastasa de neutrófilos humanos también está presente dentro de placas ateroscleróticas donde contribuye a la degradación de la matriz y el debilitamiento de la pared de los vasos asociados con las complicaciones de la formación de aneurisma y rotura de placas. Se une por otras proteasas extracelulares en estas acciones pero la amplia serie de sustratos y potencias de esta enzima acoplada con la actividad asociada con desgranulación de neutrófilos destaca esta proteasa disruptiva como diana terapéutica en la enfermedad aterosclerótica.

Métodos: este ensayo usa el Kit de Ensayo de Elastasa ENZCHEK® (Invitrogen Catalog n.º E-12056). Se preparan neutrófilos de sangre humana nueva usando una solución de dextrano al 6 % y se lisan glóbulos rojos antes de sembrar células en medio RPMI (el medio debería estar no complementado sin suero, sin antibióticos). Se prepara

una solución de reserva de 1,0 mg/ml del sustrato de elastina DQ añadiendo 1,0 ml de agua desionizada (dH2O) directamente a uno de los tres viales que contienen el sustrato liofilizado y mezclando para disolver. Se prepara tampón de reacción 1X diluyendo 6 ml del tampón de reacción 10X en 54 ml de dH2O. Se prepara una solución de trabajo de 100 µg/ml del sustrato de elastina DQ diluyendo la solución de reserva de elastina DQ diez veces en Tampón de Reacción 1X. Se prepara solución de reserva de elastasa pancreática porcina preparando una solución de reserva 100 U/ml en dH2O. Para ensayar con respecto a actividad de elastasa, se pipetean 50 µl de tampón de reacción 1X en cada pocillo de ensayo que contiene 500.000 neutrófilos/ml en un volumen de 30 µl. Se añaden 8 µl de cada polipéptido de MetRS por pocillo, y la muestra se incuba durante 20 minutos a 37 ºC. Se añaden 50 µl de solución de trabajo de elastina DQ 100 µg/ml a cada pocillo y se mezclan. Las muestras se incuban a temperatura ambiente, protegidas de la luz, durante 30 minutos. La intensidad de fluorescencia en un lector de microplacas de fluorescencia equipado con filtros de fluoresceína convencionales (ex 485/Em 535) puede medirse sobre múltiples puntos temporales.

UNIÓN A RECEPTORES DE TIPO TOLL Y ACTIVACIÓN DE NFKB (ENSAYOS D1-D4 EN LAS TABLAS DE DATOS POSTERIORES)

Antecedentes y relevancia terapéutica: los macrófagos son agentes importantes en el sistema inmunitario innato y expresan un gran repertorio de clases diferentes de receptores de reconocimiento de patrones (PRR), incluyendo la familia de receptores de tipo Toll (TLR) que son reguladores y controladores potentes de la respuesta inmunitaria.

La estimulación de TLR por patógenos microbianos y ligandos endógenos inicia cascadas de señalización que inducen la secreción de citocinas proinflamatorias y citocinas efectoras que dirigen respuestas inmunitarias adaptativas corriente abajo. Ligandos endógenos, así como componentes microbianos, son reconocidos por y pueden activar TLR, planteando la posibilidad de que estos receptores puedan ser dianas críticas para el desarrollo de nuevas terapias para múltiples enfermedades.

En consecuencia los polipéptidos de MetRS que modulan la actividad del receptor de TLR tienen utilidad terapéutica en una amplia serie de enfermedades y trastornos incluyendo por ejemplo enfermedades y trastornos inflamatorios, enfermedades autoinmunitarias, trasplante de tejidos/rechazo de órganos, prevención o tratamiento de cáncer, la modulación de la hematopoyesis e infección.

Medición de la activación de TLR en células RAW BLUE

Los macrófagos de ratón comercializados con la marca comercial células RAW-BLUE™ (Invivogen, código de Catálogo: raw-sp) expresan todos los TLR excepto TLR5 e incluyen un gen de fosfatasa alcalina embrionaria secretada (SEAP) que es inducible por factores de transcripción NF-kB y AP-1. Tras la estimulación de TLR, las células RAW-BLUE™ activan NF-kB y/o AP-1 lo que conduce a la secreción de SEAP que es medible cuando se usa el medio de detección de SEAP.

Métodos: las células RAW-BLUE™ se lavan dos veces con PBS, se tripsinizan y se resuspenden en medio de cultivo nuevo (medio de cultivo: DMEM, glucosa 4,5 g/l, suero bovino fetal inactivado por calor 10 % (30 minutos a 56 ºC), ZEOCIN™ 100 mg/ml, L-glutamina 2 mM). Las células se siembran a una concentración de 50.000 células/pocillo en una placa de 96 pocillos en un volumen total de 100 µl y se añaden polipéptidos de MetRS, controles o polipéptidos de MetRS (+LPS) a cada pocillo a las concentraciones mostradas en los experimentos perfilados posteriormente. Las células se incuban a 37 ºC en un incubador de CO_2 5 % durante 18 horas. El día experimental 2, se prepara medio de detección de SEAP (QUANTI-BLUE™) (Invivogen código de Catálogo: rep-qb1) siguiendo las instrucciones y se añaden 120 µl por pocillo a una placa de 96 pocillos de fondo plano transparente, y se añade sobrenadante celular (20 µl). Las muestras se incuban a 37 ºC durante aproximadamente 30 minutos hasta 2 horas. Se determinan los niveles de SEAP usando un espectrofotómetro y se lee la absorbancia a 650 nm.

Para detectar polipéptidos de MetRS que bloquean específicamente la activación de TLR este ensayo puede modificarse para identificar antagonistas de TLR potenciales. En este caso se añaden polipéptidos de MetRS a las células a una concentración final de aproximadamente 250 nM por pocillo (o como se especifica de otro modo en los ejemplos posteriores) 1 hora antes de añadir LPS 50 ng/ml. Las células se incuban y se detecta SEAP como se ha descrito anteriormente. Se usan pocillos de control de PBS sin LPS o con polipéptido de MetRS solo añadido para encontrar el nivel basal de estimulación de TLR en el momento de la medición. Los pocillos de control se pretratan con PBS y agonistas y antagonistas de TLR conocidos. La relación de [PBS más señal de LPS] sin fondo frente a [polipéptido de MetRS frente a señal de LPS] se usa para determinar el porcentaje de antagonismo.

Cribado de TLR humanos en células Hek293

Las células HEK293 humanas están modificadas genéticamente y se comercializan con la marca comercial células TLR HEK-Blue™ (Invivogen). Las versiones TLR2 y TLR4 de este tipo de células expresan selectivamente todos los TLR2 o TLR4 e incluyen un gen indicador de fosfatasa alcalina embrionaria secretada (SEAP) bajo el control de un promotor mínimo de IFN-beta que se fusiona con cinco sitios de unión a factores de transcripción de NF-kB y AP-1.

Con el uso de agonistas de TLR2 o 4 específicos (respectivamente), las células HEK-BLUE™ TLR2 y HEK-BLUE™ TLR4 activan NF-kB y/o AP-1 lo que conduce a la secreción de SEAP, que puede medirse cuando se usa reactivo de detección de SEAP. Las células HEK-BLUE™ TLR2 se cotransfectan con la proteína correceptora de LPS CD14 para potenciar la sensibilidad de TLR2 y mejorar la calidad de la señal. La célula parental expresa niveles endógenos de TLR1, 3, 5, 6 y también NOD1.

Métodos: Las células HEK-BLUE™-TLR2 o HEK-BLUE™-TLR4 se lavan dos veces con PBS, se tripsinizan y se resuspenden en medio de cultivo nuevo (medio de cultivo: DMEM, glucosa 4,5 g/l, suero bovino fetal inactivado por calor al 10 % (30 minutos a 56 ºC), ZEOCIN^{TM} 100 mg/ml, L-glutamina 2 mM). Las células se siembran a una concentración de 50.000 células/pocillo en una placa de 96 pocillos en un volumen total de 100 µl, y se añaden polipéptidos de MetRS, controles o polipéptidos de MetRS (+LPS) a cada pocillo a las concentraciones mostradas en los experimentos perfilados posteriormente. Las células se incuban a 37 ºC en un incubador de CO_2 5 % durante 18 horas. El día experimental 2, se prepara medio de detección de SEAP (QUANTI-BLUE™) (Invivogen código de Catálogo: rep-qb1) siguiendo las instrucciones y se añaden 120 µl por pocillo a una placa de 96 pocillos de fondo plano transparente y se añade sobrenadante celular (20 µl). Las muestras se incuban a 37 ºC durante aproximadamente 30 minutos hasta 2 horas. Se determinan los niveles de SEAP usando un espectrofotómetro y leyendo la absorbancia a 650 nm. Se pretratan pocillos de control con PBS y agonistas de TLR conocidos tales como LPS UltraPure (TLR-4) o PAM3CSK4 (TLR-2). La relación del [PBS más señal de LPS] menos el fondo con respecto a [polipéptido de MetRS más señal de LPS] se usa para determinar el porcentaje de agonismo.

LIBERACIÓN DE CITOCINAS (ENSAYOS E1-E16 EN LAS TABLAS DE DATOS POSTERIORES)

Antecedentes y relevancia terapéutica: Las citocinas son un conjunto diverso de moléculas proteicas de señalización celular pequeñas que se usan extensivamente para comunicación intercelular, y desempeñan funciones significativas en la homeostasis corporal normal, incluyendo inmunomodulación y regulación. En consecuencia los polipéptidos de MetRS que modulan la liberación, o las actividades biológicas de citocinas, tienen utilidad terapéutica en una amplia serie de enfermedades y trastornos incluyendo por ejemplo enfermedades y trastornos inflamatorios, enfermedades autoinmunitarias, trasplante de tejidos/rechazo de órganos, prevención o tratamiento del cáncer, la modulación de la hematopoyesis e infección.

Liberación de citocinas de células en cultivo

Métodos: Se siembran células de ensayo en una placa de 24 pocillos a una densidad de aproximadamente 1 millón de células/pocillo en 1 ml de medio de cultivo. Las células se tratan con polipéptido de MetRS (a las concentraciones mostradas en los ejemplos posteriores) o un volumen igual de PBS y se incuban durante una noche a 37 ºC con CO_2 5 %. Después del tratamiento celular, las muestras se centrifugan a 4 ºC en una centrífuga de cubo oscilante a 2.000 x g durante 5 minutos. El medio se retira cuidadosamente para no alterar el sedimento celular y se transfiere a un nuevo tubo. Las muestras se ensayan inmediatamente o se congelan de forma instantánea en nitrógeno líquido para análisis posterior. Se determina la liberación de citocinas (incluyendo las citocinas MIF, IL-8, IL-10, Serpina E1, GM-CSF, GRO, IL-1 alfa, IL-1beta, IL-1ra, IL-6, MCP-1, MIP-1, RANTES y TNFalfa) usando kits disponibles en el mercado (R & D Systems, Inc., MN, Estados Unidos) o mediante una organización de investigación contractual (MD Biosciences (St. Paul, MN)).

Liberación de citocinas de sangre humana completa

Métodos: Se obtiene sangre humana completa de donantes humanos normales y se recoge con heparina en tubos de recogida convencionales. Se usa sangre el mismo día que se recoge para asegurar una salud celular adecuada. La sangre se mezcla suavemente y se siembra en un volumen de 100 µl en placas de fondo en V de policarbonato de 96 pocillos. Se añaden polipéptidos de MetRS y se mezclan lentamente en sangre 2X usando un conjunto de pipetas multicanal en 50 µl. Se usan puntas de filtro para toda la experimentación y se usa EPI completo. Toda la experimentación se produce en una campana de bioseguridad dedicada que es adecuada para experimentación con sangre humana. La sangre se incuba durante una noche a 37 ºC con CO_2 5 %. Después del tratamiento celular, las muestras se centrifugan en una centrífuga de cubo oscilante a 2.000 x g durante 5 minutos. El sobrenadante se recoge para ELISA de citocinas. Se realizan ELISA como se ha descrito previamente.

Liberación de citocinas* de *PBMC

Métodos: Para aislar células mononucleares de sangre periférica se superpone suavemente sangre completa humana recién aislada sobre HISTOPAQUE®-1077 de Sigma a una relación de 1:1 en tubos cónicos de 50 ml a temperatura ambiente. Se centrifugan muestras en capas a 400 x g en una centrífuga clínica de cubo oscilante durante 30 minutos a temperatura ambiente sin freno. La capa celular blanca en la interfaz entre el plasma y el gradiente de densidad se retira después mediante pipeta. Estas células mononucleares de sangre periférica se lavan dos veces con RPMI-1640 (Invitrogen n.º 22400-105) por dilución y centrifugación durante 10 minutos a 250 x g. Las PBMC lavadas se resuspendieron en RPMI-1640 + FBS 10 % y se sembraron a $1x10^6$ células/ml.

Liberación de citocinas de sinoviocitos humanos

Antecedentes y relevancia terapéutica: Un gran número de estudios han demostrado que IL-6 e IL-8 se sobreproducen en varias enfermedades, y por tanto pueden desempeñar un papel fundamental en la patogénesis de enfermedad inflamatoria. IL-6 activa la producción de células endoteliales, lo que conduce a la liberación de IL-8 y proteína quimioatrayente de monocitos, expresión de moléculas de adhesión y reclutamiento de leucocitos a sitios inflamatorios. Estas citocinas se expresan en tipos celulares asociados con enfermedad inflamatoria, incluyendo células implicadas en la patogénesis de artritis juvenil sistémica, lupus eritematoso sistémico, enfermedad de Crohn y artritis reumatoide. Una de las acciones sistémicas más importantes de la producción de citocinas es la inducción de la respuesta de fase aguda. Las proteínas de fase aguda principalmente por el hígado e incluyen proteínas que promueven la respuesta inmunitaria a través de la activación del complemento, inducción de citocinas proinflamatorias y estimulación de quimiotaxis de neutrófilos. Como alternativa, la respuesta de fase aguda puede ser útil, y proteínas de fase aguda, tales como antagonistas de proteinasa, opsoninas y procoagulantes, ayudan a limitar la destrucción tisular resolviendo inflamación. En particular, IL-6 puede estimular la proliferación de sinoviocitos y activación de osteoclastos, lo que conduce a formación de pannus sinovial y reparación. IL-6 actúa con IL-1 para aumentar la producción de metaloproteinasas de la matriz, lo que puede contribuir a la destrucción de articulaciones y cartílagos. Sin embargo, IL-6 también puede tener efectos protectores en la articulación, como sugiere el hallazgo de que esta citocina induce la expresión del inhibidor tisular de metaloproteinasa y estimula la síntesis de proteoglucanos cuando se inyecta en las articulaciones de ratones con artritis inducida por antígenos. Se aíslan sinoviocitos de tipo fibroblasto humano-artritis reumatoide (HFLS-RA) de tejidos sinoviales obtenidos de pacientes con artritis reumatoide (AR). Se crioconservan en segundo pase y pueden cultivarse y propagarse al menos 5 duplicaciones de población. Los HFLS se han conocido durante mucho tiempo por su papel en la destrucción de articulaciones produciendo citocinas y metaloproteinasas que contribuyen a la degradación de cartílagos.

En consecuencia los polipéptidos de MetRS con la capacidad de modular el crecimiento, la diferenciación o el perfil de liberación de citocinas de sinoviocitos de tipo fibroblastos-artritis reumatoide (HFLS-RA) tienen utilidad terapéutica en una amplia serie de enfermedades, incluyendo, por ejemplo, el tratamiento de enfermedades y trastornos inflamatorios Incluyendo artritis juvenil sistémica, lupus eritematoso sistémico, enfermedad de Crohn y artritis reumatoide.

Métodos: se mantienen células adultas, HFLS-RA (Cell Applications Cat n.º 408RA-05a), en medio de cultivo de sinoviocitos (Cell Applications Cat n.º 415-50) en 15 ml de medio en matraces de 125 ml durante 1 pase antes de su uso. Las células se mantienen a 37 ºC, CO_2 5 %, en un ambiente humidificado y se utilizan en campanas de cultivo tisular certificadas para BSL2 usando técnica estéril y equipo de protección individual apropiado incluyendo gafas, guantes y batas de laboratorio. Se siembra un volumen de 80 µl de células en placa durante una noche en medio de cultivo a una densidad celular de aproximadamente 50.000 células/ml. Se añaden polipéptidos de MetRS a una concentración final de 250 nM por pocillo (o como se indica de otro modo en los ejemplos posteriores) en PBS estéril a cada pocillo después de adherencia durante una noche. Los pocillos de control contienen células no tratadas y se incuban con un volumen equivalente de PBS. Las células se exponen a proteínas o PBS en medio basal (Cell Applications Cat n.º 310-470) durante 24 horas. El sobrenadante se retira y se procesan ensayos de ELISA de IL-8, IL-6 y TNFa de acuerdo con las instrucciones del fabricante (RND Systems, Cat n.º DY206 y DY-208, kits Duo-set DY-210). La proliferación se evalúa con resazurina como se ha descrito previamente añadiendo medio nuevo que contiene resazurina a placas después de retirar el sobrenadante e incubando durante tres horas a 37 ºC. Las placas se leen en un lector de placas fluorescentes y se expresa viabilidad/proliferación en función de la fluorescencia asociada a resorufina de pocillos tratados con polipéptido de MetRS dividida por fluorescencia asociada a resorufina de pocillos tratados solamente con PBS.

Proliferación de astrocitos humanos y producción de citocinas inflamatorias

Antecedentes y relevancia terapéutica: Se derivan astrocitos humanos (AH) de corteza cerebral humana. Se crioconservan en el segundo pase y pueden cultivarse y propagarse 10 duplicaciones de población. Los AH son las células más abundantes en el sistema nervioso central y realizan muchas funciones tales como provisión de apoyo mecánico y nutrientes a neuronas, y retirada de residuos de neuronas. Además de desempeñar un papel de apoyo clínico para el funcionamiento neuronal óptimo, también proporcionan soporte bioquímico de células endoteliales que forman la barrera hematoencefálica. Estudios recientes han mostrado que los astrocitos son capaces de regular la neurogénesis instruyendo a las células madre para que adopten un destino neuronal y controlando la función de sinapsis individuales, participan activamente en la transferencia y almacenamiento de información en el cerebro. Está aumentando el reconocimiento de la importancia de los astrocitos en el funcionamiento del sistema nervioso, AH pueden actuar como modelo *in vitro* útil para explorar la diversidad de funciones de los astrocitos. Se ha mostrado que los astrocitos proliferan en respuesta a IL6 y TNFalfa. Además, estas células son capaces de realizar sus propios IL6 y TNFalfa. Por tanto, los polipéptidos de MetRS que modulan la proliferación y producción de citocinas en AH tienen utilidad terapéutica en diversas enfermedades neurológicas incluyendo neuroinflamación, neurodegeneración, tumorigénesis del cerebro e isquemia y reparación cerebral.

Métodos: Se mantienen astrocitos humanos (AH) de Cell Applications (Cat n.º 882K-05f) en medio de cultivo celular

de AH Cell Applications (Cat n.º 821-500) de acuerdo con las instrucciones del fabricante. Las células se mantienen a 37 ºC, CO_2 5 %, en un ambiente humidificado y se utilizan en campanas de cultivo tisular certificadas para BSL2 usando técnica estéril y equipo de protección individual adecuado, incluyendo gafas, guantes y batas de laboratorio. Se siembra un volumen de 80 µl de células en placas recubiertas con colágeno durante una noche en medio completo (véase anteriormente) a una densidad celular de 50.000 células/ml. Las células se lavan una vez con PBS y se añaden 80 µl de medio de cultivo sin suero a cada pocillo. Se añaden polipéptidos de MetRS a una concentración final de 250 nM por pocillo (o como se describe de otro modo en los ejemplos posteriores) en un volumen uniforme en PBS estéril a cada pocillo. Las células se exponen a polipéptidos de MetRS durante 48 horas y se retira el medio gastado para evaluación de citocinas (como se ha descrito previamente). Las células se exponen a proteínas o PBS en medio basal (Cell Applications Cat n.º 310-470) durante 48 horas. El sobrenadante se retira y se realizan ensayos de ELISA de IL-8 e IL-6 de acuerdo con las instrucciones del fabricante (RND Systems, Cat n.º DY206 y DY-208, kits Duo-set DY-210). Se evalúa la proliferación con resazurina como se ha descrito previamente añadiendo medio nuevo que contiene resazurina a placas después de retirar el sobrenadante e incubando durante tres horas a 37 ºC. Las placas se leen en un lector de placas fluorescente y se expresa la viabilidad/proliferación en función de la fluorescencia asociada a resorufina de pocillos tratados con polipéptido de MetRS dividida por fluorescencia asociada a resorufina de pocillos tratados solamente con PBS.

PROLIFERACIÓN DE CÉLULAS ENDOTELIALES MICROVASCULARES DE PULMÓN HUMANO (HLMVEC) Y PRODUCCIÓN DE CITOCINAS INFLAMATORIAS

Antecedentes y relevancia terapéutica: La vasculatura pulmonar es de gran significación fisiológica/patológica. Se reconoce ahora que es un tejido compuesto de células funcionalmente sensibles, metabólicamente activas, que interaccionan con sustratos en circulación y elementos formados de maneras que regulan la composición de la sangre arterial sistémica, afectan a funciones de órganos diana y contribuyen a trombosis, hemostasia y reacciones inmunitarias, así como metástasis tumoral. Las células endoteliales microvasculares de pulmón humano (HLMVEC) muestran expresión elevada de citocinas quimioatrayentes y moléculas de adhesión celular que proporcionan señales críticas para la migración dirigida de leucocitos al pulmón durante la lesión pulmonar aguda. Este tipo celular primario puede ser una herramienta útil para estudiar diversos aspectos de la patología y biología de la microvasculatura pulmonar *in vitro*. Se cree que la alteración de la estructura y función de la microvasculatura en respuesta a estímulos inflamatorios es un factor clave en el daño de órganos y en condiciones apropiadas puede proporcionar un estímulo para reparación. Una causa significativa de estas alteraciones vasculares es la inducción de una reacción inflamatoria que implica infiltración de leucocitos. Diversos estudios centrados en la adhesión de granulocitos al endotelio ha revelado que el reclutamiento y la emigración de leucocitos implican una cascada de adhesión bien orquestrada. La cascada de adhesión comienza cuando el granulocito se une al endotelio y comienza a rodar en la dirección de flujo de fluido a una baja velocidad. A medida que el granulocito rueda, se activa, se adhiere firmemente posteriormente al endotelio y migra a través del endotelio al espacio extravascular. Estos acontecimientos de adhesión están mediados, en parte, por interacciones moleculares que suceden entre CAM en la superficie de los granulocitos y glucoproteínas afines presentes en el endotelio. Diversos estudios ha revelado que la molécula de adhesión celular endotelial E-selectina puede interaccionar con ligandos de granulocitos que presentan glucano de tipo SLex para mediar en las etapas de unión y rodado de la cascada de adhesión. Las etapas corriente abajo de la cascada implican la interacción de molécula de adhesión intercelular expresada en el endotelio con integrinas CD 18 expresadas en granulocitos.

Por lo tanto los polipéptidos de MetRS que modulan la proliferación y/o producción de citocinas de células endoteliales microvasculares pulmonares humanas tienen utilidad terapéutica en diversas enfermedades vasculares y pulmonares incluyendo enfermedades inflamatorias y obstructivas del pulmón incluyendo por ejemplo, hipertensión pulmonar, enfermedad pulmonar obstructiva crónica, fibrosis pulmonar idiopática y asma.

Métodos: Se mantienen HLMVEC (Cell Applications, n.º de Catálogo 540-05) en medio de cultivo de células endoteliales microvasculares Cell Applications (Cat n.º 111-500). Para crecimiento apropiado, se usa una Solución de Factor de Unión que contiene colágeno (Cell Applications, n.º de Catálogo 123-100), para recubrir placas y matraces antes de sembrar células. Las células se mantienen a 37 ºC, CO_2 5 %, en un ambiente humidificado y se utilizan en campanas de cultivo tisular certificadas para BSL2 usando técnica estéril y equipo de protección individual apropiado incluyendo gafas, guantes y batas de laboratorio. Se siembra un volumen de 80 µl de células en placas recubiertas con colágeno durante una noche en medio completo (véase anteriormente) a una densidad celular de 50.000 células/ml. Las células se lavan una vez con PBS y se añaden 80 µl de medio de cultivo sin suero a cada pocillo. Se añaden polipéptidos de MetRS a una concentración final de 250 nM por pocillo (o como se describe de otro modo en los ejemplos posteriores) en un volumen uniforme en PBS estéril a cada pocillo. Las células se exponen a polipéptidos de MetRS durante 48 horas y se retira el medio gastado para ELISA para moléculas de adhesión celular y evaluación de citocinas (como se ha descrito previamente). Se miden las moléculas de adhesión celular incluyendo VCAM y/o ICAM solubles usando un kit de ELISA convencional de RND Systems (Cat n.º DY643 y DY720, respectivamente). La proliferación se evalúa con resazurina como se ha descrito previamente añadiendo medio nuevo que contiene resazurina a placas después de retirar el sobrenadante e incubando durante tres horas a 37 ºC. Las placas se leen en un lector de placas fluorescente y se expresa la viabilidad/proliferación en función de la fluorescencia asociada con resorufina de pocillos tratados con polipéptido de MetRS dividida por fluorescencia asociada a resorufina de pocillos tratados solamente con PBS.

ADHESIÓN CELULAR (ENSAYOS F1-F7 EN LAS TABLAS DE DATOS POSTERIORES)

Antecedentes y relevancia terapéutica: Las moléculas de adhesión celular (CAM) son proteínas localizadas en la superficie celular que están implicadas en la unión a otras células o con la matriz extracelular (ECM) en el proceso denominado adhesión celular. Estas proteínas son normalmente receptores transmembrana y están compuestas de tres dominios: un dominio intercelular que interacciona con el citoesqueleto, un dominio transmembrana y un dominio extracelular que interacciona con otras CAM del mismo tipo (unión homófila) o con otras CAM o la matriz extracelular (unión heterófila). La mayoría de las CAM pertenecen a cuatro familias de la proteína: superfamilia de Ig (inmunoglobulina) (CAM IgSF), las integrinas, las cadherinas y las selectinas. Las moléculas de adhesión celular de superfamilia de inmunoglobulinas (IgSF) son glucoproteínas transmembrana independientes de calcio, que incluyen: moléculas de adhesión a células neurales (NCAM), moléculas de adhesión a células intercelulares (ICAM), moléculas de adhesión células vasculares (VCAM), moléculas de adhesión a células endoteliales-plaquetas (PECAM-1), molécula de adhesión selectiva a células endoteliales (ESAM), molécula de adhesión de punto de unión (JAM), nectinas y otras moléculas de adhesión celular.

Las moléculas de adhesión celular son glucoproteínas de superficie celular que son críticas para la adhesión de leucocitos al endotelio sinusoide y transmigración y citotoxicidad en diversas enfermedades hepáticas inflamatorias. ICAM-1 desempeña un papel importante en la inflamación, y la expresión aumentada de ICAM-1 en células endoteliales se refleja en la activación de células endoteliales. ICAM-1 es particularmente importante ya que media en la adhesión endotelial firme y facilita la transmigración de leucocitos. Los estudios han mostrado que hay una regulación positiva de ICAM-1 tanto en células sinusoides como en hepatocitos en afecciones inflamatorias del hígado tales como infección vírica de hepatitis B, trastornos hepáticos autoinmunitarios, hepatitis alcohólica y rechazo de aloinjerto de hígado.

Por tanto los polipéptidos de MetRS que modulan la producción de moléculas de adhesión celular y adhesión celular a células endoteliales tienen utilidad terapéutica en diversas enfermedades inflamatorias incluyendo por ejemplo enfermedades cardiovasculares, aterosclerosis, autoinmunidad e hipertensión pulmonar.

Métodos: Se siembran células de vena umbilical humanas (ATCC, Cat n.º CRL-2873) (HUVEC) a una concentración de aproximadamente 1,2 x 10^5 células/pocillo en placas de 12 pocillos recubiertas con solución de unión a fibronectina humana en el medio ATCC sugerido y complementos y se cultiva de acuerdo con las instrucciones del fabricante. Las células se estimulan con polipéptidos de MetRS a las concentraciones indicadas, o solamente con PBS, y se incuban durante una noche en medio de cultivo. Se resuspenden células de leucemia monocítica aguda humana (THP-1 (TIB-202)) en medio sin suero RPMI/BSA a 0,1 % con calceína AM (6 µl/ml; Invitrogen Cat n.º C1430) y se incuban durante 30 minutos. Las células marcadas se recogen y se resuspenden en medio RPMI que contiene FBS a 10 %, y la densidad se ajusta a 2 x 10^6 células/ml.

Se colocan 100 µl (2 x 10^5) de células THP-1 marcadas en cada pocillo de la monocapa de HUVEC en 1 ml de medio de cultivo y se incuban durante 15 minutos. Los pocillos se lavan dos veces con PBS para retirar células no unidas, y después las células se leen por un lector de placas fluorescente con una longitud de onda de excitación de 488 nm y una longitud de onda de emisión de 530 nm.

DIFERENCIACIÓN CELULAR (ENSAYOS G1-G4 EN LAS TABLAS DE DATOS POSTERIORES)

Diferenciación y proliferación de adipocitos en células preadipocíticas humanas primarias.

Antecedentes y relevancia terapéutica: Tanto la obesidad como la lipodistrofia están asociadas habitualmente con patologías que incluyen diabetes y enfermedades cardiovasculares. Se reconoce ahora que el tejido adiposo es un órgano endocrino que secreta una amplia diversidad de factores, y la secreción desregulada afecta a la adipogénesis así como a la homeostasis de insulina/glucosa de todo el cuerpo. El exceso de tejido adiposo que conduce a obesidad se ha convertido en un riesgo grave para la salud pública. El desarrollo de tejido adiposo puede verse afectado por el fondo genéticos, el equilibrio hormonal, la dieta y la actividad física. La masa de tejido adiposo puede aumentar cuando las células grasas aumentan de tamaño debido a la mayor acumulación de triacilglicerol. Además, un aumento en el número de células grasas, que surge de la diferenciación de las células precursoras en adipocitos, también puede producirse incluso en adultos como se observa en obesidad humana grave y en roedores a los que se alimenta con una dieta alta en carbohidratos o alta en grasas. Se cree que los adipocitos específicamente surgen de células mesenquimales que experimentan el proceso de compromiso y diferenciación, adipogénesis. Las líneas celulares preadipocíticas pueden experimentar diferenciación en adipocitos tras el tratamiento con agentes adipogénicos constituidos por glucocorticoide sintético, dexametasona (DEX), isobutilmetilxantina (IBMX) e insulina, y han sido valiosas en estos estudios. Se ha establecido firmemente que el receptor γ activado por proliferador de peroxisoma (PPARγ) y la familia de proteínas de unión a potenciador CCAAT (C/EBP) de factores de transcripción desempeña funciones críticas en la diferenciación de adipocitos. Temprano durante la diferenciación de adipocitos, C/EBPβ y C/EBPδ se inducen por DEX e IBMX, respectivamente, que juntos inducen después PPARγ y C/EBPa para activar diversos marcadores de adipocitos que se requieren para la función de adipocitos. También se ha indicado que otros factores de transcripción regulan positiva o negativamente la adipogénesis y diversos factores de crecimiento y hormonas pueden afectar a la diferenciación de adipocitos

regulando la expresión de factores de transcripción adipogénicos. De hecho, además de ser el principal sitio para el almacenamiento de energía en mamíferos almacenando triacilglicerol y liberando ácidos grasos cuando se necesiten, el tejido adiposo secreta una amplia serie de moléculas que están implicadas en diversos procesos fisiológicos incluyendo respuesta inmunitaria, función vascular y homeostasis energética. Los adipocitos secretan citocinas tales como TNF-a e IL-6. Algunos de esos factores también pueden afectar al crecimiento y desarrollo de tejido adiposo por acción autocrina/paracrina.

Por tanto los polipéptidos de MetRS que tienen la capacidad de modular la diferenciación y/o proliferación de pre-adipocitos humanos normales tienen utilidad terapéutica en una amplia serie de enfermedades incluyendo, por ejemplo, el tratamiento y prevención de enfermedades metabólicas, enfermedades cardiovasculares, obesidad y lipodistrofias, así como las complicaciones a largo plazo de la diabetes.

