

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 602 602**

51 Int. Cl.:

C07K 14/81 (2006.01)

C12N 15/10 (2006.01)

C12N 15/62 (2006.01)

G01N 33/53 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.04.2009 PCT/GB2009/050380**

87 Fecha y número de publicación internacional: **12.11.2009 WO09136182**

96 Fecha de presentación y número de la solicitud europea: **16.04.2009 E 09742378 (4)**

97 Fecha y número de publicación de la concesión europea: **17.08.2016 EP 2279205**

54 Título: **Proteínas de andamiaje de Estefina A modificada**

30 Prioridad:

18.04.2008 GB 0807065

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
21.02.2017

73 Titular/es:

**AVACTA LIFE SCIENCES LIMITED (100.0%)
Unit 651, Thorp Arch Estate
Wetherby LS23 7FZ, GB**

72 Inventor/es:

**KO FERRIGNO, PAUL y
GENDRA, ELISENDA**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 602 602 T3

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

DESCRIPCIÓN

Proteínas de andamiaje de Estefina A modificada

5 La presente invención se refiere a novedosas proteínas de andamiaje ("scaffold protein") para la presentación de péptidos, tales como aptámeros peptídicos. En particular, la invención se refiere al uso de polipéptidos de Estefina A modificados y proteínas artificiales modificadas basadas en Estefina A, todos para su uso como proteínas de andamiaje y como sistemas de presentación.

Antecedentes

10 El estudio de las interacciones de proteína es vital para un entendimiento de muchos procesos biológicos, tal como los papeles de los productos génicos *in vivo* tanto en salud como en enfermedad. Los aptámeros peptídicos en particular han surgido como herramientas moleculares importantes que son útiles para tanto los aspectos básicos como los aplicados de la medicina molecular. Debido a su capacidad de unirse específicamente a, e inactivar, una proteína diana dada a nivel intracelular, proporcionan una estrategia experimental para los análisis de proteína funcional, tanto *in vitro* como *in vivo*. También se pueden usar frente a proteínas extracelulares. Así como aplicaciones en el estudio de la función proteica, por lo tanto, estas herramientas pueden ser útiles para la detección molecular, diagnósticos y/o como agentes terapéuticos. Los péptidos y aptámeros peptídicos se pueden usar libres en solución. Sin embargo, péptidos pequeños cuando están libres tenderán a formar estructuras que exponen una superficie de interacción limitada. Además, con frecuencia perderán entropía conformacional tras la asociación con moléculas diana, reduciendo la energía libre de unión y, por consiguiente, los péptidos libres con frecuencia no formarán complejos no covalentes apretados, lo cual es un problema. Además, dentro de las células los péptidos se degradan rápidamente, lo cual limita su eficacia para el estudio de interacciones proteicas *in vivo*, lo cual también es un problema.

25 En lugar de usarse en disoluciones libres, los péptidos de interés se pueden unir a soportes físicos, o presentarse en el contexto de un polipéptido más largo. Lo primero no se puede aplicar fácilmente a estudios *in vivo*. En lo último, los péptidos se insertan genéticamente en la secuencia primaria de una proteína de andamiaje estable simple. El plegamiento del andamio ("scaffold") restringe conformacionalmente el péptido, así los aptámeros peptídicos se unen a parejas con alta especificidad y afinidad. Se presenta en el contexto de un polipéptido que es importante en la presente invención. Tal presentación con frecuencia se ocasiona usando proteínas de andamiaje.

30 Los andamios de la técnica anterior han incluido nucleasa estafilocócica inactivada, proteína fluorescente verde (PFV) y tioredoxina A (TrxA), así como pliegues proteicos aislados tales como el dominio Z de la proteína A estafilocócica, "aficuerpos", anticalinas, y repeticiones de anquirina.

Proteínas de andamiaje de la técnica anterior adicionales incluyen el dominio tipo III de fibronectina ('Fn3'), proteínas de la familia de la lipocalina a partir de las que se derivan las anticalinas, proteína de unión a bilina (BBP, del inglés "Bilin Binding protein"), y otras.

35 Más recientemente (el documento WO 2006/131749) se describen varias mutaciones racionales hechas en Estefina A para mejorarla como andamio. El andamio de Estefina A modificada comprende mutaciones en los siguientes tres sitios Lys71-Leu73, V48D y G4W y se refiere como STM (del inglés "Stefin A Triple Mutant"). Se demostró que la combinación de estas tres mutaciones generó una proteína que tenía interacciones mínimas con las proteínas en células humanas, y en particular había perdido toda interacción detectable con sus parejas naturales conocidas. Sin embargo, encontramos que la intersección de los péptidos en la proteína en la posición 71 a 73 conducía a una fuerte presión de selección para truncamientos de la proteína en el extremo del péptido insertado. Aunque tales proteínas truncadas podrían presentar eficacia biológica, esta observación conduce a preocuparse de que un subconjunto de péptidos que simplemente están insertados en la posición 71 a 73 sin truncamiento no puede estar libremente disponible para la interacción con una proteína diana, lo cual es un problema. Además, la inserción de los péptidos de un único sitio inevitablemente limita el área superficial total usado para una interacción de proteína, lo cual poco a poco limita la afinidad de unión y potencialmente la especificidad.

45 Las novedosas mutaciones hechas a Estefina A y a proteínas artificiales modificadas basadas en Estefina A tal como STM (Mutante Triple de Estefina A) como se describe en la presente invención proporcionan alternativas proteínas de andamiaje mejoradas y más estables y también proporcionan sistemas de presentación que son más versátiles que los de la técnica anterior. Además, estos nuevos andamios proteicos/sistemas de presentación también son bastantes impredecibles como entidades de presentación eficaces y sólidas. Las nuevas mutaciones descritas más adelante se han hecho en diversas áreas específicas de las proteínas Estefina A/STM y sorprendentemente se ha encontrado que no afectan a la configuración de la proteína Estefina A/STM o su función potencial como proteínas de andamiaje. Además, con los andamios mejorados de la presente invención debido a la modificación por ingeniería adicional es posible proporcionar modificaciones en las que los andamios tienen inserciones múltiples algo que no era hasta ahora posible en los andamios de la técnica anterior.

Breve compendio de la descripción

Según el primer aspecto de la invención se ha proporcionado un polipéptido de Estefina A que comprende una secuencia de aminoácidos que tiene al menos un 80% de identidad con SEQ ID NO: 1; comprendiendo dicho polipéptido una mutación en la posición 4 en donde la Glicina de la Estefina A está reemplazada por Arginina;

- 5 caracterizado por que el polipéptido comprende además una inserción de péptido heterólogo, en donde
- a) dicho péptido heterólogo está insertado en la proteína en el sitio G4 de Estefina A; o
 - b) dicho péptido heterólogo está insertado en la proteína en la posición 46 a 54; o
 - c) dicho péptido heterólogo está insertado en la proteína en la posición 67 a 84; o
 - d) dicho péptido heterólogo está insertado en la proteína en la posición 82/83.

10 Adecuadamente dicho polipéptido de Estefina A comprende una inserción de péptido heterólogo adicional de modo que al menos un péptido heterólogo está insertado en la proteína en al menos dos posiciones seleccionadas entre (a), (b) y (c).

15 Adecuadamente dicho polipéptido de Estefina A comprende una inserción de péptido heterólogo adicional de modo que al menos un péptido heterólogo está insertado en la proteína en al menos tres posiciones seleccionadas entre (a), (b) y (c).

En otro aspecto, la invención se refiere a un polipéptido de Estefina A que comprende una secuencia de aminoácidos que tiene al menos un 80% de identidad con SEQ ID NO: 1; comprendiendo dicho polipéptido una mutación en la posición 4 en donde la Glicina de la Estefina A está reemplazada por Arginina;

caracterizado por que el polipéptido comprende además una inserción de péptido heterólogo, en donde

- 20
- i) dicho péptido heterólogo está insertado en la proteína en la posición 46 a 54; o
 - ii) dicho péptido heterólogo está insertado en la proteína en la posición 67 a 84; o
 - iii) dicho péptido heterólogo está insertado en la proteína en la posición 82/83.

25 Adecuadamente dicho polipéptido de Estefina A comprende una inserción de péptido heterólogo adicional de modo que al menos un péptido heterólogo está insertado en la proteína en al menos dos posiciones seleccionadas entre (i) y (ii).

En otro aspecto, la invención se refiere a un polipéptido que comprende la secuencia de aminoácidos como se expone en una cualquiera de las SEQ ID NO: 10, 13, 16, 17, 22, 23, 24 y 25.

Adecuadamente dicho péptido heterólogo de (a) está insertado próximo al sitio G4 de la Estefina A.

30 En otro aspecto, la invención se refiere al uso de un polipéptido como se describió anteriormente como proteína de andamiaje.

En otro aspecto, la invención se refiere a un método para identificar un péptido diana capaz de unirse a una estructura de interés que comprende:

- 35
- (i) proporcionar una proteína de andamiaje de estefina A como se describió anteriormente que comprende un péptido diana;
 - (ii) poner en contacto dicha proteína de andamiaje con dicha estructura de interés; y
 - (iii) hacer un seguimiento de la asociación entre el andamio y la estructura de interés, en donde la asociación de la proteína de andamiaje con la estructura de interés identifica el péptido diana como un péptido diana candidato capaz de unirse a dicha estructura.

40 En otro aspecto, la invención se refiere al uso de un polipéptido como se describió anteriormente para presentar péptido(s) sobre micromatriz(ces).

En otro aspecto, la invención se refiere al uso de un polipéptido como se describió anteriormente como agente para unir y detectar específicamente un biomarcador.

En otro aspecto, la invención se refiere al uso de un polipéptido como se describió anteriormente en un ensayo de interacción, en donde dicho polipéptido está inmovilizado.

45 En otro aspecto, la invención se refiere a un método para detectar una estructura de interés que comprende:

proporcionar una proteína de andamiaje de estefina A como se describió anteriormente que comprende un péptido diana capaz de unirse a dicha estructura de interés;

poner en contacto dicha proteína de andamiaje con dicha estructura de interés; y

5 hacer un seguimiento de la asociación entre el andamio y la estructura de interés, en donde la asociación de la proteína de andamiaje con la estructura de interés detecta dicha estructura.

Preferiblemente dicha estructura de interés comprende un biomarcador.

A lo largo de la descripción y las reivindicaciones de esta memoria, las palabras “comprender” y “contener” y variaciones de las palabras, por ejemplo, “que comprende” y “comprende”, significa “que incluye, pero no se limita a”, y esto no intenta (y no lo hace) excluir otros restos, aditivos, componentes, números enteros o etapas.

10 A lo largo de la descripción y las reivindicaciones de esta memoria, el singular abarca el plural a menos que el contexto indique lo contrario. En particular, cuando se usa el artículo indefinido, la especificación es para entenderse que contempla la pluralidad, así como la singularidad, a menos que el contexto indique lo contrario.

15 Rasgos, números enteros, características, compuestos, restos químicos o grupos descritos junto con un aspecto particular, realización o ejemplo de la invención son para entender que son aplicables a cualquier otro aspecto, realización o ejemplo descrito en la presente memoria a menos que sean incompatibles con los mismos.

20 La referencia en la presente memoria a una proteína de andamiaje es a secuencias fusionadas juntas en una proteína, el término también es sinónimo de proteína de fusión. Por una “proteína de fusión” se supone una proteína que incluye la proteína de andamiaje de la invención unida a uno o más péptidos o proteínas diferentes (es decir, “heterólogos”). La inserción de los péptidos o proteínas heterólogos posibilita que la proteína de fusión se una a una diana deseada.

25 La presente invención se basa en modificaciones novedosas que incluyen inserciones de la propia proteína Estefina A tipo natural, preferiblemente la Estefina A es una Estefina A humana, o su versión triple mutante, STM, que las vuelve formas adecuadas para su uso como proteínas de andamiaje estables mientras que de manera simultánea ventajosamente las vuelve biológicamente neutras al extirpar interacciones y actividades biológicamente significantes mediante mutación de residuos que se requieren para las interacciones naturales con o bien catepsinas u otras proteínas desconocidas.

30 Además, se prevé que el(los) sitio(s) de mutación o inserción seleccionado(s) es(son) capaz(es) de aceptar y restringir péptidos insertados para producir, por ejemplo, aptámeros peptídicos. Mientras que los estudios en humanos pueden requerir un andamio humano, el uso de por ejemplo Estefina A de ratón, puede ser ventajoso para los estudios de la biología y/o enfermedad en modelo de ratón, similarmente la Estefina A derivada de otras especies o plantas también puede ser de utilidad en esa especie específica. Por consiguiente, los andamios y los sistemas de presentación de la presente invención intentan ser útiles para cualquier especie seleccionada y la derivación de la Estefina A es dependiente de un requerimiento del usuario.

35 Se apreciará que los cambios en las secuencias de ADN que codifican los aminoácidos en el codón 4 de o bien Estefina A o su forma STM, o los cambios en los codones 46 a 54 incluidos, que codifican los aminoácidos que comprenden o restringen el bucle (“loop”) 1 de o bien la Estefina A o su forma STM, o los cambios en los codones 67 a 84 incluidos, que codifican los aminoácidos que comprenden o restringen el bucle 2 de o bien Estefina A o su forma STM, pueden ser independientes uno de otro. Es decir, que las modificaciones a la proteína Estefina A pueden ser en una de tres áreas o regiones separadas diferentes, es decir, en la posición 4 o en el bucle 1 o bucle 2 de restricción. Similarmente, las modificaciones a la forma triple mutante STM también puede ser en uno cualquiera de los tres sitios separados independientes especificados, es decir, en la posición 4 o en el bucle 1 o bucle 2 de restricción. El resto de la secuencia de Estefina A o STM estará inalterada y comprenderá las secuencias como las se exponen a continuación.

La secuencia de Estefina A humana de tipo natural se muestra a continuación como SEQ ID NO: 1:

45 MIPGGLSEAKPATPEIQEIVDKVKPQLEEKTNETYGKLEAVQYKTQVVAGTNYIYKVR
GDNKYMHLKVF~~KSL~~PGQNE~~DLV~~LTGYQVDKNK~~DEL~~TGF

La secuencia del triple mutante STM se muestra a continuación como SEQ ID NO: 2, los sitios de mutación y, por tanto, donde varía STM de la Estefina A de tipo natural están marcados en negrita y subrayados:

MIP**W**GGLSEAKPATPEIQEIVDKVKPQLEEKTNETYGKLEAVQYKTQV**D**AGTNYIYKVR
GDNKYMHLKVF**NGPP**PGQNE~~DLV~~LTGYQVDKNK~~DEL~~TGF

La referencia en la presente memoria a un “cambio mutacional” expresa que hay un cambio permanente en el material genético, el cambio mutacional puede ser por adición(adiciones) o deleciones o inserción(inserciones) o sustitución(sustituciones) al(a los) residuo(s) de aminoácidos.

5 Preferiblemente, en la realización donde el cambio mutacional sencillo es Glicina en el codón 4 de Estefina A, su sustitución se selecciona entre el grupo que consiste en G4V, G4I, G4L, G4M, G4F, G4P, G4N, G4V, G4Q, G4S, G4T, G4W, G4Y, G4R, G4H, G4K, G4D y G4E. Más preferiblemente el cambio es G4R, es decir que la Glicina está sustituida por Arginina en el codón 4.

10 Preferiblemente, en la realización donde el cambio mutacional sencillo es Triptófano en el codón 4 de STM, su sustitución se selecciona entre el grupo que comprende W4V, W4I, W4L, W4M, W4F, W4P, W4N, W4V, W4Q, W4S, W4T, W4G, W4Y, W4R, W4H, W4K, W4D y W4E. Más preferiblemente el cambio es W4R, es decir que el Triptófano está sustituido por Arginina en la posición 4.

15 Se ha encontrado que un cambio en la región 5' del marco de lectura abierto de Estefina A que codifica los 8 aminoácidos del amino-terminal de Estefina A o STM permite la introducción de un sitio de fragmentación para una endonucleasa de restricción o un sitio de recombinación guiado. Por ejemplo, se demuestra en la presente memoria que después de cambiar la secuencia de ADN para codificar, por ejemplo, una Arginina en la posición 4 (sustituir la Glicina tipo natural, o el triptófano en STM) permite la producción de una proteína sorprendentemente estable que posee las mismas características biofísicas que STM, pero donde el marco de lectura abierto ahora posee un único sitio de restricción para una enzima tal como, y sin limitación, la enzima AvrII, y así sirve una proteína de andamiaje alternativa y eficaz.

20 Preferiblemente, en la realización donde el cambio mutacional es cualquier cambio en los codones 46 a 54 incluidos que codifican los aminoácidos que comprenden o restringen el bucle 1 de la Estefina A (SEQ ID NO: 3 **QVVAGTNY**) o STM (SEQ ID NO: 4: **QVDAGTNY**) el cambio comprende, por ejemplo, **QVLASTNY** (SEQ ID NO: 5). Sorprendentemente se ha demostrado que introducir una secuencia de aminoácidos tales como, y sin limitación, Leucina, Alanina, Serina en las posiciones 48, 49 y 50, conduce a una proteína con las mismas características biofísicas que STM, y así es probablemente que sea un andamio eficaz.

25 Preferiblemente, el cambio mutacional en Estefina A es en 48-VAG-50 y en STM es en 48-LAS-50 de modo que el resultado es 48-LXS-50, en donde X es cualquier aminoácido.

30 Preferiblemente, en la realización donde el cambio mutacional es cualquier cambio en los codones 67 a 84 incluidos que codifican aminoácidos que comprenden o restringen el bucle 2 de Estefina A (SEQ ID NO: 6: **LKVFKSLPGQNE DLVLTG**) o STM (SEQ ID NO: 7: **LKVFNGPPGQNE DLVLTG**) el cambio comprende, por ejemplo, SEQ ID NO: 8: **LKVFNGPPGQNE DLVRS**. Sorprendentemente se ha demostrado que una secuencia de aminoácidos tal como Arginina seguido de una Serina (sustituir Leucina 82 y treonina 83 de Estefina A o STM) conduce a la producción de una proteína estable que, como STM, puede servir como un buen andamio para la exposición de aptámeros peptídicos.

35 Preferiblemente, el cambio mutacional en Estefina A es en 71-KSL-73 y 82-LT-83 y en STM es en 71-NPG-73 y 82-LT-83 de modo que el resultado es 71-NxP-73 y 82-RS-83, en donde X es cualquier aminoácido.

Preferiblemente, hay un cambio mutacional adicional en Estefina A y STM que está en 82-LT-83 de modo que el resultado es 82-XX-83, en donde X es cualquier aminoácido y en una realización particularmente preferida es 82-RS-83. El cambio mutacional puede ser en o bien 82 o 83 o en ambas posiciones.

40 Preferiblemente, el cambio mutacional puede ser cualquier combinación de las anteriormente mencionadas en la presente memoria con por ejemplo, y sin limitación, 82-XX-83 y una variante preferida particular tiene cambios mutacionales al menos en las posiciones 71 a 73 y/o 82-83.

45 En la presente memoria se describe un polipéptido de Estefina A modificado o proteína STM modificada que comprende dos cambios mutacionales o un oligonucleótido heterólogo que codifica una inserción de péptido en sitios seleccionados entre el grupo que comprende:

- (i) una mutación en el codón 4 en donde la Glicina de Estefina A o el Triptófano de STM está sustituido por otro aminoácido que no es Triptófano para Estefina A o Glicina para STM o por un péptido codificado por el oligonucleótido heterólogo; y/o
- (ii) cualquier cambio u oligonucleótido heterólogo que codifica una inserción de péptido en los codones 46 a 54 incluidos, que codifican los aminoácidos que comprenden o restringen el bucle 1; y/o
- (iii) cualquier cambio u oligonucleótido heterólogo que codifica una inserción de péptido en los codones 67 a 84 incluidos, que codifican aminoácidos que comprenden o restringen el bucle 2.

Se apreciará que en este aspecto, donde la Estefina A modificada puede comprender dos mutaciones, puede comprender, por ejemplo, una mutación en la posición 4 y un cambio en cualquiera de los codones 46 a 54 que tienen función de bucle 1 o pueden comprender una mutación en la posición 4 y un cambio en cualquiera de los codones 67 a 84 que tienen función de bucle 2 o puede comprender un cambio en cualquiera de los codones 46 a 54 que tienen función de bucle 1 y cambio en cualquiera de los codones 67 a 84 que tienen función de bucle 2.

Similarmente, el STM puede comprender una mutación en la posición 4 y un cambio en cualquiera de los codones 46 a 54 que tienen función de bucle 1 o puede comprender una mutación en la posición 4 y un cambio en cualquiera de los codones 67 a 84 que tienen la función de bucle 2 o puede comprender un cambio en cualquiera de los codones 46 a 54 que tienen función de bucle 1 y cambio en cualquiera de los codones 67 a 84 que tienen función de bucle 2 o puede comprender un cambio en cualquiera de los codones 67 a 84 que tienen función de bucle 2.

En la presente memoria también se describe un polipéptido de Estefina A modificado o proteína STM modificada en donde los cambios mutacionales de modificación o un oligonucleótido heterólogo que codifica un péptido se insertan en tres sitios:

- (i) una mutación en el codón 4 en donde la Glicina de Estefina A o el Triptófano de STM está sustituido por otro aminoácido que no es Triptófano para Estefina A o Glicina para STM o por un péptido codificado por el oligonucleótido heterólogo; y
- (ii) cualquier cambio u oligonucleótido heterólogo que codifica una inserción de péptido en los codones 46 a 54 incluidos, que codifican aminoácidos que comprenden o restringen el bucle 1; y
- (iii) cualquier cambio u oligonucleótido heterólogo que codifica una inserción de péptido en los codones 67 a 84 incluidos, que codifican los aminoácidos que comprenden o restringen el bucle 2.

Por tanto, en este aspecto la Estefina A y STM modificados comprenden todos los tres cambios mutacionales como se describieron anteriormente. Por tanto, las proteínas de andamiaje Estefina A o STM modificadas comprenden tres mutaciones específicas en la posición 4 y un cambio en tanto el bucle 1 como bucle 2.

En la presente memoria también se describe un polipéptido de Estefina A modificado o proteína de STM modificada que comprende una cualquiera sola o cualquier combinación de las secuencias anteriormente enumeradas, pero que terminan en o bien el residuo 73 de Estefina A o STM, o el residuo 84 de Estefina A o STM, y o bien con o sin la inserción de una nueva secuencia de aminoácidos en estas posiciones. Se ha encontrado que hay una fuerte presión de selección para un codón de terminación seguido de inserción de secuencias de aminoácidos después de NGP de STM, pero que sorprendentemente tales proteínas truncadas son tanto estables como capaces de interferir con las funciones biológicas de una proteína diana.

Por lo tanto, la presente descripción también incluye proteínas de andamiaje de Estefina A y STM modificadas truncadas o acortadas idealmente acortadas en 15 o 25 residuos en el extremo C-terminal y terminando así en o bien los residuos 73 o 84 de o bien Estefina A o STM. También se describen Estefina A o STM modificadas, truncadas o acortadas que se acortan en cualquier número entero entre 15 y 25 y así terminan en residuos entre 73 a 84 de Estefina A o STM.

La presente descripción preferiblemente incluye todas las variantes descritas, ya que cada una permite la introducción de un péptido heterólogo en uno o más sitios de la Estefina A o variante STM, por inserción de un oligonucleótido en un sitio de restricción modificado por ingeniería en los marcos de lectura abiertos que hemos creado. Por tanto, la presente descripción proporciona diversos nuevos andamios a base de:

Un único péptido heterólogo insertado en la proteína en la posición 4, con el resto de la proteína que se parece o bien a Estefina A o STM o una de las otras variantes descritas en la presente memoria.

Un único péptido heterólogo insertado en la proteína en la posición 46 a 54 y especialmente en las posiciones 48/49/50, con el resto de la proteína que se parece o bien a Estefina A o STM o una de las otras variantes descritas en la presente memoria.

Un único péptido heterólogo insertado en la proteína en la posición 67 a 84 y especialmente en la posición 71/72/73 con el resto de la proteína que se parece o bien a Estefina A o STM o una de las otras variantes descritas en la presente memoria.

Un único péptido heterólogo insertado en la proteína en la posición 67 a 84 y especialmente en la posición 82/83 con el resto de la proteína que se parece o bien a Estefina A o STM o una de las otras variantes descritas en la presente memoria.

Cualquier combinación de péptidos múltiples insertados en las posiciones 4 y/o 48/49/50 y/o 71/72/73 y/o 82/83.

Cualquier combinación de péptidos simples o múltiples insertados en las posiciones 4 y/o 48/49/50 y/o 71/72/73 y/o 82/83 que se sigue por un codón de terminación que elimina o sustituye los últimos 25 o los últimos 15 residuos de aminoácidos de o bien Estefina A o STM.

5 Una ventaja particular de las proteínas de andamiaje y las nuevas mutaciones es que posibilitan el uso de todo el bucle 1 o bucle 2, o tanto bucle 1 como bucle 2 así como el terminal amino. Juntas, estas mutaciones permitirán la exposición de superficies al menos tan grandes como las usadas por los anticuerpos. Además, cada una se puede usar individualmente, o se pueden usar por pares o en combinaciones múltiples con otras mutaciones. Las posiciones que difieren de las superficies de interacción, combinadas con las interacciones que difieren entre los péptidos insertados en los diferentes sitios, es probable que proporcionen usos novedosos del andamio, tal como
10 cuando los péptidos que podrían no estar expuestos para la interacción útil en un sitio ahora pueden ser expuestos por otro, o cuando las combinaciones de péptidos en diferentes sitios permiten que un péptido dado cambie de una conformación no interactiva a una conformación interactiva. Además, cualquiera de estas nuevas mutaciones se pueden usar en el contexto de Estefina A de longitud completa, STM de longitud completa, las proteínas de longitud completa descritas en la presente memoria, o en las versiones mutantes de cualquiera de estas proteínas donde el
15 último residuo derivado de o bien Estefina A o STM es Leu73 de SteA o sus nuevas variantes descritas en la presente memoria, o Pro73 de STM o sus nuevas variantes, o el último residuo de un péptido heterólogo insertado cuando se han trucado los últimos 15 o 25 aminoácidos de Estefina A o STM.

20 En la presente memoria también se describen ácidos nucleicos aislados que comprenden secuencias de nucleótidos que codifican las secuencias de aminoácidos de una proteína de andamiaje o polipéptidos como anteriormente se ha descrito.

25 En la presente memoria también se describe un método para identificar un péptido diana capaz de unirse a una estructura de interés que comprende proporcionar una proteína de andamiaje de proteína Estefina A modificada o STM como se describió anteriormente en la presente memoria que comprende un péptido diana; poner en contacto dicha proteína de andamiaje con dicha estructura de interés, y hacer un seguimiento de la asociación entre el andamio y la estructura de interés, en donde la asociación de la proteína de andamiaje con la estructura de interés identifica el péptido diana como péptido diana candidato capaz de unirse a dicha estructura.