Métodos: Se mantienen HPAd (preadipocitos humanos) (Cell Application Cat n.º 803sD) de acuerdo con las instrucciones del proveedor. Para el cultivo, las células se descongelan rápidamente, y se transfieren inmediatamente a 15 ml de medio de cultivo de adipocitos (Cell Application Cat n.º 811M-250) y se siembran en un matraz tratado con cultivo tisular estéril convencional. El medio se reemplaza con medio de cultivo de adipocitos nuevo cada dos días hasta que la célula es >60 % confluente. Las células se cultivan a 37 ºC, CO_2 5 %, en un ambiente humidificado y se utilizan en campanas de cultivo tisular certificadas BSL2 usando técnica estéril y equipo de protección individual apropiado, incluyendo gafas, guantes y batas de laboratorio. Las células se siembran en placas de ensayo tratadas con cultivo tisular de 96 pocillos de paredes negras de fondo transparente para diferenciación a una concentración de aproximadamente 50.000 células/ml. Se añaden polipéptidos de MetRS a una concentración final de 250 nM por pocillo (o como se indica de otro modo en los Ejemplos posteriores) a cada pocillo de ensayo. Todas las células se mantienen en medio de cultivo durante 2 días con la excepción de los controles positivos que se estimulan con medios de diferenciación adipogénicos (Cell Applications Cat n.º 811D-250). Las células se exponen a polipéptidos de MetRS durante 48 horas. Se miden moléculas de adhesión celular que incluyen VCAM y/o ICAM solubles usando un kit ELISA convencional de RND Systems (Cat n.º DY643 y DY720, respectivamente). La proliferación se evalúa con resazurina como se ha descrito previamente añadiendo medio nuevo que contiene resazurina a placas después de retirar sobrenadante e incubando durante tres horas a 37 ºC. Las placas se leen en un lector de placas fluorescente y se expresa la viabilidad/proliferación en función de la fluorescencia asociada con resorufina de pocillos tratados con polipéptido de MetRS dividida por la fluorescencia asociada con resorufina de pocillos tratados solamente con PBS. Se añade medio nuevo y se mantiene la diferenciación durante 16 días después del intercambio de medio inicial, intercambiando el medio nuevo cada dos días para mantener la salud celular. El día 15, las células se colocan en medios sin suero. El día 16, se evalúa la diferenciación a adipocitos maduros con tinción de Nile Red (Invitrogen, concentración de 3 µM final) y se cuantifica con un lector de placas fluorescente con las longitudes de onda apropiadas. Para realizar este ensayo las células se fijan con paraformaldehído al 10 %, se lavan en PBS y se permeabilizan en PBS que contiene BSA 0,5 % y Triton X-100 0,1 %. La proliferación celular se evalúa con una medida de la intensidad en un lector fluorescente con colorante Hoechst 33432 a una concentración de 1 µg/ml final, como se ha descrito previamente. Se expresa adipogénesis como intensidad de señal de Nile Red. Se usa señal de colorante Hoechst para evaluar el número celular.

Diferenciación y proliferación de células de músculo esquelético humano.

Antecedentes y relevancia terapéutica: el desarrollo del músculo esquelético es un proceso multietapa que implica la determinación de células mesodérmicas pluripotenciales para dar lugar a mioblastos, retirada de los mioblastos del ciclo celular y diferenciación en células musculares, y finalmente crecimiento y maduración de fibras de músculo esquelético. La diferenciación de músculo esquelético implica alineamientos de mioblastos, elongación y fusión en miotubos multinucleados, junto con la inducción de genes específicos de músculo reguladores y estructurales. En el nivel molecular, el compromiso miogénico y la expresión génica específica de músculo implican la familia de proteínas MyoD de hélice-bucle-hélice específica de músculo esquelético (bHLH), que incluye MyoD, miogenina, myf-5 y MRF4 y el factor de unión a potenciador de miocito 2 (MEF2). La actividad de unión a ADN de proteínas de la familia MyoD está atenuada por Id, que forma complejos con productos génicos de E2a en células en proliferación y está regulado negativamente cuando se induce su diferenciación. La decisión de diferenciar en miotubos está influida negativamente por varios factores. Se sabe que el tratamiento de mioblastos con suero bovino fetal, factor de crecimiento de fibroblastos básico 2 o factor de crecimiento transformante β1 inhibe la diferenciación de mioblastos. La miogénesis también se regula negativamente por oncogenes tales como c-myc, c-jun, c-fos, H-ras y E1a. Hay muy poca información con respecto a la señalización que se desencadena en el mioblasto tras retirar el suero lo que conduce a la inducción de la expresión génica de la familia de MyoD y la diferenciación muscular. La diferenciación miogénica parece depender de la activación de integrinas presentes en la membrana plasmática de mioblastos lo que sugiere la actuación de una ruta bioquímica “de fuera a dentro” en la que la integrina es la especie molecular corriente arriba. Las interacciones del factor de crecimiento de tipo insulina (IGF)-I y -II con sus receptores también son reguladores positivos de la diferenciación del músculo esquelético.

En consecuencia, los polipéptidos de MetRS con la capacidad de modular el desarrollo muscular tienen utilidad terapéutica en una amplia serie de enfermedades incluyendo por ejemplo el tratamiento de enfermedad metabólica, caquexia, diversas afecciones de debilitamiento muscular, así como enfermedad musculoesquelética en la que la atrofia muscular desempeña un papel clave en la patogénesis y sintomatología. Las células de músculo esquelético

humano (HSkMC) pueden experimentar diferenciación para mostrar miofilamentos de actina y miosina. Las HSkMC se han usado en el estudio de enfermedades musculares genéticas tales como hipertermia maligna. HSkMC también tienen el potencial de actuar como un injerto cardíaco, reparar daño al corazón y por lo tanto los polipéptidos de MetRS con la capacidad de modular el desarrollo muscular también tienen utilidad como reguladores de la miogénesis *in vitro* e *in vivo*.

Métodos: para evaluar el papel potencial de polipéptidos de MetRS en este proceso, se empleó un ensayo convencional de diferenciación de células del músculo esquelético. Para este ensayo, se aíslan células de músculo esquelético adulto humano (HSkMC, Cell Application Cat n.º 150-05f) de donantes humanos sanos de músculo esquelético de las extremidades. Las células se mantienen en medio de cultivo de HSkMC (Cell Applications, Cat n.º 151-500). Estas células pueden cultivarse y propagarse durante al menos 15 duplicaciones de población. Para diferenciación, las células se mantienen en medio de cultivo durante un pase y después se siembran a 50.000 células por ml de medio en placas tratadas con TC de paredes negras de fondo transparente de 96 pocillos tratadas con colágeno a 100 µl por pocillo. Se permite que las células se adhieran durante una noche. Se añaden polipéptidos de MetRS en PBS, o solamente PBS, a cada pocillo a una concentración final de proteína de 250 nM (o como se indica de otro modo en los ejemplos posteriores). Los pocillos de control recibieron el mismo volumen de Medio de Diferenciación (Cell Applications Cat n.º 151D-250) en este momento. Las células se incuban con proteína o medio de diferenciación durante 48 horas. A las 48 horas, se recoge sobrenadante de cultivo celular de todos los pocillos y se añade medio de diferenciación a un volumen de 150 µl a la placa completa con la excepción de pocillos de control que se mantienen en medio de cultivo solamente. Se utiliza sobrenadante para evaluar la producción de citocinas incluyendo IL6 e IL8 como se ha descrito previamente. La proliferación se evalúa con resazurina como se ha descrito previamente añadiendo medio nuevo que contiene resazurina a placas después de retirada del sobrenadante e incubando durante tres horas a 37 ºC. Las células se supervisan bajo el microscopio y se intercambia medio por medio de diferenciación nuevo cada 2 días. El día 10, se retira el medio, y las células se fijan con paraformaldehído al 10 % durante 30 minutos. Las células se permeabilizan con Triton X-100 0,1 % en PBS durante 15 minutos y las células se tiñen con faloidina marcada con TR y Hoechst 33432 (como se ha descrito previamente) para definir la actina y los núcleos respectivamente. Se usa la intensidad nuclear para determinar la proliferación celular en cada pocillo y se usa la intensidad de faloidina para determinar el contenido de actina total. Las células también se tiñen con anticuerpo de alfa actina de músculo esquelético (GenTex Cat n.º GTX101362). Se toman fotos digitales usando un microscopio fluorescente así como inspecciones visuales y puntuación de todos los pocillos.

Diferenciación y proliferación de células madre mesenquimales de médula ósea humana.

Antecedentes y relevancia terapéutica: las células madre mesenquimales (MSC) son células madre multipotentes que pueden diferenciarse en diversos tipos celulares, incluyendo osteoblastos, condrocitos, miocitos, adipocitos, células de islotes beta pancreáticos y potencialmente células neuronales. Muchos acontecimientos diferentes contribuyen al compromiso de las MSC a otros linajes incluyendo la coordinación de una red compleja de factores de transcripción, cofactores e intermedios de señalización de numerosas vías. Las MSC tienen un gran interés terapéutico porque representan una población de células con el potencial para tratar una amplia serie de enfermedades agudas y degenerativas.

Además los polipéptidos de MetRS con la capacidad de modular la diferenciación de MSC en diferentes rutas del desarrollo tienen utilidad terapéutica significativa para permitir la modulación *in vitro* o *in vivo* de hematopoyesis, neurogénesis, miogénesis, osteogénesis y adipogénesis, así como en una amplia serie de trastornos y enfermedades incluyendo por ejemplo respuestas inflamatorias, autoinmunidad, cáncer, degeneración neuronal, distrofia muscular, osteoporosis y lipodistrofia. Las MSC humanas están inmunoprivilegiadas, y representan un tipo celular ventajoso para trasplante alogénico, reduciendo el riesgo de rechazo y complicaciones del trasplante. Recientemente, también ha habido avances significativos en el uso de células madre mesenquimales autólogas para regenerar tejidos humanos, incluyendo cartílago y menisco, tendones y fracturas óseas. Muchos estudios también han investigado el uso de MSC para terapia génica, incluyendo trasplante de MSC transfectadas con factor de crecimiento endotelial vascular para la mejora de la función cardíaca después de MI en ratas, MSC como vehículos para el suministro de interferón β a tumores en ratones y terapia génica con MSC que expresan BMP para promover la formación de hueso. En consecuencia debido al gran interés en las MSC como productos terapéuticos directos y modificados, así como el potencial de polipéptidos de MetRS para actuar como agentes terapéuticos para regular la diferenciación de MSC *in vivo*, los polipéptidos de MetRS se ensayaron como inductores potenciales de proliferación y diferenciación de MSC.

Métodos: se mantienen hMSC (células del estroma de la médula ósea) (Cell Application Cat n.º 492-05f) de acuerdo con las instrucciones del proveedor. Para cultivar, las células se descongelan rápidamente, y se transfieren inmediatamente a 15 ml de medio de cultivo de células del estroma de la médula ósea (Cell Application Cat n.º 419-500) y se siembran en un matraz tratado con cultivo tisular estéril convencional. Se reemplaza medio con medio de cultivo de células del estroma de la médula ósea nuevo cada dos días hasta que las células son confluyentes en > 60 %. Las células se cultivan a 37 ºC, CO_2 5 %, en un ambiente humidificado y se utilizan en campanas de cultivo tisular certificadas para BSL2 usando técnica estéril y equipo de protección individual apropiado incluyendo gafas, guantes y batas de laboratorio. Las células se siembran en placas de ensayo tratadas con cultivo tisular de 96

pocillos de paredes negras de fondo transparente para diferenciación a una concentración de 50.000 células/ml. Se añaden proteínas derivadas de ARNt sintetasa a una concentración final de 250 nM por pocillo (o como se especifica de otro modo en los Ejemplos posteriores) a cada pocillo de ensayo. Todas las células se mantienen en medio de cultivo durante 2 días con la excepción de los controles positivos, que se estimularon con medio de diferenciación osteogénico o condrogénico (StemPro, Invitrogen, Cat n.º A10072-01 y A10071-01, respectivamente). Las células se exponen a polipéptidos de MetRS durante 48 horas. Se mide VCAM soluble usando un kit de ELISA convencional de RND Systems (Cat n.º DY643). La proliferación se evalúa con resazurina como se ha descrito previamente añadiendo medio nuevo que contiene resazurina a placas después de retirar el sobrenadante e incubando durante tres horas a 37 ºC. Las placas se leen en un lector de placas fluorescente y se expresa la viabilidad/proliferación en función de fluorescencia asociada a resorufina de pocillos tratados con polipéptido de MetRS dividida por la fluorescencia asociada a resorufina de pocillos tratados solamente con PBS. Después de una evaluación de la viabilidad celular, se retira resazurina con dos intercambios de medio y se añade medio de diferenciación 0,5X a todos los pocillos. La diferenciación se supervisa por inspecciones visuales de todos los pocillos durante 10 días después de intercambio de medio con medio nuevo intercambiado cada dos días para mantener la salud celular. La diferenciación se evaluó con tinción con fosfatasa alcalina usando tinción ELF-97 (Invitrogen Cat n.º E6601) al día 10 después del primer intercambio de diferenciación. (Yang *et al*., Nature Protocols (6) 187-213 (2011) doi:10.1038/nprot.2010.189).

Proliferación y diferenciación de células de músculo liso de arteria pulmonar humana (hPASMC).

Antecedentes y relevancia terapéutica: las células de músculo liso de la arteria pulmonar (PASMC) en vasos sanguíneos de pulmón adulto humano normales son principalmente estáticas, no migratorias y están comprometidas principalmente a ejecutar su función contráctil en el pulmón. Sin embargo, las PASMC no están diferenciadas terminalmente y poseen la capacidad de modular su fenotipo y salir de su estado estático en respuesta a cambios en las señales ambientales locales. Este estado de diferenciación puede producirse en el desarrollo, lesión tisular y remodelación de vasos en respuesta a cambios en la demanda tisular. La hipertensión pulmonar (HP) se asocia con diversas afecciones subyacentes incluyendo un aumento de la resistencia vascular pulmonar periférica como resultado de tono vascular aumentado y contractilidad de PASMC y remodelación vascular. La remodelación vascular implica crecimiento de PASMC, síntesis de material de matriz y alteraciones en las interacciones célula-célula y célula-matriz en las paredes de arterias pulmonares (AP) pequeñas, lo que conduce a aumento del grosor del componente de músculo liso de la pared vascular y muscularización anómala de las AP distantes, no muscularizadas normalmente. Este proceso contribuye a reducción del diámetro del lumen y aumento de la resistencia periférica. Aunque el papel preciso de las PASMC en la causa inicial de la enfermedad es controvertido, los cambios que se producen desempeñan un papel clave en las consecuencias clínicas de la enfermedad. Una etapa crucial en el estudio de la diferenciación celular es identificar un conjunto de genes específicos de células o selectivos de células que contribuyen a la función o las funciones diferenciadas de la célula. Se han identificado diversos genes de células de músculo liso (SMC) que actúan como marcadores útiles del estado relativo de diferenciación o maduración de las SMC vasculares, tales como alfa-actina de SM, MHC de SM, h1-calponina, SM22-alfa, desmina, metavinculina, smoothelina y otras. El marcador más ampliamente usado es alfa actina de SM, parcialmente debido a la disponibilidad comercial de varios anticuerpos de muy alta afinidad y altamente selectivos para esta proteína. Sigue siendo una cuestión no resuelta si los cambios en las PASMC resultan de sus características inherentes o de la desregulación de acontecimientos moleculares que gobiernan el crecimiento de PASMC. Sin embargo la determinación de las señales reguladoras y el control de la desregulación potencial proporciona información terapéutica significativa para tratar diversas enfermedades vasculares y pulmonares incluyendo hipertensión pulmonar, enfermedades vasculares.

Por lo tanto los polipéptidos de MetRS que tienen la capacidad de modular la diferenciación y/o proliferación de PASMC humanas normales derivadas de humanos adultos tienen utilidad terapéutica en diversas enfermedades vasculares y pulmonares incluyendo enfermedades pulmonares inflamatorias y obstructivas incluyendo por ejemplo hipertensión pulmonar, enfermedad pulmonar obstructiva crónica, fibrosis pulmonar idiopática y asma.

Métodos: se mantienen HPASMC (Cell Application Cat n.º 352-05a) en medio de cultivo de HPASMC (Cell Applications Cat n.º 352-05a) en 15 ml de medio en matraces de 125 ml durante 1 pase antes de su uso. Las células se mantienen a 37 ºC, CO_2 5 %, en un ambiente humidificado y se utilizan en campanas de cultivo tisular certificadas para BSL2 usando técnica estéril y equipo de protección individual apropiado incluyendo gafas, guantes y batas de laboratorio. Se siembra un volumen de 80 µl de células en colágeno aplicado durante una noche en medio de cultivo a una densidad celular de 50.000 células/ml. Se añadieron polipéptidos de MetRS en PBS estéril a cada pocillo a una concentración final de 250 nM (o como se especifica de otro modo en los Ejemplos posteriores). Los pocillos de control contuvieron solamente un volumen equivalente de PBS. Se incubaron muestras de control positivas con medio de diferenciación de HPASMC proporcionado por el proveedor (Cell Applications Cat n.º 311D-250). Las células se exponen a polipéptidos de MetRS o PBS en medio basal (Cell Applications Cat n.º 310-470) durante 48 horas seguido de un cambio de medio a medio de diferenciación para la placa completa. El sobrenadante se recoge y se utiliza para evaluar la producción de citocinas incluyendo IL6 e IL8 como se ha descrito previamente. La proliferación se evalúa con resazurina como se ha descrito previamente añadiendo medio nuevo que contiene resazurina a placas después de retirar el sobrenadante e incubando durante tres horas a 37 ºC. Las células se supervisan durante 10 días con un intercambio de medio cada dos días. La diferenciación se evalúa después de

fijación como se ha descrito anteriormente, y permeabilización con Triton X-100 0,1 %, cuantificando la tinción para tinción de actina-alfa de músculo liso usando un anticuerpo anti SMA alfa (GeneTex Cat n.º GTX101362) y un anticuerpo secundario conjugado con Alexa 405. La proliferación se evalúa con tinción de Hoechst después de fijación celular en formaldehído 10 % durante 30 minutos. El colorante Hoechst se lee usando un lector de placas fluorescente de lectura de fondo con una longitud de onda de excitación (Ex) de 405 nm, y una longitud de onda de emisión (Em) de 450 nm. La tinción de actina total se evalúa mediante el uso de un colorante de faloidina marcado con Alexa-488 (Invitrogen Cat n.º A12379).

ANÁLISIS DE LA UNIÓN DE POLIPÉPTIDOS DE METRS CON CÉLULAS (ENSAYOS H1-H10 EN LAS TABLAS DE DATOS POSTERIORES)

Antecedentes y relevancia terapéutica: la unión de polipéptidos de MetRS con tipos celulares específicos demuestra que el tipo celular en cuestión expresa receptores específicos para el polipéptido de MetRS en cuestión. Dependiendo del tipo celular en cuestión, la unión celular implica un papel potencial para el polipéptido de MetRS en la regulación de la actividad o el comportamiento de la célula, o tipos similares de célula, *in vivo*. Los ejemplos específicos de dichas funciones reguladoras incluyen por ejemplo la unión y modulación de linfocitos B y linfocitos T (inmunomodulación/quimiotaxis/autoinmunidad/inflamación); Células HepG2 (control del metabolismo, captación o metabolismo del colesterol); THP-1, jurkat, células Raji (inmunomodulación/quimiotaxis/autoinmunidad/inflamación), plaquetas (trombopoyesis), adipocitos 3T3L1 (lipogénesis/metabolismo) y mioblastos de ratón C2C12 (miogénesis, osteogénesis).

Unión a células sanguíneas:

Métodos: se recoge sangre en tubos de EDTA de donantes sanos. Se colocan 2 ml de sangre completa en tubos Falcon de FACS de 5 ml. Se añaden 2 ml de tampón de tinción (PBS + FBS 2 %), se agitan vorticialmente durante 3-5 segundos, se centrifugan durante 5 minutos a 300 x g. El sobrenadante se aspira, el lavado se repite y el sedimento se resuspende en 2 ml de tampón de tinción.

Se transfieren 100 µl de sangre lavada para limpiar tubos de muestras de FACS de 5 ml. Se añaden polipéptidos de MetRS marcados con His6 o V5-His6 a tubos a las concentraciones indicadas en los experimentos específicos destacados posteriormente y se incuban en hielo durante 45 minutos. Después de la incubación, se añaden anticuerpos para los diferentes marcadores de superficie de tipo celular (BD Pharmigen Cat. n.º 560910, 555398, 555415, 340953, 560361) y anticuerpo anti marcador de V5 marcado con FITC (V5-FITC, Invitrogen Cat n.º R96325) o anticuerpo anti His6 marcado con FITC (AbCam Cat n.º ab1206) a tubos, incubados en oscuridad en hielo durante 30 minutos. Después de la incubación se añaden 2 ml de solución de lisado de FACS BD (cat n.º 349202) a los tubos. Las muestras se agitan vorticialmente y se colocan en hielo durante 15 minutos. Las muestras se lavan con 1 x 2 ml de PBS y se resuspenden en 2 ml de formaldehído 2 % en PBS antes del análisis de FACS. Los polipéptidos de MetRS que se unen a más del 25 % de una población celular, en la que el anticuerpo solo no tiene señal significativa, se consideran un acierto.

Ensayos de unión de plaquetas: se transfirieren 50 µl de sangre lavada a tubos de muestras de FACS de 5 ml limpios, se añaden polipéptidos de MetRS marcados con His6 o V5-His6 a tubos a las concentraciones indicadas en los experimentos específicos perfilados posteriormente y los tubos se colocan en hielo durante 45 minutos. Se añaden a cada tubo 20 µl de anticuerpo pan plaquetario de CD61 (BD Pharmigen, Cat n.º 555754) y 0,5 µl de anticuerpo marcado anti V5-FITC (Invitrogen, R96325) o anticuerpo anti His6 marcado con FITC (AbCam Cat n.º ab1206). Los tubos se colocan en hielo y se protegen de la luz durante 30 minutos. Las muestras se llevan a un volumen total en 2 ml de formaldehído 1 % en PBS y se analizan mediante citometría de flujo en un periodo de 24 horas. Los polipéptidos de MetRS que se unen a más del 25 % de una población celular, en los que el anticuerpo solo no tiene señal significativa, se consideran un acierto.

Unión a células en cultivo: se colocan aproximadamente 1 x 10^6 células en 100 µl de medio RPMI completo en tubos FACS de 5 ml. Se añaden polipéptidos de MetRS marcados con His6 o V5-His6 a tubos a las concentraciones indicadas en los experimentos específicos perfilados posteriormente y los tubos se colocan en hielo durante 45 minutos. Las muestras celulares se lavan dos veces con 1 ml de tampón de tinción (PBS + FBS 2 %) y después se añaden 0,5 µl de anticuerpo anti V5-FITC (Invitrogen R96325) o anticuerpo anti His6 marcado con FITC (AbCam Cat n.º ab1206) en tampón de tinción con IgG humana 200 µg/ml, y las muestras se incuban en hielo, protegidas de la luz, durante 30 minutos. Las muestras se lavan dos veces con 1 ml de tampón de tinción, y después se llevan a un volumen total en 2 ml de formaldehído 1 % en PBS y se analizan por citometría de flujo en un periodo de 24 horas. Los polipéptidos de MetRS que se unen a más del 25 % de una población celular, en los que el anticuerpo solo no tiene señal significativa, se consideran un acierto.

ESTUDIOS ANIMALES: MODULACIÓN DE LA HEMATOPOYESIS Y CITOCINAS EN CIRCULACIÓN

Antecedentes y relevancia terapéutica: la hematopoyesis (como alternativa hemopoyesis) es la formación de componentes celulares sanguíneos. Todos los componentes celulares sanguíneos derivan de células madre hematopoyéticas (HSC) que residen en la médula del hueso (médula ósea) y tienen la capacidad única de dar lugar

a todos los tipos celulares sanguíneos maduros diferentes. Las HSC se auto renuevan: cuando proliferan, al menos algunas de sus células descendientes permanecen como HSC, de modo que el grupo de células madre no se reduce. Las otras descendientes de HSC (células progenitoras mieloides y linfoides) sin embargo pueden comprometerse a cualquiera de las rutas de diferenciación alternativas que conducen a la producción de uno o más tipos específicos de células sanguíneas, pero no pueden en sí mismas auto renovarse. Un cambio en los componentes sanguíneos en respuesta a la exposición a un polipéptido de MetRS sugiere por lo tanto que el polipéptido de MetRS es capaz de modular la hematopoyesis, y regular el desarrollo de células madre hematopoyéticas.

Todas las células sanguíneas pueden dividirse en tres linajes; Células eritroides, linfocitos y mielocitos.

Las células eritroides son los glóbulos rojos portadores de oxígeno. Tanto los reticulocitos como los eritrocitos son funcionales y se liberan a la sangre. En consecuencia un recuento de reticulocitos estima la tasa de eritropoyesis, y un cambio en el recuento de glóbulos rojos sugiere que un polipéptido de MetRS modula la eritropoyesis.

Los linfocitos son la piedra angular del sistema inmunitario adaptativo. Derivan de progenitores linfoides comunes. El linaje linfoide está compuesto principalmente por linfocitos T y linfocitos B (tipos de glóbulos blancos). En consecuencia un cambio en el recuento o la composición de glóbulos blancos en respuesta a exposición a un polipéptido de MetRS sugiere que el polipéptido de MetRS modula la linfopoyesis.

Los mielocitos, que incluyen granulocitos, megacariocitos y macrófagos, y derivan de progenitores mieloides comunes, están implicados en diversas funciones, incluyendo inmunidad innata, inmunidad adaptativa y coagulación sanguínea. En consecuencia un cambio en el recuento o la composición de células mieloides en respuesta a exposición a un polipéptido de MetRS sugiere que el polipéptido de MetRS modula la mielopoyesis. La misma lógica puede usarse para establecer si los polipéptidos de MetRS modulan la granulopoyesis, midiendo cambios en el número de granulocitos en respuesta a exposición a los polipéptidos de MetRS. Un papel para el polipéptido de MetRS en la modulación de megacariocitopoyesis puede inferirse por un cambio en la composición o número de megacariocitos o plaquetas en la sangre.

La liberación de citocinas en ratones de tipo silvestre, o en diversos sistemas de inflamación de modelos animales, proporciona una evaluación inicial de la capacidad potencial de los polipéptidos de MetRS para modular las respuestas inflamatorias. El papel de los polipéptidos de MetRS en la modulación de procesos inflamatorios crónicos agudos por ejemplo, puede evaluar fácilmente el uso de un modelo de ratón de obesidad inducida por dieta (DIO). El modelo de DIO se centra en alimentar a roedores con una dieta alta en grasas durante varios meses lo que conduce a aumento de la obesidad, resistencia a insulina y disfunción del sistema inmunitario. Una consecuencia particular de esta desregulación del sistema inmunitario da como resultado producción aumentada de citocinas proinflamatorias en animales con DIO lo que conduce a una afección de inflamación sistémica crónica. Existe un volumen creciente de pruebas que sugieren que la inflamación de grado bajo contribuya al desarrollo y mantenimiento de obesidad y un fenotipo diabético que observa de forma similar en la afección humana denominada síndrome metabólico. Como tal, la capacidad de los polipéptidos de MetRS para modular el sistema inmunitario y restaurar el equilibrio homeostático hacia una resolución de este estado inflamatorio crónico sería particularmente beneficiosa en numerosas enfermedades y trastornos incluyendo pero sin limitación al tratamiento y prevención de los síntomas y efectos secundarios de enfermedad metabólica, diabetes, enfermedades cardiovasculares, aterosclerosis, obesidad, así como diversas enfermedades y trastornos autoinmunitarios, incluyendo por ejemplo, esclerosis múltiple, trastornos vasculares y alérgicos.

Métodos: se obtienen ratones de control de tipo silvestre machos (C57BL/6) o ratones con obesidad inducida por dieta (C57BL/6NHsd) de Harlan (Indianápolis, IN) y se alojan individualmente. Se alimenta a ratones DIO con una dieta alta en grasas (Cat. n.º TD.06414-60 % kcal de grasa) y se alimenta a ratones de control con una dieta normal (Cat. n.º 2018S-18 % kcal de grasa). Los ratones DIO se alimentan con una dieta alta en grasas comenzando a las 6 semanas de edad durante un total de 10 semanas. Se permite que los ratones tanto DIO como de control se alimenten y beban a voluntad. A las 16 semanas de edad, los ratones se clasifican y separan aleatoriamente en grupos de 5 animales basándose en el peso. El día 2, los ratones se pesan y se toma una muestra de sangre de la vena de la cola (100 µl) para análisis de recuento de sangre completo (CBC) antes del tratamiento. El día 1, los ratones se pesan y se les inyecta por vía intravenosa a través de la vena de la cola vehículo (PBS) o polipéptidos de MetRS individuales a 10 mg/kg. Cuatro horas después de la inyección, se toman muestras de sangre de vena facial de los ratones (150-200 µl) para análisis de citocinas posterior. Los días 2, 3 y 4, se dosifica por vía intravenosa a los ratones como el día 1. El día 5, los ratones se pesan, se sacrifican y se recoge sangre por punción cardíaca para recuento sanguíneo completo (análisis de CBC) (plasma-EDTA) y examen de citocinas (suero).

Análisis de CBC y citocinas: Se analizan recuentos de sangre completa de extracciones sanguíneas antes de inyecciones (día -2) y 24 horas después de la inyección final (día 5). Los valores de CBC se evalúan con respecto a recuentos de glóbulos blancos totales y morfología de glóbulos rojos general. Los glóbulos blancos se caracterizan adicionalmente por porcentaje total y fraccionario de neutrófilos, linfocitos, monocitos, eosinófilos y basófilos. El análisis de glóbulos rojos incluyó mediciones de hemoglobina (dL), hematocrito (%), volumen corpuscular medio (fL), hemoglobina corpuscular media, concentración de hemoglobina corpuscular media (%) y recuento de plaquetas total

(10^3/µl). Se realiza análisis de CBC por Antech Diagnostics (Fishers, IN).

Se examinan los niveles de citocinas de circulación a las 4 horas después de la inyección (día 1) y 24 horas después de la inyección final (día 5). El suero se aísla, se congela instantáneamente y se envía a Rules Based Medicine (Austin, TX) para perfiles de multianalitos. Se analizan muestras de suero usando el panel RodentMap que abarca 59 biomarcadores únicos incluyendo Apo A-1, CD40, CD40-L, CRP, ET-1, eotaxina, EGF, Factor VII, fibrinógeno, FGF-9, FGF-básico, GST-α GCP-2, GM-CSF, KC/GROα, haptoglobina, IgA, IFNγ, IP-10, IL-1α, IL-1β, IL-10, IL-11, IL-12p70, I1-17A, IL-18 IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, LIF, linfotactina, M-CSF-1, MIP-1α, MIP-1β, MIP-1γ, MIP-2 , MCP-3, MCP-1, MCP-3, MCP-5, MPO, mioglobina, SAP, SGOT, SCF, RANTES, TPO, factor tisular, TIMP-1, TNF-α, VCAM- 1, VEGF-A, y vWF. Un cambio en los niveles de citocinas se contó como un acierto si la citocina aumentaba en al menos 2 veces o se reducía al menos 50 % en comparación con controles de vehículo.

EJEMPLO 1 IDENTIFICACIÓN DE FRAGMENTOS PROTEOLÍTICOS Y PRODUCTOS DE CORTE Y EMPALME ALTERNATIVO DE METRS USANDO UNA PLATAFORMA DE ANÁLISIS DE MIGRACIÓN Y TOPOGRAFÍA DE PROTEÍNAS

Para identificar fragmentos de MetRS de líneas celulares, medios acondicionados y tejidos, se prepararon muestras de la siguiente manera:

Macrófagos de ratón (RAW 264.7), citosol y medios acondicionados: Las células se tratan con medio DMEM sin suero a una densidad de 15 x 10^6 células/matraz. Después de 48 horas se recogen y procesan los medios acondicionados y sedimentos celulares. Se separan 200 µg de proteína de fracciones proteómicas secretadas y citosólicas por SDS-PAGE y se preparan cortes de gel para análisis por espectrometría de masas.