En un aspecto adicional la proteína de andamiaje se selecciona entre el grupo que comprende:

SEQ ID NO: 9 (SDM): MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYKVRAG DNKYMHLKVF NGPPGQNE DL VRSGYQVDKN KDELDTGF*

SEQ ID NO:10 (SQM): MIPRGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYKVRAG DNKYMHLKVF NGPPGQNE DL VRSGYQVDKN KDELDTGF*

SEQ ID NO:11 (SUC): MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVVAG TNYIYKVRAG DNKYMHLKVF NGPPGQNE DL VRSGYQVDKN KDELDTGF*

SEQ ID NO:12 (SUM): MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYIKVRAG DNKYMHLKVF KSLPGQNEEDL VLTGYQVDKN KDELDTGF*

SEQ ID NO:13 (SUN): MIPRGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVVAG TNYIYIKVRAG DNKYMHLKVF KSLPGQNEEDL VLTGYQVDKN KDELDTGF*

SEQ ID NO: 14 (SDM-): MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYIKVRAG DNKYMHLKVF NGPPGQNEEDL VRS*

SEQ ID NO: 15 (SDM--): MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYIKVRAG DNKYMHLKVF NGP*

SEQ ID NO:16 (SQM-): MIPRGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYIKVRAG DNKYMHLKVF NGPPGQNEEDL VRS*

SEQ ID NO:17 (SQM--): MIPRGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYIKVRAG DNKYMHLKVF NGP*

SEQ ID NO:18 (SUC-):MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVVAG TNYIYIKVRAG DNKYMHLKVF NGPPGQNEEDL VRS*

SEQ ID NO:19 (SUC--):MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVVAG TNYIYIKVRAG DNKYMHLKVF NGP*

SEQ ID NO:20 (SUM-): MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYIKVRAG DNKYMHLKVF KSLPGQNEEDL VLT*

SEQ ID NO:21 (SUM--): MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYIKVRAG DNKYMHLKVF KSL*

SEQ ID NO:22 (SUN-): MIPRGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVVAG TNYIYIKVRAG DNKYMHLKVF KSLPGQNEEDL VLT*

SEQ ID NO:23 (SUN--): MIPRGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVVAG TNYIYIKVRAG DNKYMHLKVF KSL*

SEQ ID NO:24 (SQT): MIPRGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLAS TNYIYIKVRAG DNKYMHLKVF NGPPGQNADR VLTGYQVDKN KDELDTGF*

SEQ ID NO:25 (SQL): MIPRGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA
VQYKTQVLALAS TNYIYIKVRAG DNKYMHLKVF NGPPGQNADR VLTGYQVDKN KDELDTGF*

5 En un aspecto adicional se ha proporcionado el uso de las proteínas de andamiaje de la presente invención como agente seleccionado entre el grupo que comprende compuestos de diagnóstico, compuestos terapéuticos, biomarcadores, agentes para unirse a y específicamente detectar biomarcadores, plantillas de diseño de fármaco racionalizado, dianas o reactivos para el descubrimiento de fármaco, sustitutos de anticuerpo y herramientas de investigación.

En un otro aspecto adicional se ha proporcionado el uso de las proteínas de andamiaje de la presente invención como proteína de fusión.

10 Los rasgos preferidos descritos anteriormente en la presente memoria se aplican *mutatis mutandis* a todos y cada uno de los aspectos.

15 Tomados en conjunto, los resultados descritos más adelante y anteriormente demuestran que, los andamios descritos en la presente memoria están dispuestos a modificarse por ingeniería en localizaciones múltiples, con cada cambio solo o en combinación que es sorprendentemente bien tolerado y que algún efecto de desestabilización de las mutaciones aparentemente se magnifica por inserciones en el terminal amino y en el bucle 2. Por lo tanto, estos sitios no se pueden usar de manera rutinaria, pero sorprendentemente algunos insertos son tolerados lo que nos permitirá usarlos para mejora la afinidad y la especificidad de unión de los aptámeros peptídicos en, por ejemplo, el SQM-bucle 1. Además, el bucle 1 parece capaz de exponer un intervalo de péptidos con muy poco problema. Esto es muy sorprendente, ya que es el bucle más corto.

Descripción detallada de la invención

Una "delección" se refiere a un cambio en una secuencia de aminoácidos o nucleótidos debido a la ausencia de uno o más residuos de aminoácidos o nucleótidos. Los términos "inserción" o "adición" se refieren a cambios en una secuencia de aminoácidos o nucleótidos que dan como resultado la adición de uno o más residuos de aminoácidos, respectivamente, a una molécula o representación de la misma, en comparación con una secuencia de referencia, por ejemplo, la secuencia encontrada en la molécula que se da de manera natural. Una "sustitución" se refiere a la sustitución de uno o más aminoácidos o nucleótidos por diferentes aminoácidos o nucleótidos, respectivamente.

Para mejorar Estefina A o STM como andamio, es deseable que sea capaz de insertar péptidos heterólogos en sitios alternativos, y/o en sitios múltiples. Para hacer esto se requiere alterar el marco de lectura abierto que codifica o bien Estefina A o STM, para introducir los sitios de reconocimiento de endonucleasa de restricción en los que se podría insertar oligonucleótidos que codifican péptidos heterólogos. La alteración del marco de lectura abierto casi conduce inevitablemente a una alteración de la secuencia de aminoácidos que comprende la proteína expresada. Dado que las proteínas han desarrollado una combinación óptima de función y estabilidad, el resultado más probable (y más frecuentemente observado) de un cambio a la secuencia de aminoácidos de una proteína es una pérdida de estructura secundaria y de ahí de estabilidad. En la presente invención las nuevas proteínas de andamiaje retuvieron la estabilidad (Véase los Ejemplos y las Figuras).

Para establecer si las alteraciones hechas en el nivel de ADN (marco de lectura abierto) que también alteran la secuencia de aminoácidos de las nuevas variantes de Estefina A/STM conducen a estabilidad disminuida de la proteína, todas las variantes descritas en la presente memoria se expresaron en *E. coli* y se usó dicroísmo circular para comparar su composición de estructura secundaria con la de Estefina A. Se encontró que todas las proteínas se expresan igualmente bien en *E. coli*, generalmente a aproximadamente 28 mg de proteína variante/ml de cultivo bacteriano (Figura 1). Las proteínas se purificaron cerca de homogeneidad por cromatografía de afinidad usando Ni-agarosa, y las preparaciones purificadas diluidas a 0,3 mg de proteína/ml. Cada muestra resultante se sometió a análisis por dicroísmo circular. Esto implica hacer un barrido de la proteína a través de un intervalo de longitudes de ondas cerca de UV, de modo que la elipticidad de la luz está afectada por elementos de la estructura secundaria (hélice alfa o hebras beta) de la proteína. A mayor proporción de la estructura secundaria, mayor es el efecto de la elipticidad. Debido a que el efecto está afectado por la concentración de proteína, las proteínas se diluyeron a 0,3 mg/ml inmediatamente antes del análisis. Debido a que el efecto es proporcional al número de enlaces peptídicos en la proteína a analizar, se muestra la elipticidad molar, la cual tiene en cuenta este efecto. Los datos se muestran en la Figura 3. Estos datos muestran que la proporción de la estructura secundaria se conserva entre STM y las nuevas variantes, y que la presencia de los insertos en STM no afecta de manera adversa a su estructura. Se indicó que dos variantes (SUN y SQM) parecen mostrar estructura aumentada en comparación con los otros. Esto puede ser atribuible a la adquisición de la estructura secundaria en la cola amino-terminal que está presente en todas estas proteínas, y se conducirían por la sustitución de Glicina (Estefina A) o Triptófano (STM) en la posición 4 por Arginina, ya que esto es el único cambio que es común a SUN y SQM, y estos son las únicas variantes que poseen esta alteración.

Andamio

Como es bien conocido en la técnica, el término "andamio" se refiere a una proteína que puede exponer péptidos diana a disolvente sin que su propia estructura se deforme por el péptido diana. Con respecto a la exposición de péptido a disolvente, esto se puede ensayar usando experimentos de inmunoprecipitación. Por ejemplo, un indicio de que un péptido está siendo expuesto al disolvente se puede obtener por su disponibilidad a un anticuerpo capaz de reconocerlo. Por tanto, para ensayar la capacidad de una proteína de andamiaje de exponer un péptido a disolvente, se expresará el andamio que comprende el péptido y se usará un anticuerpo que reconoce el péptido para tratar de inmunoprecipitar la fusión andamio-péptido. Si esta proteína se puede inmunoprecipitar o capturar sobre el anticuerpo, esto demuestra que el péptido estaba expuesto al disolvente como es requerido por una proteína de andamiaje. Otro, o alternativo, indicio de que un péptido está siendo expuesto al disolvente se puede obtener por estudios de fosforilación. Al incorporar un sitio aceptor de fosfato dentro del péptido diana y, a continuación, poner en contacto la fusión andamio-péptido con la quinasa cognada en condiciones permisivas de fosforilación, entonces se puede verificar la exposición del péptido a disolvente. La fosforilación del péptido indica correcta exposición a disolvente. Con respecto a la resistencia de la proteína de andamiaje a ser deformada por el péptido diana que porta, esto se puede ensayar usando técnicas tales como dicroísmo circular o estabilidad térmica. Específicamente, un análisis de dicroísmo circular de una proteína de andamiaje sin péptido diana insertado en él debería ser básicamente el mismo que las características del dicroísmo circular de la misma proteína de andamiaje cuando porta un péptido diana. Esto proporciona una demostración de que la presencia del péptido diana en la proteína de andamiaje no ha comprometido o deformado la estructura de la proteína de andamiaje que lo porta. Otro modo de ensayar esta resistencia a la deformación por el péptido diana es estudiando la estabilidad térmica de la proteína de andamiaje con y sin péptido diana insertado.

Una proteína de andamiaje debe ser capaz de aceptar un inserto peptídico. Preferiblemente el inserto peptídico es de 36 aminoácidos o menos, preferiblemente de 20 aminoácidos o menos. Preferiblemente el inserto peptídico diana es de 12 aminoácidos o menos.

Una proteína de andamiaje debe ser de estructura conocida. Por “estructura conocida” se supone que la estructura de cristal o una estructura de solución (estructura de NMR) debe ser conocida.

Rasgos preferidos de proteínas de andamiaje según la presente invención

5 Preferiblemente una proteína de andamiaje restringe el péptido diana. La presencia de un efecto de restricción en una proteína de andamiaje se puede demostrar comparando la afinidad de una entidad que se une al péptido diana cuando el péptido diana está en la proteína de andamiaje con la afinidad cuando el péptido no está en la proteína de andamiaje. Una diferencia en estas dos afinidades indica que la proteína de andamiaje está restringiendo el péptido para asumir una conformación tridimensional particular. Preferiblemente una proteína de andamiaje restringe un péptido de manera que demuestra una afinidad de unión aumentada cuando está presente en el contexto de la proteína de andamiaje. En otras palabras, preferiblemente la proteína de andamiaje disminuye el coste entrópico de la unión y así aumenta la afinidad medida cuando se compara con la unión de un péptido libre.

En algunas realizaciones, la restricción puede ser proporcionada por una fusión sencilla de N-terminal o C-terminal al péptido diana.

15 Preferiblemente una proteína de andamiaje proporciona el péptido diana con una estabilidad aumentada *in vivo*. Este efecto se puede demostrar por comparación de la expresión del péptido diana en el contexto de la proteína de andamiaje con la expresión del péptido diana sobre sí mismo. Preferiblemente, el péptido diana muestra estabilidad aumentada en el contexto de la proteína de andamiaje.

20 Una proteína de andamiaje es preferiblemente biológicamente neutra. Por “biológicamente neutra” se supone que se han suprimido las interacciones con otras proteínas conocidas. Además, alguna capacidad de señalización poseída por la proteína es preferiblemente eliminada. Por tanto, una proteína de andamiaje preferida según la presente invención es la proteína de andamiaje de STM.

25 La neutralidad biológica es una ventaja de la presente invención puesto que no existe en la mayoría de las proteínas de andamiaje de la técnica anterior. Por ejemplo, Tioredoxina A actúa como un negativo dominante de las rutas redox naturales en células. Además, se sabe que inhibe P53 y se sabe que inhibe las rutas de señalización de BCL6. De manera ventajosa, las proteínas de andamiaje de la presente invención no interfieren con las rutas de señalización que se dan de manera natural.

Una proteína de andamiaje debería ser pequeña. Por “pequeña” se supone menos de 25 kDa, preferiblemente menos de 13 kDa. Lo más preferiblemente una proteína de andamiaje debería ser menor de 110 aa (excluyendo el inserto peptídico diana).

30 Preferiblemente una proteína de andamiaje según la presente invención será conformacionalmente estable. Por “conformacionalmente estable” se supone que no deberían darse cambios conformacionales. Preferiblemente una proteína de andamiaje no tiene región bisagra. Preferiblemente una proteína de andamiaje no tiene dominio PH. Preferiblemente una proteína de andamiaje no tiene dominio SH3. Preferiblemente una proteína de andamiaje no tiene dominio SH2. Preferiblemente una proteína de andamiaje no tiene dominio “WW”. Preferiblemente una proteína de andamiaje no tiene dominio “WD”. Preferiblemente una proteína de andamiaje no tiene repeticiones HEAT. Preferiblemente una proteína de andamiaje no tiene dominio rico en prolina. Preferiblemente una proteína de andamiaje no tiene modificación postraducción en células. Preferiblemente una proteína de andamiaje no tiene otro dominio conocido para facilitar cambios conformacionales.

40 Una proteína de andamiaje según la presente invención preferiblemente no tiene dominios de interacción proteína-proteína. Una proteína se considerará que no tiene dominios de interacción proteína-proteína si estos han mutado para volverlos no funcionales.