Tejido de páncreas de ratón: El páncreas de tres ratones se pica, se homogeneiza por dounce y se somete a ultrasonidos en PBS con inhibidores de proteasa. El proteoma citosólico se aísla por centrifugación y se separan 200 µg de proteína por SDS-PAGE y se preparan cortes de gel para análisis por espectrometría de masas.

Tejido de hígado de ratón: Se pican tres hígados de ratón, se homogeneizan por dounce y se someten a ultrasonidos en PBS con inhibidores de proteasa. El proteoma citosólico se aísla por centrifugación y se separan 200 µg de proteína por SDS-PAGE y se preparan cortes de gel para el análisis por espectrometría de masas.

Se analizan digestiones en gel por espectrómetro de masas de trampa iónica LTQ XL (ThermoFisher) equipado con sistema ultimate 3000 µLC (Dionex). Las muestras se cargan en primer lugar en PepTrap (michrom) durante 10 minutos con Acetonitrilo a 5 % en ácido fórmico 0,1 % usando automostrador Dionex. Después las muestras se analizan con una columna capilar de sílice fusionado de 100 µm (diámetro interno) que contiene 10 cm de resina C18 (michrom). Los péptidos se eluyen de la columna en espectrómetro de masas con un caudal de 0,45 µl/min usando un gradiente lineal de acetonitrilo 5-33,5 % en ácido fórmico 0,1 % en un periodo de 110 minutos.

LTQ se opera en un modo de cribado dependiente de datos de modo que un cribado EM completa se sigan de siete exploraciones EM/EM de los siete iones más abundantes. Se permite la exclusión dinámica con recuento repetido igual a 1, duración repetida igual a 20 segundos, el tamaño de la lista de exclusión es de 300 y la duración de la exclusión es de 60 segundos.

Después de análisis de CL-EM/EM, los datos en bruto se buscan con BioWorks3.3.1 (SEQUEST) usando una variante de señuelo/diana concatenada de la base de datos IPI de ratón. Los datos SEQUEST se filtran y se clasifican con DTASelect. **Las Tablas 1** y **4** muestran secuencias identificadas de esta manera.

EJEMPLO 2 IDENTIFICACIÓN DE VARIANTES DE CORTE Y EMPALME USANDO SECUENCIACIÓN PROFUNDA

Se identifican variantes de corte y empalme de la metionil ARNt sintetasa usando secuenciación de alto rendimiento de bibliotecas de ADNc enriquecidas con respecto a transcritos de metionil ARNt sintetasa. Los moldes de ADNc se preparan a partir de extractos de ARN totales de tejidos tales como cerebros adultos y fetales humanos y enriquecidos con respecto a transcritos de metionil ARNt sintetasa usando secuencias de cebadores específicas para todos los exones indicados de todas las aminoacil ARNt sintetasas humanas indicadas y sus proteínas asociadas.

Se obtienen ARN totales humanos de Clontech. Para muestras tisulares de ratón y línea celular, se extraen ARN totales usando Kit de Extracción de ARN II (MN). Se digiere ADN genómico en las muestras de ARN total por ADNasa I. Para obtener ARN mensajeros (ARNm) maduros, las muestras de ARN se enriquecen dos veces mediante unión de poliA + ARN y digestión de ARN sin recubrimiento 5' por exonucleasa dependiente de 5'-fosfato. Se sintetiza ADN complementario (ADNc) a partir de ARN maduros usando cebadores que hibridan con secuencias exónicas de genes de aminoacilARNt sintetasa. Un transcriptoma enriquecido con respecto a genes de aminoacilARNt sintetasa se amplifica mediante PCR múltiple usando el ARNc específico de exón de aminoacilARNt sintetasa y diferentes combinaciones de cebadores de exon-aminoacilARNt sintetasa. Los productos de PCR de

transcriptoma enriquecidos con aminoacilARNt sintetasa bicatenarios se reparan enzimáticamente en ambos extremos antes de añadir salientes A, a los extremos 3' de los fragmentos reparados. Después se añaden adaptadores de secuenciación y secuencias índice a los productos de PCR de transcriptoma enriquecido con aminoacilARNt sintetasa para generar bibliotecas de ADNc para secuenciación profunda con Kit de Secuenciación Múltiple de Illumina. Brevemente, los productos de PCR de transcriptoma enriquecido con aminoacilARNt sintetasa con salientes 3'-A se ligan con los oligonucleótidos adaptadores de InPE proporcionados en los kits. Se añaden secuencias índices a los productos de PCR con adaptadores de InPE. Para obtener suficientes fragmentos de ADN para secuenciación profunda, los productos de PCR con secuencias índices se amplifican adicionalmente por PCR. Se agrupan bibliotecas de ADNc enriquecidas con aminoacilARNt sintetasa con diferentes índices y se secuencian usando una máquina de secuenciación de ADN Illumina para obtener lecturas finales de 50 de pares de bases. Las lecturas de secuenciación se mapean en genoma humano o de ratón para identificación de acontecimientos de corte y empalme alternativos. Se usa software "Splicemap" (disponible para descarga pública en http://www-stat.stanford.edu/~kinfai/SpliceMap/) para identificar puntos de unión de corte y empalme.

Se realiza secuenciación profunda de estos ADNc para generar aproximadamente 1 millón de lecturas de secuenciación de aproximadamente 50 nucleótidos de longitud. Las secuencias específicas para exones de las aminoacilARNt sintetasas se consultan frente a puntos de unión de exones indicados y nuevos puntos de unión de exones se identifican como acontecimientos de corte y empalme alternativos.

Las columnas en las **Tablas 2** y **5** marcadas como "exón 5'" y "exón 3'" indican, cuando están presentes, qué exones se fusionan entre sí en la secuencia de ADNc. Las **Tablas 2** y **5** muestran las secuencias que se identificaron para acontecimientos de corte y empalme alternativo, transcritos que contienen dichos acontecimientos de corte y empalme y los polipéptidos expresados por esos transcritos. Se identifican variantes de corte y empalme alternativo identificadas por secuenciación profunda en las **Tablas 2** y **5** como las que tienen números mayores que cero en las columnas marcadas como "Lecturas de secuenciación" en el cerebro adulto o fetal humano.

EJEMPLO 3 IDENTIFICACIÓN DE POLIPÉPTIDOS DE METRS USANDO BIOINFORMÁTICA

Se identifican fragmentos proteicos de MetRS (péptidos de resectina o apendacrina) usando bioinformática. Se alinean secuencias de aminoácidos de metionil ARNt sintetasa humana de longitud completa con la secuencia de aminoácidos de longitud completa de su ortólogo de la bacteria *Escherichia coli* usando un programa tal como FASTA (disponible en el sitio web http: //fasta.bioch.virginia. Edu/fasta_www2/fasta_www.cgi) o el programa BLASTP del NCBI (disponible en el sitio web http://www.ncbi.nlm.nih.gov/blast/Blast.cgi?PROGRAM=blastp&BLAST_PROGRAMS=blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthom). Se identifican secuencias de resectina de las proteínas humanas como secuencias que abarcan regiones en las que hay huecos en la secuencia bacteriana en el alineamiento, o regiones con baja homología entre las dos especies. El péptido, y secuencias de ADN correspondientes en las **Tablas 3 y 6** incluyen ejemplos identificados de esta manera.

EJEMPLO 4 EXPRESIÓN DIFERENCIAL DE POLIPÉPTIDOS DE METRS IDENTIFICADOS POR ESPECTROMETRÍA DE MASAS

La técnica de PROTOMAP se usa como se ha descrito en el Ejemplo 1 para comparar la expresión diferencial de las metionil ARNt sintetasas en diferentes tejidos/tipos celulares (refiérase a las **Tablas 1** y **4** para secuencias y comparaciones): se compara la expresión de resectina de aminoacil ARNt sintetasa entre tejido del hígado de ratón y tejido de páncreas de ratón. Se compara la expresión de resectina de aminoacil ARNt sintetasa entre el citosol RAW264.7 y el medio acondicionado de células RAW264.7 recogidas después de 48 horas de privación de suero.

EJEMPLO 5 EXPRESIÓN DIFERENCIAL DE POLIPEPTIDOS DE METRS IDENTIFICADOS POR SECUENCIACIÓN PROFUNDA

Para ensayar con respecto a expresión diferencial de acontecimientos de corte y empalme, la secuenciación profunda se realiza para ADNc preparados a partir de tejidos diferentes.

La expresión de acontecimientos de corte y empalme alternativos específicos para metionil ARNt sintetasas es inesperada e indica importancia biológica. La variación en el número relativo de lecturas vistas en la secuenciación profunda de diferentes muestras de transcriptoma indica que acontecimientos de corte y empalme alternativos aminoacil ARNt sintetasas están regulados diferencialmente y no son solamente artefactos debidos a la manipulación de la muestra.

EJEMPLO 6 CRIBADO DE ANTICUERPOS

Para facilitar el descubrimiento de anticuerpos que presentan unión preferente con fragmentos de metionil ARNt sintetasa (p.ej., afinidad ≥ 10 veces mayor en comparación con la enzima de longitud completa parental), se explora una biblioteca de presentación en fagos de anticuerpos humanos mediante AbD Serotec (una división de MORPHOSYS™, Martinsried/Planegg, Alemania) usando técnicas de enriquecimiento de afinidad (selección). Los

anticuerpos enriquecidos después de múltiples ciclos de cribado con los fragmentos de metionil ARNt sintetasas se caracterizan posteriormente por ELISA con respecto a reactividad con los fragmentos, y con la enzima de longitud completa parental. Se caracterizan adicionalmente clones que demuestran unión preferente (por ejemplo, afinidad ≥ 10 veces mayor) con los fragmentos de metionil ARNt sintetasa.

Si no se consigue la especificidad necesaria al final de este proceso, se usan estrategias de resta, tales como etapas de preadsorción con la enzima de longitud completa y/o contracribado, para eliminar anticuerpos de reacción cruzada y conducir el proceso de selección hacia el epítopo o los epítopos únicos en los fragmentos de metionil ARNt sintetasa.

EJEMPLO 7 IDENTIFICACIÓN DE VARIANTES DE CORTE Y EMPALME USANDO PCR SISTEMÁTICA

Se transcriben de forma inversa moldes de ADNc para reacciones de PCR a partir de extractos de ARN totales de tejidos o células (p.ej., cerebro humano, IMR-32 y HEK293T). Se realizan reacciones de PCR usando cebadores específicos de metionil ARNt sintetasa, emparejando un cebador directo (FP1) diseñado para hibridar con la región no traducida 5' o exones en la mitad 5' del gen con un cebador inverso (RP1) diseñado para hibridar con exones en la mitad de 3' del gen o la 3'UTR. Se analizan productos de ADN amplificados por electroforesis en gel de agarosa para identificar productos de PCR que son de un tamaño diferente que el fragmento amplificado de los transcritos canónicos. Estos productos de PCR diferentes se escinden y se purifican a partir del gel y se ligan en un vector de clonación convencional para análisis de secuencia de ADN. Se identifican variantes de corte y empalme alternativas como secuencias diferentes de los transcritos canónicos. Se muestran variantes de corte y empalme identificadas por este enfoque de PCR sistemática en las **Tablas 2** y **5**.

EJEMPLO 8 OPTIMIZACIÓN DE CODONES DE POLINUCLEÓTIDOS DE METRS SELECCIONADOS

Se seleccionan polipéptidos de MetRS representativos (resumidos en la **Tabla E2**) para caracterización bioquímica, biofísica y funcional adicional basada en uno o más de los siguientes criterios, i) la identificación de fragmentos proteolíticos de polipéptidos de MetRS, ii) la identificación de las variantes de corte y empalme de polipéptidos de MetRS, iii) la identificación de polipéptidos de MetRS por análisis bioinformático, iv) pruebas de expresión diferencial de polipéptidos de MetRS específicos, v) la estructura de dominio de la proteína de MetRS, vi) el tamaño del polipéptido de MetRS y vii) la minimización de secuencias duplicativas similares.

Tabla E2
Sumario de Polipéptidos de MetRS Seleccionados para Optimización de Codones y Expresión Bacteriana

Nombre del polipéptido de MetRS	SEQ ID NO para Polipéptidos de MetRS marcados con epítopos	SEQ ID NO para Polinucleótidos de MetRS	Restos de proteína MetRS	Localización de marcador epitópico	Método de clonación/síntesis usado
$MetRS^{N1}$	**SEQ ID NO 68**	**SEQ ID NO 90**	1-231	N-terminal	1
$MetRS^{N1}$	**SEQ ID NO 69**	**SEQ ID NO 90**	1-231	C-terminal	1
$MetRS^{N2}$	**SEQ ID NO 70**	**SEQ ID NO 91**	1-263	N-terminal	1
$MetRS^{N2}$	**SEQ ID NO 71**	**SEQ ID NO 91**	1-263	C-terminal	1
$MetRS^{N3}$	**SEQ ID NO 72**	**SEQ ID NO 92**	1-67+48 aa	N-terminal	1
$MetRS^{N3}$	**SEQ ID NO 73**	**SEQ ID NO 92**	1-67+48 aa	C-terminal	1
$MetRS^{N4}$	**SEQ ID NO 74**	**SEQ ID NO 93**	1-364+63 aa	N-terminal	1
$MetRS^{N4}$	**SEQ ID NO 75**	**SEQ ID NO 93**	1-364+63 aa	C-terminal	1
$MetRS^{N6}$	**SEQ ID NO 76**	**SEQ ID NO 94**	1-545+2 aa	N-terminal	1
$MetRS^{N6}$	**SEQ ID NO 77**	**SEQ ID NO 94**	1-545+2 aa	C-terminal	1
$MetRS^{N9}$	**SEQ ID NO 78**	**SEQ ID NO 95**	1-197	N-terminal	1
$MetRS^{N9}$	**SEQ ID NO 79**	**SEQ ID NO 95**	1-197	C-terminal	1
$MetRS^{N10}$	**SEQ ID NO 80**	**SEQ ID NO 96**	1-214	N-terminal	1
$MetRS^{N10}$	**SEQ ID NO 81**	**SEQ ID NO 96**	1-214	C-terminal	1
$MetRS^{N11}$	**SEQ ID NO 82**	**SEQ ID NO 97**	1-221	N-terminal	1
$MetRS^{N11}$	**SEQ ID NO 83**	**SEQ ID NO 97**	1-221	C-terminal	1
$MetRS^{N12}$	**SEQ ID NO 84**	**SEQ ID NO 98**	1-67+47 aa	N-terminal	1
$MetRS^{N12}$	**SEQ ID NO 85**	**SEQ ID NO 98**	1-67+47 aa	C-terminal	1
$MetRS^{N13}$	**SEQ ID NO 86**	**SEQ ID NO 99**	1-221+55 aa	N-terminal	1
$MetRS^{N13}$	**SEQ ID NO 87**	**SEQ ID NO 99**	1-221+55 aa	C-terminal	1
$MetRS^{N14}$	**SEQ ID NO 88**	**SEQ ID NO 100**	1-257+16 aa	N-terminal	1
$MetRS^{N14}$	**SEQ ID NO 89**	**SEQ ID NO 100**	1-257+16 aa	C-terminal	1
$MetRS^{C2}$	**SEQ ID NO 165**	**SEQ ID NO 173**	773-900	N-terminal	1
$MetRS^{C2}$	**SEQ ID NO 166**	**SEQ ID NO 173**	773-900	C-terminal	1
$MetRS^{C3}$	**SEQ ID NO 167**	**SEQ ID NO 174**	552-900	N-terminal	1
$MetRS^{C3}$	**SEQ ID NO 168**	**SEQ ID NO 174**	552-900	C-terminal	1
$MetRS^{C4}$	**SEQ ID NO 169**	**SEQ ID NO 175**	8 aa +586-900	N-terminal	1
$MetRS^{C4}$	**SEQ ID NO 170**	**SEQ ID NO 175**	8 aa +586-900	C-terminal	1

$MetRS^{C5}$	**SEQ ID NO 171**	**SEQ ID NO 176**	846-900	N-terminal	1
$MetRS^{C5}$	**SEQ ID NO 172**	**SEQ ID NO 176**	846-900	C-terminal	1

Se sintetizan polinucleótidos que codifican los polipéptidos de MetRS seleccionados enumerados en la **Tabla E2**, junto con el marcador epitópico N o C-terminal apropiado, y se clonan como se describe en la sección de Materiales y Métodos Generales usando la metodología de síntesis génica enumerada en la **Tabla E2**.

EJEMPLO 9 EXPRESIÓN Y PURIFICACIÓN BACTERIANA A PEQUEÑA ESCALA

Los polipéptidos de MetRS enumerados en la **Tabla E2** se expresan en *E. coli*, como se describe en la sección de Materiales y Métodos Generales. La expresión relativa de polipéptidos de MetRS localizados en cuerpos de inclusión y solubles se resume en la **Tabla E3** a continuación.

Tabla E3
Sumario de Características de Expresión Bacteriana de Polipéptidos de MetRS

Polipéptido de MetRS	Localización de Marcador Epitópico	Cantidad de Proteínas Recuperada de Fracción Soluble	Cantidad de Proteínas Recuperada de Cuerpos de Inclusión
$MetRS^{N1}$	N-terminal	+	+
$MetRS^{Ni}$	C-terminal	+	+
$MetRS^{N2}$	N-terminal	+	+
$MetRS^{N2}$	C-terminal	+	+
$MetRS^{N3}$	N-terminal	+	+
$MetRS^{N3}$	C-terminal	+	+
$MetRS^{N4}$	N-terminal	+	+
$MetRS^{N4}$	C-terminal	+	+
$MetRS^{N6}$	N-terminal	+	+
$MetRS^{N6}$	C-terminal	+	+
$MetRS^{N9}$	N-terminal	+	++
$MetRS^{N9}$	C-terminal	+	+
$MetRS^{N10}$	N-terminal	+	+
$MetRS^{N10}$	C-terminal	+	+
$MetRS^{N11}$	N-terminal	+	+
$MetRS^{N11}$	C-terminal	+	+
$MetRS^{N12}$	N-terminal	+	+
$MetRS^{N12}$	C-terminal	+	+
$MetRS^{N13}$	N-terminal	+	+
$MetRS^{N13}$	C-terminal	+	+
$MetRS^{N14}$	N-terminal	+	+
$MetRS^{N14}$	C-terminal	+	+
$MetRS^{C2}$	N-terminal	+	ND
$MetRS^{C2}$	C-terminal	+	ND
$MetRS^{C3}$	N-terminal	+	ND
$MetRS^{C3}$	C-terminal	+	ND
$MetRS^{C4}$	N-terminal	+	ND
$MetRS^{C4}$	C-terminal	+	ND
$MetRS^{C5}$	N-terminal	++	ND
$MetRS^{C5}$	C-terminal	++	ND
"+" representa 0-1 mg/l de expresión de polipéptido de MetRS "++" representa 1-5 mg/l de expresión de polipéptido de MetRS; "+++" representa 5-10 mg/l de expresión de polipéptido de MetRS; "++++" representa 10-15 mg/l de expresión de polipéptido de MetRS; "+++++" representa ≥ 15 mg/l de expresión de polipéptido de MetRS;			
ND: no determinada			

Sorprendentemente, los datos de expresión de proteínas demuestran la existencia de al menos un dominio proteico que muestra alto nivel de expresión de proteína soluble cuando se expresa en *E. coli*. Específicamente, los datos demuestran que el polipéptido MetRS $MetRS^{C5}$ (aminoácidos 846-900) define el límite de un dominio proteico nuevo que está altamente expresado en *E. coli*.

EJEMPLO 10 PRODUCCIÓN A GRAN ESCALA DE POLIPÉPTIDOS DE METRS

Se preparan polipéptidos de MetRS representativos en cantidades mayores para permitir la caracterización funcional y biofísica adicional. Los polipéptidos de MetRS enumerados en la **Tabla E4** se expresan en *E. coli* en cultivo a gran escala como se describe en la sección de Materiales y Métodos Generales. Los rendimientos, y las características biofísicas específicas, para cada proteína soluble expresada se resumen a continuación en la **Tabla E4**.

Tabla E4 Sumario de rendimiento y caracterización biofísica de Polipéptidos de PheRSa representativos								
Polipéptido de MetRS	Localización de marcador Epitópico	Rendimiento [mg/l] (1)	Pureza [%]	Endotoxina [UE/mg]	Peso Molecular	Concentración de reserva de trabajo [mg/ml]	Estabilidad [porcentaje de recuperación] (2)	Agregación [DLS]
$MetRS^{C5}$	N-terminal	0,5	80	13,4	ND	2,6	ND	ND
Notas (1) : Rendimiento determinado midiendo la recuperación de proteínas después de la última etapa de purificación (2) : Determinado como un porcentaje de recuperación de material no agregado después de 1 semana a 25 ºC								
Clave:								
ND: No Determinado								

Los resultados de estos estudios establecen que proteínas de MetRS representativas de la familia de $MetRS1^{C5}$ de proteínas de MetRS, muestran rendimientos de expresión y características de solubilidad de proteínas iniciales razonables.

EJEMPLO 11 PERFIL DE TRANSCRIPCIÓN DE POLIPÉPTIDOS DE METRS REPRESENTATIVOS

Para ensayar con respecto a la capacidad de los polipéptidos de MetRS para modular la expresión génica, se incubaron polipéptidos de MetRS seleccionados con células madre mesenquimales o células del músculo esquelético humano durante los tiempos y a las concentraciones mostradas en la **Tabla E5**.

Tabla E5 Perfil de Transcripción de Polipéptidos de MetRS Representativos en Células Madre Mesenquimales (MSC) o Células de Músculo Esquelético Humano (HSkMC)						
Descripción de la Muestra de Ensayo		**Tipo Celular y Tiempo de Exposición**				
Polipéptidos de MetRS	**Localización de Marcador Epitópico**	**Concentración nM**	**MSC 24 horas**	**MSC 72 horas**	**HSkMC 24 horas**	**HSkMC 72 horas**
$MetRS^{C5}$	N-terminal	250	0	1	10	8
$MetRS^{C5}$	C-terminal	250	1	7	8	7
Controles						
Promedio entre todos los polipéptidos de MetRS explorados			3	5	6	7
Cóctel de Osteogénesis			17	20	11	16
Cóctel de Condrogénesis			17	19	14	19
Cóctel de Adipogénesis			19	15	16	18
Ctrl Pos de SKMC			11	8	5	4
No tratado			0	0	1	1

En la **Tabla E5**, los números en cada columna representan el número de genes que se modularon, bien positivamente o bien negativamente, al menos 4 veces en comparación con los genes de control, como se describe en la sección de métodos generales. Los datos muestran que formas específicas de los polipéptidos de MetRS ensayados tienen la sorprendente capacidad de regular la transcripción, y por lo tanto modulan potencialmente el destino de desarrollo o estado de diferenciación cuando se añaden a Células Madre Mesenquimales (MSC) y/o Células de Músculo Esquelético Humano (HSkMC). Las células sombreadas con números en negrita en la tabla representan ejemplos en los que el polipéptido de MetRS muestra una influencia significativa en la regulación de la transcripción génica en las líneas celulares y los tiempos indicados en la tabla.

Se concluye que $MetRS1^{C5}$ parece ser un regulador importante de la expresión génica de células mesenquimales y/o células del músculo esquelético humano.

EJEMPLO 12 PERFILES FUNCIONALES DE POLIPÉPTIDOS DE METRS

Para ensayar con respecto a la capacidad de los polipéptidos de MetRS para modular una serie de procesos fenotípicos, se incubaron polipéptidos de MetRS seleccionados con los tipos celulares, y las condiciones proporcionadas en la sección de métodos generales, y las **Tablas E5 y E6**.

Tabla E6 Clave para Ensayos y criterios para indicar un acierto	
Ensayos de proliferación	
Fuente y tipo celular	Número de Ensayo
Células de leucemia megacariocítica humana/Mo7e	A1
Células de leucemia promielocítica aguda humana/HL60	A2
Linfoblasto humano (línea celular de cáncer)/RPMI8226	A3
Células madre mesenquimales humanas/hMSC	A4
Astrocitos humanos	A5
Células de aspirado de Médula Ósea humana/Células de Médula Ósea	A6
Células de aspirado de Médula Ósea humana/Células de Médula Ósea (Cultivo a Largo Plazo)	A7
Sinoviocito Humano/HFLS-SynRA	A8
Células preadipocíticas humanas/hPAD	A9
Célula de músculo liso de arteria pulmonar humana/hPASMC	A10
Célula de músculo esquelético humano/hSKMC	A11
Se realizó análisis de datos para ensayos de proliferación dividiendo el valor numérico en el pocillo de ensayo por el valor de PBS promedio para la placa de ensayo. Se consideró que los polipéptidos de MetRS eran proliferativos si estaban a más de 3 DT de distancia del valor de PBS en la dirección positiva. Se consideró que un polipéptido de MetRS derivado de ARNt sintetasa era citotóxico si estaba a más de 3 DT de distancia del valor de PBS en la dirección negativa. Se utilizó un compuesto citotóxico como un control negativo y el valor promedio para esto estuvo siempre a más de 3 DT de distancia del valor promedio de PBS.	
Ensayos de diferenciación y fenotipo celular	
Descripción del Ensayo	**Número de Ensayo**
Captación de LDL acetilada por hepatocito humano (Células HepG2C3a)	B1
Se realizó análisis de datos para ensayo de captación de ac-LDL dividiendo el valor numérico en el pocillo de ensayo por el valor de PBS promedio para la placa de ensayo. Se consideró que los polipéptidos de MetRS eran un modulador de la captación de ac-LDL si estaban a más de 2 DT de distancia del valor de PBS en la dirección positiva o negativa. Se realizó una comprobación visual para confirmar los resultados del lector de placas usando un microscopio fluorescente.	
Ensayos de Neutrófilos Humanos	
Descripción del Ensayo	**Número de Ensayo**
Elastasa de Neutrófilo	C1
Estallido oxidativo de Neutrófilo (agonista)	C2
Estallido oxidativo de Neutrófilo (antagonista)	C3
Se realizó análisis de datos para ensayos de neutrófilos dividiendo el valor numérico en el pocillo de ensayo por el valor de PBS promedio para la placa de ensayo. Se consideró que los polipéptidos de MetRS eran un modulador de la producción de elastasa o biología de estallido oxidativo de neutrófilos si estaban a más de 2 DT de distancia del valor de PBS en la dirección positiva o negativa.	
Modulación de Receptores de Tipo Toll (TLR)	
Descripción del Ensayo	**Número de Ensayo**
Activación de TLR en células RAW BLUE	D1
Antagonismo de TLR en células RAW BLUE	D2
Activación de hTLR2	D3
Activación de hTLR4	D4
Se realiza análisis de datos para ensayos de modulación de TLR dividiendo el valor numérico en el pocillo de ensayo por el valor de PBS promedio para la placa de ensayo. Se consideró que los polipéptidos de MetRS eran un modulador de la biología específica de TLR si estaban más de 3 DT de distancia del valor de PBS en la dirección positiva o negativa. Los controles positivos, incluyendo el LPS y el reactivo de detección, fueron siempre significativamente distintos y >3 DT del valor promedio del PBS.	
Liberación de Citocina	
Descripción del Ensayo	**Número de Ensayo**
Producción de citocinas por sinoviocitos humanos (liberación de IL6)	E1
Producción de citocinas por células de músculo liso de arteria pulmonar humana (hPASMC) (liberación de IL6)	E2
Producción de citocinas por células de músculo esquelético humano (hSKMC) (liberación de IL6)	E3
Producción de citocinas por astrocitos humanos (liberación de IL6)	E4
Liberación de IL6 de sangre completa	E5
Producción de citocinas por células de músculo liso de arteria pulmonar humana (hPASMC) (liberación de IL8) Incubación de 72 h	E6
Producción de IL8	
Descripción del Ensayo	**Número de Ensayo**

Producción de citocinas por sinoviocitos humanos (liberación de IL8)	E7
Producción de citocinas por células de músculo liso de arteria pulmonar humana (hPASMC) (liberación de IL8)	E8
Producción de citocinas por células de músculo esquelético humano (hSKMC) (liberación de IL8)	E9
Producción de citocinas por astrocitos humanos (liberación de IL8)	E10
Liberación de IL8 por hepatocitos humanos (células HepG2C3a)	E11
Células de leucemia promielocítica aguda humana / HL60 (liberación de IL8)	E12
Linfoblasto humano (línea celular de cáncer) / RPMI8226 (liberación de IL8)	E13
Producción de TNF alfa	
Producción de citocinas por sinoviocitos humanos (liberación de TNF alfa)	E14
Liberación de TNF alfa de sangre completa	E15
Liberación de IL10	
Células de leucemia promielocítica aguda humana / HL60 (liberación de IL10)	E16
Células mononucleares de sangre primaria humana (liberación de IL10)	E17
Se realizaron análisis de datos para ensayos de liberación de citocinas dividiendo el valor numérico en el pocillo de ensayo por el valor de PBS promedio para la placa de ensayo. Se consideró que los polipéptidos de MetRS eran un modulador de la producción de citocinas o biología relacionada con citocinas si el valor medido estaba a más de 2 DT de distancia de valor de PBS en la dirección positiva o negativa. Se procesó un patrón proteico (específico para cada kit de ensayo) en cada placa para asegurar la buena calidad de ensayo. Solamente se eligieron ensayos con curvas patrón de proteínas que tenían un valor $R^2 > 0,9$ para análisis de datos	
Adhesión celular y quimiotaxis	
Descripción del Ensayo	**Número de Ensayo**
Adhesión celular de monocito THP1/ Célula endotelial de vena umbilical humana (HUVEC)	F1
Hepatocito humano (células HepG2C3a) (liberación de ICAM)	F2
Regulación de la adhesión celular de células endoteliales microvasculares pulmonares humanas (HLMVEC) (liberación de ICAM)	F3
Regulación de la adhesión celular de células endoteliales de la vena umbilical humana (HUVEC) (liberación de VCAM)	F4
Regulación de la adhesión celular de células madre mesenquimales humanas (hMSC) (Liberación de VCAM)	F5
Regulación de la adhesión celular de células del músculo esquelético humano (hSKMC) (Liberación de VCAM)	F6
Regulación de la adhesión celular de células de músculo liso de arteria pulmonar humana (hPASMC) (liberación de VCAM)	F7
Se realizó análisis de datos para ensayos de regulación de adhesión celular dividiendo el valor numérico en el pocillo de ensayo por el valor de PBS promedio para la placa de ensayo. Se consideró que los polipéptidos de MetRS eran un modulador de la adhesión celular o un regulador de la biología relacionada con la adhesión celular si se obtuvo un valor a más de 2 DT de distancia del valor de PBS en la dirección positiva o negativa. En el caso de los ensayos de ELISA, se procesó un patrón proteico (específico para cada kit de ensayo) en cada placa para asegurar la buena calidad de ensayo. Solamente se seleccionaron ensayos con curvas patrón de proteínas que tenían un valor R2 > 0,9 para análisis de datos.	
Diferenciación Celular	
Descripción del Ensayo	**Número de Ensayo**
Diferenciación de células preadipocíticas humanas (hPAD)	G1
Diferenciación de células de músculo esquelético humano (hSKMC)	G2
Diferenciación de células madre mesenquimales humanas (hMSC)	G3
Diferenciación de células de músculo liso de arteria pulmonar humana (hPASMC)	G4
Se realizó análisis de datos para ensayos de diferenciación celular dividiendo el valor numérico en el pocillo de ensayo por el valor de PBS promedio para la placa de ensayo. Los ensayos de diferenciación se puntuaron basándose en la intensidad de fluorescencia de anticuerpos particulares como se describe en la sección de métodos. Se consideró que los polipéptidos de MetRS eran un modulador de la diferenciación celular si un valor de intensidad para un marcador específico de diferenciación estaba a más de 2 DT de distancia del valor de PBS en la dirección positiva o negativa en un pocillo tratado dado. Para el análisis de hSKMC, se tomaron fotos digitales de todos los pocillos y las fotos se puntuaron con ocultación por tres personas usando un sistema de puntuación de 4 puntos en el que una puntuación de “4” indicó tinción de actina de músculo esquelético intensa y formación de miotubos evidente y una puntuación de "1" indicaba una falta de diferenciación o una deleción de la diferenciación. Se usó el valor promedio de la puntuación visual y solamente se consideraron aciertos pocillos con un valor promedio de ≥ 3. Los pocillos tratados con control de diferenciación en este ensayo normalmente puntuaron ≥ 2, mientras que los pocillos tratados con PBS puntuaron ≤ 2.	
Unión Celular	
Descripción del Ensayo	**Número de Ensayo**

PBMC	H1
Linfocito T primario	H2
Linfocito B primario	H3
Monocito primario	H4
HepG2	H5
3T3L1	H6
C2C12	H7
THP1	H8
Jurkat	H9
Raji	H10
Se consideró que los polipéptidos de MetRS se unían con un tipo celular particular si la intensidad de fluorescencia unida a célula media estaba a más de 2 DT de distancia de los valores del control de reactivo para ese tipo celular.	

Tabla E7
Resultados de Estudios de Perfiles Funcional de Polipéptidos de MetRS

Polipéptidos de MetRS	Localización de marcador Epitópico	Concentración [nM]	Aciertos de Ensayo
$MetRS^{N9}$	N-terminal	250	A2, A3, A5, A8, A9, A10 (Apoptosis) C1, C2, C3 (Activación de neutrófilos) D1, D2 (Modulación de receptores de tipo Toll) E1, E8 (Liberación de citocinas) F7 (Adhesión celular y quimiotaxis) G1, G4, G3 (Diferenciación)
$MetRS^{C5}$	N- terminal	250	E6 (Liberación de Citocinas) G2 (Diferenciación)
$MetRS^{C5}$	C-terminal	250	A7, A8 (Proliferación) C1 (Activación de neutrófilos) G2, G4 (Diferenciación)

Se concluye que $MetRS1^{N9}$ y $MetRS1^{C5}$ parecen ser reguladores importantes de la apoptosis, liberación de citocinas, activación de neutrófilos, adhesión celular y quimiotaxis, diferenciación. Cuando se ven en el contexto de los estudios de perfiles transcripcionales, los datos de cribado fenotípicos demuestran que los polipéptidos de MetRS $MetRS1^{N9}$ (aminoácidos 1-197) y $MetRS1^{C5}$ (aminoácidos 846-900) definen los límites de dos dominios proteicos nuevos que están altamente activos en una amplia serie de ensayos de cribado fenotípicos.

En consecuencia se concluye que los polipéptidos de MetRS que comprenden los aminoácidos 1-197 y 846-900 de la metionil ARNt sintetasa definen los límites aproximados (es decir en el intervalo de aproximadamente +/- 5 aminoácidos) de dos dominios de polipéptidos de MetRS altamente activos, nuevos, que son i) altamente funcionalmente activos, ii) pueden prepararse fácilmente y producirse en *E. coli*, y iii) muestran características de estabilidad y agregación de proteínas favorables.

LISTADO DE SECUENCIAS

<110> aTyr Pharma Inc.
Pangu Biopharma Limited
Greene, Leslie Ann
Chiang, Kyle P.
Hong, Fei
Vasserot, Alain P.
Lo, Wing-Sze
Watkins, Jeffry D.
Mendlein, John D.
Quinn, Cheryl L.

<120> DESCUBRIMIENTO INNOVADOR DE COMPOSICIONES TERAPÉUTICAS, DE DIAGNÓSTICO Y DE ANTICUERPOS RELACIONADAS CON FRAGMENTOS DE PROTEÍNAS DE METIONIL-ARNT SINTETASAS

<130> 120161.462pc

<140> PCT
<141> 03-05-2011

<150> US 61/330.596
<151 > 03-05-2010

<150> US 61/331.282
<151 > 04-05-2010

<160> 176

<170> FastSEQ para Windows Versión 4.0

<210> 1
<211> 77
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 1

aggaggtaaa acatatgcat catcatcatc atcacggtaa gcctatccct aaccctttgc 60
tcggtctcga ttctacg 77

<210> 2
<211> 12
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 2
taatgactcg ag 12

<210> 3
<211> 17
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 3
aggagataaa acatatg 17

<210> 4
<211> 14
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 4
aggaggtaaa acat 14

<210> 5
<211> 14
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 5

aggagataaa acat 14

<210> 6
<211> 15
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 6
gaaggagata tacat 15

<210> 7
<211> 72
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 7

ggtaagccta tccctaaccc tctcctcggt ctcgattcta cgcaccacca tcatcaccat 60
taatgactcg ag 72

<210> 8
<211> 72
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 8

catatgcatc atcatcatca tcacggtaag cctatcccta accctctcct cggtctcgat 60
tctacgggat cc 72

<210> 9
<211> 12
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 9
ctcgagtaat ga 12

<210> 10
<211> 12
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 10
catatgggat cc 12

<210> 11

<211> 72
<212> ADN
<213> Secuencia artificial

<220>
<223> Oligonucleótido

<400> 11

ctcgagggta agcctatccc taaccctctc ctcggtctcg attctacgca ccaccaccac 60
caccactaat ga 72

<210> 12
<211> 231
<212> PRT
<213> *Homo sapiens*

<400> 12

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

<210> 13
<211> 693
<212> ADN
<213> *Homo sapiens*

<400> 13

atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagccccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gag 693

<210> 14
<211> 19
<212> PRT
<213> *Mus musculus*

<400> 14

Leu Phe Val Ser Glu Gly Ser Pro Gly Ser Leu Pro Val Leu Ala Ala
1 5 10 15
Ala Ala Arg

<210> 15
<211> 22
<212> PRT
<213> *Mus musculus*

<400> 15

Gly Arg Ala Glu Leu Leu Ile Ser Thr Val Gly Pro Glu Glu Cys Val
1 5 10 15
Val Pro Phe Leu Thr Arg
20

<210> 16
<211> 20
<212> PRT
<213> *Mus musculus*

<400> 16

Val Pro Val Leu Gln Leu Asp Ser Gly Asn Tyr Leu Phe Ser Ala Ser
1 5 10 15
Ala Ile Cys Arg
20

<210> 17
<211> 137
<212> PRT
<213> *Mus musculus*

<400> 17

```
Tyr Phe Phe Leu Leu Cys Gly Trp Glu Gln Asp Asp Leu Thr Asn Gln
 1               5                  10                  15
Trp Leu Glu Trp Glu Ala Thr Glu Leu Gln Pro Val Leu Ser Ala Ala

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

<210> 18
<211> 12
<212> PRT
<213> *Mus musculus*

<400> 18

```
Gln Pro Gln Pro Gln Pro Pro Pro Pro Glu Gly Arg
 1               5                  10
```

<210> 19
<211> 214
<212> PRT
<213> *Mus musculus*

<400> 19

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

<210> 20
<211> 115
<212> PRT
<213> *Homo sapiens*

<400> 20

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

<210> 21
<211> 348
<212> ADN
<213> *Homo sapiens*

<400> 21

atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg gtattctatg tctggtttga tgccactatt ggctatctgt 240
ccatcacagc caactacaca gaccagtggg agagatggtg gaagaaccca gagcaagtgg 300
acctgtatca gttcatggcc aaagacaatg ttcctttcca tagcttag 348

<210> 22
<211> 427
<212> PRT
<213> *Homo sapiens*

<400> 22

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

Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu Glu Glu Glu
    210                 215                 220
Leu Ala Thr Leu Ser Glu Glu Glu Ile Ala Met Ala Val Thr Ala Trp
225                 230                 235                 240
Glu Lys Gly Leu Glu Ser Leu Pro Pro Leu Arg Pro Gln Gln Asn Pro
                245                 250                 255
Val Leu Pro Val Ala Gly Glu Arg Asn Val Leu Ile Thr Ser Ala Leu
            260                 265                 270
Pro Tyr Val Asn Asn Val Pro His Leu Gly Asn Ile Ile Gly Cys Val
        275                 280                 285
Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser Arg Leu Arg Gln Trp Asn
    290                 295                 300
Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly Thr Ala Thr Glu Thr
305                 310                 315                 320
Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile Cys Asp Lys Tyr
                325                 330                 335
His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn Ile Ser Phe Asp
            340                 345                 350
Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys Ser Leu Ser Val
        355                 360                 365
Lys Ser Ala Asp His Ala Leu Trp Cys Ser Arg Ala Ser Thr Cys Phe
    370                 375                 380
Trp Thr Cys Leu Ser Trp Arg Ser Asp Trp Arg Ser Gly Trp Gly Gly
385                 390                 395                 400
His Cys Leu Ala Val Thr Gly His Pro Met Pro Ser Leu Ser Pro Val
                405                 410                 415
Leu Gly Phe Gly Met Ala Ser Ser His Ala Ala
            420                 425
```

<210> 23
<211> 1284
<212> ADN
<213> *Homo sapiens*

<400> 23

atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagccccgg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttct ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gccccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaaa ggaatgtgct catcaccagt gccctccctt acgtcaacaa tgtcccccac 840
cttgggaaca tcattggttg tgtgctcagt gccgatgtct ttgccaggta ctctcgcctc 900
cgccagtgga acaccctcta tctgtgtggg acagatgagt atggtacagc aacagagacc 960
aaggctctgg aggagggact aacccccccag gagatctgcg acaagtacca catcatccat 1020
gctgacatct accgctggtt taacatttcg tttgatattt ttggtcgcac caccactcca 1080
cagcagacca aaagcctcag tgtaaagtct gccgatcatg ccctgtggtg cagtcgagcc 1140
agcacctgtt tctggacctg cctaagctgg agaagcgact ggaggagtgg ttggggagga 1200
cattgcctgg cagtgactgg acacccaatg cccagtttat cacccgttct tggcttcggg 1260
atggcctcaa gccacgctgc ataa 1284

<210> 24
<211> 868
<212> PRT
<213> *Homo sapiens*

<400> 24

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

```
                500                 505                 510
Lys Val Asp Leu Tyr Gln Phe Met Ala Lys Asp Asn Val Pro Phe His
        515                 520                 525
Ser Leu Val Phe Pro Cys Ser Ala Leu Gly Ala Glu Asp Asn Tyr Thr
    530                 535                 540
Leu Val Ser His Leu Ile Ala Thr Glu Tyr Leu Asn Tyr Glu Asp Gly
545                 550                 555                 560
Lys Phe Ser Lys Ser Arg Gly Val Gly Val Phe Gly Asp Met Ala Gln
                565                 570                 575
Asp Thr Gly Ile Pro Ala Asp Ile Trp Arg Phe Tyr Leu Leu Tyr Ile
            580                 585                 590
Arg Pro Glu Gly Gln Asp Ser Ala Phe Ser Trp Thr Asp Leu Leu Leu
        595                 600                 605
Lys Asn Asn Ser Glu Leu Leu Asn Asn Leu Gly Asn Phe Ile Asn Arg
    610                 615                 620
Ala Gly Met Phe Val Ser Lys Phe Phe Gly Gly Tyr Val Pro Glu Met
625                 630                 635                 640
Val Leu Thr Pro Asp Asp Gln Arg Leu Leu Ala His Val Thr Leu Glu
                645                 650                 655
Leu Gln His Tyr His Gln Leu Leu Glu Lys Val Arg Ile Arg Asp Ala
            660                 665                 670
Leu Arg Ser Ile Leu Thr Ile Ser Arg His Gly Asn Gln Tyr Ile Gln
        675                 680                 685
Val Asn Glu Pro Trp Lys Arg Ile Lys Gly Ser Glu Ala Asp Arg Gln
    690                 695                 700
Arg Ala Gly Thr Val Thr Gly Leu Ala Val Asn Ile Ala Ala Leu Leu
705                 710                 715                 720
Ser Val Met Leu Gln Pro Tyr Met Pro Thr Val Ser Ala Thr Ile Gln
                725                 730                 735
Ala Gln Leu Gln Leu Pro Pro Pro Ala Cys Ser Ile Leu Leu Thr Asn
            740                 745                 750
Phe Leu Cys Thr Leu Pro Ala Gly His Gln Ile Gly Thr Val Ser Pro
        755                 760                 765
Leu Phe Gln Lys Leu Glu Asn Asp Gln Ile Glu Ser Leu Arg Gln Arg
    770                 775                 780
Phe Gly Gly Gly Gln Ala Lys Thr Ser Pro Lys Pro Ala Val Val Glu
785                 790                 795                 800
Thr Val Thr Thr Ala Lys Pro Gln Gln Ile Gln Ala Leu Met Asp Glu
                805                 810                 815
Val Thr Lys Gln Gly Asn Ile Val Arg Glu Leu Lys Ala Gln Lys Ala
            820                 825                 830
Asp Lys Asn Glu Val Ala Ala Glu Val Ala Lys Leu Leu Asp Leu Lys
        835                 840                 845
Lys Gln Leu Ala Val Ala Glu Gly Lys Pro Pro Glu Ala Pro Lys Gly
    850                 855                 860
Lys Lys Lys Lys
865
```

<210> 25
<211> 2607
<212> ADN
<213> *Homo sapiens*

<400> 25

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600

tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gccccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaaa ggaatgtgct catcaccagt gccctccctt acgtcaacaa tgtcccccac 840
cttgggaaca tcattggttg tgtgctcagt gccgatgtct ttgccaggta ctctcgcctc 900
cgccagtgga acaccctcta tctgtgtggg acagatgagt atggtacagc aacagagacc 960
aaggctctgg aggagggact aacccccccag gagatctgcg acaagtacca catcatccat 1020
gctgacatct accgctggtt taacatttcg tttgatattt ttggtcgcac caccactcca 1080
cagcagacca aaatcaccca ggacattttc cagcagttgc tgaaacgagg ttttgtgctg 1140
caagatactg tggagcaact gcgatgtgag cactgtgctc gcttcctggc tgaccgcttc 1200
gtggagggcg tgtgtccctt ctgtggctat gaggaggctc ggggtgacca gtgtgacaag 1260
tgtggcaagc tcatcaatgc tgtcgagctt aagaagcctc agtgtaaagt ctgccgatca 1320
tgccctgtgg tgcagtcgag ccagcacctg tttctggacc tgcctaagct ggagaagcga 1380
ctggaggagt ggttggggag gacattgcct ggcagtgact ggacacccaa tgcccagttt 1440
atcacccgtt cttggcttcg ggatggcctc aagccacgct gcataacccg agacctcaaa 1500
tggggaaccc ctgtaccctt agaaggtttt gaagacaagg tggacctgta tcagttcatg 1560
gccaaagaca atgttccttt ccatagctta gtctttcctt gctcagccct aggagctgag 1620
gataactata ccttggtcag ccacctcatt gctacagagt acctgaacta tgaggatggg 1680
aaattctcta agagccgcgg tgtgggagtg tttggggaca tggcccagga cacggggatc 1740
cctgctgaca tctggcgctt ctatctgctg tacattcggc ctgagggcca ggacagtgct 1800
ttctcctgga cggacctgct gctgaagaat aattctgagc tgcttaacaa cctgggcaac 1860
ttcatcaaca gagctgggat gtttgtgtct aagttctttg ggggctatgt gcctgagatg 1920
gtgctcaccc ctgatgatca gcgcctgctg gcccatgtca ccctggagct ccagcactat 1980
caccagctac ttgagaaggt tcggatccgg gatgccttgc gcagtatcct caccatatct 2040
cgacatggca accaatatat tcaggtgaat gagccctgga agcggattaa aggcagtgag 2100
gctgacaggc aacgggcagg aacagtgact ggcttggcag tgaatatagc tgccttgctc 2160
tctgtcatgc ttcagcctta catgcccacg gttagtgcca caatccaggc ccagctgcag 2220
ctcccacctc cagcctgcag tatcctgctg acaaacttcc tgtgtacctt accagcagga 2280
caccagattg gcacagtcag tcccttgttc caaaaattgg aaaatgacca gattgaaagt 2340
ttaaggcagc gctttggagg gggccaggca aaaacgtccc cgaagccagc agttgtagag 2400
actgttacaa cagccaagcc acagcagata caagcgctga tggatgaagt gacaaaacaa 2460
ggaaacattg tccgagaact gaaagcacaa aaggcagaca agaacgaggt tgctgcggag 2520
gtggcgaaac tcttggatct aaagaaacag ttggctgtag ctgaggggaa acccccctgaa 2580
gcccctaaag gcaagaagaa aaagtaa                                      2607
```

<210> 26
<211> 547
<212> PRT
<213> *Homo sapiens*

<400> 26

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

```
Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu Glu Glu Glu
    210                 215                 220
Leu Ala Thr Leu Ser Glu Glu Glu Ile Ala Met Ala Val Thr Ala Trp
225                 230                 235                 240
Glu Lys Gly Leu Glu Ser Leu Pro Pro Leu Arg Pro Gln Gln Asn Pro
                245                 250                 255
Val Leu Pro Val Ala Gly Glu Arg Asn Val Leu Ile Thr Ser Ala Leu
            260                 265                 270
Pro Tyr Val Asn Asn Val Pro His Leu Gly Asn Ile Ile Gly Cys Val
        275                 280                 285
Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser Arg Leu Arg Gln Trp Asn
    290                 295                 300
Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly Thr Ala Thr Glu Thr
305                 310                 315                 320
Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile Cys Asp Lys Tyr
                325                 330                 335
His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn Ile Ser Phe Asp
            340                 345                 350
Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys Ile Thr Gln Asp
        355                 360                 365
Ile Phe Gln Gln Leu Leu Lys Arg Gly Phe Val Leu Gln Asp Thr Val
    370                 375                 380
Glu Gln Leu Arg Cys Glu His Cys Ala Arg Phe Leu Ala Asp Arg Phe
385                 390                 395                 400
Val Glu Gly Val Cys Pro Phe Cys Gly Tyr Glu Glu Ala Arg Gly Asp
                405                 410                 415
Gln Cys Asp Lys Cys Gly Lys Leu Ile Asn Ala Val Glu Leu Lys Lys
            420                 425                 430
Pro Gln Cys Lys Val Cys Arg Ser Cys Pro Val Val Gln Ser Ser Gln
        435                 440                 445
His Leu Phe Leu Asp Leu Pro Lys Leu Glu Lys Arg Leu Glu Glu Trp
    450                 455                 460
Leu Gly Arg Thr Leu Pro Gly Ser Asp Trp Thr Pro Asn Ala Gln Phe
465                 470                 475                 480
Ile Thr Arg Ser Trp Leu Arg Asp Gly Leu Lys Pro Arg Cys Ile Thr
                485                 490                 495
Arg Asp Leu Lys Trp Gly Thr Pro Val Pro Leu Glu Gly Phe Glu Asp
            500                 505                 510
Lys Val Phe Tyr Val Trp Phe Asp Ala Thr Ile Gly Tyr Leu Ser Ile
        515                 520                 525
Thr Ala Asn Tyr Thr Asp Gln Trp Glu Arg Trp Trp Lys Asn Pro Glu
    530                 535                 540
Gln Ser Thr
545
```

<210> 27
<211> 1644
<212> ADN
<213> *Homo sapiens*

<400> 27

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660

gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gccccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaaa ggaatgtgct catcaccagt gccctccctt acgtcaacaa tgtcccccac 840
cttgggaaca tcattggttg tgtgctcagt gccgatgtct ttgccaggta ctctcgcctc 900
cgccagtgga acacccctcta tctgtgtggg acagatgagt atggtacagc aacagagacc 960
aaggctctgg aggagggact aacccccccag gagatctgcg acaagtacca catcatccat 1020
gctgacatct accgctggtt taacatttcg tttgatattt ttggtcgcac caccactcca 1080
cagcagacca aaatcaccca ggacattttc cagcagttgc tgaaacgagg ttttgtgctg 1140
caagatactg tggagcaact gcgatgtgag cactgtgctc gcttcctggc tgaccgcttc 1200
gtggagggcg tgtgtcccctt ctgtggctat gaggaggctc ggggtgacca gtgtgacaag 1260
tgtggcaagc tcatcaatgc tgtcgagctt aagaagcctc agtgtaaagt ctgccgatca 1320
tgccctgtgg tgcagtcgag ccagcacctg tttctggacc tgcctaagct ggagaagcga 1380
ctggaggagt ggttggggag gacattgcct ggcagtgact ggacacccaa tgcccagttt 1440
atcacccgtt cttggcttcg ggatggcctc aagccacgct gcataacccg agacctcaaa 1500
tggggaaccc ctgtacccctt agaaggtttt gaagacaagg tattctatgt ctggtttgat 1560
gccactattg gctatctgtc catcacagcc aactacacag accagtggga gagatggtgg 1620
aagaacccag agcaaagtac ctga                                        1644
```

<210> 28
<211> 750
<212> PRT
<213> *Homo sapiens*

<400> 28

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

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

<210> 29
<211> 2253
<212> ADN
<213> *Homo sapiens*

<400> 29

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagccccgg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240

actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gcccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaaa ggaatgtgct catcaccagt gccctccctt acgtcaacaa tgtcccccac 840
cttgggaaca tcattggttg tgtgctcagt gccgatgtct ttgccaggta ctctcgcctc 900
cgccagtgga acaccctcta tctgtgtggg acagatgagt atggtacagc aacagagacc 960
aaggctctgg aggagggact aacccccccag gagatctgcg acaagtacca catcatccat 1020
gctgacatct accgctggtt taacatttcg tttgatattt ttggtcgcac caccactcca 1080
cagcagacca aaatcaccca ggacattttc cagcagttgc tgaaacgagg ttttgtgctg 1140
caagatactg tggagcaact gcgatgtgag cactgtgctc gcttcctggc tgaccgcttc 1200
gtggagggcg tgtgtccctt ctgtggctat gaggaggctc ggggtgacca gtgtgacaag 1260
tgtggcaagc tcatcaatgc tgtcgagctt aagaagcctc agtgtaaagt ctgccgatca 1320
tgccctgtgg tgcagtcgag ccagcacctg tttctggacc tgcctaagct ggagaagcga 1380
ctggaggagt ggttggggag gacattgcct ggcagtgact ggacacccaa tgcccagttt 1440
atcacccgtt cttggcttcg ggatggcctc aagccacgct gcataacccg agacctcaaa 1500
tggggaaccc ctgtaccctt agaaggtttt gaagacaagg tattctatgt ctggtttgat 1560
gccactattg gctatctgtc catcacagcc aactacacag accagtggga gagatggtgg 1620
aagaacccag agcaagtgga cctgtatcag ttcatggcca aagacaatgt tcctttccat 1680
agcttagtct ttccttgctc agccctagga gctgaggata actatacctt ggtcagccac 1740
ctcattgcta cagagtacct gaactatgag gatgggaaat tctctaagag ccgcggtgtg 1800
ggagtgtttg gggacatggc ccaggacacg gggatccctg ctgacatctg gcgcttctat 1860
ctgctgtaca ttcggcctga gggccaggac agtgctttct cctggacgga cctgctgctg 1920
aagaataatt ctgagctgct taacaacctg ggcaacttca tcaacagagc tgggatgttt 1980
gtgtctaagt tctttggggg ctatgtgcct gagatggtgc tcacccctga tgatcagcgc 2040
ctgctggccc atgtcaccct ggagctccag cactatcacc agctacttga gaaggttcgg 2100
atccgggatg ccttgcgcag tatcctcacc atatctcgac atggcaacca atatattcag 2160
gtgaatgagc cctggaagcg gattaaaggc agtgaggctg acaggtcagt cccttgttcc 2220
aaaaattgga aaatgaccag attgaaagtt taa                              2253
```

<210> 30
<211> 869
<212> PRT
<213> *Homo sapiens*

<400> 30

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

```
Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu Glu Glu Glu
    210                 215                 220
Leu Ala Thr Leu Ser Glu Glu Glu Ile Ala Met Ala Val Thr Ala Trp
225                 230                 235                 240
Glu Lys Gly Leu Glu Ser Leu Pro Pro Leu Arg Pro Gln Gln Asn Pro
                245                 250                 255
Val Leu Pro Val Ala Gly Glu Arg Asn Val Leu Ile Thr Ser Ala Leu
            260                 265                 270
Pro Tyr Val Asn Asn Val Pro His Leu Gly Asn Ile Ile Gly Cys Val
        275                 280                 285
Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser Arg Leu Arg Gln Trp Asn
    290                 295                 300
Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly Thr Ala Thr Glu Thr
305                 310                 315                 320
Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile Cys Asp Lys Tyr
                325                 330                 335
His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn Ile Ser Phe Asp
            340                 345                 350
Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys Ile Thr Gln Asp
        355                 360                 365
Ile Phe Gln Gln Leu Leu Lys Arg Gly Phe Val Leu Gln Asp Thr Val
    370                 375                 380
Glu Gln Leu Arg Cys Glu His Cys Ala Arg Phe Leu Ala Asp Arg Phe
385                 390                 395                 400
Val Glu Gly Val Cys Pro Phe Cys Gly Tyr Glu Glu Ala Arg Gly Asp
                405                 410                 415
Gln Cys Asp Lys Cys Gly Lys Leu Ile Asn Ala Val Glu Leu Lys Lys
            420                 425                 430
Pro Gln Cys Lys Val Cys Arg Ser Cys Pro Val Val Gln Ser Ser Gln
        435                 440                 445
His Leu Phe Leu Asp Leu Pro Lys Leu Glu Lys Arg Leu Glu Glu Trp
    450                 455                 460
Leu Gly Arg Thr Leu Pro Gly Ser Asp Trp Thr Pro Asn Ala Gln Phe
465                 470                 475                 480
Ile Thr Arg Ser Trp Leu Arg Asp Gly Leu Lys Pro Arg Cys Ile Thr
                485                 490                 495
Arg Asp Leu Lys Trp Gly Thr Pro Val Pro Leu Glu Gly Phe Glu Asp
            500                 505                 510
Lys Val Phe Tyr Val Trp Phe Asp Ala Thr Ile Gly Tyr Leu Ser Ile
        515                 520                 525
Thr Ala Asn Tyr Thr Asp Gln Trp Glu Arg Trp Trp Lys Asn Pro Glu
    530                 535                 540
Gln Val Asp Leu Tyr Gln Phe Met Ala Lys Asp Asn Val Pro Phe His
545                 550                 555                 560
Ser Leu Val Phe Pro Cys Ser Ala Leu Gly Ala Glu Asp Asn Tyr Thr
                565                 570                 575
Leu Val Ser His Leu Ile Ala Thr Glu Tyr Leu Asn Tyr Glu Asp Gly
            580                 585                 590
Lys Phe Ser Lys Ser Arg Gly Val Gly Val Phe Gly Asp Met Ala Gln
        595                 600                 605
Asp Thr Gly Ile Pro Ala Asp Ile Trp Arg Phe Tyr Leu Leu Tyr Ile
    610                 615                 620
Arg Pro Glu Gly Gln Asp Ser Ala Phe Ser Trp Thr Asp Leu Leu Leu
625                 630                 635                 640
Lys Asn Asn Ser Glu Leu Leu Asn Asn Leu Gly Asn Phe Ile Asn Arg
                645                 650                 655
Ala Gly Met Phe Val Ser Lys Phe Phe Gly Gly Tyr Val Pro Glu Met
            660                 665                 670
Val Leu Thr Pro Asp Asp Gln Arg Leu Leu Ala His Val Thr Leu Glu
        675                 680                 685
Leu Gln His Tyr His Gln Leu Leu Glu Lys Val Arg Ile Arg Asp Ala
    690                 695                 700
Leu Arg Ser Ile Leu Thr Ile Ser Arg His Gly Asn Gln Tyr Ile Gln
```

```
705                 710                 715                 720
Val Asn Glu Pro Trp Lys Arg Ile Lys Gly Ser Glu Ala Asp Arg Gln
            725                 730                 735
Arg Ala Gly Thr Val Thr Gly Leu Ala Val Asn Ile Ala Ala Leu Leu
        740                 745                 750
Ser Val Met Leu Gln Pro Tyr Met Pro Thr Val Ser Ala Thr Ile Gln
    755                 760                 765
Ala Gln Leu Gln Leu Pro Pro Pro Ala Cys Ser Ile Leu Leu Thr Asn
770                 775                 780
Phe Leu Cys Thr Leu Pro Ala Gly His Gln Ile Gly Thr Val Ser Pro
785                 790                 795                 800
Leu Phe Gln Lys Leu Glu Asn Asp Gln Ile Glu Ser Leu Arg Gln Arg
            805                 810                 815
Phe Gly Gly Gly Gln Gly Asn Ile Val Arg Glu Leu Lys Ala Gln Lys
        820                 825                 830
Ala Asp Lys Asn Glu Val Ala Ala Glu Val Ala Lys Leu Leu Asp Leu
    835                 840                 845
Lys Lys Gln Leu Ala Val Ala Glu Gly Lys Pro Pro Glu Ala Pro Lys
850                 855                 860
Gly Lys Lys Lys Lys
865
```

<210> 31
<211> 2610
<212> ADN
<213> *Homo sapiens*

<400> 31

atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gcccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaaa ggaatgtgct catcaccagt gccctccctt acgtcaacaa tgtcccccac 840
cttgggaaca tcattggttg tgtgctcagt gccgatgtct ttgccaggta ctctcgcctc 900
cgccagtgga acacccctcta tctgtgtggg acagatgagt atggtacagc aacagagacc 960
aaggctctgg aggagggact aacccccccag gagatctgcg acaagtacca catcatccat 1020
gctgacatct accgctggtt taacatttcg tttgatattt ttggtcgcac caccactcca 1080
cagcagacca aaatcaccca ggacattttc cagcagttgc tgaaacgagg ttttgtgctg 1140
caagatactg tggagcaact gcgatgtgag cactgtgctc gcttcctggc tgaccgcttc 1200
gtggagggcg tgtgtccctt ctgtggctat gaggaggctc ggggtgacca gtgtgacaag 1260
tgtggcaagc tcatcaatgc tgtcgagctt aagaagcctc agtgtaaagt ctgccgatca 1320
tgccctgtgg tgcagtcgag ccagcacctg tttctggacc tgcctaagct ggagaagcga 1380
ctggaggagt ggttggggag gacattgcct ggcagtgact ggacacccaa tgcccagttt 1440
atcacccgtt cttggcttcg ggatggcctc aagccacgct gcataacccg agacctcaaa 1500
tggggaaccc ctgtaccctt agaaggtttt gaagacaagg tattctatgt ctggtttgat 1560
gccactattg gctatctgtc catcacagcc aactacacag accagtggga gagatggtgg 1620
aagaacccag agcaagtgga cctgtatcag ttcatggcca aagacaatgt tcctttccat 1680
agcttagtct ttccttgctc agccctagga gctgaggata actatacctt ggtcagccac 1740
ctcattgcta cagagtacct gaactatgag gatgggaaat tctctaagag ccgcggtgtg 1800
ggagtgtttg gggacatggc ccaggacacg gggatccctg ctgacatctg gcgcttctat 1860
ctgctgtaca ttcggcctga gggccaggac agtgctttct cctggacgga cctgctgctg 1920
aagaataatt ctgagctgct taacaacctg ggcaacttca tcaacagagc tgggatgttt 1980
gtgtctaagt tctttggggg ctatgtgcct gagatggtgc tcacccctga tgatcagcgc 2040
ctgctggccc atgtcaccct ggagctccag cactatcacc agctacttga gaaggttcgg 2100
atccgggatg ccttgcgcag tatcctcacc atatctcgac atggcaacca atatattcag 2160

gtgaatgagc cctggaagcg gattaaaggc agtgaggctg acaggcaacg ggcaggaaca 2220
gtgactggct tggcagtgaa tatagctgcc ttgctctctg tcatgcttca gccttacatg 2280
cccacggtta gtgccacaat ccaggcccag ctgcagctcc cacctccagc ctgcagtatc 2340
ctgctgacaa acttcctgtg taccttacca gcaggacacc agattggcac agtcagtccc 2400
ttgttccaaa aattggaaaa tgaccagatt gaaagtttaa ggcagcgctt tggaggggcc 2460
cagggaaaca ttgtccgaga actgaaagca caaaaggcag acaagaacga ggttgctgcg 2520
gaggtggcga aactcttgga tctaaagaaa cagttggctg tagctgaggg gaaacccccct 2580
gaagccccta aaggcaagaa gaaaaagtaa 2610

<210> 32
<211> 114
<212> PRT
<213> *Homo sapiens*

<400> 32

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

<210> 33
<211> 345
<212> ADN
<213> *Homo sapiens*

<400> 33

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg aggagctgag tgccctgcac agctggttcc agacactgag 240
tacccaggaa ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc 300
tctccggcct tacctccaaa agcagcccca gcccagcccc gctga                  345
```