45 Preferiblemente una proteína de andamiaje según la presente invención no tiene modificaciones postraducción. Por tanto, preferiblemente una proteína de andamiaje según la presente invención no tiene sitio de glicosilación. Esto es una ventaja sobre las proteínas de andamiaje de la técnica anterior tales como distrofina debido a que las modificaciones postraducción pueden interferir con interacciones o crear ellas mismas interacciones falsas.

50 Tal como se indicó anteriormente, las proteínas de andamiaje no se deberían deformar por el inserto peptídico. Sobre este criterio, la proteína fluorescente verde no debería ser considerada una proteína de andamiaje debido a que al menos un tercio de los péptidos diana insertados suprimen la fluorescencia de la proteína fluorescente verde. Esto es una demostración de que el inserto peptídico diana está deformando la estructura de la proteína. Por lo tanto, no es una proteína de andamiaje según la presente invención puesto que una proteína de andamiaje preferiblemente no se debería deformar por el inserto peptídico diana.

55 La tioredoxina A (TrxA) es una proteína de andamiaje de la técnica anterior. TrxA es pequeña y es estable. Sin embargo, la inserción de los péptidos diana en TrxA se da entre dos residuos de cisteína. Las proteínas de andamiaje según la presente invención ventajosamente evitan esta disposición debido a que los residuos de cisteína en TrxA pueden someterse a unión a disulfuro reversible que puede alterar la conformación de la proteína de

andamiaje y puede afectar la conformación del péptido diana expuesto. Por tanto, preferiblemente el sitio de inserción para el péptido diana no está entre dos residuos de cisteína en la proteína de andamiaje.

Consideraciones de diseño

Las proteínas de andamiaje preferiblemente tienen uno o más de los siguientes rasgos:

- 5 1) ser de estructura conocida, permitiendo de ese modo una elección informada del sitio para la inserción o sustitución peptídica;
- 2) ser lo suficiente estable para restringir el plegamiento de un amplio intervalo de péptidos;
- 3) ser biológicamente neutra, es decir, carecer de interacciones con proteínas celulares que podrían aportar a fenotipo; y
- 10 4) ser capaz de plegarse similarmente, preferiblemente idénticamente en tanto ambientes procariotas como eucariotas, de manera que los datos obtenidos en un sistema pueden informar de experimentos realizados en el otro.

15 La presente invención proporciona un andamio ajustado a los requerimientos de la tecnología de aptámero peptídico. Las proteínas de andamiaje de la presente invención preferiblemente poseen todos los criterios anteriormente definidos: la estructura de Estefina A parental es conocida; el andamio modificado por ingeniería es estable y tolera la inserción de al menos un péptido sin perder su estabilidad biofísica; es capaz de exponer un intervalo amplio de péptidos para la interacción funcional; y no solamente todas las interacciones biológicas conocidas han sido modificadas por ingeniería sin parar.

Aplicaciones adicionales

20 El lector experto en la técnica apreciará que el uso de aptámeros peptídicos en micromatrices es particularmente ventajoso cuando aquellos aptámeros peptídicos están expuestos en la proteína de andamiaje según la presente invención. La tecnología de micromatriz de la técnica anterior depende fuertemente de los anticuerpos. Sin embargo, los anticuerpos pueden perder especificidad cuando se unen a la matriz. Además, las proteínas recombinantes usadas en las micromatrices pueden proporcionar información de que esas proteínas están presentes, pero no pueden proporcionar información sobre a lo que se están unidas. En cambio, usar aptámeros peptídicos presentados en proteínas de andamiaje según la presente invención ventajosamente puede proporcionar mucha más información cuando se interroga una matriz. Por ejemplo, tras la observación de una pareja de unión, la información contextual ventajosamente se deriva cuando se usa una proteína de andamiaje para presentar el aptámero. Esta ventaja se caracteriza como la diferencia entre una genoteca nativa y una informada. Por tanto, en otro aspecto de la invención se refiere al uso de estas nuevas proteínas de andamiaje para presentar péptidos sobre las micromatrices.

35 La proteína de andamiaje según la presente invención está basada en la secuencia de Estefina A. Por "basada en la secuencia de Estefina A" se supone que la proteína de andamiaje debería poseer al menos 30 de los 98 residuos de aminoácidos de Estefina A, preferiblemente 25% de la secuencia de aminoácidos de Estefina A, preferiblemente 30%, 40%, 50%, 60% o 70% de la secuencia de aminoácidos de la Estefina A, preferiblemente 80%, preferiblemente 85%, preferiblemente 90%, preferiblemente 95% o incluso más de la secuencia de Estefina A. Lo más preferiblemente la proteína de andamiaje tendrá la secuencia de Estefina A o STM o una de las nuevas variantes descritas en la presente memoria y comprende uno o más de los cambios mutacionales anteriormente descritos.

40 La capacidad de los aptámeros peptídicos de interrumpir las interacciones proteína-proteína *in vivo* puede permitir la rápida identificación de novedosas guías de fármaco. Además, el uso de pequeña(s) molécula(s) de fármaco candidata(s) para interrumpir la interacción proteína-proteína ventajosamente está facilitada por la presente invención.

45 Se pueden emplear ventajosamente el uso de los insertos peptídicos que comprenden sitios de modificación postraduccionales tales como sitio(s) de fosforilación. Esto es beneficioso en interacciones de disección que son variadas según el estado de fosforilación del péptido diana. Además, permite la identificación de aptámeros peptídicos candidatos que se unen de una manera dependiente de la fosforilación.

50 En algunas realizaciones, se puede desear introducir enlaces disulfuro a ambos lados del inserto peptídico diana, por ejemplo, modificando por ingeniería un residuo de cisteína a cada lado del inserto peptídico diana. Esto puede ser útil si el andamio está siendo usado exclusivamente en un ajuste. A este respecto, hay que indicar que las cistatinas de la familia II usan un enlace disulfuro para formar elementos de la estructura secundaria que corresponden a una región preferida de la inserción. En el contexto de la presente invención esto se puede conseguir, por ejemplo, por la adición de una cisteína sencilla en el terminal C del polipéptido de andamiaje, o dentro del péptido diana como en el extremo C-terminal del péptido diana, y además de un segundo residuo de cisteína insertado en una segunda localización como en el terminal N del andamio o en el extremo N-terminal del péptido diana, permitiendo así entrecruzamiento entre los dos. Sin embargo, se prefiere evitar de esta manera la restricción

covalente de los péptidos. Por tanto, en andamios preferidos de la presente invención, preferiblemente el péptido diana no está flanqueado por los residuos diana.

En conjunto se apreciará que los diferentes andamios pueden forzar una predisposición sobre los péptidos que exponen, de manera que el estudio de los péptidos diana puede implicar ventajosamente péptidos y/o genotecas expuestas en más de un andamio, para maximizar la probabilidad de éxito.

Los andamios de la invención permiten a los investigadores extender las observaciones *in vitro* al ambiente celular y viceversa, así como permitir la identificación *in vitro* o la creación de herramientas que se pueden usar dentro de las células sin preocupación por los patrones de plegamiento o el estado de oxidación de los enlaces disulfuro.

Los aptámeros peptídicos basados en los andamios de la presente invención son herramientas que se pueden usar para validar las dianas de fármaco que se pueden usar como componentes de ensayos de diagnóstico o pronóstico o incluso forman la base para compuestos principales para el tratamiento de enfermedad humana. Los andamios de la invención, ventajosamente basados en una proteína humana de longitud completa, se pueden usar como compuestos terapéuticos biológicos y/o en terapia génica.

Péptido diana

El término "péptido diana" como se usa en la presente memoria se refiere a un péptido de interés. El péptido diana preferiblemente es un péptido heterólogo. Por heterólogo se supone un péptido que se separa de su contexto normal, preferiblemente un péptido que tiene una secuencia no encontrada normalmente en la secuencia de la proteína de andamiaje que lo porta, lleva o presenta. Si el péptido tiene una secuencia que se da en otra parte en la secuencia de la proteína de andamiaje, entonces para que sea "heterólogo" esa secuencia estará fuera del contexto, es decir, no estará ocupando su posición que se da de manera natural (dirección) dentro del polipéptido de la proteína de andamiaje. En este contexto, "posición" significa posición dentro de la cadena de aminoácidos lineal más que posición en el espacio tridimensional relativo a otros residuos de aminoácidos. El péptido diana puede ser artificial, por ejemplo, generado por la construcción de una genoteca de péptidos para la incorporación dentro de la proteína de andamiaje. En estas realizaciones, se considera que el(los) péptido(s) es(son) "heterólogo(s)" para el fin de la invención.

Los aptámeros peptídicos son péptidos restringidos y expuestos por una proteína de andamiaje que se usan para estudiar la función proteica en las células. Algunos son capaces de interrumpir las interacciones proteína-proteína y algunos son capaces de constituir los módulos de reconocimiento que permiten la creación de un kit de herramientas molecular para el análisis intracelular de la función proteica.

La capacidad para diseñar o identificar moléculas pequeñas que se unen específicamente y con alta afinidad a una proteína dada es una etapa limitante en muchos experimentos, incluyendo el desarrollo de los micromatrices de proteína, el análisis de proteínas en el contexto de células vivas y la validación de las dianas candidatas de fármaco. En la naturaleza, las interacciones proteína-proteína se pueden mediar por superficies pequeñas de proteínas plegadas. Esto ha conducido al uso de pequeñas superficies peptídicas expuestas dentro del contexto de una proteína estable, denominada el andamio, como módulos de reconocimiento de proteína. Tales reactivos, denominados en la presente memoria aptámeros peptídicos, se han usado para interrumpir la actividad proteica biológica en una variedad de sistemas.

Los aptámeros peptídicos son más fácilmente repartidos y más estables en células que los péptidos libres y sus resultados de plegamiento restringido en un coste entrópico inferior de unión y de ahí afinidad aumentada para las proteínas diana. La modificación por ingeniería de los aptámeros peptídicos los permite proporcionar la funcionalidad de reconocimiento en el diseño de un kit de herramienta molecular aunque este potencial aún tiene que ser completamente realizado. La afinidad de los aptámeros peptídicos para sus dianas oscila entre 10^{-6} a 5×10^{-9} M en comparación con 10^{-7} a 10^{-11} M para interacciones anticuerpo/diana. Al usar las inserciones múltiples para aumentar el área superficial de la interacción, se espera que los aptámeros peptídicos sean capaces de igualar o posiblemente exceder de manera ventajosa las afinidades de unión de los anticuerpos. No obstante, los aptámeros peptídicos claramente son capaces de interrumpir las interacciones proteína-proteína *in vivo*. Los cribados de aptámero peptídico se realizan en levadura o en células de mamífero, lo cual los distingue de los cribados de presentación en fago ("phage display") de genotecas de péptido o anticuerpo realizadas frente a proteína expresada de forma procariota potencialmente mal plegada.

Aunque el andamio más ampliamente usado es la proteína tioredoxina en *Escherichia coli* (TrxA), se han usado un número de otras proteínas. El éxito de esta tecnología depende de la solidez del andamio, aún un tercio de los péptidos pueden desestabilizar PFV, mientras que muchos aptámeros peptídicos basados en TrxA no se expresan establemente en células humanas cultivadas, sugiriendo que este andamio tampoco puede ser lo suficiente rígido para exponer péptidos sin llegar a estar parcialmente desplegado. Los péptidos tomados fuera del contexto de un andamio y colocado en otro frecuentemente pierden la capacidad de interactuar con sus proteínas diana, elevando la posibilidad de que los cribados para interactores restringidos con una diana dada puedan fallar a menos que se use un andamio apropiado. Finalmente, las actividades biológicas de los andamios usados para exponer péptidos han sido rigurosamente caracterizadas en la técnica anterior, conduciendo a preocuparse de que cualquier fenotipo

observado cuando se expresa un aptámero peptídico podría, al menos en parte, ser debido a un efecto del andamio y no al péptido insertado. Por lo tanto, hemos producido un andamio sólido, versátil, biológicamente neutro para la exposición de péptidos restringidos. Buscamos una proteína que se podría expresar de manera estable en una variedad de sistemas experimentales mientras exponen péptidos que son capaces de interactuar funcionalmente con un intervalo amplio de dianas. Tal andamio básicamente mejora la tecnología del aptámero peptídico aumentando su solidez. Además, al ampliar el repertorio de andamios disponibles, la presente invención ventajosamente aumenta la probabilidad de que se obtendrán correspondencias en cribados frente a un número mayor de proteínas diana usando genotecas en andamios múltiples en cribados simultáneos frente a cada diana.

Estefina A

La Estefina A es el miembro fundador de la familia de cistatina de los inhibidores proteicos de la catepsinas de cisteína, las cuales son peptidasas lisosomales de la familia de la papaína. El subgrupo de estefina de la familia de la cistatina son proteínas de dominio sencillo relativamente pequeñas (alrededor de 100 aminoácidos). No Reciben modificación postraducción conocida, y carecen de enlaces disulfuro, sugiriendo que serán capaces de plegarse idénticamente en una amplia variedad de ambientes extra- e intracelulares. La propia SteA (Estefina A) es una proteína monomérica, de cadena sencilla, de dominio sencillo de 98 aminoácidos. La estructura de SteA se ha resuelto, facilitando la mutación racional de SteA en el andamio de STM. La única actividad biológica conocida de las cistatinas es la inhibición de la actividad de la catepsina, lo cual nos permitió ensayar exhaustivamente la actividad biológica residual de nuestras proteínas modificadas por ingeniería. Por tanto, revelamos que la modificación por ingeniería de proteína de la SteA nativa puede producir variantes que son útiles como andamios de aptámero peptídico. La técnica anterior del aptámero peptídico ha sido obstaculizada por dificultades en la identificación de la actividad biológica en ensayos basados en célula, causadas al menos en parte por el funcionamiento subóptimo de los diversos andamios existentes. La presente invención proporciona un andamio útil que será de gran beneficio a los que buscan estudiar las interacciones proteína-proteína *in vitro* e *in vivo*.