<210> 34
<211> 276
<212> PRT
<213> *Homo sapiens*

<400> 34

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

```
                100                 105                 110
Val Leu Gly Ser Val Arg Arg Ala Leu Thr His Ile Asp His Ser Leu
        115                 120                 125
Ser Arg Gln Asn Cys Pro Phe Leu Ala Gly Glu Thr Glu Ser Leu Ala
    130                 135                 140
Asp Ile Val Leu Trp Gly Ala Leu Tyr Pro Leu Leu Gln Asp Pro Ala
145                 150                 155                 160
Tyr Leu Pro Glu Glu Leu Ser Ala Leu His Ser Trp Phe Gln Thr Leu
                165                 170                 175
Ser Thr Gln Glu Pro Cys Gln Arg Ala Ala Glu Thr Val Leu Lys Gln
            180                 185                 190
Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu Val Ala Cys
    210                 215                 220
Gly Trp Arg Lys Glu Cys Ala His His Gln Cys Pro Pro Leu Arg Gln
225                 230                 235                 240
Gln Cys Pro Pro Pro Trp Glu His His Trp Leu Cys Ala Gln Cys Arg
                245                 250                 255
Cys Leu Cys Gln Val Leu Ser Pro Pro Pro Val Glu His Pro Leu Ser
            260                 265                 270
Val Trp Asp Arg
        275
```

<210> 35
<211> 831
<212> ADN
<213> *Homo sapiens*

<400> 35

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagccccgg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggttgcct gtggctggag aaaggaatgt gctcatcacc agtgccctcc cttacgtcaa 720
caatgtcccc caccttggga acatcattgg ttgtgtgctc agtgccgatg tctttgccag 780
gtactctcgc ctccgccagt ggaacaccct ctatctgtgt gggacagatg a          831
```

<210> 36
<211> 273
<212> PRT
<213> *Homo sapiens*

<400> 36

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

            100                 105                 110
Val Leu Gly Ser Val Arg Arg Ala Leu Thr His Ile Asp His Ser Leu
        115                 120                 125
Ser Arg Gln Asn Cys Pro Phe Leu Ala Gly Glu Thr Glu Ser Leu Ala
    130                 135                 140
Asp Ile Val Leu Trp Gly Ala Leu Tyr Pro Leu Leu Gln Asp Pro Ala
145                 150                 155                 160
Tyr Leu Pro Glu Glu Leu Ser Ala Leu His Ser Trp Phe Gln Thr Leu
                165                 170                 175
Ser Thr Gln Glu Pro Cys Gln Arg Ala Ala Glu Thr Val Leu Lys Gln
            180                 185                 190
Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu Glu Glu Glu
    210                 215                 220
Leu Ala Thr Leu Ser Glu Glu Glu Ile Ala Met Ala Val Thr Ala Trp
225                 230                 235                 240
Glu Lys Gly Leu Glu Ser Leu Pro Pro Leu Arg Pro Gln Gln Asn Pro
                245                 250                 255
Val Trp Thr Cys Ile Ser Ser Trp Pro Lys Thr Met Phe Leu Ser Ile
            260                 265                 270
Ala
```

<210> 37
<211> 822
<212> ADN
<213> *Homo sapiens*

<400> 37

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gccccccgctg cggccccagc agaatccagt gtggacctgt 780
atcagttcat ggccaaagac aatgttcctt tccatagctt ag                    822
```

<210> 38
<211> 876
<212> PRT
<213> *Homo sapiens*

<400> 38

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

```
                100                 105                 110
Val Leu Gly Ser Val Arg Arg Ala Leu Thr His Ile Asp His Ser Leu
        115                 120                 125
Ser Arg Gln Asn Cys Pro Phe Leu Ala Gly Glu Thr Glu Ser Leu Ala
    130                 135                 140
Asp Ile Val Leu Trp Gly Ala Leu Tyr Pro Leu Leu Gln Asp Pro Ala
145                 150                 155                 160
Tyr Leu Pro Glu Glu Leu Ser Ala Leu His Ser Trp Phe Gln Thr Leu
                165                 170                 175
Ser Thr Gln Glu Pro Cys Gln Arg Ala Ala Glu Thr Val Leu Lys Gln
            180                 185                 190
Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu Glu Glu Glu
    210                 215                 220
Leu Ala Thr Leu Ser Glu Glu Glu Ile Ala Met Ala Val Thr Ala Trp
225                 230                 235                 240
Glu Lys Gly Leu Glu Ser Leu Pro Pro Leu Arg Pro Gln Gln Asn Pro
                245                 250                 255
Val Leu Pro Val Ala Gly Glu Arg Asn Val Leu Ile Thr Ser Ala Leu
            260                 265                 270
Pro Tyr Val Asn Asn Val Pro His Leu Gly Asn Ile Ile Gly Cys Val
        275                 280                 285
Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser Arg Leu Arg Gln Trp Asn
    290                 295                 300
Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly Thr Ala Thr Glu Thr
305                 310                 315                 320
Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile Cys Asp Lys Tyr
                325                 330                 335
His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn Ile Ser Phe Asp
            340                 345                 350
Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys Ile Thr Gln Asp
        355                 360                 365
Ile Phe Gln Gln Leu Leu Lys Arg Gly Phe Val Leu Gln Asp Thr Val
    370                 375                 380
Glu Gln Leu Arg Cys Glu His Cys Ala Arg Phe Leu Ala Asp Arg Phe
385                 390                 395                 400
Val Glu Gly Val Cys Pro Phe Cys Gly Tyr Glu Glu Ala Arg Gly Asp
                405                 410                 415
Gln Cys Asp Lys Cys Gly Lys Leu Ile Asn Ala Val Glu Leu Lys Lys
            420                 425                 430
Pro Gln Cys Lys Val Cys Arg Ser Cys Pro Val Val Gln Ser Ser Gln
        435                 440                 445
His Leu Phe Leu Asp Leu Pro Lys Leu Glu Lys Arg Leu Glu Glu Trp
    450                 455                 460
Leu Gly Arg Thr Leu Pro Gly Ser Asp Trp Thr Pro Asn Ala Gln Phe
465                 470                 475                 480
Ile Thr Arg Ser Trp Leu Arg Asp Gly Leu Lys Pro Arg Cys Ile Thr
                485                 490                 495
Arg Asp Leu Lys Trp Gly Thr Pro Val Pro Leu Glu Gly Phe Glu Asp
            500                 505                 510
Lys Val Phe Tyr Val Trp Phe Asp Ala Thr Ile Gly Tyr Leu Ser Ile
        515                 520                 525
Thr Ala Asn Tyr Thr Asp Gln Trp Glu Arg Trp Trp Lys Asn Pro Glu
    530                 535                 540
Gln Val Asp Leu Tyr Gln Phe Met Ala Lys Asp Asn Val Pro Phe His
545                 550                 555                 560
Ser Leu Val Phe Pro Cys Ser Ala Leu Gly Ala Glu Asp Asn Tyr Thr
                565                 570                 575
Leu Val Ser His Leu Ile Ala Thr Glu Tyr Leu Asn Tyr Glu Asp Gly
            580                 585                 590
Lys Phe Ser Lys Ser Arg Gly Val Gly Val Phe Gly Asp Met Ala Gln
        595                 600                 605
Asp Thr Gly Ile Pro Ala Asp Ile Trp Arg Phe Tyr Leu Leu Tyr Ile
    610                 615                 620
```

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

<210> 39
<211> 2631
<212> ADN
<213> *Homo sapiens*

<400> 39

atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gccccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaaa ggaatgtgct catcaccagt gccctccctt acgtcaacaa tgtcccccac 840
cttgggaaca tcattggttg tgtgctcagt gccgatgtct ttgccaggta ctctcgcctc 900
cgccagtgga acacccctcta tctgtgtggg acagatgagt atggtacagc aacagagacc 960
aaggctctgg aggagggact aacccccccag gagatctgcg acaagtacca catcatccat 1020
gctgacatct accgctggtt taacatttcg tttgatattt ttggtcgcac caccactcca 1080
cagcagacca aaatcaccca ggacattttc cagcagttgc tgaaacgagg ttttgtgctg 1140
caagatactg tggagcaact gcgatgtgag cactgtgctc gcttcctggc tgaccgcttc 1200
gtggagggcg tgtgtccctt ctgtggctat gaggaggctc ggggtgacca gtgtgacaag 1260
tgtggcaagc tcatcaatgc tgtcgagctt aagaagcctc agtgtaaagt ctgccgatca 1320
tgccctgtgg tgcagtcgag ccagcacctg tttctggacc tgcctaagct ggagaagcga 1380
ctggaggagt ggttggggag gacattgcct ggcagtgact ggacacccaa tgcccagttt 1440
atcacccgtt cttggcttcg ggatggcctc aagccacgct gcataacccg agacctcaaa 1500

tggggaaccc ctgtaccctt agaaggtttt gaagacaagg tattctatgt ctggtttgat 1560
gccactattg gctatctgtc catcacagcc aactacacag accagtggga gagatggtgg 1620
aagaacccag agcaagtgga cctgtatcag ttcatggcca aagacaatgt tcctttccat 1680
agcttagtct ttccttgctc agccctagga gctgaggata actatacctt ggtcagccac 1740
ctcattgcta cagagtacct gaactatgag gatgggaaat tctctaagag ccgcggtgtg 1800
ggagtgtttg gggacatggc ccaggacacg gggatccctg ctgacatctg gcgcttctat 1860
ctgctgtaca ttcggcctga gggccaggac agtgctttct cctggacgga cctgctgctg 1920
aagaataatt ctgagctgct taacaacctg ggcaacttca tcaacagagc tgggatgttt 1980
gtgtctaagt tctttggggg ctatgtgcct gagatggtgc tcaccccctga tgatcagcgc 2040
ctgctggccc atgtcaccct ggagctccag cactatcacc agctacttga gaaggttcgg 2100
atccgggatg ccttgcgcag tatcctcacc atatctcgac atggcaacca atatattcag 2160
gtgaatgagc cctggaagcg gattaaaggc agtgaggctg acaggcaacg ggcaggaaca 2220
gtgactggct tggcagtgaa tatagctgcc ttgctctctg tcatgcttca gccttacatg 2280
cccacggtta gtgccacaat ccaggcccag ctgcagctcc cacctccagc ctgcagtatc 2340
ctgctgacaa acttcctgtg taccttacca gcaggacacc agattggcac agcaaaaacg 2400
tccccgaagc cagcagttgt agagactgtt acaacagcca agccacagca gatacaagcg 2460
ctgatggatg aagtgacaaa acaaggaaac attgtccgag aactgaaagc acaaaaggca 2520
gacaagaacg aggttgctgc ggaggtggcg aaactcttgg atctaaagaa acagttggct 2580
gtagctgagg ggaaaccccc tgaagccctt aaaggcaaga agaaaaagta a 2631

<210> 40
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 40
tcttctccac tagtgcaatc tgccggtatt ctatgtctgg tttgatgcca 50

<210> 41
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 41

Phe Ser Thr Ser Ala Ile Cys Arg Tyr Ser Met Ser Gly Leu Met Pro
1 5 10 15

<210> 42
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 42
gcaccaccac tccacagcag accaaaagcc tcagtgtaaa gtctgccgat 50

<210> 43
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 43

Thr Thr Thr Pro Gln Gln Thr Lys Ser Leu Ser Val Lys Ser Ala Asp
1 5 10 15

<210> 44
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 44
acccttagaa ggttttgaag acaaggtgga cctgtatcag ttcatggcca 50

<210> 45
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 45

Pro Leu Glu Gly Phe Glu Asp Lys Val Asp Leu Tyr Gln Phe Met Ala
1 5 10 15

<210> 46
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 46
gagatggtgg aagaacccag agcaaagtac ctgaactatg aggatgggaa 50

<210> 47
<211> 10
<212> PRT
<213> *Homo sapiens*

<400> 47

Arg Trp Trp Lys Asn Pro Glu Gln Ser Thr
1 5 10

<210> 48
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 48
ggattaaagg cagtgaggct gacaggtcag tcccttgttc caaaaattgg 50

<210> 49
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 49

Ile Lys Gly Ser Glu Ala Asp Arg Ser Val Pro Cys Ser Lys Asn Trp
1 5 10 15

<210> 50
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 50
aaggcagcgc tttggagggg gccagggaaa cattgtccga gaactgaaag 50

<210> 51
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 51

Arg Gln Arg Phe Gly Gly Gly Gln Gly Asn Ile Val Arg Glu Leu Lys
1 5 10 15

<210> 52
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 52
tcttctccac tagtgcaatc tgccgaggag ctgagtgccc tgcacagctg 50

<210> 53
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 53

Phe Ser Thr Ser Ala Ile Cys Arg Gly Ala Glu Cys Pro Ala Gln Leu
1 5 10 15

<210> 54
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 54
aagggctgtc accaatgagc ctgaggttgc ctgtggctgg agaaaggaat 50

<210> 55
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 55

Arg Ala Val Thr Asn Glu Pro Glu Val Ala Cys Gly Trp Arg Lys Glu
1 5 10 15

<210> 56
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 56
cgctgcggcc ccagcagaat ccagtgtgga cctgtatcag ttcatggcca 50

<210> 57
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 57

Leu Arg Pro Gln Gln Asn Pro Val Trp Thr Cys Ile Ser Ser Trp Pro
1 5 10 15

<210> 58
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 58
accagcagga caccagattg gcacagcaaa aacgtccccg aagccagcag 50

<210> 59
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 59

Pro Ala Gly His Gln Ile Gly Thr Ala Lys Thr Ser Pro Lys Pro Ala
1 5 10 15

<210> 60
<211> 263
<212> PRT
<213> *Homo sapiens*

<400> 60

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

<210> 61
<211> 789
<212> ADN
<213> *Homo sapiens*

<400> 61

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atattttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720

gagaagggcc tagaaagttt gcccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaa                                                        789
```

<210> 62

<211> 197
<212> PRT
<213> *Homo sapiens*

<400> 62

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

<210> 63
<211> 591
<212> ADN
<213> *Homo sapiens*

<400> 63

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc t          591
```

<210> 64
<211> 214
<212> PRT
<213> *Homo sapiens*

<400> 64

```
Met Arg Leu Phe Val Ser Asp Gly Val Pro Gly Cys Leu Pro Val Leu
 1               5                  10                  15
Ala Ala Ala Gly Arg Ala Arg Gly Arg Ala Glu Val Leu Ile Ser Thr
            20                  25                  30
Val Gly Pro Glu Asp Cys Val Val Pro Phe Leu Thr Arg Pro Lys Val

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

<210> 65
<211> 642
<212> ADN
<213> *Homo sapiens*

<400> 65

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gg                    642
```

<210> 66
<211> 221
<212> PRT
<213> *Homo sapiens*

<400> 66

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

Ser Arg Gln Asn Cys Pro Phe Leu Ala Gly Glu Thr Glu Ser Leu Ala
    130                 135                 140
Asp Ile Val Leu Trp Gly Ala Leu Tyr Pro Leu Leu Gln Asp Pro Ala
145                 150                 155                 160
Tyr Leu Pro Glu Glu Leu Ser Ala Leu His Ser Trp Phe Gln Thr Leu
                165                 170                 175
Ser Thr Gln Glu Pro Cys Gln Arg Ala Ala Glu Thr Val Leu Lys Gln
            180                 185                 190
Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu
    210                 215                 220
```

<210> 67
<211> 663
<212> ADN
<213> *Homo sapiens*

<400> 67

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gag                                                               663
```

<210> 68
<211> 251
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 68

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

Gln Asp Pro Ala Tyr Leu Pro Glu Glu Leu Ser Ala Leu His Ser Trp
            180                 185                 190
Phe Gln Thr Leu Ser Thr Gln Glu Pro Cys Gln Arg Ala Ala Glu Thr
        195                 200                 205
Val Leu Lys Gln Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys
    210                 215                 220
Gln Pro Gln Pro Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro
225                 230                 235                 240
Glu Glu Glu Glu Leu Ala Thr Leu Ser Glu Glu
                245                 250
```

<210> 69
<211> 251
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 69

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

<210> 70
<211> 283
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 70

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

<210> 71
<211> 283
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 71

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

Ser Arg Gln Asn Cys Pro Phe Leu Ala Gly Glu Thr Glu Ser Leu Ala
    130                 135                 140
Asp Ile Val Leu Trp Gly Ala Leu Tyr Pro Leu Leu Gln Asp Pro Ala
145                 150                 155                 160
Tyr Leu Pro Glu Glu Leu Ser Ala Leu His Ser Trp Phe Gln Thr Leu
                165                 170                 175
Ser Thr Gln Glu Pro Cys Gln Arg Ala Ala Glu Thr Val Leu Lys Gln
            180                 185                 190
Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu Glu Glu Glu
    210                 215                 220
Leu Ala Thr Leu Ser Glu Glu Glu Ile Ala Met Ala Val Thr Ala Trp
225                 230                 235                 240
Glu Lys Gly Leu Glu Ser Leu Pro Pro Leu Arg Pro Gln Gln Asn Pro
                245                 250                 255
Val Leu Pro Val Ala Gly Glu Gly Lys Pro Ile Pro Asn Pro Leu Leu
            260                 265                 270
Gly Leu Asp Ser Thr His His His His His His
        275                 280
```

<210> 72
<211> 135
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 72

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

<210> 73
<211> 135
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 73

```
Met Arg Leu Phe Val Ser Asp Gly Val Pro Gly Cys Leu Pro Val Leu
 1               5                  10                  15
Ala Ala Ala Gly Arg Ala Arg Gly Arg Ala Glu Val Leu Ile Ser Thr


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

<210> 74
<211> 447
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 74

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

Ile Gly Cys Val Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser Arg Leu
305                 310                 315                 320
Arg Gln Trp Asn Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly Thr
                325                 330                 335
Ala Thr Glu Thr Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile
            340                 345                 350
Cys Asp Lys Tyr His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn
        355                 360                 365
Ile Ser Phe Asp Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys
    370                 375                 380
Ser Leu Ser Val Lys Ser Ala Asp His Ala Leu Trp Cys Ser Arg Ala
385                 390                 395                 400
Ser Thr Cys Phe Trp Thr Cys Leu Ser Trp Arg Ser Asp Trp Arg Ser
                405                 410                 415
Gly Trp Gly Gly His Cys Leu Ala Val Thr Gly His Pro Met Pro Ser
            420                 425                 430
Leu Ser Pro Val Leu Gly Phe Gly Met Ala Ser Ser His Ala Ala
        435                 440                 445
```

<210> 75
<211> 447

<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 75

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

```
        275                 280                 285
Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser Arg Leu Arg Gln Trp Asn
    290                 295                 300
Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly Thr Ala Thr Glu Thr
305                 310                 315                 320
Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile Cys Asp Lys Tyr
                325                 330                 335
His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn Ile Ser Phe Asp
            340                 345                 350
Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys Ser Leu Ser Val
        355                 360                 365
Lys Ser Ala Asp His Ala Leu Trp Cys Ser Arg Ala Ser Thr Cys Phe
    370                 375                 380
Trp Thr Cys Leu Ser Trp Arg Ser Asp Trp Arg Ser Gly Trp Gly Gly
385                 390                 395                 400
His Cys Leu Ala Val Thr Gly His Pro Met Pro Ser Leu Ser Pro Val
                405                 410                 415
Leu Gly Phe Gly Met Ala Ser Ser His Ala Ala Gly Lys Pro Ile Pro
            420                 425                 430
Asn Pro Leu Leu Gly Leu Asp Ser Thr His His His His His His
        435                 440                 445
```

<210> 76
<211> 567
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 76

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

```
Val Thr Ala Trp Glu Lys Gly Leu Glu Ser Leu Pro Pro Leu Arg Pro
            260                 265                 270
Gln Gln Asn Pro Val Leu Pro Val Ala Gly Glu Arg Asn Val Leu Ile
        275                 280                 285
Thr Ser Ala Leu Pro Tyr Val Asn Asn Val Pro His Leu Gly Asn Ile
    290                 295                 300
Ile Gly Cys Val Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser Arg Leu
305                 310                 315                 320
Arg Gln Trp Asn Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly Thr
                325                 330                 335
Ala Thr Glu Thr Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile
            340                 345                 350
Cys Asp Lys Tyr His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn
        355                 360                 365
Ile Ser Phe Asp Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys
    370                 375                 380
Ile Thr Gln Asp Ile Phe Gln Gln Leu Leu Lys Arg Gly Phe Val Leu
385                 390                 395                 400
Gln Asp Thr Val Glu Gln Leu Arg Cys Glu His Cys Ala Arg Phe Leu
                405                 410                 415
Ala Asp Arg Phe Val Glu Gly Val Cys Pro Phe Cys Gly Tyr Glu Glu
            420                 425                 430
Ala Arg Gly Asp Gln Cys Asp Lys Cys Gly Lys Leu Ile Asn Ala Val
        435                 440                 445
Glu Leu Lys Lys Pro Gln Cys Lys Val Cys Arg Ser Cys Pro Val Val
    450                 455                 460
Gln Ser Ser Gln His Leu Phe Leu Asp Leu Pro Lys Leu Glu Lys Arg
465                 470                 475                 480
Leu Glu Glu Trp Leu Gly Arg Thr Leu Pro Gly Ser Asp Trp Thr Pro
                485                 490                 495
Asn Ala Gln Phe Ile Thr Arg Ser Trp Leu Arg Asp Gly Leu Lys Pro
            500                 505                 510
Arg Cys Ile Thr Arg Asp Leu Lys Trp Gly Thr Pro Val Pro Leu Glu
        515                 520                 525
Gly Phe Glu Asp Lys Val Phe Tyr Val Trp Phe Asp Ala Thr Ile Gly
    530                 535                 540
Tyr Leu Ser Ile Thr Ala Asn Tyr Thr Asp Gln Trp Glu Arg Trp Trp
545                 550                 555                 560
Lys Asn Pro Glu Gln Ser Thr
                565
```

<210> 77
<211> 567
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 77

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

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

<210> 78

<211> 217
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 78

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

<210> 79
<211> 217
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 79

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

145                 150                 155                 160
Tyr Leu Pro Glu Glu Leu Ser Ala Leu His Ser Trp Phe Gln Thr Leu
                165                 170                 175
Ser Thr Gln Glu Pro Cys Gln Arg Ala Ala Glu Thr Val Leu Lys Gln
            180                 185                 190
Gln Gly Val Leu Ala Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu
        195                 200                 205
Asp Ser Thr His His His His His His
    210                 215
```

<210> 80
<211> 234
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 80

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

<210> 81
<211> 234
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 81

```
Met Arg Leu Phe Val Ser Asp Gly Val Pro Gly Cys Leu Pro Val Leu
 1               5                  10                  15
```

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

<210> 82
<211> 241
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 82

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

        195                 200                 205
Val Leu Lys Gln Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys
    210                 215                 220
Gln Pro Gln Pro Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro
225                 230                 235                 240
Glu
```

<210> 83
<211> 241
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 83

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

<210> 84
<211> 134
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 84

```
Met His His His His His His Gly Lys Pro Ile Pro Asn Pro Leu Leu
 1               5                  10                  15

Gly Leu Asp Ser Thr Arg Leu Phe Val Ser Asp Gly Val Pro Gly Cys
            20                  25                  30
Leu Pro Val Leu Ala Ala Ala Gly Arg Ala Arg Gly Arg Ala Glu Val
        35                  40                  45
Leu Ile Ser Thr Val Gly Pro Glu Asp Cys Val Val Pro Phe Leu Thr
    50                  55                  60
Arg Pro Lys Val Pro Val Leu Gln Leu Asp Ser Gly Asn Tyr Leu Phe
65                  70                  75                  80
Ser Thr Ser Ala Ile Cys Arg Gly Ala Glu Cys Pro Ala Gln Leu Val
                85                  90                  95
Pro Asp Thr Glu Tyr Pro Gly Thr Met Ser Ala Ser Cys Arg Asp Cys
            100                 105                 110
Thr Glu Thr Ala Arg Cys Pro Gly Ser Pro Ala Leu Pro Pro Lys Ala
        115                 120                 125
Ala Pro Ala Gln Pro Arg
    130
```

<210> 85
<211> 134
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 85

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

<210> 86
<211> 296
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 86

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

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

<210> 87
<211> 296
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 87

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

Ser Thr Gln Glu Pro Cys Gln Arg Ala Ala Glu Thr Val Leu Lys Gln
            180                 185                 190
Gln Gly Val Leu Ala Leu Arg Pro Tyr Leu Gln Lys Gln Pro Gln Pro
        195                 200                 205
Ser Pro Ala Glu Gly Arg Ala Val Thr Asn Glu Pro Glu Val Ala Cys
    210                 215                 220
Gly Trp Arg Lys Glu Cys Ala His His Gln Cys Pro Pro Leu Arg Gln
225                 230                 235                 240
Gln Cys Pro Pro Pro Trp Glu His His Trp Leu Cys Ala Gln Cys Arg
                245                 250                 255
Cys Leu Cys Gln Val Leu Ser Pro Pro Pro Val Glu His Pro Leu Ser
            260                 265                 270
Val Trp Asp Arg Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp
        275                 280                 285
Ser Thr His His His His His His
    290                 295
```

<210> 88
<211> 293
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 88

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

                290
```

<210> 89
<211> 293
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 89

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

<210> 90
<211> 702
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 90

```
cgcttattcg tgtccgacgg cgtcccaggc tgcttaccag tgctggcggc ggctggccgc 60
gcacgtggtc gcgcggaggt cctgatcagc accgttggcc cggaggactg tgtggttccg 120
```

```
tttctgaccc gcccgaaggt cccggtgctg caattggact cgggcaacta cctgtttagc 180
acctccgcga tttgccgcta cttcttcctg ctgagcggtt gggaacaaga cgatctgacc 240
aatcaatggc tggagtggga ggccaccgag ctgcagccgg cactgtctgc agcgttgtac 300
tatctggtcg tgcaaggtaa gaaaggcgaa gatgtgctgg gtagcgtccg tcgtgccttg 360
acgcatatcg accacagcct gtcccgtcaa aattgtccgt tcctggctgg tgaaaccgag 420
agcttggcgg atattgttct gtggggcgcg ctgtatcctc tgctgcagga tccagcctat 480
ctgccggaag agttgtctgc cctgcacagc tggtttcaaa cgctgagcac ccaggaaccg 540
tgccagcgtg ccgcagaaac tgttctgaaa cagcaaggtg ttctggcact gcgtccgtac 600
ctgcagaaac agcctcagcc gagcccagct gagggtcgtg cggtgacgaa cgagccggaa 660
gaagaagagc tggcgacgct gagcgaagag taatgactcg ag                     702
```

<210> 91
<211> 798
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 91

```
cgcttattcg tgtccgacgg cgtcccaggc tgcttaccag tgttggcggc tgcaggtcgt 60
gcgcgcggtc gcgccgaggt cctgattagc accgtgggtc cggaagattg tgtggtccct 120
ttcctgaccc gtccgaaagt tccggtcctg cagctggatt ccggcaatta cttgtttagc 180
accagcgcaa tttgtcgcta cttctttctg ctgagcggtt gggagcaaga tgacctgacg 240
aaccagtggc tggagtggga ggccacggag ctgcagccgg cgctgagcgc agccctgtat 300
tacttggttg tgcagggtaa gaaaggcgag gacgtgctgg gcagcgttcg tcgtgcgctg 360
acccacatcg accattctct gagccgtcaa aattgcccgt ttctggcggg tgagactgag 420
agcctggcag acatcgtgct gtggggcgcg ttgtatccac tgctgcaaga tccggcttac 480
ctgccggaag agctgagcgc cctgcactcg tggttccaaa cgctgtctac tcaggaaccg 540
tgccaacgtg ctgccgaaac cgttctgaag caacagggtg ttctggcgct gcgcccgtat 600
ttgcagaaac agccacaacc gagcccggca gaaggccgtg cggtcaccaa cgagccggaa 660
gaagaagagc tggcgaccct gtctgaagaa gaaatcgcga tggcagttac ggcgtgggag 720
aagggtctgg agagcctgcc gccgttgcgt ccacagcaaa acccggtgct gccggtcgca 780
ggtgagtaat gactcgag                                              798
```

<210> 92
<211> 354
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 92

```
agattattcg ttagcgacgg cgtcccaggc tgcctcccag tactggcggc agcgggccgc 60
gcccgtggcc gcgccgaggt gttgatttct acggttggcc cagaagattg cgttgtcccg 120
tttctgacgc gtccgaaggt gcctgtcctg cagctggaca gcggtaacta cctgttcagc 180
acctccgcta tctgccgtta tagcatgagc ggtctgatgc cgctgctggc aatctgtccg 240
agccaaccga ccacccaaac gtcgggtcgt gacggtggtc gcacccagtc caagtggact 300
tgtattagca gctggccgaa aaccatgttc ctgagcatcg cgtaatgact cgag       354
```

<210> 93
<211> 1290
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 93

```
agattattcg tcagcgacgg cgtcccaggc tgcctcccag tgctggcggc agcaggccgc 60
gctcgcggcc gtgcggaggt gctgattagc accgttggcc cggaggattg cgtggtcccg 120

ttcctgaccc gtccgaaagt gccggtgctg cagctggaca gcggtaacta cctgtttagc 180
acgtccgcga tctgtcgcta tttcttcctg ctgagcggtt gggagcaaga cgacctgacc 240
aatcagtggc tggagtggga ggcgaccgag ctgcaaccgg cgctgagcgc agctctgtac 300
tatttggttg tgcaaggtaa gaaaggtgag gatgtgctgg gcagcgtccg tcgtgccttg 360
acccatatcg accatagcct gagccgccag aattgtccgt ttctggcggg tgaaaccgag 420
tccctggcgg atatcgtgct gtggggtgcg ctgtacccgc tgctgcagga cccggcgtac 480
ttgccggaag aactgtccgc tctgcactcc tggttccaaa ccttgagcac gcaagagccg 540
tgccaacgtg ccgcggagac tgtcctgaag caacaaggtg ttctggcact gcgtccgtac 600
ctgcagaaac agccgcagcc gagccccggcc gagggtcgtg ctgttaccaa tgagcctgaa 660
gaagaagagc tggcgacgct gagcgaagaa gaaattgcaa tggccgtgac tgcgtgggaa 720
aagggtctgg aatctctgcc gccactgcgc ccgcaacaaa accctgttct gccagtcgcc 780
ggtgagcgta acgttctgat taccagcgca ctgccgtatg tgaacaacgt tccgcacctg 840
ggcaacatta tcggttgtgt cctgagcgcc gatgttttcg cacgctatag ccgtctgcgt 900
cagtggaata cgctgtattt gtgcggtact gacgagtacg gtacggctac cgaaaccaaa 960
gcgctggaag agggtctgac gccgcaggag atctgcgata agtatcacat cattcatgcg 1020
gatatttacc gctggttcaa tatcagcttt gacattttcg gccgtacgac gacccctcag 1080
cagaccaaga gcctgagcgt caaatcggcc gaccacgcat tgtggtgtag ccgtgcatct 1140
acctgctttt ggacctgcct gtcgtggcgc agcgattggc gttctggttg gggcggtcat 1200
tgtctggcgg ttacgggcca cccaatgccg agcttgagcc cggtcttggg ctttggcatg 1260
gcgagctctc acgcggctta atgactcgag                                  1290
```

<210> 94
<211> 1650
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 94