Secuencias de Estefina A

Un andamio "basado en" Estefina A tiene una secuencia que está derivada de estefina A. Preferiblemente la secuencia derivada de estefina A comprende la secuencia de tipo natural de estefina A, que comprende preferiblemente una o más de las modificaciones (mutaciones) descritas en la presente memoria. Para el experto en la técnica estará claro que se pueden hacer modificaciones menores a la secuencia andamio sin apartarse de la invención. En particular, la invención se refiere a secuencias de aminoácidos y/o secuencias de nucleótidos que tienen al menos 25%, 35%, 45%, 55% o 60% de identidad con las correspondientes secuencias mostradas en la presente memoria, preferiblemente al menos 70%, preferiblemente al menos 80%, preferiblemente al menos 85%, preferiblemente al menos 90%, preferiblemente al menos 92%, preferiblemente al menos 94%, preferiblemente al menos 95%, preferiblemente al menos 96%, preferiblemente al menos 97%, preferiblemente al menos 98%, preferiblemente al menos 99% de identidad, o incluso más, sin embargo en cada caso, las variaciones de secuencia se consideran "menores" si no afectan adversamente a la capacidad del andamio a exponer el péptido diana a disolvente, y no restaurar o generar funciones biológicas tales como las que son poseídas por la estefina A de tipo natural pero las cuales están suprimidas en cambios mutacionales de la presente invención.

Además, las modificaciones menores también pueden incluir deleciones o adiciones pequeñas a la estefina A o secuencias derivadas de estefina A descritas en la presente memoria, tal como la adición o deleción de 10 aminoácidos o menos al polipéptido derivado de la estefina A. Por tanto, la invención se refiere a secuencias de aminoácidos que tienen una adición o deleción total con respecto a las secuencias de estefina A o STM descritas en la presente memoria de 40 aminoácidos o menos, preferiblemente 30 aminoácidos o menos, preferiblemente 20 aminoácidos o menos, preferiblemente 15 aminoácidos o menos, más preferiblemente 10 aminoácidos o menos, preferiblemente 9 aminoácidos o menos, preferiblemente 8 aminoácidos o menos, preferiblemente 7 aminoácidos o menos, preferiblemente 6 aminoácidos o menos, preferiblemente 5 aminoácidos o menos, preferiblemente 4 aminoácidos o menos, preferiblemente 3 aminoácidos o menos, preferiblemente 2 aminoácidos o menos, preferiblemente 1 aminoácido. La adición o deleción total es el factor importante, de manera que una diferencia de 9 o menos puede significar una deleción de 9 aminoácidos, o tres deleciones de tres aminoácidos cada una, dos adiciones de tres aminoácidos y una deleción de tres aminoácidos, etc. La invención también se refiere a las correspondientes variantes de ácido nucleico. En cada caso, las variaciones de secuencia se consideran modificaciones "menores" si no afectan adversamente a la capacidad del andamio de presentar el péptido diana al disolvente, y no restauran o generan funciones biológicas tales como las que poseen la estefina A de tipo natural.

Mutaciones de Estefina A y STM

En el contexto de tratar los sitios de mutación, "cerca de" significa dentro de 7 aminoácidos, preferiblemente dentro de 5 aminoácidos, preferiblemente dentro de 3 aminoácidos, preferiblemente dentro de 2 aminoácidos, preferiblemente en el aminoácido propuesto o uno de los dos aminoácidos vecinos.

En el contexto de las inserciones, se prefiere que a nivel de ácido nucleico se introduzca(n) el(los) sitio(s) de restricción, preferiblemente sitio(s) de restricción único(s), para facilitar las futuras intersecciones. Estas enseñanzas y conocimientos generales comunes en la técnica de la tecnología de ácido nucleico recombinante posibilitan al

experto introducir el(los) sitio(s) de restricción relevante(s) mientras que se conserva las características claves del andamio. Por "único" se supone único en la secuencia codificadora de la proteína de andamiaje. Se pueden usar sitios no únicos, pero los sitios únicos son preferidos para facilitar la inserción y manipulación de las construcciones. Cuando se usan dos o más sitios, por ejemplo, para facilitar la separación y sustitución de la secuencia de cualquiera de los codones 67 a 84 del bucle 1 de SteA, preferiblemente cada uno de los dos o más sitios es único. Sin embargo, si los dos o más sitios son idénticos ventajosamente se puede simplificar las operaciones de eliminación y sustitución, por ejemplo, implicando solamente un tratamiento de enzima de restricción sencillo. Estas elecciones están bien dentro de la capacidad del experto que maneja la invención. En una realización preferida, se introducen dos sitios idénticos para la eliminación y sustitución del bucle. Preferiblemente los sitios de restricción usados en las secuencias que codifican los cambios mutacionales son diferentes de manera que las inserciones o modificaciones en cada una de estas cuatro localizaciones en la secuencia codificadora se pueden hacer usando una enzima de restricción diferente para facilitar la manipulación.

Mutación posición 4

El término "mutación posición 4" se usa en la presente memoria para describir la mutación alrededor de, preferiblemente cerca de o preferiblemente en, el sitio G4 de estefina A, o el sitio W4 de STM, la mutación se refiere a adición(adiciones) o inserción(inserciones) o sustitución(sustituciones) al(a los) residuo(s) de aminoácidos del terminal amino de SteA o STM. Preferiblemente tales mutaciones, son proximales a Pro3, preferiblemente proximales a G4 (Estefina A) o W4 (STM). Preferiblemente tales mutaciones están cerca de, o preferiblemente en, el Pro3 de Estefina A o STM humana. Lo más preferido es la sustitución del residuo 4 con R.

En una realización preferida, el sitio de la posición 4 se usa como un sitio de inserción primaria, secundaria o terciaria además de los otros cambios mutacionales a bucle 1 y/o 2 como se describió anteriormente en la presente memoria. La presencia de R en lugar de G aumenta la accesibilidad de la superficie de reconocimiento (de unión a diana) puesto que R es un aminoácido positivamente cargado y, por tanto, previene un bucle helicoidal alfa que cubre el sitio de reconocimiento. Además, el cambio desestabilizó el aptámero cuando está solo, pero se estabiliza una vez que el aptámero se une a una diana.

Mutación en cualquiera de los codones 46 a 54

El término "mutación en cualquiera de los codones 46 a 54" se usa en la presente memoria para describir la mutación alrededor, preferiblemente cerca de o preferiblemente en, el sitio VAG de SteA o el sitio DAG de STM. El sitio VAG son los residuos 48 a 50 del sitio QVVAG que está en los residuos 46 a 50 de SteA humana. El sitio DAG son los residuos 48 a 50 del sitio QVDAG que está en los residuos 46 a 50 de STM.

Preferiblemente esto se refiere a la(s) adición(adiciones) o inserción(inserciones) o sustitución(sustituciones) alrededor, preferiblemente cerca de o preferiblemente en los sitios VAG/DAG. Preferiblemente esto se refiere a las adiciones o las inserciones en los sitios VAG/DAG.

En una realización preferida, el sitio 46 a 54 se usa como un sitio de inserción primaria, secundaria o terciaria en combinación con los cambios mutacionales anteriormente descritos en la presente memoria.

En una realización preferida la mutación en los sitios VAG/DAG es LAS.

Los experimentos han demostrado que las modificaciones D48L y G50S conducen a expresión aumentada en un sistema bacteriano.

Mutación en cualquiera de los codones 67 a 84

El término "mutación en cualquiera de los codones 67 a 84" se usa en la presente memoria para describir la mutación alrededor o preferiblemente cerca de o preferiblemente en el bucle L37-L80 de la Estefina A humana o el bucle P73-L80 de STM.

El término se puede referir a la(s) adición(adiciones) o inserción(inserciones), o la sustitución en estos sitios.

En una realización, la mutación puede comprender la sustitución del bucle completo entre L73 y L80 o P73 y L80 con cualquier secuencia peptídica, preferiblemente con una variedad de diferentes secuencias peptídicas diana (preferiblemente solamente una por molécula de andamiaje de estefina), es decir, una genoteca.

A un nivel de ácido nucleico, las mutaciones preferidas son aquellas que dan como resultado un sitio de restricción para la inserción en el bucle, y más preferiblemente dos sitios de restricción para la sustitución de la secuencia codificadora de este bucle. Particularmente preferido son los sitios de restricción, son los sitios de restricción RsrII.

En una realización preferida, el sitio del bucle 2 se usa como sitio de inserción primaria, secundaria o terciaria en combinación con los cambios mutacionales anteriormente descritos en la presente memoria.

Dos nuevos andamios según la presente invención donde los cambios mutacionales se han modificado por ingeniería en NGP (SQM tiene L82R y T83S y SQT tiene E78A y L80R) presentan cada uno alta expresión en *E. coli*, lo cual es lo más inesperado ya que es una diferencia significativa desde la proteína madre. Tanto SQM como SQT tienen una estructura estable cuando se muestra por difracción circular, lo cual es inesperado ya que es una diferencia significativa desde la proteína madre.

Los péptidos insertados en estos aptámeros están disponibles a los disolventes como se muestra por los experimentos de unión a anticuerpo y ventajosamente estas proteínas retienen su unión y función cuando se adhiere a una superficie sólida. Además, tienen un área superficial aumentada debido a las posiciones para tres insertos y, por lo tanto, dan mayor afinidad de unión y los experimentos han demostrado que SQM se pliega correctamente con un conjunto específico de insertos peptídicos y no forma dímeros que puedan enmascarar sitios de unión, esto es una diferencia de STM y así proporciona ventajas significativas sobre la técnica anterior. Además, con los andamios de la presente invención se pueden hacer genotecas de aptámero peptídico usando, por ejemplo, SQM y así se pueden identificar los aptámeros que tienen el potencial de interactuar con las dianas en el tejido humano debido a las superficies múltiples de unión.

15 Inserciones

Preferiblemente, los insertos están cerca de o preferiblemente en el bucle L73-L80 de la Estefina A humana o el bucle P73-L80 de STM y más preferiblemente con dos residuos LeuAla codificados por la secuencia de anillamiento y, por tanto, la proteína de andamiaje es dos residuos más larga que la Estefina A original.

Mutaciones de combinación

Preferiblemente una proteína de andamiaje se basa en la Estefina A o STM y comprende al menos una de las mutaciones anteriormente descritas. Preferiblemente, la proteína de andamiaje comprende al menos dos o todas las tres mutaciones anteriormente descritas. Preferiblemente una proteína de andamiaje posee todas las tres mutaciones anteriormente descritas, con el resto de la proteína que se parece o bien a Estefina A o STM o una de las otras variantes. Además o alternativamente, cuando la mutación terminal está en la posición 72/73 o 82/83 se sigue por un codón de terminación que elimina o sustituye los últimos 25 o los últimos 15 aminoácidos de o bien Estefina A o STM. Los péptidos diana ventajosamente se pueden insertar en cualquiera de los tres sitios de mutación preferidos. En la realización altamente preferida las proteínas de andamiaje basadas en Estefina A/STM permiten el uso de tres superficies en total. Estas son las superficies definidas en la posición 4, bucle 1 y bucle 2 (Figura 2).

30 Fase sólida y micromatrices

Tal como se indicó anteriormente, la invención encuentra aplicación en micromatrices. En realizaciones en fase sólida tal como las realizaciones en micromatrices, las proteínas de andamiaje de la invención preferiblemente se modifican por ingeniería para facilitar su asociación o adhesión al sustrato de fase sólida para el ensayo. Preferiblemente esto es mediante el pegado a un revestimiento de oro, o mediante la asociación con biotina. Para modificar por ingeniería el andamio para que se pegue al revestimiento de oro, preferiblemente se introducen uno o más residuos de Cys en el terminal C o N de la proteína de andamiaje. Para modificar por ingeniería el andamio para inmovilización por adhesión a biotina, preferiblemente se introducen una o más copias de un dominio de unión a biotina de ocho aminoácidos ("etiqueta strep") en dicho andamio. La inmovilización puede ser por uno o más de estos o cualquier otro medio adecuado. Preferiblemente se inmoviliza la proteína de andamiaje de la invención. Preferiblemente las proteínas de andamiaje de la invención se modifican por ingeniería para la inmovilización. Preferiblemente los ensayos de interacción según la presente invención se llevan a cabo usando proteínas de andamiaje inmovilizadas.

Ventajas adicionales de la invención

Las proteínas de andamiaje basadas en Estefina A son superiores al uso de péptidos debido a que se pueden usar *in vivo*. Además, emplear sistemas recombinantes es más barato que trabajar con péptidos sintéticos. Además, la construcción de genotecas es más barato que usar genotecas sintéticas por la misma razón, y también debido a que están racionalmente diseñados usando la manipulación de ácido nucleico. Esto reduce la dependencia de química complicada para la síntesis peptídica.

Las proteínas de andamiaje basadas en Estefina A son superiores a la técnica anterior como presentación en fago puesto que son internas a la célula, mientras que la presentación en fago depende de la interacción extracelular. Además, las proteínas de andamiaje de la presente invención se pueden usar para trabajar sobre dianas nativas en lugar de dianas recombinantes. Esto tiene una ventaja adicional que permite el examen de proteínas modificadas después de la traducción que se someterán correctamente a fosforilación o glicosilación o de otra manera modificadas después de la traducción *in vivo* pero las cuales probablemente no se formarían correctamente si se produce *in vitro*.

Una ventaja adicional de las proteínas de andamiaje según la presente invención es que permiten la interrogación del espectro que se da de manera natural de las variantes de empalme y las variantes de modificación postraduccion que se producen *in vivo* sin tener que fabricar individualmente cada una de ellas y disponerlas o de lo contrario compartimentarlas para analizar.