```
agattattcg tcagcgacgg cgtcccaggc tgcctcccag tgctggccgc ggcaggccgt 60
gcgcgcggtc gcgcagaagt tttgatcagc accgtcggtc cggaggactg cgtggttccg 120
ttcctgaccc gcccgaaagt cccagtgttg cagctggata gcggtaacta cctgtttagc 180
acgtccgcga tctgtcgtta tttcttcttg ctgagcggtt gggagcagga cgacctgacc 240
aatcaatggt tggagtggga ggcgaccgag ctgcagccgg ccctgtctgc ggcattgtac 300
tacctggttg tccagggtaa gaaaggcgaa gatgtgctgg gcagcgtccg ccgtgcgctg 360
acccacatcg atcactctct gtcgcgccaa aattgcccgt tcctggcggg tgaaacggag 420
agcctggcag acattgtgct gtggggtgct ttgtatccgc tgctgcaaga cccggcctat 480
ctgccggaag aactgagcgc tctgcacagc tggtttcaaa cgctgagcac gcaggaaccg 540
tgtcagcgtg cggcggagac tgttctgaag caacagggtg ttctggcgct gcgtccgtac 600
ctgcaaaagc aaccgcagcc gagccctgcg gagggtcgtg cagttacgaa cgaaccggaa 660
gaagaagagc tggcgaccct gagcgaagaa gagattgcaa tggcagtcac ggcgtgggaa 720
aagggtctgg agtcgttgcc gccgctgcgc ccacagcaaa atccggttct gccggtcgct 780
ggcgagcgta acgtgctgat taccagcgcc ctgccgtacg tgaacaacgt gccgcatctg 840
ggtaacatca ttggctgcgt cctgagcgcg gatgtctttg cgcgttatag ccgcctgcgt 900
caatggaata ccctgtacct gtgcggtacc gacgagtatg gtaccgctac cgaaactaag 960
gcattggaag agggcctgac cccgcaggag atttgcgata aataccatat cattcatgcc 1020
gacatctacc gttggtttaa catcagcttc gatatctttg gccgcacgac cacgccgcaa 1080
caaaccaaaa ttacccaaga tattttccaa cagctgctga agcgcggctt tgttctgcag 1140
gacaccgttg agcagctgcg ttgcgagcac tgcgcacgtt ttctggccga ccgtttcgtg 1200
gagggtgtct gtccgttctg cggttatgaa gaagcccgtg gtgaccagtg tgacaaatgt 1260
ggcaaactga tcaatgctgt cgaactgaag aaacctcagt gcaaagtgtg tcgtagctgc 1320
cctgttgtgc aaagcagcca acacctgttc ctggatctgc caaagctgga gaagcgtctg 1380
gaagagtggc tgggccgtac cctgccgggc agcgattgga ccccgaatgc tcagttcatt 1440
acgcgctcct ggctgcgcga tggtctgaaa ccgcgttgta tcactcgtga cctgaaatgg 1500
ggtacgccgg ttccactgga gggctttgag gataaagtgt tctatgtgtg gtttgatgcg 1560
actattggtt acttgtccat caccgcgaat tacacggacc agtgggagcg ttggtggaag 1620
aacccggagc agtccaccta atgactcgag                                  1650
```

<210> 95
<211> 600
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 95

```
agattattcg tcagcgacgg cgtcccaggc tgcctcccag tgctggcggc cgcgggccgt 60
gcacgtggcc gtgcagaggt gttgattagc acggttggtc cggaagattg cgtcgttccg 120
tttctgacgc gtccgaaagt cccggtgctg caactggata gcggtaacta cctgttcagc 180
accagcgcaa tctgtcgcta tttctttctg ctgagcggct gggagcaaga cgacctgacc 240
aatcaatggt tggagtggga ggcgaccgaa ctgcaaccgg ccctgagcgc tgcgctgtat 300
tacttggtgg tgcagggtaa gaaaggtgag gatgtcctgg gttccgttcg tcgcgcgttg 360
acccacatcg atcacagcct gtcgcgtcaa aactgtccgt tcctggcagg tgaaaccgag 420
tctctggccg acattgtcct gtggggcgcg ctgtatccgc tgctgcagga cccggcatac 480
ctgcctgaag aactgagcgc gctgcattct tggtttcaga cgctgagcac ccaggagcca 540
tgccagcgcg cggctgaaac tgttctgaag caacagggtg tgttggccta atgactcgag 600
```

<210> 96
<211> 651
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 96

```
agattattcg ttagcgacgg cgtcccaggc tgcctcccag tactggcagc ggccggtcgt 60
gcgcgcggcc gtgcggaggt gttgatcagc accgttggcc cggaggattg cgtggtccct 120
ttcctgactc gtccgaaagt cccggtcctg cagttggact cgggcaatta cctgtttagc 180
accagcgcca tttgccgtta cttctttctg ctgagcggtt gggagcaaga cgatctgacg 240
aaccaatggc tggagtggga agcaaccgag ctgcaaccgg ctctgtccgc tgcgctgtac 300
tatctggtcg ttcagggtaa gaagggtgag gacgttctgg gctccgttcg tcgtgcgctg 360
acgcatatcg atcacagcct gagccgtcaa aactgtccgt tcctggcggg tgaaacggaa 420
agcctggccg acattgttct gtggggtgcg ttgtacccgt tgctgcagga tccggcatat 480
ttgccggaag aactgagcgc actgcactcg tggtttcaga ccctgagcac gcaagagccg 540
tgtcaacgcg ctgcggaaac cgtgctgaaa cagcagggtg tgctggcact gcgcccttat 600
ctgcagaaac agccgcaacc atctccggcg gagggccgct aatgactcga g          651
```

<210> 97
<211> 672
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 97

```
agattattcg ttagcgacgg cgtcccaggc tgccttccag tattggcggc ggcgggtcgc 60
gcgcgtggcc gtgctgaagt gctgattagc accgtgggtc ctgaggactg tgttgtgccg 120
ttcttgaccc gtccgaaagt tccggtgttg cagctggata gcggcaatta cctgtttagc 180
acctccgcaa tctgccgtta cttcttcctg ctgtctggtt gggagcaaga tgacttgacc 240
aatcagtggc tggaatggga ggcaacggag ctgcaacctg ccctgagcgc agcgctgtat 300
tatctggtcg tgcagggtaa gaaaggcgag gatgttctgg gtagcgtccg tcgcgccctg 360
acgcacatcg accattcgct gtcccgccaa aactgtccgt ttctggcggg cgagactgag 420
agcctggcag acattgtcct gtggggtgcc ctgtacccgt tgctgcagga tccggcgtat 480
ctgccggaag aactgagcgc gctgcactct tggtttcaga ccctgagcac ccaagaacca 540
tgccagcgtg ctgctgaaac ggttctgaag caacaaggcg tcctggcgct gcgtccgtac 600
ctgcagaaac agccgcaacc gagcccggcc gagggtcgcg cagttacgaa cgaaccggag 660
taatgactcg ag                                                     672
```

<210> 98
<211> 351
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 98

```
agattattcg tcagcgacgg cgtcccaggc tgcctcccag tactggcggc tgccggtcgc 60
gcacgtggtc gcgcagaagt tctgatttcc accgtcggcc cagaggattg tgtcgtgccg 120
ttcctgaccc gtcctaaagt gccggtgctg cagctggaca gcggcaacta cttgtttagc 180
acctctgcga tctgccgtgg cgcggagtgc ccggcacaac tggttccgga caccgagtat 240
ccgggtacta tgtcggcgag ctgccgtgat tgtacggaaa cggcccgttg tccgggtagc 300
ccggctctgc cgcctaaggc ggcgccggca cagccgcgct aatgactcga g          351
```

<210> 99
<211> 837
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 99

```
agattattcg tcagcgacgg cgtcccaggc tgcctcccag tactggccgc ggcgggtcgt 60
gcacgtggcc gcgctgaggt tttgatctcc accgtgggtc ctgaggattg tgtcgtgcct 120
tttctgacgc gcccgaaggt tccggttctg cagttggata gcggcaatta tttgttcagc 180
accagcgcaa tttgccgtta tttcttcttg ctgtctggtt gggagcagga cgatctgacc 240
aatcagtggc tggagtggga agccaccgag ttgcaaccgg ccctgagcgc tgcgctgtat 300
tacctggtgg tccaaggtaa gaaaggtgag gacgttctgg gtagcgtccg tcgtgcactg 360
acgcatatcg accacagcct gtctcgtcaa aactgtccgt ttctggcggg cgaaacggag 420
agcctggcag acattgttct gtggggcgcg ctgtacccgc tgctgcaaga cccggcgtac 480
ctgccggaag aactgtccgc cctgcatagc tggtttcaga ccctgagcac gcaggaaccg 540
tgccagcgcg cggcagagac tgtgctgaag caacaaggtg tcctggcgct gcgcccttac 600
ctgcagaaac aaccgcagcc gagcccggcc gagggtcgtg ctgtgaccaa cgaaccagaa 660
gttgcgtgcg gctggcgtaa agaatgcgcg caccatcaat gcccgccgct gcgtcagcaa 720
tgccctccgc cgtgggagca ccactggctg tgtgctcagt gccgctgtct gtgtcaggtc 780
ctgtcgccgc cgccggtcga gcacccgttg agcgtgtggg atcgttaatg actcgag    837
```

<210> 100
<211> 828
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 100

```
agattattcg tcagcgacgg cgtcccaggc tgcctcccag tactggcagc cgccggtcgc 60
gcgcgtggtc gtgcggaggt tctgattagc accgtgggcc cggaggactg tgtggttccg 120
ttcctgacgc gtccgaaagt tcctgtgctg cagctggata gcggcaacta tttgtttagc 180
accagcgcga tttgccgtta cttctttctg ctgtccggtt gggagcaaga cgatttgacg 240
aatcaatggc tggagtggga agcaactgag ctgcaaccag cgctgtcggc ggcgctgtac 300
tatttggtcg ttcagggtaa gaaaggtgag gatgtcctgg gtagcgtccg ccgtgcactg 360
actcacatcg accacagctt gtctcgtcag aactgcccgt tcctggcagg tgaaacggag 420
tccctggccg acatcgtgct gtggggcgcg ctgtacccgt tgctgcagga tccggcgtac 480
ctgccggaag agctgagcgc cctgcatagc tggtttcaga ccctgagcac ccaagagccg 540
tgccaacgtg cagcagaaac ggtcctgaag cagcaaggcg ttctggcgct gcgcccgtat 600
ctgcaaaagc agccgcagcc gtccccagcc gaaggccgcg ctgtcaccaa tgagccggaa 660
gaagaagagc tggctacgct gagcgaagaa gaaattgcta tggccgtgac cgcgtgggag 720
aaaggtctgg agagcttgcc gccgctgcgt cctcaacaga acccggtttg gacctgtatc 780
tctagctggc cgaaaaccat gttcctgagc atcgcgtaat gactcgag              828
```

<210> 101
<211> 707
<212> PRT
<213> *Homo sapiens*

<400> 101

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

```
Leu Thr Pro Asp Asp Gln Arg Leu Leu Ala His Val Thr Leu Glu Leu
                485                 490                 495
Gln His Tyr His Gln Leu Leu Glu Lys Val Arg Ile Arg Asp Ala Leu
            500                 505                 510
Arg Ser Ile Leu Thr Ile Ser Arg His Gly Asn Gln Tyr Ile Gln Val
        515                 520                 525
Asn Glu Pro Trp Lys Arg Ile Lys Gly Ser Glu Ala Asp Arg Gln Arg
    530                 535                 540
Ala Gly Thr Val Thr Gly Leu Ala Val Asn Ile Ala Ala Leu Leu Ser
545                 550                 555                 560
Val Met Leu Gln Pro Tyr Met Pro Thr Val Ser Ala Thr Ile Gln Ala
                565                 570                 575
Gln Leu Gln Leu Pro Pro Pro Ala Cys Ser Ile Leu Leu Thr Asn Phe
            580                 585                 590
Leu Cys Thr Leu Pro Ala Gly His Gln Ile Gly Thr Val Ser Pro Leu
        595                 600                 605
Phe Gln Lys Leu Glu Asn Asp Gln Ile Glu Ser Leu Arg Gln Arg Phe
    610                 615                 620
Gly Gly Gly Gln Ala Lys Thr Ser Pro Lys Pro Ala Val Val Glu Thr
625                 630                 635                 640
Val Thr Thr Ala Lys Pro Gln Gln Ile Gln Ala Leu Met Asp Glu Val
                645                 650                 655
Thr Lys Gln Gly Asn Ile Val Arg Glu Leu Lys Ala Gln Lys Ala Asp
            660                 665                 670
Lys Asn Glu Val Ala Ala Glu Val Ala Lys Leu Leu Asp Leu Lys Lys
        675                 680                 685
Gln Leu Ala Val Ala Glu Gly Lys Pro Pro Glu Ala Pro Lys Gly Lys
    690                 695                 700
Lys Lys Lys
705
```

<210> 102
<211> 2124
<212> ADN
<213> *Homo sapiens*

<400> 102

```
ggtgtcctgg ctctccggcc ttacctccaa aagcagcccc agcccagccc cgctgaggga 60
agggctgtca ccaatgagcc tgaggaggag gagctggcta ccctatctga ggaggagatt 120
gctatggctg ttactgcttg ggagaagggc ctagaaagtt tgcccccgct gcggccccag 180
cagaatccag tgttgcctgt ggctggagaa aggaatgtgc tcatcaccag tgccctccct 240
tacgtcaaca atgtccccca ccttgggaac atcattggtt gtgtgctcag tgccgatgtc 300
tttgccaggt actctcgcct ccgccagtgg aacaccctct atctgtgtgg gacagatgag 360
tatggtacag caacagagac caaggctctg gaggagggac taacccccca ggagatctgc 420
gacaagtacc acatcatcca tgctgacatc taccgctggt ttaacatttc gtttgatatt 480
tttggtcgca ccaccactcc acagcagacc aaaatcaccc aggacatttt ccagcagttg 540
ctgaaacgag gttttgtgct gcaagatact gtggagcaac tgcgatgtga gcactgtgct 600
cgcttcctgg ctgaccgctt cgtggagggc gtgtgtccct tctgtggcta tgaggaggct 660
cggggtgacc agtgtgacaa gtgtggcaag ctcatcaatg ctgtcgagct taagaagcct 720
cagtgtaaag tctgccgatc atgccctgtg gtgcagtcga gccagcacct gtttctggac 780
ctgcctaagc tggagaagcg actggaggag tggttgggga ggacattgcc tggcagtgac 840
tggacaccca atgcccagtt tatcacccgt tcttggcttc gggatggcct caagccacgc 900
tgcataaccc gagacctcaa atggggaacc cctgtaccct tagaaggttt tgaagacaag 960
gtattctatg tctggtttga tgccactatt ggctatctgt ccatcacagc caactacaca 1020
gaccagtggg agagatggtg gaagaaccca gagcaagtgg acctgtatca gttcatggcc 1080
aaagacaatg ttcctttcca tagcttagtc tttccttgct cagccctagg agctgaggat 1140
aactatacct tggtcagcca cctcattgct acagagtacc tgaactatga ggatgggaaa 1200
ttctctaaga gccgcggtgt gggagtgttt ggggacatgg cccaggacac ggggatccct 1260
gctgacatct ggcgcttcta tctgctgtac attcggcctg agggccagga cagtgctttc 1320
tcctggacgg acctgctgct gaagaataat tctgagctgc ttaacaacct gggcaacttc 1380
atcaacagag ctgggatgtt tgtgtctaag ttctttgggg gctatgtgcc tgagatggtg 1440
ctcacccctg atgatcagcg cctgctggcc catgtcaccc tggagctcca gcactatcac 1500
cagctacttg agaaggttcg gatccgggat gccttgcgca gtatcctcac catatctcga 1560
catggcaacc aatatattca ggtgaatgag ccctggaagc ggattaaagg cagtgaggct 1620

gacaggcaac gggcaggaac agtgactggc ttggcagtga atatagctgc cttgctctct 1680
gtcatgcttc agccttacat gcccacggtt agtgccacaa tccaggccca gctgcagctc 1740
ccacctccag cctgcagtat cctgctgaca aacttcctgt gtaccttacc agcaggacac 1800
cagattggca cagtcagtcc cttgttccaa aaattggaaa atgaccagat tgaaagttta 1860
aggcagcgct ttggaggggg ccaggcaaaa acgtccccga agccagcagt tgtagagact 1920
gttacaacag ccaagccaca gcagatacaa gcgctgatgg atgaagtgac aaaacaagga 1980
aacattgtcc gagaactgaa agcacaaaag gcagacaaga acgaggttgc tgcggaggtg 2040
gcgaaactct tggatctaaa gaaacagttg gctgtagctg aggggaaacc ccctgaagcc 2100
cctaaaggca agaagaaaaa gtaa                                        2124
```

<210> 103
<211> 28
<212> PRT
<213> *Mus musculus*

<400> 103

```
Thr Val Ser Asn Glu Leu Glu Glu Glu Glu Leu Ala Thr Leu Ser Glu
 1               5                  10                  15
Glu Asp Ile Val Thr Ala Val Ala Ala Trp Glu Lys
            20                  25
```

<210> 104
<211> 22
<212> PRT
<213> *Mus musculus*

<400> 104

Gly Leu Glu Ser Leu Pro Pro Leu Lys Leu Gln Gln His Pro Val Leu
1 5 10 15
Pro Val Pro Gly Glu Arg
20

<210> 105
<211> 10
<212> PRT
<213> *Mus musculus*

<400> 105

Asn Val Leu Ile Thr Ser Ala Leu Pro Tyr
1 5 10

<210> 106
<211> 22
<212> PRT
<213> *Mus musculus*

<400> 106

Val Asn Asn Val Pro His Leu Gly Asn Ile Ile Gly Cys Val Leu Ser
1 5 10 15
Ala Asp Val Phe Ala Arg
20

<210> 107
<211> 3
<212> PRT
<213> *Mus musculus*

<400> 107

Tyr Cys Arg
1

<210> 108
<211> 22
<212> PRT
<213> *Mus musculus*

<400> 108

Leu Arg Gln Trp Asn Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly
1 5 10 15
Thr Ala Thr Glu Thr Lys
20

<210> 109
<211> 9
<212> PRT
<213> *Mus musculus*

<400> 109

Ala Met Glu Glu Gly Leu Thr Pro Arg
1 5

<210> 110
<211> 15
<212> PRT
<213> *Mus musculus*

<400> 110

Glu Ile Cys Asp Lys Tyr His Ala Ile His Ala Asp Ile Tyr Arg
1 5 10 15

<210> 111
<211> 11
<212> PRT
<213> *Mus musculus*

<400> 111

Trp Phe Gly Ile Ser Phe Asp Thr Phe Gly Arg
1 5 10

<210> 112
<211> 8
<212> PRT
<213> *Mus musculus*

<400> 112

Thr Thr Thr Pro Gln Gln Thr Lys
1 5

<210> 113
<211> 8
<212> PRT
<213> *Mus musculus*

<400> 113

Ile Thr Gln Asp Ile Phe Gln Arg
1 5

<210> 114
<211> 27
<212> PRT
<213> *Mus musculus*

<400> 114

Leu Leu Thr Arg Gly Phe Val Leu Arg Asp Thr Val Glu Gln Leu Arg
1 5 10 15
Cys Glu Arg Cys Ala Arg Phe Leu Ala Asp Arg
20 25

<210> 115
<211> 15
<212> PRT

<213> *Mus musculus*

<400> 115

```
Phe Val Glu Gly Val Cys Pro Phe Cys Gly Tyr Glu Glu Ala Arg
 1               5                  10                  15
```

<210> 116
<211> 31
<212> PRT
<213> *Mus musculus*

<400> 116

```
Gly Asp Gln Cys Asp Arg Cys Gly Lys Leu Ile Asn Ala Ile Glu Leu
 1               5                  10                  15
Lys Lys Pro Gln Cys Lys Ile Cys Arg Ser Cys Pro Val Val Arg
            20                  25                  30
```

<210> 117
<211> 11
<212> PRT
<213> *Mus musculus*

<400> 117

```
Ser Ser Gln His Leu Phe Leu Asp Leu Pro Lys
 1               5                  10
```

<210> 118
<211> 98
<212> PRT
<213> *Mus musculus*

<400> 118

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

<210> 119
<211> 25
<212> PRT
<213> *Mus musculus*

<400> 119

```
Asp Asn Val Pro Phe His Gly Leu Val Phe Pro Cys Ser Val Leu Gly
 1               5                  10                  15
Ala Glu Asp Asn Tyr Thr Leu Val Lys
            20                  25
```

<210> 120
<211> 14
<212> PRT
<213> *Mus musculus*

<400> 120

```
His Ile Ile Ala Thr Glu Tyr Leu Asn Tyr Glu Asp Gly Lys
 1               5                  10
```

<210> 121
<211> 30
<212> PRT
<213> *Mus musculus*

<400> 121

```
Phe Ser Lys Ser Arg Gly Ile Gly Val Phe Gly Asp Met Ala Lys Asp
 1               5                  10                  15
Thr Gly Ile Pro Ala Asp Ile Trp Arg Phe Tyr Leu Leu Tyr
            20                  25                  30
```

<210> 122
<211> 33
<212> PRT
<213> *Mus musculus*

<400> 122

```
Ile Arg Pro Glu Gly Gln Asp Ser Ala Phe Ser Trp Thr Asp Leu Leu
 1               5                  10                  15
Ile Lys Asn Asn Ser Glu Leu Leu Asn Asn Leu Gly Asn Phe Ile Asn
            20                  25                  30
Arg
```

<210> 123
<211> 57
<212> PRT
<213> *Mus musculus*

<400> 123

```
Ala Gly Met Phe Val Ser Lys Phe Phe Gly Gly Cys Val Pro Glu Met
 1               5                  10                  15
Ala Leu Thr Pro Asp Asp Arg Arg Leu Val Ala His Val Ser Trp Glu
            20                  25                  30
Leu Gln His Tyr His Gln Leu Leu Glu Lys Val Arg Ile Arg Asp Ala
        35                  40                  45
Leu Arg Ser Ile Leu Thr Ile Ser Arg
    50                  55
```

<210> 124
<211> 14

<212> PRT
<213> *Mus musculus*

<400> 124

```
His Gly Asn Gln Tyr Ile Gln Val Asn Glu Pro Trp Lys Arg
 1               5                  10
```

<210> 125
<211> 67
<212> PRT
<213> *Mus musculus*

<400> 125

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

<210> 126
<211> 10
<212> PRT
<213> *Mus musculus*

<400> 126

```
Ile Gly Thr Val Ser Pro Leu Phe Gln Lys
 1               5                  10
```

<210> 127
<211> 10
<212> PRT
<213> *Mus musculus*

<400> 127

```
Leu Glu Asn Asp Gln Ile Glu Asn Leu Arg
 1               5                  10
```

<210> 128
<211> 9
<212> PRT
<213> *Mus musculus*

<400> 128

```
Gln Arg Phe Gly Gly Gly Gln Ala Lys
 1               5
```

<210> 129
<211> 28

<212> PRT
<213> *Mus musculus*

<400> 129

```
Gly Ser Pro Lys Pro Ala Ala Val Glu Ala Val Thr Ala Ala Gly Ser
 1               5                  10                  15
Gln His Ile Gln Thr Leu Thr Asp Glu Val Thr Lys
            20                  25
```

<210> 130
<211> 30
<212> PRT
<213> *Mus musculus*

<400> 130

```
Gln Gly Asn Val Val Arg Glu Leu Lys Ala Gln Lys Ala Asp Lys Asn
 1               5                  10                  15
Gln Val Ala Ala Glu Val Ala Lys Leu Leu Asp Leu Lys Lys
            20                  25                  30
```

<210> 131
<211> 14
<212> PRT
<213> *Mus musculus*

<400> 131

```
Gln Leu Ala Leu Ala Glu Gly Lys Pro Ile Glu Thr Pro Lys
 1               5                  10
```

<210> 132
<211> 681
<212> PRT
<213> *Mus musculus*

<400> 132

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

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

<210> 133
<211> 349
<212> PRT
<213> *Homo sapiens*

<400> 133

```
Met Ala Lys Asp Asn Val Pro Phe His Ser Leu Val Phe Pro Cys Ser
 1               5                  10                  15
Ala Leu Gly Ala Glu Asp Asn Tyr Thr Leu Val Ser His Leu Ile Ala
            20                  25                  30
Thr Glu Tyr Leu Asn Tyr Glu Asp Gly Lys Phe Ser Lys Ser Arg Gly

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

<210> 134
<211> 1050
<212> ADN
<213> *Homo sapiens*

<400> 134

atggccaaag acaatgttcc tttccatagc ttagtctttc cttgctcagc cctaggagct 60
gaggataact ataccttggt cagccacctc attgctacag agtacctgaa ctatgaggat 120
gggaaattct ctaagagccg cggtgtggga gtgtttgggg acatggccca ggacacgggg 180
atccctgctg acatctggcg cttctatctg ctgtacattc ggcctgaggg ccaggacagt 240
gctttctcct ggacggacct gctgctgaag aataattctg agctgcttaa caacctgggc 300
aacttcatca acagagctgg gatgtttgtg tctaagttct ttgggggcta tgtgcctgag 360
atggtgctca cccctgatga tcagcgcctg ctggcccatg tcaccctgga gctccagcac 420
tatcaccagc tacttgagaa ggttcggatc cgggatgcct tgcgcagtat cctcaccata 480
tctcgacatg gcaaccaata tattcaggtg aatgagccct ggaagcggat taaaggcagt 540
gaggctgaca ggcaacgggc aggaacagtg actggcttgg cagtgaatat agctgccttg 600
ctctctgtca tgcttcagcc ttacatgccc acggttagtg ccacaatcca ggcccagctg 660
cagctcccac ctccagcctg cagtatcctg ctgacaaact tcctgtgtac cttaccagca 720
ggacaccaga ttggcacagt cagtcccttg ttccaaaaat tggaaaatga ccagattgaa 780
agtttaaggc agcgctttgg agggggccag gcaaaaacgt ccccgaagcc agcagttgta 840
gagactgtta caacagccaa gccacagcag atacaagcgc tgatggatga agtgacaaaa 900
caaggaaaca ttgtccgaga actgaaagca caaaaggcag acaagaacga ggttgctgcg 960
gaggtggcga aactcttgga tctaaagaaa cagttggctg tagctgaggg gaaacccct 1020
gaagccccta aaggcaagaa gaaaaagtaa 1050

<210> 135
<211> 323
<212> PRT
<213> *Homo sapiens*

<400> 135

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

<210> 136
<211> 972
<212> ADN
<213> *Homo sapiens*

<400> 136

```
atggtggaag aacccagagc aaagtacctg aactatgagg atgggaaatt ctctaagagc 60
cgcggtgtgg gagtgtttgg ggacatggcc caggacacgg ggatccctgc tgacatctgg 120
cgcttctatc tgctgtacat tcggcctgag ggccaggaca gtgctttctc ctggacggac 180
ctgctgctga agaataattc tgagctgctt aacaacctgg gcaacttcat caacagagct 240
gggatgtttg tgtctaagtt ctttgggggc tatgtgcctg agatggtgct caccctgat 300
gatcagcgcc tgctggccca tgtcaccctg gagctccagc actatcacca gctacttgag 360
aaggttcgga tccgggatgc cttgcgcagt atcctcacca tatctcgaca tggcaaccaa 420
tatattcagg tgaatgagcc ctggaagcgg attaaaggca gtgaggctga caggcaacgg 480
```

```
gcaggaacag tgactggctt ggcagtgaat atagctgcct tgctctctgt catgcttcag 540
ccttacatgc ccacggttag tgccacaatc caggcccagc tgcagctccc acctccagcc 600
tgcagtatcc tgctgacaaa cttcctgtgt accttaccag caggacacca gattggcaca 660
gtcagtccct tgttccaaaa attggaaaat gaccagattg aaagtttaag gcagcgcttt 720
ggaggggccc aggcaaaaac gtccccgaag ccagcagttg tagagactgt tacaacagcc 780
aagccacagc agatacaagc gctgatggat gaagtgacaa aacaaggaaa cattgtccga 840
gaactgaaag cacaaaaggc agacaagaac gaggttgctg cggaggtggc gaaactcttg 900
gatctaaaga aacagttggc tgtagctgag gggaaacccc ctgaagcccc taaaggcaag 960
aagaaaaagt aa                                                    972
```

<210> 137
<211> 55
<212> PRT
<213> *Homo sapiens*

<400> 137

```
Met Asp Glu Val Thr Lys Gln Gly Asn Ile Val Arg Glu Leu Lys Ala
 1               5                  10                  15
Gln Lys Ala Asp Lys Asn Glu Val Ala Ala Glu Val Ala Lys Leu Leu
            20                  25                  30
Asp Leu Lys Lys Gln Leu Ala Val Ala Glu Gly Lys Pro Pro Glu Ala
        35                  40                  45
Pro Lys Gly Lys Lys Lys Lys
    50                  55
```

<210> 138
<211> 168
<212> ADN
<213> *Homo sapiens*

<400> 138

```
atggatgaag tgacaaaaca aggaaacatt gtccgagaac tgaaagcaca aaaggcagac 60
aagaacgagg ttgctgcgga ggtggcgaaa ctcttggatc taaagaaaca gttggctgta 120
gctgagggga aacccccctga agcccctaaa ggcaagaaga aaaagtaa              168
```

<210> 139
<211> 773
<212> PRT
<213> *Homo sapiens*

<400> 139

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

```
            180                 185                 190
Thr Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile Cys Asp Lys
        195                 200                 205
Tyr His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn Ile Ser Phe
    210                 215                 220
Asp Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys Ile Thr Gln
225                 230                 235                 240
Asp Ile Phe Gln Gln Leu Leu Lys Arg Gly Phe Val Leu Gln Asp Thr
                245                 250                 255
Val Glu Gln Leu Arg Cys Glu His Cys Ala Arg Phe Leu Ala Asp Arg
            260                 265                 270
Phe Val Glu Gly Val Cys Pro Phe Cys Gly Tyr Glu Glu Ala Arg Gly
        275                 280                 285
Asp Gln Cys Asp Lys Cys Gly Lys Leu Ile Asn Ala Val Glu Leu Lys
    290                 295                 300
Lys Pro Gln Cys Lys Val Cys Arg Ser Cys Pro Val Val Gln Ser Ser
305                 310                 315                 320
Gln His Leu Phe Leu Asp Leu Pro Lys Leu Glu Lys Arg Leu Glu Glu
                325                 330                 335
Trp Leu Gly Arg Thr Leu Pro Gly Ser Asp Trp Thr Pro Asn Ala Gln
            340                 345                 350
Phe Ile Thr Arg Ser Trp Leu Arg Asp Gly Leu Lys Pro Arg Cys Ile
        355                 360                 365
Thr Arg Asp Leu Lys Trp Gly Thr Pro Val Pro Leu Glu Gly Phe Glu
    370                 375                 380
Asp Lys Val Phe Tyr Val Trp Phe Asp Ala Thr Ile Gly Tyr Leu Ser
385                 390                 395                 400
Ile Thr Ala Asn Tyr Thr Asp Gln Trp Glu Arg Trp Trp Lys Asn Pro
                405                 410                 415
Glu Gln Val Asp Leu Tyr Gln Phe Met Ala Lys Asp Asn Val Pro Phe
            420                 425                 430
His Ser Leu Val Phe Pro Cys Ser Ala Leu Gly Ala Glu Asp Asn Tyr
        435                 440                 445
Thr Leu Val Ser His Leu Ile Ala Thr Glu Tyr Leu Asn Tyr Glu Asp
    450                 455                 460
Gly Lys Phe Ser Lys Ser Arg Gly Val Gly Val Phe Gly Asp Met Ala
465                 470                 475                 480
Gln Asp Thr Gly Ile Pro Ala Asp Ile Trp Arg Phe Tyr Leu Leu Tyr
                485                 490                 495
Ile Arg Pro Glu Gly Gln Asp Ser Ala Phe Ser Trp Thr Asp Leu Leu
            500                 505                 510
Leu Lys Asn Asn Ser Glu Leu Leu Asn Asn Leu Gly Asn Phe Ile Asn
        515                 520                 525
Arg Ala Gly Met Phe Val Ser Lys Phe Phe Gly Gly Tyr Val Pro Glu
    530                 535                 540
Met Val Leu Thr Pro Asp Asp Gln Arg Leu Leu Ala His Val Thr Leu
545                 550                 555                 560
Glu Leu Gln His Tyr His Gln Leu Leu Glu Lys Val Arg Ile Arg Asp
                565                 570                 575
Ala Leu Arg Ser Ile Leu Thr Ile Ser Arg His Gly Asn Gln Tyr Ile
            580                 585                 590
Gln Val Asn Glu Pro Trp Lys Arg Ile Lys Gly Ser Glu Ala Asp Arg
        595                 600                 605
Gln Arg Ala Gly Thr Val Thr Gly Leu Ala Val Asn Ile Ala Ala Leu
    610                 615                 620
Leu Ser Val Met Leu Gln Pro Tyr Met Pro Thr Val Ser Ala Thr Ile
625                 630                 635                 640
Gln Ala Gln Leu Gln Leu Pro Pro Pro Ala Cys Ser Ile Leu Leu Thr
                645                 650                 655
Asn Phe Leu Cys Thr Leu Pro Ala Gly His Gln Ile Gly Thr Val Ser
            660                 665                 670
Pro Leu Phe Gln Lys Leu Glu Asn Asp Gln Ile Glu Ser Leu Arg Gln
        675                 680                 685
Arg Phe Gly Gly Gly Gln Ala Lys Thr Ser Pro Lys Pro Ala Val Val
    690                 695                 700
```

```
Glu Thr Val Thr Thr Ala Lys Pro Gln Gln Ile Gln Ala Leu Met Asp
705                 710                 715                 720
Glu Val Thr Lys Gln Gly Asn Ile Val Arg Glu Leu Lys Ala Gln Lys
                725                 730                 735
Ala Asp Lys Asn Glu Val Ala Ala Glu Val Ala Lys Leu Leu Asp Leu
            740                 745                 750
Lys Lys Gln Leu Ala Val Ala Glu Gly Lys Pro Pro Glu Ala Pro Lys
        755                 760                 765
Gly Lys Lys Lys Lys
    770
```