- 5 Una aplicación adicional de la invención está en el uso de micropalancas (“microcantilevers”) como una lectura de la interacción con proteínas de andamiaje basadas en Estefina A. Además, las proteínas de andamiaje de la presente invención son particularmente adecuadas para su uso con las lecturas de tipo transistor electroquímicos y de película fina.

- 10 Otra ventaja adicional del andamio de la presente invención es que los aptámeros peptídicos de la presente invención se pueden sustituir por anticuerpos y los resultados han demostrado que pueden incluso funcionar mejor ya que, por ejemplo, CDK2 se detectó más rápidamente usando aptámeros peptídicos que los anticuerpos. Por consiguiente, el uso de los aptámeros peptídicos en lugar de anticuerpos significa que se necesitará usar menos animales en la producción de sondas moleculares lo cual ofrece ventajas significativas a la investigación científica.

- 15 Ahora, la presente invención se describirá a modo de ejemplo, en lo cual la referencia se hará a las siguientes figuras.

Breve descripción de los dibujos

La Figura 1 muestra la expresión y la solubilidad de STM y las variantes en *E. coli*; la Figura 1A muestra las variantes SUN, SUM y STM, la Figura 1B muestra las variantes SUC y SDM y la Figura 1C muestra la variante SQM.

- 20 La Figura 2 muestra una representación de la estructura de solución de NMR de las variantes STM en la posición de codón 4, las posiciones de codón 48 a 50 en el bucle 1, la posición de codones 67 a 84 en el bucle 2 y los codones en las posiciones 91 a 92 creado usando el programa informático Cn3D y IDVD coordinadas PDB (Martin et al., 1995 “The three-dimensional solution structure de human stefin A”, *J. Mol. Biol.*, vol. 246 pp. 331-43). Se indican las regiones que se mutaron para producir las proteínas de Estefina A modificadas.

- 25 La Figura 3 muestra el análisis de espectroscopia de dicroísmo circular (CD, del inglés “circular dichroism”) para SDM, SUC, AUM, SUN, SQM, STM W4R y un gráfico de referencia de STM que mide las diferencias en la absorción de luz polarizada a mano izquierda frente a la luz polarizada a mano derecha que surge debido a la asimetría estructural para ilustrar la proporción de la estructura secundaria conservada entre STM y las nuevas variantes.

La Figura 4 muestra el análisis de espectroscopia de dicroísmo circular (CD) para STM, SDM, SQM, SUM, SUN, SUC, pep6M, pep9M y pep10M seguido de almacenamiento prolongado.

- 30 La Figura 5 muestra el análisis de espectroscopia de dicroísmo circular (CD) para preparaciones independientes de STM, SDM, SQM, SUM, SUN, SUC, pep6M, pep9M y pep10M.

La Figura 6 muestra el análisis de espectroscopia de dicroísmo circular (CD) de SQT y variantes de las mismas etiquetadas con epítipo: - SQT-AUI(1), SQT-AUI(2), SQT-HA(2), SQT-myc(1), SQT-myc(2), SQT-AUI(1),AUI(2), SQT-AUI(1),HA(2) y SQT-AUI(1),myc(2).

- 35 La Figura 7 muestra el análisis de espectroscopia de dicroísmo circular (CD) de las variantes de SQM etiquetadas con epítipo: - SQM-myc(1), SQM-AUI(2), SQM-myc(n)AUI(1), SQM-AUI(1),HA(2), SQM-myc(n),AUI(1),HA(2), SQM-HA(n-term), SQM-myc(2), SQM-(21aleatorio-bucle 1), SQM-AUI(1), SQM(AUIx2, bucle 1), SQM-HA(n),AUI(1), SQM-HA(n),myc(2), SQM-AUI(1),myc(2), SQM-HAx2(n),AUI(1),myc(2), SQM-HA(n),AUI(1),myc(2) y un aptámero peptídico diseñado pep22(Trx).

- 40 La Figura 8A muestra la inmunoprecipitación del péptido AUI por anticuerpo anti-AUI, la Figura 8B muestra inmunoprecipitación de SQM-myc (bucle 1) y la Figura 8C muestra inmunoprecipitación de SQT-HA (bucle 2).

La Figura 9 muestra los resultados de un experimento de micromatriz con aptámeros peptídicos con diferentes epítipos en diferentes posiciones.

- 45 La Figura 10 muestra mediciones de la resonancia de plasmón superficial (SPR, del inglés “Surface Plasmon Resonance”) de interacciones anticuerpo/epítipo en el contexto del andamio de SQM. La Figura 10 A muestra una comparación de soluciones 33 nM de Anti-Myc (rojo), Anti-Ha (azul) y Anti-Au1 (verde) en pH 7,3, tampón fosfato 10 mM que están inmovilizadas sobre una monocapa (negra) de SQM quimioadsorbido (Nt Ha, L1 Au1, L2 Myc). La Figura 10B muestra la respuesta de SPR a concentración variable de Anti-Myc adecuado a la función de las cinéticas de saturación $M=(CxBmax)/(C+K)$, dando lugar a una constante de equilibrio de $50 \times 10^{-9} M$.

- 50 La Figura 11, los resultados de un micromatriz que comprende 864 aptámeros peptídicos aleatorios insertados en el bucle 1 (arriba) o el bucle 2 (centro) o el terminal de aminoácidos (abajo).

La Figura 12 muestra el uso de los aptámeros peptídicos en SQM (SQM-pep6) y SQT (SQT-pep10m) para detectar la expresión endógena de cdk2 en lisados duplicados de células humanas (HeLa).

Ejemplo 1

5 Con referencia a la Figura 2 se muestra una estructura tridimensional de Estefina A y tres sitios en Estefina A que están mutados para generar las nuevas proteínas de andamiaje de la presente invención. Estos sitios están: en la posición G4 de Estefina A o W4 de STM; cualquiera de los codones 46 a 54 incluidos del bucle 1 de restricción y particularmente en posiciones de codón 48 a 50; cualquiera de los codones y; cualquiera de los codones 67 a 84 incluidos del bucle 2 de restricción y particularmente mutaciones de 70-73. La Estefina A modificada o los polipéptidos de STM para su uso como proteínas de andamiaje se producen por mutación de la secuencia de la Estefina A como se describe. Las proteínas resultantes basadas en Estefina A pero que poseen los cambios mutacionales específicos se dan en las secuencias anteriormente descritas en la presente memoria.

Ejemplo 2

15 La Figura 1 muestra la expresión de STM y las variantes ilustrativas en *E. coli*. Los marcos de lectura abiertos para STM y las variantes descritas en la presente memoria se clonaron en una versión del vector de expresión pET30a+ de *E. coli* que se ha modificado por ingeniería para incluir funcionalidades adicionales en la cola del amino-terminal, tal como un residuo de cisteína (presente en todas las variantes mostradas) o una etiqueta StrepII (en solamente STM). Los 8 aminoácidos adicionales de la etiqueta StrepII insertada justifican la ligera diferencia en la migración de la proteína STM en comparación con las otras variantes. Las células de *E. coli* que llevan estas construcciones de expresión se dejaron crecer en ausencia (-) o presencia (+) de isopropil- β -D-tiogalactopiranosido (IPTG), lo cual induce la expresión de STM y las proteínas variantes (destacado por *). Después de 3 horas de inducción a 37°C, las células se sometieron a lisis por sonicación, y el lisado celular total (T) o la porción soluble (S) recuperados después de una centrifugación durante 10 minutos a 16.000xg se cargaron sobre geles de poli(acrilamida) al 15%. Las proteínas se visualizaron por tinción con Coomassie. Por lo que se determinó por este método, se podría recuperar el 100% de cada proteína mostrada (SUN, SUM, STM, SUC, SDM y SQM) en la fracción soluble, indicando que las proteínas variantes se mantienen capaces de plegarse en *E. coli*.

Ejemplo 3

30 En la producción de las nuevas proteínas de andamiaje basadas en Estefina A y/o STM se ha empleado un enfoque racional al diseño de un nuevo péptido. Es deseable para las nuevas proteínas de andamiaje de la presente invención poseer cualidades que un andamio ideal necesitaría poseer para ser ampliamente útil para los estudios *in vitro* e *in vivo* y aplicar estos criterios al diseño de un nuevo andamio.

Partiendo de la pequeña Estefina A inhibidora de proteasa intracelular estable o STM, hemos modificado por ingeniería un número de andamios biológicamente neutros que retienen la conformación estable de la proteína madre. Esperamos que las nuevas proteínas de andamiaje modificadas sean capaces de exponer péptidos que se unan a dianas de interés, tanto en el contexto de interactores conocidos como en un cribado de genoteca. 35 Herramientas moleculares basadas en los andamios encuentran aplicación en un amplio intervalo de estudios de rutas biológicas, y en la validación de dianas de fármaco. SteA es una proteína de dominio sencillo monomérica de 98 aminoácidos que no recibe modificación postraducción conocida y carece de enlaces disulfuro. SteA muestra extraordinaria termoestabilidad con una transición reversible observada a 90,8°C y entalpía de plegamiento de 490 kJ/mol, todos los rasgos importantes de un andamio basado en SteA.

Ejemplo 4

40 Se transformaron plásmidos de expresión de la variante STM (usando todos pET30a+) en *E. coli*. Se inocularon colonias sencillas en cultivos de toda la noche para dejarlas crecer a 37°C con agitación (250 rpm en un agitador orbital). A la siguiente mañana, 0,5 ml de cada cultivo de toda la noche se inoculó en 500 ml de medio fresco, complementado con kanamicina para mantener la selección para los plásmidos pET30. Se indujo la expresión de la proteína variante una vez que los cultivos alcanzaron media fase logarítmica (DO600 aproximadamente 0,6 a 0,8). 45 Los cultivos se dejaron crecer durante unas 3 horas más, aún a 37°C con agitación. Se recogieron las células de *E. coli* por centrifugación y se sometieron a lisis usando una Presa Francesa. Los lisados se aclararon por centrifugación y las proteínas variantes STM se purificaron a partir de los sobrenadantes resultantes usando cromatografía de afinidad de quelato de Ni. Para esto, se usó 0,5 ml de Ni-NTA agarosa (QIAGEN) por 20 ml de lisado. La resina se centrifugó en tubos Falcon de 50 ml a 700 g durante 2 min y se desechó el sobrenadante. La resina se lavó tres veces con 2,5 ml de 1x tampón de Calibración/Lavado, resuspendiendo la resina en tampón y, a continuación, centrifugando a 700 g durante 2 min a 4°C y separando el sobrenadante. El lisado se combinó con la resina de afinidad de metal lavada y se incubó sobre rodillos durante 2 h a 4°C. Una alícuota del lisado se retuvo durante el posterior análisis. La resina se separó del lisado por centrifugación a 700 g durante 5 min a 4°C y se separó el sobrenadante. 50 Se retuvo una alícuota adicional del lisado para el análisis posterior de eficacia de unión. La resina se lavó seis veces resuspendiendo las perlas en 10 ml de Tampón de Lavado (NaH₂PO₄ 50 mM, NaCl 300 mM, imidazol 20 mM, pH 7,4), centrifugando a continuación a 700 g durante 2 min a 4°C y separando el sobrenadante. La resina se incubó con 1 ml de Tampón de Elución (NaH₂PO₄ 50 mM, NaCl 300 mM, Imidazol 150

mM, pH 7,4) durante 10 min a temperatura ambiente. La suspensión se centrifugó a 700 g durante 5 min y se retuvo el sobrenadante. Esta etapa se repitió dos veces más para producir dos fracciones de elución adicionales. Las fracciones con las concentraciones de proteína más altas se retuvieron y almacenaron a 4°C. Para el dicroísmo circular, las muestras se intercambiaron con tampón en tampón fosfato 50 mM pH 7,4. Las muestras se diluyeron a aproximadamente 0,3 mg/ml en el día de análisis, y se tomaron mediciones exactas de proteína usando un espectrofotómetro NanoDrop. Cuando se indica en las figuras, se añadió 1 mM de DTT inmediatamente antes del análisis. Se recogieron los espectros de CD desde 200 a 260 nm usando un espectropolarímetro Jasco J715. Los espectros se normalizaron a la concentración de proteína molar y se trazaron las elipticidades molares residuales para minimizar los artefactos entre las muestras.

10 Ejemplo 5

Para establecer si las nuevas proteínas de andamiaje de la presente invención basadas en la Estefina A y STM son conformacionalmente estables y si las alteraciones hechas a nivel de ADN (marco de lectura abierto) que alteraron la secuencia de aminoácidos de las nuevas variantes de Estefina A/STM conducen a estabilidad disminuida de la proteína, todas las variantes descritas en la presente memoria se expresaron y sometieron a dicroísmo circular para comparar su composición estructural secundaria con la de la Estefina A. En referencia a la Figura 1 se puede ver que se encontró que todas las proteínas se expresan igualmente bien en *E. coli*, generalmente a aproximadamente 28 mg de proteína variante/ml de cultivo bacteriano. Posteriormente, se purificaron las proteínas a cerca de homogeneidad por cromatografía de afinidad usando Ni-agarosa, y se diluyeron las preparaciones purificadas a 0,3 mg/ml inmediatamente antes del análisis de dicroísmo circular. Tal como se mencionó anteriormente, el análisis de dicroísmo circular implica el barrido de la proteína a través de un intervalo de longitudes de onda cerca de UV, de modo que la elipticidad de la luz está afectada por los elementos estructurales secundarios (hélice alfa o hebra beta) de la proteína.

En referencia a la Figura 3 se puede ver que la proporción de la estructura secundaria se conserva entre STM y las nuevas variantes de la presente invención convenientemente referidas como SDM, SQM, SUM, SUN y SUC (dadas anteriormente en la presente memoria como SEQ ID NO: 9 a 13) y que la presencia de insertos en STM no afecta de manera adversa su estructura. Particularmente digno de mención es que dos variantes (SUN y SQM) parecían mostrar estructura aumentada en comparación con otras proteínas de andamiaje de ensayo. Esto puede ser atribuible a la adquisición de estructura secundaria en la cola amino-terminal que está presente en todas estas proteínas, y se conducirá por la sustitución de Glicina (Estefina A) o Triptófano (STM) en la posición 4 por Arginina, ya que esto es el único cambio que es común a SUN y SQM, y estas son las únicas variantes que poseen esta alteración. Cuando se consideran los cambios mínimos en SUN, SUM, SUC y SDM, los resultados mostraron que la localización del más importante punto de inflexión en 218 nm en gran medida no está afectada por cada cambio (Figura 3), indicando que, a diferencia de la expectativa de que las alteraciones de aminoácidos generalmente se espera que desestabilicen las proteínas, la proporción de la estructura secundaria en los derivados de Estefina A no se cambia por las alteraciones de aminoácidos. Al contrario, la profundidad de la inflexión se altera significativamente entre las variantes (Figura 3).

Ejemplo 6

Se investigó el efecto del almacenamiento. La Figura 4 muestra que, después de almacenamiento a 4°C durante dos semanas en un tampón fosfato a pH 7,4 de las reservas concentradas de todas las variantes de proteína de andamiaje ensayadas, la proporción de la estructura secundaria se retiene cuando se analizan diluciones frescas de las mismas muestras (comparación con Figura 5). Esto es altamente inesperado ya que la mayoría de las proteínas están completamente desnaturalizadas, o se han perdido debido a la adsorción del recipiente de almacenamiento, por esta vez a menos que se añada una gran cantidad de proteína vehículo. Esta etapa, es decir, la adición de una proteína vehículo es indeseable, ya que el objetivo es usar preparaciones altamente purificadas de aptámeros peptídicos para minimizar señales no específicas en ensayos de diagnóstico y analíticos que resultarán de la presencia de proteínas irrelevantes como la proteína vehículo. En conclusión, las nuevas variantes de proteína de andamiaje de Estefina A se pueden almacenar en un tampón de fosfato sencillo, sin aparentes efectos adversos. Esta observación significa que es probable ayudar en la aplicación industrial de las nuevas variantes de la proteína de andamiaje de Estefina A.

50 Ejemplo 7

También se realizó dicroísmo circular sobre variantes STM, SDM, SQM, SUM, SUN, SUC más tres aptámeros peptídicos pep6M, pep9M y pep10M (Figura 6). Se añadió DTT para prevenir los residuos de cisteína presentes en las colas amino-terminales de las proteínas variantes expresadas a partir de la formación de enlaces de disulfuro intermoleculares que podrían afectar al experimento. La presencia o ausencia de DTT no cambió la estructura secundaria observada, aunque los espectros obtenidos en su ausencia fueron más fáciles de interpretar, ya que el propio DTT contribuye a la señal en el intervalo cerca de UV. Estos resultados muestran que DTT no afecta a la estructura secundaria de las proteínas de andamiaje de la presente invención.

Ejemplo 8

5 Se valoró la capacidad como sistema de presentación de proteína de andamiaje de la variante SQM (SEQ ID NO: 10) que comprende un cambio mutacional en la posición 4 además de los cambios tanto en bucle 1 como 2, es decir, sitios de inserción múltiple en un andamio sencillo. Los péptidos (HA, AU1 o MYC) se insertaron en las posiciones 4, 48 y 72/82 respectivamente y se generaron datos de dicroísmo circular (Figuras 7). Los datos de análisis de dicroísmo circular indican que los nuevos sitios de inserción no son solamente capaces de exponer péptidos para la interacción si no que lo hacen así sin pérdida significativa de estructura.

Ejemplo 9

10 Se usaron dos métodos para determinar el efecto de los cambios de aminoácidos sobre la estructura de andamio. El primer método es toscamente para determinar los niveles de expresión relativos de las proteínas modificadas por ingeniería en *E. coli*, con el fundamento de que la mayoría de los cambios de aminoácidos son probablemente para desestabilizar la proteína. La Tabla 1 de a continuación da los rendimientos de expresión de diversas variantes andamio a partir de cultivo bacteriano, el rendimiento se da como mg de proteína purificada por litro de cultivo.

Tabla 1

ANDAMIO	RENDIMIENTO	ANDAMIO	RENDIMIENTO	INSERCIÓN
STM	59	SQM-myc	11	L2
SUN	63	SQM-AU1-	103	L1
SUM	48	SQM-AU1 HA	206	L1+L2
SUC	71	SQM-AU1 myc	14	L1+L2
SDM	57	SQM HA	19	N
SQM	58	SQM HA-myc	4	N+L2
STM Pep2	19	SQM HA AU1-	2	N+L1
SQM Pep2	9	SQM HA AU1 myc	2	N+L1+L2
STM Pep6	36	SQM myc AU1 -	11	N+L1
SQM Pep6	2	SQM myc AU1 HA	7	N+L1+L2
STM Pep9	117, 61			
SQM Pep9	21			
STM Pep10m	76, 40			
SQM Pep10m	2			
STMA7	23			
SQMA7	41			
STMA48	52, 43			
SQMA48	42			
STMA58	62, 61			
SQMA58	70			
SQT	45			

15 El segundo método usado era para valorar directamente la proporción de elementos de la estructura secundaria en cada proteína por dicroísmo circular (CD). Cuando se compararon los rendimientos de proteína de *E. coli*, se encontró que cada cambio individual de Estefina A en SUN, SUM y SUC tenía poco efecto, o conducía a rendimiento aumentado en comparación con STM (Tabla 1). Esto era cierto también cuando se combinaban dos mutaciones (SDM) o todos los tres cambios juntos en una proteína, SQM. De hecho, el rendimiento promedio de la proteína a

partir de 4 preparaciones de SQM era de 58 \pm 29 mg de andamio purificado/litro de cultivo bacteriano, mientras que el rendimiento a partir de los tres cultivos de STM era de 59 \pm 41 mg de andamio/litro de cultivo.

5 Los espectros de CD para las diferentes variantes de STM se obtuvieron entre 200 nm y 260 nm. Un cambio en la forma del gráfico del espectro de CD se considera que refleja cambios en el contenido α -helicoidal y de β -lamina de la proteína plegada. Todas las variantes de STM mostraron espectros de CD similares con un punto de inflexión a aproximadamente 218 nm que indica la similitud estructural sin efectos alterantes mediante las modificaciones introducidas (Figura 3). Sin embargo, las diferencias observadas en la amplitud de los espectros de CD también se observaron con SDM, SUC y SUM que muestran una curva más plana, y SQM y SUN que muestran una curva más profunda en comparación con STM (Figura 3). Una posible explicación debería ser una estabilidad mejorada del plegamiento, es decir, el contenido con el paso del tiempo promedio de proteína correctamente plegada es mayor en soluciones de SQM y SUN en comparación con STM dando como resultado una lectura de elipticidad mayor. En base a los espectros de CD en combinación con los rendimientos de proteína determinados (Tabla 1) se concluye que es probable que la variante SQM sea una proteína de andamiaje adecuada.

Ejemplo 10

15 Las investigaciones se condujeron para valorar si SQM sería capaz de exponer péptidos para la interacción ensayando si las etiquetas de epítipo simple serían capaces de ser reconocidas por sus anticuerpos cognados cuando se exponen en los nuevos sitios creados en SQM. Se seleccionaron tres epítopos peptídicos (etiquetas AU1, HA y MYC) que difieren en características tanto de longitud como fisicoquímicas. Estos péptidos se insertaron individualmente o en diversas combinaciones en las posiciones disponibles (el terminal N, bucle 1 o bucle 2) en el andamio. Inicialmente, la etiqueta HA se insertó en el sitio amino-terminal, la etiqueta AU1 (el péptido más corto) en el bucle 1 y la etiqueta Myc en el bucle 2. Sorprendentemente, las inserciones de la etiqueta HA en el terminal amino solamente fueron escasamente toleradas, con rendimientos de proteína disminuidos aproximadamente 2,5 veces en comparación con el andamio vacío. Similarmente, la inserción Myc en el bucle 2 dio como resultado descenso en >5 veces en el rendimiento proteico en *E. coli*. Al contrario, la inserción de la etiqueta AU1 en el bucle 1 no desestabilizó SQM, y de hecho puede que incremente los rendimientos (Tabla 1).

Tomados en conjunto, los datos de expresión de proteína indican que SQM es capaz de exponer péptidos a partir de tres sitios – terminal amino, bucle 1 y bucle 2. De estos, el nuevo sitio bucle 1 parece ser el más ampliamente útil.

Ejemplo 11

30 Los experimentos se condujeron para valorar los efectos de la inserción de péptidos cortos derivados de aptámeros peptídicos existentes en el bucle 2 del andamio de STM previamente descrito. Estos péptidos eran de diversas longitudes, siendo el largo de 10 residuos (A48, A52 y A58), 17 residuos (A7) y 22 residuos (tándem A52). De estos, solamente A7 afectó a la estructura secundaria de la proteína STM (datos no mostrados). Sorprendentemente, cuando los mismos péptidos se insertaban en el bucle 2 de SQM, la proporción de la estructura secundaria en cada aptámero peptídico resultante era igual (datos no mostrados). Esto indica que SQM es más capaz de tolerar inserciones peptídicas que STM. Además, los experimentos se condujeron para valorar los efectos de la inserción de las etiquetas de epítipo en los tres sitios como se describió anteriormente (Tabla 1) sobre la proporción de la estructura secundaria en los aptámeros peptídicos resultantes. Los resultados inesperadamente mostraron que incluso aquellas inserciones que disminuyeron los rendimientos de proteína no interrumpieron apreciablemente la estructura secundaria de los aptámeros peptídicos resultantes (Figura 7). Sin embargo, esto indicó que la presencia de un péptido en el sitio N-terminal cambia la forma de la curva, empujando el punto de inflexión desde 218 nm hacia 209 nm (Figura 7, SQM-Ha). Para establecer si esto puede reflejar un efecto general de las inserciones en este sitio sobre la estructura del andamio, se analizó el espectro para una variedad de aptámeros peptídicos derivados de SQM con insertos en el bucle 1 y/o el bucle 2, todo con un inserto en el terminal amino. Consecuentemente se encontró que estas proteínas poseían menos estructura secundaria que las correspondientes proteínas que carecían de un inserto en el sitio N-terminal (Figura 7).

Ejemplo 12

Al haber determinado que es posible, dentro de los límites, insertar péptidos modelos en cada una de las tres posiciones de la presente invención, se encontró que se podría inmunoprecipitar el aptámero peptídico insertado. En referencia a la Figura 8A, se muestra la inmunoprecipitación del péptido AU1 insertado en el bucle 1 indicando que el péptido AU1 in el bucle 1 se podría unir por un anticuerpo anti-AU1 suficientemente bien de modo que el aptámero peptídico se podría inmunoprecipitar. Las Figuras 8B y 8C también muestran que los epítopos en el bucle 1 de SQM (AU1 y MYC) y en el bucle 2 de SQT (HA) se pueden reconocer por anticuerpo cognado, el cual puede inmunoprecipitar la variante de andamio etiquetada con epítipo. Esto se confirmó con otras etiquetas epítipo que se reconocen igualmente bien usando un formato de micromatriz (Figura 9). Los aptámeros peptídicos (Pep2, pep6, pep9 o pep10m) se inmovilizaron en el bucle 2 de SQM o SQT sobre un portaobjetos de microscopio de vidrio usando la química de amina y se sondaron con CDK2 activa comercial, soluble y no marcada, (New England Biolabs). Después de un exhaustivo lavado, se detectó la CDK2 unida a aptámeros peptídicos inmovilizados usando un anticuerpo anti-CDK2 y un anticuerpo secundario marcado, de lo cual se obtuvo una imagen usando un escáner de micromatriz de ADN estándar. Los datos mostraron que SQM puede exponer pep6 mejor que cualquier otro

andamio, mientras que SQT puede exponer pep10m mejor que cualquier variante de andamio y en un intervalo clínicamente relevante. El andamio SQT vacío da una señal baja que fijamos como el ruido de fondo. Los aptámeros peptídicos con diferentes epítomos a diferentes posiciones purificados de *E. coli* se imprimieron sobre portaobjetos de vidrio con o bien un quelato de níquel (que debería capturar la etiqueta de hexahistidina de cada aptámero peptídico para permitir la orientación controlada o simplemente con poli-L-lisina. Todas las impresiones empezaron y finalizaron con manchas de tampón de impresión y manchas de SQM libres que siempre dan no señal y actúan como controles negativos. En todos los casos la absorción de anticuerpo era específica. Estos resultados se han confirmado con SPR (Figura 10). Los anticuerpos tienen la afinidad más alta para la diana SQM en el siguiente orden: Anti-Myc, Anti-Ha y Anti-Au1, con cambios en la concentración superficial de 0,4, 0,13 y 0,02 $\mu\text{M cm}^{-2}$ observados tras la inyección de soluciones 33 nM (véase más adelante la Tabla 2). Estos valores indican las amplitudes cualitativas relativas de las constantes de equilibrio K. El valor K que se ha calculado para el anticuerpo anti-Myc para que esté en el orden de aproximadamente 50 nM, está de acuerdo con eso para los estudios similares hechos con anti-cistatina. Los valores K para Anti-Ha y Anti-Au1 no se han determinado con exactitud debido a su comportamiento no ideal en el ensayo de SPR. El Anti-Ha muestra una punta de asociación inmediata e intensa, seguido de un componente de asociación más lento secundario, con comportamiento similar visto en la disociación. Tal comportamiento indica procesos múltiples que se dan y pueden resultar de la unión a conformaciones alternativas múltiples del terminal amino de andamio, o incluso de impurezas en la muestra comercial. Alternativamente, puede ser el resultado de constantes de asociación y disociación cinética muy rápidas. AU1 se comporta similarmente a Myc con una curva de asociación bien resuelta, pero la magnitud absoluta del anticuerpo inmovilizado se ve que drásticamente se reduce en alrededor 2 órdenes de magnitud. En los casos de Ha y Au1 los valores k son claramente menores que 50×10^{-9} M.