<210> 140
<211> 2322
<212> ADN
<213> *Homo sapiens*

<400> 140

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ggagctgagt 120
gccctgcaca gctggttcca gacactgagt acccaggaac catgtcagcg agctgcagag 180
actgtactga aacagcaagg tgtcctggct ctccggcctt acctccaaaa gcagccccag 240
cccagccccg ctgagggaag ggctgtcacc aatgagcctg aggaggagga gctggctacc 300
ctatctgagg aggagattgc tatggctgtt actgcttggg agaagggcct agaaagtttg 360
cccccgctgc ggccccagca gaatccagtg ttgcctgtgg ctggagaaag gaatgtgctc 420
atcaccagtg ccctccctta cgtcaacaat gtcccccacc ttgggaacat cattggttgt 480
gtgctcagtg ccgatgtctt tgccaggtac tctcgcctcc gccagtggaa caccctctat 540
ctgtgtggga cagatgagta tggtacagca acagagacca aggctctgga ggagggacta 600
acccccccagg agatctgcga caagtaccac atcatccatg ctgacatcta ccgctggttt 660
aacatttcgt ttgatatttt tggtcgcacc accactccac agcagaccaa aatcacccag 720
gacattttcc agcagttgct gaaacgaggt tttgtgctgc aagatactgt ggagcaactg 780
cgatgtgagc actgtgctcg cttcctggct gaccgcttcg tggagggcgt gtgtcccttc 840
tgtggctatg aggaggctcg gggtgaccag tgtgacaagt gtggcaagct catcaatgct 900
gtcgagctta agaagcctca gtgtaaagtc tgccgatcat gccctgtggt gcagtcgagc 960
cagcacctgt ttctggacct gcctaagctg gagaagcgac tggaggagtg gttggggagg 1020
acattgcctg gcagtgactg gacacccaat gcccagttta tcacccgttc ttggcttcgg 1080
gatggcctca agccacgctg cataacccga gacctcaaat ggggaacccc tgtaccctta 1140
gaaggttttg aagacaaggt attctatgtc tggtttgatg ccactattgg ctatctgtcc 1200
atcacagcca actacacaga ccagtgggag agatggtgga agaacccaga gcaagtggac 1260
ctgtatcagt tcatggccaa agacaatgtt cctttccata gcttagtctt tccttgctca 1320
gccctaggag ctgaggataa ctataccttg gtcagccacc tcattgctac agagtacctg 1380
aactatgagg atgggaaatt ctctaagagc cgcggtgtgg gagtgtttgg ggacatggcc 1440
caggacacgg ggatccctgc tgacatctgg cgcttctatc tgctgtacat tcggcctgag 1500
ggccaggaca gtgctttctc ctggacggac ctgctgctga agaataattc tgagctgctt 1560
aacaacctgg gcaacttcat caacagagct gggatgtttg tgtctaagtt ctttgggggc 1620
tatgtgcctg agatggtgct caccctgat gatcagcgcc tgctggccca tgtcaccctg 1680
gagctccagc actatcacca gctacttgag aaggttcgga tccgggatgc cttgcgcagt 1740
atcctcacca tatctcgaca tggcaaccaa tatattcagg tgaatgagcc ctggaagcgg 1800
attaaaggca gtgaggctga caggcaacgg gcaggaacag tgactggctt ggcagtgaat 1860
atagctgcct tgctctctgt catgcttcag ccttacatgc ccacggttag tgccacaatc 1920
caggcccagc tgcagctccc acctccagcc tgcagtatcc tgctgacaaa cttcctgtgt 1980
accttaccag caggacacca gattggcaca gtcagtccct tgttccaaaa attggaaaat 2040
gaccagattg aaagtttaag gcagcgcttt ggaggggggcc aggcaaaaac gtccccgaag 2100
ccagcagttg tagagactgt tacaacagcc aagccacagc agatacaagc gctgatggat 2160
gaagtgacaa aacaaggaaa cattgtccga gaactgaaag cacaaaaggc agacaagaac 2220
gaggttgctg cggaggtggc gaaactcttg gatctaaaga aacagttggc tgtagctgag 2280
gggaaacccc ctgaagcccc taaaggcaag aagaaaaagt aa                    2322
```

<210> 141
<211> 666
<212> PRT

<213> *Homo sapiens*

<400> 141

Met Ala Val Thr Ala Trp Glu Lys Gly Leu Glu Ser Leu Pro Pro Leu
5

```
 1               5                  10                  15
Arg Pro Gln Gln Asn Pro Val Leu Pro Val Ala Gly Glu Arg Asn Val
            20                  25                  30
Leu Ile Thr Ser Ala Leu Pro Tyr Val Asn Asn Val Pro His Leu Gly
        35                  40                  45
Asn Ile Ile Gly Cys Val Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser
    50                  55                  60
Arg Leu Arg Gln Trp Asn Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr
65                  70                  75                  80
Gly Thr Ala Thr Glu Thr Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln
                85                  90                  95
Glu Ile Cys Asp Lys Tyr His Ile Ile His Ala Asp Ile Tyr Arg Trp
            100                 105                 110
Phe Asn Ile Ser Phe Asp Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln
        115                 120                 125
Thr Lys Ile Thr Gln Asp Ile Phe Gln Gln Leu Leu Lys Arg Gly Phe
    130                 135                 140
Val Leu Gln Asp Thr Val Glu Gln Leu Arg Cys Glu His Cys Ala Arg
145                 150                 155                 160
Phe Leu Ala Asp Arg Phe Val Glu Gly Val Cys Pro Phe Cys Gly Tyr
                165                 170                 175
Glu Glu Ala Arg Gly Asp Gln Cys Asp Lys Cys Gly Lys Leu Ile Asn
            180                 185                 190
Ala Val Glu Leu Lys Lys Pro Gln Cys Lys Val Cys Arg Ser Cys Pro
        195                 200                 205
Val Val Gln Ser Ser Gln His Leu Phe Leu Asp Leu Pro Lys Leu Glu
    210                 215                 220
Lys Arg Leu Glu Glu Trp Leu Gly Arg Thr Leu Pro Gly Ser Asp Trp
225                 230                 235                 240
Thr Pro Asn Ala Gln Phe Ile Thr Arg Ser Trp Leu Arg Asp Gly Leu
                245                 250                 255
Lys Pro Arg Cys Ile Thr Arg Asp Leu Lys Trp Gly Thr Pro Val Pro
            260                 265                 270
Leu Glu Gly Phe Glu Asp Lys Val Phe Tyr Val Trp Phe Asp Ala Thr
        275                 280                 285
Ile Gly Tyr Leu Ser Ile Thr Ala Asn Tyr Thr Asp Gln Trp Glu Arg
    290                 295                 300
Trp Trp Lys Asn Pro Glu Gln Val Asp Leu Tyr Gln Phe Met Ala Lys
305                 310                 315                 320
Asp Asn Val Pro Phe His Ser Leu Val Phe Pro Cys Ser Ala Leu Gly
                325                 330                 335
Ala Glu Asp Asn Tyr Thr Leu Val Ser His Leu Ile Ala Thr Glu Tyr
            340                 345                 350
Leu Asn Tyr Glu Asp Gly Lys Phe Ser Lys Ser Arg Gly Val Gly Val
        355                 360                 365
Phe Gly Asp Met Ala Gln Asp Thr Gly Ile Pro Ala Asp Ile Trp Arg
    370                 375                 380
Phe Tyr Leu Leu Tyr Ile Arg Pro Glu Gly Gln Asp Ser Ala Phe Ser
385                 390                 395                 400
Trp Thr Asp Leu Leu Leu Lys Asn Asn Ser Glu Leu Leu Asn Asn Leu
                405                 410                 415
Gly Asn Phe Ile Asn Arg Ala Gly Met Phe Val Ser Lys Phe Phe Gly
            420                 425                 430
Gly Tyr Val Pro Glu Met Val Leu Thr Pro Asp Asp Gln Arg Leu Leu
        435                 440                 445
Ala His Val Thr Leu Glu Leu Gln His Tyr His Gln Leu Leu Glu Lys
    450                 455                 460
Val Arg Ile Arg Asp Ala Leu Arg Ser Ile Leu Thr Ile Ser Arg His
465                 470                 475                 480
Gly Asn Gln Tyr Ile Gln Val Asn Glu Pro Trp Lys Arg Ile Lys Gly
                485                 490                 495
Ser Glu Ala Asp Arg Gln Arg Ala Gly Thr Val Thr Gly Leu Ala Val
            500                 505                 510
Asn Ile Ala Ala Leu Leu Ser Val Met Leu Gln Pro Tyr Met Pro Thr
        515                 520                 525
```

```
Val Ser Ala Thr Ile Gln Ala Gln Leu Gln Leu Pro Pro Pro Ala Cys
    530                 535                 540
Ser Ile Leu Leu Thr Asn Phe Leu Cys Thr Leu Pro Ala Gly His Gln
545                 550                 555                 560
Ile Gly Thr Val Ser Pro Leu Phe Gln Lys Leu Glu Asn Asp Gln Ile
                565                 570                 575
Glu Ser Leu Arg Gln Arg Phe Gly Gly Gly Gln Ala Lys Thr Ser Pro
            580                 585                 590
Lys Pro Ala Val Val Glu Thr Val Thr Thr Ala Lys Pro Gln Gln Ile
        595                 600                 605
Gln Ala Leu Met Asp Glu Val Thr Lys Gln Gly Asn Ile Val Arg Glu
    610                 615                 620
Leu Lys Ala Gln Lys Ala Asp Lys Asn Glu Val Ala Ala Glu Val Ala
625                 630                 635                 640
Lys Leu Leu Asp Leu Lys Lys Gln Leu Ala Val Ala Glu Gly Lys Pro
                645                 650                 655
Pro Glu Ala Pro Lys Gly Lys Lys Lys Lys
            660                 665
```

<210> 142
<211> 2001
<212> ADN
<213> *Homo sapiens*

<400> 142

```
atggctgtta ctgcttggga gaagggccta gaaagtttgc ccccgctgcg gccccagcag 60
aatccagtgt tgcctgtggc tggagaaagg aatgtgctca tcaccagtgc cctcccttac 120
gtcaacaatg tcccccacct tgggaacatc attggttgtg tgctcagtgc cgatgtcttt 180
gccaggtact ctcgcctccg ccagtggaac accctctatc tgtgtgggac agatgagtat 240
ggtacagcaa cagagaccaa ggctctggag gagggactaa ccccccagga gatctgcgac 300
aagtaccaca tcatccatgc tgacatctac cgctggttta acatttcgtt tgatattttt 360
ggtcgcacca ccactccaca gcagaccaaa atcacccagg acattttcca gcagttgctg 420
aaacgaggtt ttgtgctgca agatactgtg gagcaactgc gatgtgagca ctgtgctcgc 480
ttcctggctg accgcttcgt ggagggcgtg tgtcccttct gtggctatga ggaggctcgg 540
ggtgaccagt gtgacaagtg tggcaagctc atcaatgctg tcgagcttaa gaagcctcag 600
tgtaaagtct gccgatcatg ccctgtggtg cagtcgagcc agcacctgtt tctggacctg 660
cctaagctgg agaagcgact ggaggagtgg ttggggagga cattgcctgg cagtgactgg 720
acacccaatg cccagtttat cacccgttct tggcttcggg atggcctcaa gccacgctgc 780
ataacccgag acctcaaatg gggaacccct gtacccttag aaggttttga agacaaggta 840
ttctatgtct ggtttgatgc cactattggc tatctgtcca tcacagccaa ctacacagac 900
cagtgggaga gatggtggaa gaacccagag caagtggacc tgtatcagtt catggccaaa 960
gacaatgttc ctttccatag cttagtcttt ccttgctcag ccctaggagc tgaggataac 1020
tataccttgg tcagccacct cattgctaca gagtacctga actatgagga tgggaaattc 1080
tctaagagcc gcggtgtggg agtgtttggg gacatggccc aggacacggg gatccctgct 1140
gacatctggc gcttctatct gctgtacatt cggcctgagg gccaggacag tgctttctcc 1200
tggacggacc tgctgctgaa gaataattct gagctgctta acaacctggg caacttcatc 1260
aacagagctg ggatgtttgt gtctaagttc tttgggggct atgtgcctga gatggtgctc 1320
acccctgatg atcagcgcct gctggcccat gtcaccctgg agctccagca ctatcaccag 1380
ctacttgaga aggttcggat ccgggatgcc ttgcgcagta tcctcaccat atctcgacat 1440
ggcaaccaat atattcaggt gaatgagccc tggaagcgga ttaaaggcag tgaggctgac 1500
aggcaacggg caggaacagt gactggcttg gcagtgaata tagctgcctt gctctctgtc 1560
atgcttcagc cttacatgcc cacggttagt gccacaatcc aggcccagct gcagctccca 1620
cctccagcct gcagtatcct gctgacaaac ttcctgtgta ccttaccagc aggacaccag 1680
attggcacag tcagtccctt gttccaaaaa ttggaaaatg accagattga aagtttaagg 1740
cagcgctttg gaggggccca ggcaaaaacg tccccgaagc cagcagttgt agagactgtt 1800
acaacagcca agccacagca gatacaagcg ctgatggatg aagtgacaaa acaaggaaac 1860
attgtccgag aactgaaagc acaaaaggca gacaagaacg aggttgctgc ggaggtggcg 1920
aaactcttgg atctaaagaa acagttggct gtagctgagg ggaaaccccc tgaagcccct 1980
aaaggcaaga agaaaaagta a                                          2001
```

<210> 143
<211> 647
<212> PRT
<213> *Homo sapiens*

<400> 143

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

```
            500                 505                 510
Thr Ile Gln Ala Gln Leu Gln Leu Pro Pro Pro Ala Cys Ser Ile Leu
        515                 520                 525
Leu Thr Asn Phe Leu Cys Thr Leu Pro Ala Gly His Gln Ile Gly Thr
    530                 535                 540
Val Ser Pro Leu Phe Gln Lys Leu Glu Asn Asp Gln Ile Glu Ser Leu
545                 550                 555                 560
Arg Gln Arg Phe Gly Gly Gly Gln Ala Lys Thr Ser Pro Lys Pro Ala
                565                 570                 575
Val Val Glu Thr Val Thr Thr Ala Lys Pro Gln Gln Ile Gln Ala Leu
            580                 585                 590
Met Asp Glu Val Thr Lys Gln Gly Asn Ile Val Arg Glu Leu Lys Ala
        595                 600                 605
Gln Lys Ala Asp Lys Asn Glu Val Ala Ala Glu Val Ala Lys Leu Leu
    610                 615                 620
Asp Leu Lys Lys Gln Leu Ala Val Ala Glu Gly Lys Pro Pro Glu Ala
625                 630                 635                 640
Pro Lys Gly Lys Lys Lys Lys
                645
```

<210> 144
<211> 1944
<212> ADN
<213> *Homo sapiens*

<400> 144

```
atgagcctga ggttgcctgt ggctggagaa aggaatgtgc tcatcaccag tgccctccct 60
tacgtcaaca atgtccccca ccttgggaac atcattggtt gtgtgctcag tgccgatgtc 120
tttgccaggt actctcgcct ccgccagtgg aacaccctct atctgtgtgg gacagatgag 180
tatggtacag caacagagac caaggctctg gaggagggac taacccccca ggagatctgc 240
gacaagtacc acatcatcca tgctgacatc taccgctggt ttaacatttc gtttgatatt 300
tttggtcgca ccaccactcc acagcagacc aaaatcaccc aggacatttt ccagcagttg 360
ctgaaacgag gttttgtgct gcaagatact gtggagcaac tgcgatgtga gcactgtgct 420
cgcttcctgg ctgaccgctt cgtggagggc gtgtgtccct tctgtggcta tgaggaggct 480
cggggtgacc agtgtgacaa gtgtggcaag ctcatcaatg ctgtcgagct taagaagcct 540
cagtgtaaag tctgccgatc atgccctgtg gtgcagtcga gccagcacct gtttctggac 600
ctgcctaagc tggagaagcg actggaggag tggttgggga ggacattgcc tggcagtgac 660
tggacaccca atgcccagtt tatcacccgt tcttggcttc gggatggcct caagccacgc 720
tgcataaccc gagacctcaa atggggaacc cctgtaccct tagaaggttt tgaagacaag 780
gtattctatg tctggtttga tgccactatt ggctatctgt ccatcacagc caactacaca 840
gaccagtggg agagatggtg gaagaaccca gagcaagtgg acctgtatca gttcatggcc 900
aaagacaatg ttcctttcca tagcttagtc tttccttgct cagccctagg agctgaggat 960
aactatacct tggtcagcca cctcattgct acagagtacc tgaactatga ggatgggaaa 1020
ttctctaaga gccgcggtgt gggagtgttt ggggacatgg cccaggacac ggggatccct 1080
gctgacatct ggcgcttcta tctgctgtac attcggcctg agggccagga cagtgctttc 1140
tcctggacgg acctgctgct gaagaataat tctgagctgc ttaacaacct gggcaacttc 1200
atcaacagag ctgggatgtt tgtgtctaag ttctttgggg gctatgtgcc tgagatggtg 1260
ctcacccctg atgatcagcg cctgctggcc catgtcaccc tggagctcca gcactatcac 1320
cagctacttg agaaggttcg gatccgggat gccttgcgca gtatcctcac catatctcga 1380
catggcaacc aatatattca ggtgaatgag ccctggaagc ggattaaagg cagtgaggct 1440
gacaggcaac gggcaggaac agtgactggc ttggcagtga atatagctgc cttgctctct 1500
gtcatgcttc agccttacat gcccacggtt agtgccacaa tccaggccca gctgcagctc 1560
ccacctccag cctgcagtat cctgctgaca aacttcctgt gtaccttacc agcaggacac 1620
cagattggca cagtcagtcc cttgttccaa aaattggaaa atgaccagat tgaaagttta 1680
aggcagcgct ttggaggggg ccaggcaaaa acgtccccga agccagcagt tgtagagact 1740
gttacaacag ccaagccaca gcagatacaa gcgctgatgg atgaagtgac aaaacaagga 1800
aacattgtcc gagaactgaa agcacaaaag gcagacaaga acgaggttgc tgcggaggtg 1860
gcgaaactct tggatctaaa gaaacagttg gctgtagctg aggggaaacc ccctgaagcc 1920
cctaaaggca agaagaaaaa gtaa                                        1944
```

<210> 145
<211> 832
<212> PRT
<213> *Homo sapiens*

<400> 145

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

```
                500                 505                 510
Leu Ile Ala Thr Glu Tyr Leu Asn Tyr Glu Asp Gly Lys Phe Ser Lys
        515                 520                 525
Ser Arg Gly Val Gly Val Phe Gly Asp Met Ala Gln Asp Thr Gly Ile
    530                 535                 540
Pro Ala Asp Ile Trp Arg Phe Tyr Leu Leu Tyr Ile Arg Pro Glu Gly
545                 550                 555                 560
Gln Asp Ser Ala Phe Ser Trp Thr Asp Leu Leu Leu Lys Asn Asn Ser
                565                 570                 575
Glu Leu Leu Asn Asn Leu Gly Asn Phe Ile Asn Arg Ala Gly Met Phe
            580                 585                 590
Val Ser Lys Phe Phe Gly Gly Tyr Val Pro Glu Met Val Leu Thr Pro
        595                 600                 605
Asp Asp Gln Arg Leu Leu Ala His Val Thr Leu Glu Leu Gln His Tyr
    610                 615                 620
His Gln Leu Leu Glu Lys Val Arg Ile Arg Asp Ala Leu Arg Ser Ile
625                 630                 635                 640
Leu Thr Ile Ser Arg His Gly Asn Gln Tyr Ile Gln Val Asn Glu Pro
                645                 650                 655
Trp Lys Arg Ile Lys Gly Ser Glu Ala Asp Arg Gln Arg Ala Gly Thr
            660                 665                 670
Val Thr Gly Leu Ala Val Asn Ile Ala Ala Leu Leu Ser Val Met Leu
        675                 680                 685
Gln Pro Tyr Met Pro Thr Val Ser Ala Thr Ile Gln Ala Gln Leu Gln
    690                 695                 700
Leu Pro Pro Pro Ala Cys Ser Ile Leu Leu Thr Asn Phe Leu Cys Thr
705                 710                 715                 720
Leu Pro Ala Gly His Gln Ile Gly Thr Val Ser Pro Leu Phe Gln Lys
                725                 730                 735
Leu Glu Asn Asp Gln Ile Glu Ser Leu Arg Gln Arg Phe Gly Gly Gly
            740                 745                 750
Gln Ala Lys Thr Ser Pro Lys Pro Ala Val Val Glu Thr Val Thr Thr
        755                 760                 765
Ala Lys Pro Gln Gln Ile Gln Ala Leu Met Asp Glu Val Thr Lys Gln
    770                 775                 780
Gly Asn Ile Val Arg Glu Leu Lys Ala Gln Lys Ala Asp Lys Asn Glu
785                 790                 795                 800
Val Ala Ala Glu Val Ala Lys Leu Leu Asp Leu Lys Lys Gln Leu Ala
                805                 810                 815
Val Ala Glu Gly Lys Pro Pro Glu Ala Pro Lys Gly Lys Lys Lys Lys
            820                 825                 830
```

<210> 146
<211> 2499
<212> ADN
<213> *Homo sapiens*

<400> 146

```
atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gcccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaaa ggaatgtgct catcaccagt gccctccctt acgtcaacaa tgtcccccac 840
cttgggaaca tcattggttg tgtgctcagt gccgatgtct ttgccagaat cacccaggac 900
attttccagc agttgctgaa acgaggtttt gtgctgcaag atactgtgga gcaactgcga 960

tgtgagcact gtgctcgctt cctggctgac cgcttcgtgg agggcgtgtg tcccttctgt 1020
ggctatgagg aggctcgggg tgaccagtgt gacaagtgtg gcaagctcat caatgctgtc 1080
gagcttaaga agcctcagtg taaagtctgc cgatcatgcc ctgtggtgca gtcgagccag 1140
cacctgtttc tggacctgcc taagctggag aagcgactgg aggagtggtt ggggaggaca 1200
ttgcctggca gtgactggac acccaatgcc cagtttatca cccgttcttg gcttcgggat 1260
ggcctcaagc cacgctgcat aacccgagac ctcaaatggg gaacccctgt acccttagaa 1320
ggttttgaag acaaggtatt ctatgtctgg tttgatgcca ctattggcta tctgtccatc 1380
acagccaact acacagacca gtgggagaga tggtggaaga acccagagca agtggacctg 1440
tatcagttca tggccaaaga caatgttcct ttccatagct tagtctttcc ttgctcagcc 1500
ctaggagctg aggataacta taccttggtc agccacctca ttgctacaga gtacctgaac 1560
tatgaggatg ggaaattctc taagagccgc ggtgtgggag tgtttgggga catggcccag 1620
gacacgggga tccctgctga catctggcgc ttctatctgc tgtacattcg gcctgagggc 1680
caggacagtg ctttctcctg gacggacctg ctgctgaaga ataattctga gctgcttaac 1740
aacctgggca acttcatcaa cagagctggg atgtttgtgt ctaagttctt tgggggctat 1800
gtgcctgaga tggtgctcac ccctgatgat cagcgcctgc tggcccatgt caccctggag 1860
ctccagcact atcaccagct acttgagaag gttcggatcc gggatgcctt gcgcagtatc 1920
ctcaccatat ctcgacatgg caaccaatat attcaggtga atgagccctg gaagcggatt 1980
aaaggcagtg aggctgacag gcaacgggca ggaacagtga ctggcttggc agtgaatata 2040
gctgccttgc tctctgtcat gcttcagcct tacatgccca cggttagtgc cacaatccag 2100
gcccagctgc agctcccacc tccagcctgc agtatcctgc tgacaaactt cctgtgtacc 2160
ttaccagcag gacaccagat tggcacagtc agtcccttgt tccaaaaatt ggaaaatgac 2220
cagattgaaa gtttaaggca gcgctttgga gggggccagg caaaaacgtc cccgaagcca 2280
gcagttgtag agactgttac aacagccaag ccacagcaga tacaagcgct gatggatgaa 2340
gtgacaaaac aaggaaacat tgtccgagaa ctgaaagcac aaaaggcaga caagaacgag 2400
gttgctgcgg aggtggcgaa actcttggat ctaaagaaac agttggctgt agctgagggg 2460
aaaccccctg aagcccctaa aggcaagaag aaaaagtaa                        2499
```

<210> 147
<211> 875
<212> PRT
<213> *Homo sapiens*

<400> 147

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

```
Val Leu Pro Val Ala Gly Glu Arg Asn Val Leu Ile Thr Ser Ala Leu
            260                 265                 270
Pro Tyr Val Asn Asn Val Pro His Leu Gly Asn Ile Ile Gly Cys Val
        275                 280                 285
Leu Ser Ala Asp Val Phe Ala Arg Tyr Ser Arg Leu Arg Gln Trp Asn
    290                 295                 300
Thr Leu Tyr Leu Cys Gly Thr Asp Glu Tyr Gly Thr Ala Thr Glu Thr
305                 310                 315                 320
Lys Ala Leu Glu Glu Gly Leu Thr Pro Gln Glu Ile Cys Asp Lys Tyr
                325                 330                 335
His Ile Ile His Ala Asp Ile Tyr Arg Trp Phe Asn Ile Ser Phe Asp
            340                 345                 350
Ile Phe Gly Arg Thr Thr Thr Pro Gln Gln Thr Lys Ile Thr Gln Asp
        355                 360                 365
Ile Phe Gln Gln Leu Leu Lys Arg Gly Phe Val Leu Gln Asp Thr Val
    370                 375                 380
Glu Gln Leu Arg Cys Glu His Cys Ala Arg Phe Leu Ala Asp Arg Phe
385                 390                 395                 400
Val Glu Gly Val Cys Pro Phe Cys Gly Tyr Glu Glu Ala Arg Gly Asp
                405                 410                 415
Gln Cys Asp Lys Cys Gly Lys Leu Ile Asn Ala Val Glu Leu Lys Leu
            420                 425                 430
Glu Lys Arg Leu Glu Glu Trp Leu Gly Arg Thr Leu Pro Gly Ser Asp
        435                 440                 445
Trp Thr Pro Asn Ala Gln Phe Ile Thr Arg Ser Trp Leu Arg Asp Gly
    450                 455                 460
Leu Lys Pro Arg Cys Ile Thr Arg Asp Leu Lys Trp Gly Thr Pro Val
465                 470                 475                 480
Pro Leu Glu Gly Phe Glu Asp Lys Val Phe Tyr Val Trp Phe Asp Ala
                485                 490                 495
Thr Ile Gly Tyr Leu Ser Ile Thr Ala Asn Tyr Thr Asp Gln Trp Glu
            500                 505                 510
Arg Trp Trp Lys Asn Pro Glu Gln Val Asp Leu Tyr Gln Phe Met Ala
        515                 520                 525
Lys Asp Asn Val Pro Phe His Ser Leu Val Phe Pro Cys Ser Ala Leu
    530                 535                 540
Gly Ala Glu Asp Asn Tyr Thr Leu Val Ser His Leu Ile Ala Thr Glu
545                 550                 555                 560
Tyr Leu Asn Tyr Glu Asp Gly Lys Phe Ser Lys Ser Arg Gly Val Gly
                565                 570                 575
Val Phe Gly Asp Met Ala Gln Asp Thr Gly Ile Pro Ala Asp Ile Trp
            580                 585                 590
Arg Phe Tyr Leu Leu Tyr Ile Arg Pro Glu Gly Gln Asp Ser Ala Phe
        595                 600                 605
Ser Trp Thr Asp Leu Leu Leu Lys Asn Asn Ser Glu Leu Leu Asn Asn
    610                 615                 620
Leu Gly Asn Phe Ile Asn Arg Ala Gly Met Phe Val Ser Lys Phe Phe
625                 630                 635                 640
Gly Gly Tyr Val Pro Glu Met Val Leu Thr Pro Asp Asp Gln Arg Leu
                645                 650                 655
Leu Ala His Val Thr Leu Glu Leu Gln His Tyr His Gln Leu Leu Glu
            660                 665                 670
Lys Val Arg Ile Arg Asp Ala Leu Arg Ser Ile Leu Thr Ile Ser Arg
        675                 680                 685
His Gly Asn Gln Tyr Ile Gln Val Asn Glu Pro Trp Lys Arg Ile Lys
    690                 695                 700
Gly Ser Glu Ala Asp Arg Gln Arg Ala Gly Thr Val Thr Gly Leu Ala
705                 710                 715                 720
Val Asn Ile Ala Ala Leu Leu Ser Val Met Leu Gln Pro Tyr Met Pro
                725                 730                 735
Thr Val Ser Ala Thr Ile Gln Ala Gln Leu Gln Leu Pro Pro Pro Ala
            740                 745                 750
Cys Ser Ile Leu Leu Thr Asn Phe Leu Cys Thr Leu Pro Ala Gly His
        755                 760                 765
Gln Ile Gly Thr Val Ser Pro Leu Phe Gln Lys Leu Glu Asn Asp Gln
```

```
    770                 775                 780
Ile Glu Ser Leu Arg Gln Arg Phe Gly Gly Gly Gln Ala Lys Thr Ser
785                 790                 795                 800
Pro Lys Pro Ala Val Val Glu Thr Val Thr Thr Ala Lys Pro Gln Gln
                805                 810                 815
Ile Gln Ala Leu Met Asp Glu Val Thr Lys Gln Gly Asn Ile Val Arg
            820                 825                 830
Glu Leu Lys Ala Gln Lys Ala Asp Lys Asn Glu Val Ala Ala Glu Val
        835                 840                 845
Ala Lys Leu Leu Asp Leu Lys Lys Gln Leu Ala Val Ala Glu Gly Lys
    850                 855                 860
Pro Pro Glu Ala Pro Lys Gly Lys Lys Lys Lys
865                 870                 875
```

<210> 148
<211> 2628
<212> ADN
<213> *Homo sapiens*

<400> 148

atgagactgt tcgtgagtga tggcgtcccg ggttgcttgc cggtgctggc cgccgccggg 60
agagcccggg gcagagcaga ggtgctcatc agcactgtag gcccggaaga ttgtgtggtc 120
ccgttcctga cccggcctaa ggtccctgtc ttgcagctgg atagcggcaa ctacctcttc 180
tccactagtg caatctgccg atatttttt ttgttatctg gctgggagca agatgacctc 240
actaaccagt ggctggaatg ggaagcgaca gagctgcagc cagctttgtc tgctgccctg 300
tactatttag tggtccaagg caagaagggg gaagatgttc ttggttcagt gcggagagcc 360
ctgactcaca ttgaccacag cttgagtcgt cagaactgtc ctttcctggc tggggagaca 420
gaatctctag ccgacattgt tttgtgggga gccctatacc cattactgca agatcccgcc 480
tacctccctg aggagctgag tgccctgcac agctggttcc agacactgag tacccaggaa 540
ccatgtcagc gagctgcaga gactgtactg aaacagcaag gtgtcctggc tctccggcct 600
tacctccaaa agcagcccca gcccagcccc gctgagggaa gggctgtcac caatgagcct 660
gaggaggagg agctggctac cctatctgag gaggagattg ctatggctgt tactgcttgg 720
gagaagggcc tagaaagttt gcccccgctg cggccccagc agaatccagt gttgcctgtg 780
gctggagaaa ggaatgtgct catcaccagt gccctccctt acgtcaacaa tgtcccccac 840
cttgggaaca tcattggttg tgtgctcagt gccgatgtct ttgccaggta ctctcgcctc 900
cgccagtgga acacccctcta tctgtgtggg acagatgagt atggtacagc aacagagacc 960
aaggctctgg aggagggact aacccccccag gagatctgcg acaagtacca catcatccat 1020
gctgacatct accgctggtt taacatttcg tttgatattt ttggtcgcac caccactcca 1080
cagcagacca aaatcaccca ggacattttc cagcagttgc tgaaacgagg ttttgtgctg 1140
caagatactg tggagcaact gcgatgtgag cactgtgctc gcttcctggc tgaccgcttc 1200
gtggagggcg tgtgtccctt ctgtggctat gaggaggctc ggggtgacca gtgtgacaag 1260
tgtggcaagc tcatcaatgc tgtcgagctt aagctggaga agcgactgga ggagtggttg 1320
gggaggacat tgcctggcag tgactggaca cccaatgccc agtttatcac ccgttcttgg 1380
cttcgggatg gcctcaagcc acgctgcata acccgagacc tcaaatgggg aacccctgta 1440
cccttagaag gttttgaaga caaggtattc tatgtctggt ttgatgccac tattggctat 1500
ctgtccatca cagccaacta cacagaccag tgggagagat ggtggaagaa cccagagcaa 1560
gtggacctgt atcagttcat ggccaaagac aatgttcctt tccatagctt agtctttcct 1620
tgctcagccc taggagctga ggataactat accttggtca gccacctcat tgctacagag 1680
tacctgaact atgaggatgg gaaattctct aagagccgcg gtgtgggagt gtttggggac 1740
atggcccagg acacggggat ccctgctgac atctggcgct tctatctgct gtacattcgg 1800
cctgagggcc aggacagtgc tttctcctgg acggacctgc tgctgaagaa taattctgag 1860
ctgcttaaca acctgggcaa cttcatcaac agagctggga tgtttgtgtc taagttcttt 1920
gggggctatg tgcctgagat ggtgctcacc cctgatgatc agcgcctgct ggcccatgtc 1980
accctggagc tccagcacta tcaccagcta cttgagaagg ttcggatccg ggatgccttg 2040
cgcagtatcc tcaccatatc tcgacatggc aaccaatata ttcaggtgaa tgagccctgg 2100
aagcggatta aaggcagtga ggctgacagg caacgggcag gaacagtgac tggcttggca 2160
gtgaatatag ctgccttgct ctctgtcatg cttcagcctt acatgcccac ggttagtgcc 2220
acaatccagg cccagctgca gctcccacct ccagcctgca gtatcctgct gacaaacttc 2280
ctgtgtacct taccagcagg acaccagatt ggcacagtca gtcccttgtt ccaaaaattg 2340
gaaaatgacc agattgaaag tttaaggcag cgctttggag ggggccaggc aaaaacgtcc 2400
ccgaagccag cagttgtaga gactgttaca acagccaagc cacagcagat acaagcgctg 2460
atggatgaag tgacaaaaca aggaaacatt gtccgagaac tgaaagcaca aaaggcagac 2520
aagaacgagg ttgctgcgga ggtggcgaaa ctcttggatc taaagaaaca gttggctgta 2580
gctgagggga aacccccctga agcccctaaa ggcaagaaga aaaagtaa 2628

<210> 149
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 149
tcttctccac tagtgcaatc tgccggtatt ctatgtctgg tttgatgcca 50

<210> 150
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 150
gcaccaccac tccacagcag accaaaagcc tcagtgtaaa gtctgccgat 50

<210> 151
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 151
gagatggtgg aagaacccag agcaaagtac ctgaactatg aggatgggaa 50

<210> 152
<211> 15
<212> PRT
<213> *Homo sapiens*

<400> 152

Met Val Glu Glu Pro Arg Ala Lys Tyr Leu Asn Tyr Glu Asp Gly
1 5 10 15

<210> 153
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 153
ggattaaagg cagtgaggct gacaggtcag tcccttgttc caaaaattgg 50

<210> 154
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 154
ctcatcagca ctgtaggccc ggaagaggag ctgagtgccc tgcacagctg 50

<210> 155
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 155

Leu Ile Ser Thr Val Gly Pro Glu Glu Glu Leu Ser Ala Leu His Ser
1 5 10 15

<210> 156
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 156
tcttctccac tagtgcaatc tgccgaggag ctgagtgccc tgcacagctg 50

<210> 157
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 157
aagggctgtc accaatgagc ctgaggttgc ctgtggctgg agaaaggaat 50

<210> 158

<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 158

Gly Leu Ser Pro Met Ser Leu Arg Leu Pro Val Ala Gly Glu Arg Asn
1 5 10 15

<210> 159
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 159
tgctcagtgc cgatgtcttt gccagaatca cccaggacat ttccagcag 50

<210> 160
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 160

Leu Ser Ala Asp Val Phe Ala Arg Ile Thr Gln Asp Ile Phe Gln Gln
1 5 10 15

<210> 161
<211> 50
<212> ADN
<213> *Homo sapiens*

<400> 161
gctcatcaat gctgtcgagc ttaagctgga gaagcgactg gaggagtggt 50

<210> 162
<211> 16
<212> PRT
<213> *Homo sapiens*

<400> 162

Leu Ile Asn Ala Val Glu Leu Lys Leu Glu Lys Arg Leu Glu Glu Trp
1 5 10 15

<210> 163
<211> 128
<212> PRT
<213> *Homo sapiens*

<400> 163

Leu Pro Pro Pro Ala Cys Ser Ile Leu Leu Thr Asn Phe Leu Cys Thr

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

<210> 164
<211> 387
<212> ADN
<213> *Homo sapiens*

<400> 164

```
ctcccacctc cagcctgcag tatcctgctg acaaacttcc tgtgtacctt accagcagga 60
caccagattg gcacagtcag tcccttgttc caaaaattgg aaaatgacca gattgaaagt 120
ttaaggcagc gctttggagg gggccaggca aaaacgtccc cgaagccagc agttgtagag 180
actgttacaa cagccaagcc acagcagata caagcgctga tggatgaagt gacaaaacaa 240
ggaaacattg tccgagaact gaaagcacaa aaggcagaca agaacgaggt tgctgcggag 300
gtggcgaaac tcttggatct aaagaaacag ttggctgtag ctgaggggaa acccccctgaa 360
gcccctaaag gcaagaagaa aaagtaa                                      387
```

<210> 165
<211> 149
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 165

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

<210> 166
<211> 149
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 166

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

<210> 167
<211> 369
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 167

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

        195                 200                 205
Thr Val Thr Gly Leu Ala Val Asn Ile Ala Ala Leu Leu Ser Val Met
    210                 215                 220
Leu Gln Pro Tyr Met Pro Thr Val Ser Ala Thr Ile Gln Ala Gln Leu
225                 230                 235                 240
Gln Leu Pro Pro Pro Ala Cys Ser Ile Leu Leu Thr Asn Phe Leu Cys
                245                 250                 255
Thr Leu Pro Ala Gly His Gln Ile Gly Thr Val Ser Pro Leu Phe Gln
            260                 265                 270
Lys Leu Glu Asn Asp Gln Ile Glu Ser Leu Arg Gln Arg Phe Gly Gly
        275                 280                 285
Gly Gln Ala Lys Thr Ser Pro Lys Pro Ala Val Val Glu Thr Val Thr
    290                 295                 300
Thr Ala Lys Pro Gln Gln Ile Gln Ala Leu Met Asp Glu Val Thr Lys
305                 310                 315                 320
Gln Gly Asn Ile Val Arg Glu Leu Lys Ala Gln Lys Ala Asp Lys Asn
                325                 330                 335
Glu Val Ala Ala Glu Val Ala Lys Leu Leu Asp Leu Lys Lys Gln Leu
            340                 345                 350
Ala Val Ala Glu Gly Lys Pro Pro Glu Ala Pro Lys Gly Lys Lys Lys
        355                 360                 365
Lys
```

<210> 168
<211> 369
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 168

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

Gly His Gln Ile Gly Thr Val Ser Pro Leu Phe Gln Lys Leu Glu Asn
                245                 250                 255
Asp Gln Ile Glu Ser Leu Arg Gln Arg Phe Gly Gly Gly Gln Ala Lys
            260                 265                 270
Thr Ser Pro Lys Pro Ala Val Val Glu Thr Val Thr Thr Ala Lys Pro
        275                 280                 285
Gln Gln Ile Gln Ala Leu Met Asp Glu Val Thr Lys Gln Gly Asn Ile
    290                 295                 300
Val Arg Glu Leu Lys Ala Gln Lys Ala Asp Lys Asn Glu Val Ala Ala
305                 310                 315                 320
Glu Val Ala Lys Leu Leu Asp Leu Lys Lys Gln Leu Ala Val Ala Glu
                325                 330                 335
Gly Lys Pro Pro Glu Ala Pro Lys Gly Lys Lys Lys Lys Gly Lys Pro
            340                 345                 350
Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr His His His His His
        355                 360                 365
His
```

<210> 169
<211> 343
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 169

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

        275                 280                 285
Met Asp Glu Val Thr Lys Gln Gly Asn Ile Val Arg Glu Leu Lys Ala
    290                 295                 300
Gln Lys Ala Asp Lys Asn Glu Val Ala Ala Glu Val Ala Lys Leu Leu
305                 310                 315                 320
Asp Leu Lys Lys Gln Leu Ala Val Ala Glu Gly Lys Pro Pro Glu Ala
                325                 330                 335
Pro Lys Gly Lys Lys Lys Lys
            340
```

<210> 170
<211> 343
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 170

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

<210> 171
<211> 75
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x N terminal

<400> 171

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

<210> 172
<211> 75
<212> PRT
<213> Secuencia artificial

<220>
<223> Polipéptido de metionil ARNt sintetasa con marcador de afinidad His 6x C terminal

<400> 172

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

<210> 173
<211> 396
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 173

```
ttgcctccac cggcgtgcag catcttgctg accaacttcc tgtgcaccct gccggccggt 60
caccaaatcg gtaccgtgag cccgctgttc caaaagctgg agaatgacca aattgagagc 120
ctgcgtcaac gttttggtgg tggccaggcc aaaacgtccc cgaaaccggc agttgtcgaa 180
accgttacca cggcaaaacc gcaacagatt caggcactga tggacgaggt cacgaaacag 240
ggcaacatcg tgcgcgaatt gaaagcgcag aaagcggata agaacgaggt ggctgctgag 300
gtcgcgaagc tgctggattt gaagaaacag ctggcggttg cggaaggcaa accaccggaa 360
gcgcctaaag gtaagaagaa gaagtaatga ctcgag                           396
```

<210> 174
<211> 1056
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 174

```
gcgaaagaca acgttccatt tcacagcctg gttttcccgt gtagcgcgct gggtgcagaa 60
gataactata ccctggtttc tcacctgatc gccaccgagt acctgaacta tgaggatggc 120
aaatttagca agagccgcgg cgtcggcgtt ttcggtgaca tggctcagga cacgggtatt 180
ccggcggaca tctggcgctt ctacctgctg tacattcgtc cggagggtca ggatagcgca 240
tttagctgga ccgacttgct gctgaagaat aatagcgagc tgttgaacaa tttgggtaac 300
tttatcaacc gtgcgggtat gtttgtgtcc aaattcttcg gcggttacgt cccggagatg 360
gttctgactc cggatgatca gcgtctgctg gcgcacgtca ccttggaatt gcagcactac 420
caccagctgc tggagaaggt ccgcatccgc gacgctctgc gtagcattct gacgatcagc 480
cgtcatggta atcaatatat ccaagttaac gaaccgtgga aacgtatcaa aggtagcgag 540
gcggaccgtc agcgcgcggg caccgtgacc ggcctggccg ttaatattgc ggcgctgttg 600
agcgtgatgc tgcagccgta tatgccgact gtgagcgcca ccatccaggc acagctgcaa 660
ctgccgccgc cagcttgcag cattctgctg accaatttcc tgtgcacctt gccggcaggc 720
catcaaattg gtacggtgtc tccactgttc cagaagctgg aaaacgacca aattgagtcc 780
ctgcgtcaac gttttggtgg tggtcaagcg aaaacgtcgc ctaaaccggc ggtggtcgaa 840
acggttacga ccgcaaaacc gcaacaaatc caggctctga tggatgaggt gaccaaacaa 900
ggcaatattg ttcgtgaact gaaggcccag aaagcagaca agaacgaggt cgccgcggaa 960
gtcgcaaaac tgctggattt gaagaaacag ctggcagttg cggaaggtaa gccgccggag 1020
gcgccgaagg gcaagaagaa gaaataatga ctcgag                           1056
```

<210> 175
<211> 978
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 175

```
gttgaagaac ctcgtgcaaa atacctgaac tatgaggacg gtaaattcag caagagccgc 60
ggtgttggtg tgtttggcga tatggcccag gacaccggta ttccggcgga catttggcgt 120
ttctatctgc tgtacattcg tccggaaggc caggattccg cattcagctg gaccgacctg 180
ctgctgaaga acaacagcga actgctgaac aacctgggta attttattaa tcgcgccggt 240
atgtttgttt ctaaattctt tggcggttat gttccggaga tggttctgac gccagatgat 300
cagcgtctgc tggcgcatgt tactttggaa ctgcagcact atcatcagtt gctggagaaa 360
gtccgcatcc gcgacgcgct gcgttcgatc ttgacgatca gccgccacgg taaccaatac 420
attcaagtga acgagccgtg gaaacgtatc aagggttctg aagcggaccg tcagcgtgca 480
ggcaccgtga ccggtctggc ggtgaatatc gcggcgttgt tgtccgtgat gctgcagccg 540
tacatgccga cggtgagcgc caccatccaa gcacaattgc agctgccgcc gccggcgtgc 600
agcatcctgc tgaccaattt tctgtgtacg ctgccggcgg gccaccaaat tggcaccgtc 660
agcccactgt tccaaaagct ggagaatgat cagattgaaa gcctgcgtca gcgtttcggc 720
ggcggtcaag ctaaaacgag ccctaagccg gccgtggtcg aaaccgtcac gaccgcaaaa 780
ccgcaacaga ttcaagccct gatggacgag gtcaccaaac agggcaacat cgtgcgtgag 840
ctgaaagcac aaaaggctga caagaatgag gttgcagcgg aagttgcgaa actgctggat 900
ctgaagaagc agctggcggt cgccgagggt aaaccgcctg aggctccgaa aggtaagaag 960
aagaaataat gactcgag                                              978
```

<210> 176
<211> 174
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido de metionil ARNt sintetasa con codones optimizados

<400> 176

```
gacgaggtta ccaaacaggg taacatcgtg cgtgaactga aagcccaaaa ggctgacaag 60
aacgaggtcg cagctgaagt tgcgaaactg ttggatctga agaaacagct ggcggtggca 120
gaaggcaaac caccggaggc gccgaaaggt aagaagaaga aataatgact cgag       174
```

REIVINDICACIONES

1. Una composición terapéutica, que comprende un polipéptido de aminoacil-ARNt sintetasa (AARS) aislado de aproximadamente 50 a aproximadamente 100 aminoácidos de longitud y que comprende una secuencia de aminoácidos que es al menos el 95 %, 98 % o 100 % idéntica a SEQ ID NO: 137, en donde el polipéptido de AARS tiene al menos una actividad de señalización extracelular y una solubilidad de al menos aproximadamente 5 mg/ml, y en donde la composición tiene una pureza de al menos aproximadamente el 95 % en proteínas y menos de aproximadamente 10 UE de endotoxina/mg de proteína.

2. La composición terapéutica de la reivindicación 1, en la que el polipéptido de AARS está fusionado con un polipéptido heterólogo seleccionado del grupo que consiste en marcadores de purificación, marcadores epitópicos, secuencias de dirección, péptidos señal, secuencias de traslocación de membrana y modificadores PK.

3. La composición terapéutica de las reivindicaciones 1 o 2, en la que dicho polipéptido de AARS consiste en la secuencia de aminoácidos expuesta en SEQ ID NO: 137 o difiere de la secuencia de aminoácidos expuesta en SEQ ID NO: 137 por sustitución, deleción y/o adición de aproximadamente 1, 2, 3, 4 o 5 aminoácidos.

4. La composición terapéutica de la reivindicación 3, en la que al menos un resto o un sustrato sólido se unen de forma covalente o no covalente a dicho polipéptido, donde el al menos un resto es un marcador detectable o un polímero soluble en agua.

5. Un anticuerpo que muestra especificidad de unión por un polipéptido de AARS aislado de SEQ ID NO: 137, en donde la afinidad del anticuerpo por el polipéptido de AARS es aproximadamente 10X más fuerte que su afinidad por un polipéptido de AARS de longitud completa correspondiente.

6. Una composición terapéutica, que comprende un anticuerpo de la reivindicación 5, en la que el anticuerpo tiene una afinidad de al menos aproximadamente 1 nM por el polipéptido.

7. Un método para identificar un compuesto que se une específicamente a un polipéptido de aminoacil-ARNt sintetasa (AARS) como se expone en SEQ ID NO: 137, que comprende

a) combinar en condiciones adecuadas el polipéptido de AARS con al menos un compuesto de ensayo, y
b) detectar la unión del polipéptido de AARS al compuesto de ensayo, identificando de este modo un compuesto que se une específicamente al polipéptido de AARS.

8. Una proteína de aminoacil-ARNt sintetasa (AARS) de longitud completa modificada técnicamente que comprende un sitio de proteólisis heterólogo para permitir la generación proteolítica de un polipéptido de AARS como se expone en SEQ ID NO: 137.

9. Una composición terapéutica de cualquiera de las reivindicaciones 1-4, para su uso en el tratamiento de caquexia, un estado de debilitamiento o una enfermedad musculoesquelética.

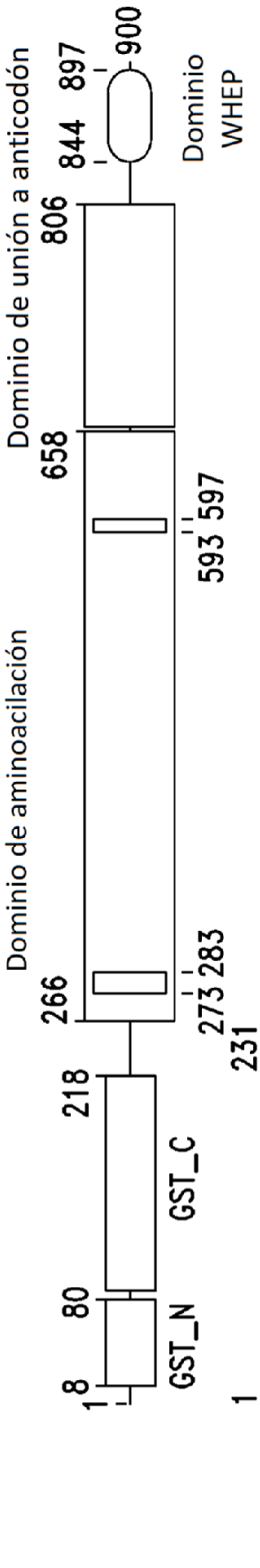
MetRS1N1
1
231
273 283
593 597
658
806
844 897 900
1 8
80
218
266
GST_N
GST_C
Dominio de aminoacilación
Dominio de unión a anticodón
Dominio WHEP

FIG. 1A

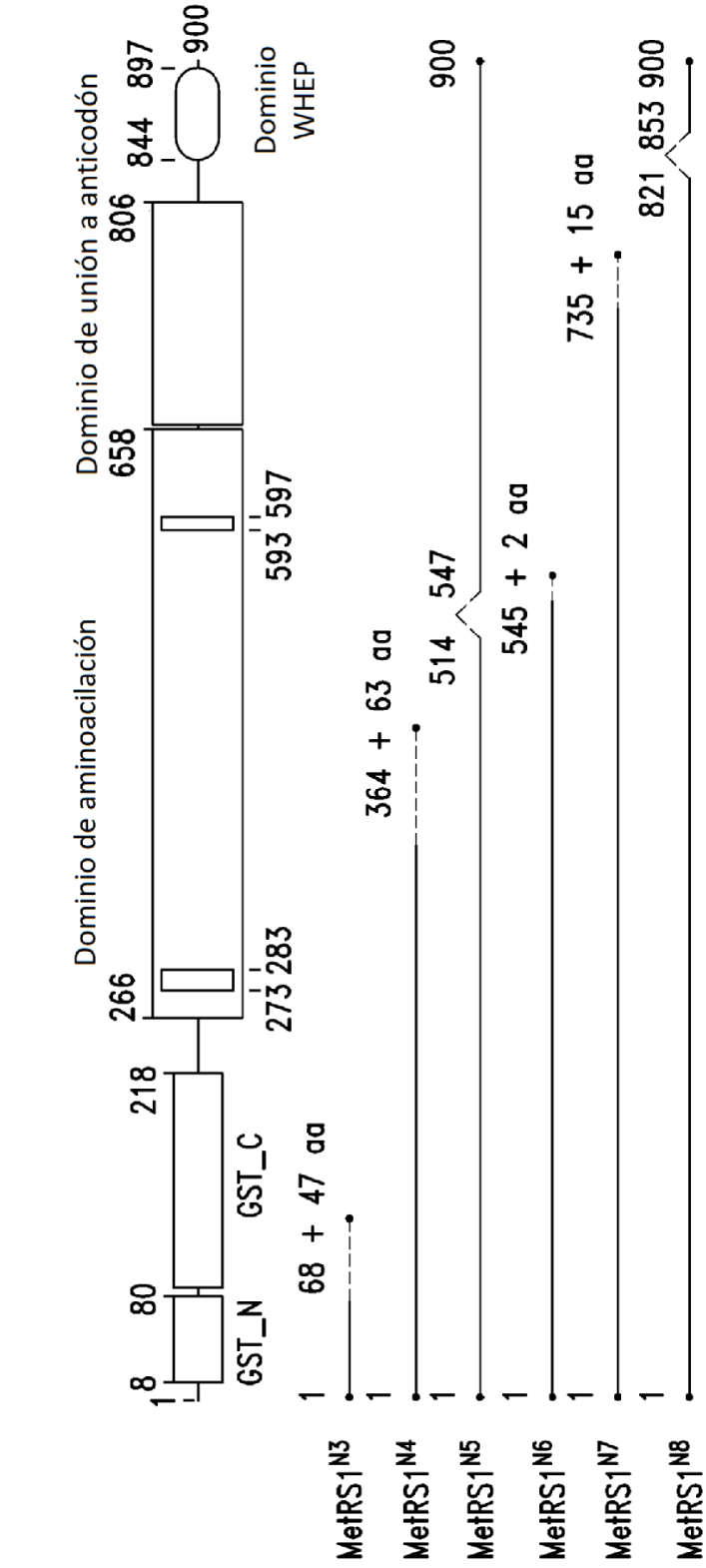
Dominio de aminoacilación
Dominio de unión a anticodón
Dominio WHEP
GST_N
GST_C
1
8
80
218
266
273
283
593
597
658
806
844
897
900
68 + 47 aa
364 + 63 aa
514
547
545 + 2 aa
735 + 15 aa
821
853
MetRS1N3
MetRS1N4
MetRS1N5
MetRS1N6
MetRS1N7
MetRS1N8

FIG. 1B

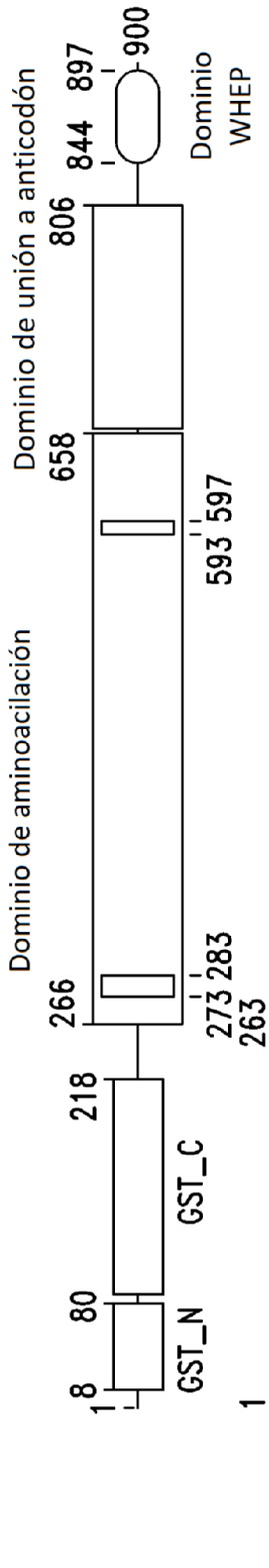
MetRS1N2
1
263
273 283
593 597
658
806
844 897 900
1 8
80
218
266
GST_N
GST_C
Dominio de aminoacilación
Dominio de unión a anticodón
Dominio WHEP

FIG. 1C

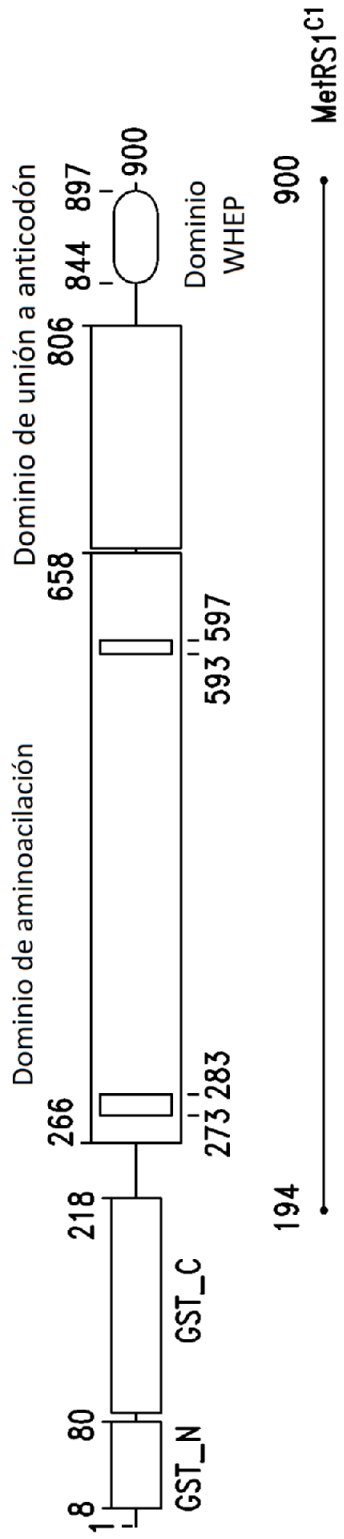
MetRS1C1
900
194
Dominio de aminoacilación
Dominio de unión a anticodón
Dominio WHEP
GST_N
GST_C
1
8
80
218
266
273
283
593
597
658
806
844
897
900

FIG. 2A

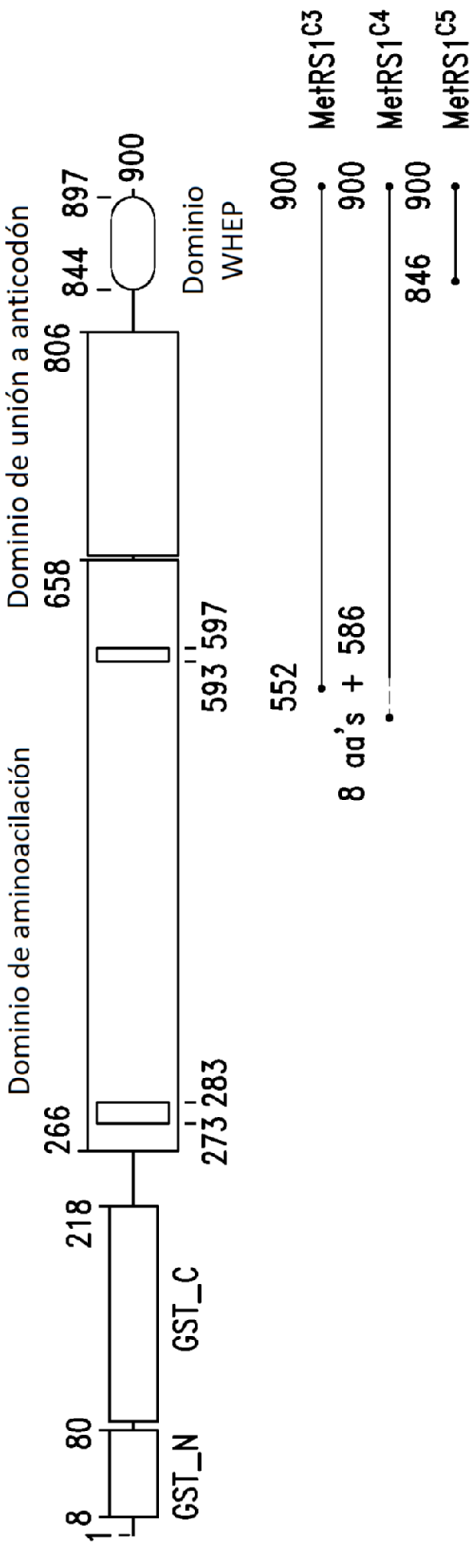
1
8
80
GST_N
218
GST_C
266
273 283
Dominio de aminoacilación
658
593 597
Dominio de unión a anticodón
806
844 897
900
Dominio WHEP
552
900
8 aa's + 586
MetRS1C3
MetRS1C4
846 900
MetRS1C5

FIG. 2B

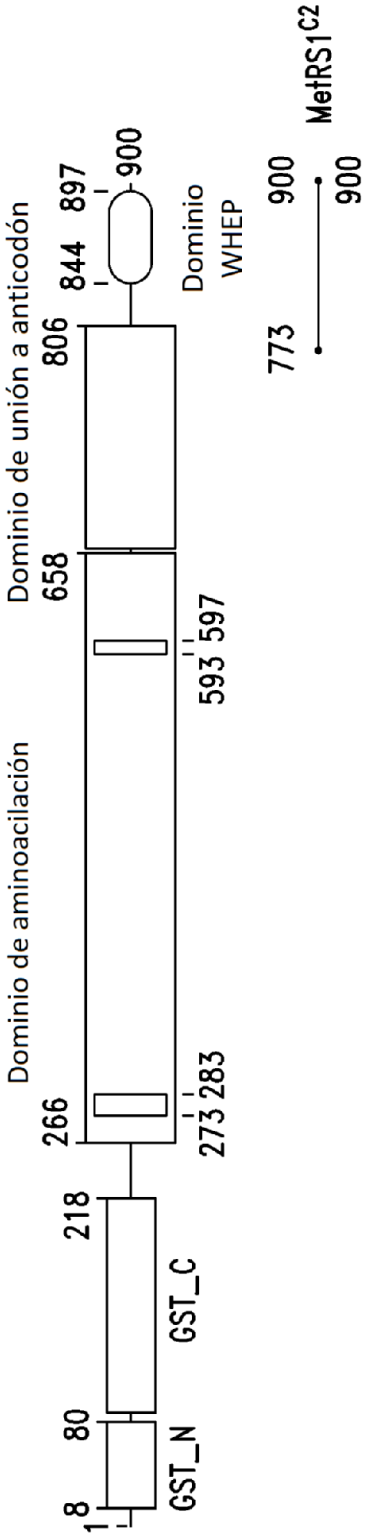
MetRS1C2
773
900
900
Dominio
WHEP
Dominio de unión a anticodón
806
844
897
900
658
593
597
Dominio de aminoacilación
266
273
283
218
1
8
80
GST_N
GST_C

FIG. 2C