La Tabla 2 muestra las respuestas relativas de los anticuerpos a aptámeros SQM covalentemente unidos que contienen todas las tres etiquetas de epítomos.

	Respuesta de SPR			K
	/Mdeg	/ng mm^{-2}	/pmol cm^{-2}	
SQM	135	1,1	5	
Myc 33 nM	79	0,64	0,4	50×10^{-9}
Ha 33 nM	25	0,20	0,13	$<50 \times 10^{-9}$
Au1 33 nM	3	0,03	0,02	$<50 \times 10^{-9}$

Los resultados de las muestras son cualitativamente idénticos entre las dos superficies y no se han observado diferencias fundamentales, aunque la señal-ruido y reproducibilidad es mejor sobre las superficies de amina. La similitud entre superficies indica que en tanto las orientaciones aleatorias como en escenarios de orientación controlados, todos los tres bucles están abiertos a los alrededores y son accesibles. No hay evidencia de epítomos idénticos colocados en el N-terminal, bucle 1 y 2 que tengan afinidades de unión variables. Los experimentos también se repitieron cuando los anticuerpos se añadieron simultáneamente en mezclas a las concentraciones anteriormente indicadas. En todos los casos se ven resultados comparables, indicando que los bucles están bien separados uno de otros y se comportan independientemente. No se ve evidencia de que los bucles estén bloqueados tras la unión del anticuerpo a bucles adyacentes. Esto indica que las superficies que proponemos usar para la exposición de péptido cada una se puede usar independientemente o también en combinación con los otros, y que el andamio modificado por ingeniería es capaz de exponer un área superficial inesperadamente grande para la interacción, compatible con la unión de tres moléculas de anticuerpo simultáneamente, lo cual se podría extrapolar para reconocer una proteína o complejo de multiproteína de tanto como 450 kDa.

Ejemplo 13

Se analizaron los perfiles de expresión de 864 péptidos aleatorios insertados en cada uno de los tres sitios. Para este experimento, se dejaron crecer cultivos a pequeña escala de los aptámeros peptídicos aleatorios en placas de 96 pocillos, se purificaron los aptámeros peptídicos en alta producción (es decir, sin optimizar los protocolos de expresión o purificación para cada pocillo) y se salpicó un volumen igual de cada aptámero peptídico en un portaobjeto de microscopio de vidrio, creando una pequeña micromatriz. A continuación, se sondó la micromatriz con cualquiera de los dos anticuerpos que reconocen la proteína de andamiaje. Entonces, la intensidad de señal obtenida con el anticuerpo sería proporcional a la cantidad de aptámero peptídico a cada rasgo de la matriz. Los resultados mostraron que las inserciones en el bucle 1 y bucle 2 generalmente se expresaron bien, mientras que casi el 50% de los insertos amino-terminales no se expresaron significativamente tan bien (Figura 11). Los resultados mostraron que el 68% de los 384 aptámeros peptídicos que ensayamos con insertos aleatorios de 6 aminoácidos de largo en el bucle 1 de SQM estaban bien expresados, y el 16% se podían expresar a un nivel inferior, mientras que la expresión del 15% no se podía detectar. El 76% de los 384 aptámeros peptídicos que

ensayamos con insertos aleatorios de 12 aminoácidos de longitud en el bucle 2 de SQM estaban bien expresados, mientras que el 14% se podía expresar a un nivel inferior y el 10% no se podía detectar en absoluto. Finalmente, de los 192 aptámeros peptídicos con insertos aleatorios en el terminal amino, solamente el 35% se expresaron bien, mientras que el 32% se expresaron a un nivel inferior y el 32% no se podían detectar.

- 5 En conclusión, aunque muchos péptidos diferentes se pueden exponer por inserciones en el sito amino-terminal, estos también pueden ser con frecuencia perjudiciales a la estabilidad del andamio. Por consiguiente, se propone que los andamios y los sitios novedosos de la presente invención se pueden usar para producir aptámeros peptídicos que interactúan con la proteína diana usando un área de superficie mayor y así con mayor afinidad y especificidad.

10 Ejemplo 14

Se emprendió modificación por ingeniería adicional del bucle 2 en un intento de rescatar el efecto de desestabilización de la nueva mutación en la posición 82 a 83 en comparación con STM. Por consiguiente, se modificó por ingeniería un nuevo andamio que posee los mismos cambios que SQM en el terminal amino y bucle 1, también tiene 71-NGP-73 pero ahora sustituye la secuencia 78-EDL-80 de tipo natural (Estefina A) con o bien 78-ADR-80 o 77-SDRL-80, o a 78-NTD-80. Cada uno de estos cambios se diseñaron para permitirnos el uso de dos sitios RsrII para introducir oligonucleótidos que codifican péptidos en el bucle 2. De estos, la versión con 78-ADR-80 resultó ser la más flexible. Se diseñó el nuevo andamio como SQT (SEQ ID NO: 24). Se encontró que los rendimientos de expresión de proteína en *E. coli* para SQT se disminuían en comparación con los de para SQM (Tabla 3). Sin embargo, los rendimientos para las versiones de SQT con insertos en el bucle 2 se mejoraron consecuentemente en comparación con los insertos de las mismas secuencias en el bucle 2 de SQM (Tabla 3). Estas diferencias eran más pronunciadas con AU1 en el bucle 2, donde el nivel de expresión era menor que 1 mg de aptámero peptídico a partir de 1 l de cultivo bacteriano cuando se usa SQM, pero era 45 mg/l de cultivo con SQT. La Tabla 3 muestra el rendimiento del péptido del aptámero peptídico que se expresa como mg de proteína purificada obtenida bajo condiciones estándar a partir de 1 litro de cultivo bacteriano.

25

Tabla 3

INSERTO	SQM	SQT
NINGUNO	19	15
Pep2	9	16
Pep6	2	3
Pep9	21	105
Pep10m	2	11
myc	11	65
AUI (bucle 1)	103	25
AUI (bucle 2)	1	36
AUI, AUI	nd	49
HA (bucle 2)	nd	32
AUI, HA	205	36
myc (bucle 1)	7	34
myc (bucle 2)	11	49
AUI, myc	14	52

30

Una observación sorprendente fue que los rendimientos de expresión para insertos en el bucle 1 solo podía ser mayor para SQM que para SQT. Por ejemplo, 25 mg de aptámeros peptídicos por litro de cultivo se obtuvieron para SQT-AU1-bucle 1, pero 103 mg por litro para SQM-AU1-bucle 1, aunque esto no es universal, ya que el efecto se invertía para los aptámeros peptídicos con el epítipo Myc en bucle 1 (Tabla 3). Los aptámeros peptídicos con doble inserciones en bucle 1 y bucle 2 de SQT generalmente se expresaron bien, y mostraron rendimientos mejorados

sobre las mismas combinaciones en SQM, aunque obtuvimos nuestro mayor rendimiento de un aptámero peptídico hasta la fecha cuando se insertó la etiqueta AU1 en el bucle 1 y la etiqueta HA en el bucle 2 de SQM (Tabla 3).

5 Cuando se consideran los efectos de los insertos de péptido alternativos, se encontró que SQT era capaz de exponer los péptidos de unión a CDK modelos (pep2, 6, 9 y 10m) o el epítipo Myc. Generalmente, SQT era más capaz de aceptar estos péptidos que SQM. Sin embargo, sorprendentemente, SQT parecía poseer poca (pep2, pep9) o no (pep6, pep10m) ventaja sobre STM. Estos datos sugieren que la mínima modificación por ingeniería del andamio y la inserción peptídica en un bucle se acomoda más fácilmente que las alteraciones más complejas, incluso si estas están diseñadas para minimizar la alteración de la estructura secundaria.

Ejemplo 15

10 Para mejorar la eficacia de la construcción de genoteca usando el bucle 1 de SQT, insertamos en el sitio *NheI* un oligonucleótido de secuencia **5'-AGGCCTTGATCACCATGGACTAGCA-3'** (SEQ ID NO:26), usando un sitio *NheI* en el extremo 5' y un sitio *SpeI* en el extremo 3', el último se pierde cuando se anilla al sitio *NheI* cortado y esto también da como resultado la pérdida del sitio *NheI* 3' en el marco de lectura abierto de STM. El oligonucleótido insertado lleva 3 nuevos sitios de restricción (*StuI*, *BclI* y *NcoI*) que se pueden usar para la inserción de oligonucleótidos codificadores de péptido. La secuencia final del andamio modificado en esta región, incluyendo el nuevo conector es: ADN:

15 **5'-AAGTGCTAGCAGGCCTTGATCACCATGGAC TAGCAAGCACAAATTA-3'** (SEQ ID NO: 27).

Proteína: (terminal N) **44-KTQVLAgldhhgla STNYIKVRAG-** (terminal C) con los aminoácidos codificados por el oligonucleótido insertado dado en caso inferior. Indicar que al usar los sitios *NheI* y *NcoI* del conector para insertar oligonucleótidos dará como resultado insertos que finalizan con dos residuos (LeuAla) codificados por la secuencia anillamiento *SpeI/NheI*, y la proteína de andamiaje SQT es así dos residuos más larga que la Estefina A original. Llamamos a esto SQL andamio alargado. Para preguntar si la estrategia mejora la construcción de genoteca, insertamos un oligonucleótido aleatorio flanqueado por sitios 5'-*NheI* y 3'-*NcoI* en los sitios correspondientes en SQL. Todos los 20 de los transformantes que secuenciamos mostraron inserción correcta de un oligonucleótido sencillo en cada clon. Al contrario, cuando usamos el sitio *NheI* solo y clonamos y secuenciamos 24 de los transformantes, solamente 8 eran como se requería (insertos sencillo, orientación correcta). Otros nueve clones eran plásmidos vacíos (indicando etapas de digestión y desfosforilación ineficaces), 3 eran concatámeros (cuando los oligonucleótidos se ligan uno a otro antes de la ligación en el andamio ORF) y 4 estaban en la orientación errónea. Estos números muestran un enorme mejoramiento en el uso de la inserción orientada sobre los obtenidos cuando oligonucleótidos aleatorios flanqueados en ambos extremos con un sitio *NheI* se insertan en el sitio *NheI* de SQT. Por consiguiente, hemos implementado una estrategia similar con el bucle 2, al cambiar la secuencia de nucleótido del marco de lectura abierto para crear un sitio *XmaI* en lugar del sitio *RsrII* de SQT. Esto no cambia la secuencia de aminoácidos de SQT.

20

25

30

Ejemplo 16

35 Se usaron los dos aptámeros peptídicos que muestran la afinidad aparente más alta para CDK2 a partir del experimento de micromatriz en un protocolo de "inmunotransferencia" (también denominado una transferencia tipo Western, véase Figura 12), excepto que se usaron los aptámeros peptídicos (SQM-pep6 y SQT-pep10m) en lugar de anticuerpos. Los anticuerpos similarmente se omitieron de la etapa de desarrollo de la "inmunotransferencia" ya que los aptámeros peptídicos poseen una etiqueta S que se puede detectar usando la proteína S fusionada a la peroxidasa de rábano picante (HRP, del inglés "Horse Radish Peroxidase"). Como control negativo, las transferencias paralelas se sondaron con el andamio vacío (se muestra SQT) y se desarrollaron usando el protocolo de proteína S. Como control positivo, las transferencias paralelas se sondaron con anticuerpo anti-cdk2, que era capaz de detectar una proteína del mismo peso molecular relativo, aunque se requirieron tiempos de exposición mayores cuando se usó el anticuerpo a la misma concentración que los aptámeros peptídicos. La Figura 12 muestra dos lisados de las células de cultivo de tejido humano que se sometieron a electroforesis en gel de poliacrilamida de desnaturalización antes de ser transferidas a membranas de nitrocelulosa o PVDF y se sondaron con andamio, aptámeros peptídicos o andamio como se describieron anteriormente. Los datos muestran que los aptámeros peptídicos no solamente pueden sustituirse por anticuerpos en este protocolo, sino que en realidad funcionan mejor ya que se detectaba CDK2 más rápidamente usando aptámeros peptídicos en lugar de anticuerpos. Por tanto, ventajosamente, el uso de aptámeros peptídicos en lugar de anticuerpos significa que se necesitará usar menos animales en la producción de sondas moleculares.

40

45

50

REIVINDICACIONES

1. Un polipéptido de Estefina A que comprende la secuencia de aminoácidos que tiene al menos 80% de identidad con SEQ ID NO: 1; comprendiendo dicho polipéptido una mutación en la posición 4 en donde la Glicina de Estefina A está sustituida por Arginina (G4R);
- 5 caracterizado por que el polipéptido comprende además una inserción de péptido heterólogo, en donde
 - a) dicho péptido heterólogo está insertado en la proteína en el sitio G4R de Estefina A; o
 - b) dicho péptido heterólogo está insertado en la proteína en la posición 46 a 54; o
 - c) dicho péptido heterólogo está insertado en la proteína en la posición 67 a 84.
- 10 2. Un polipéptido de Estefina A según la reivindicación 1 que comprende una inserción de péptido heterólogo adicional de modo que al menos un péptido heterólogo está insertado en la proteína en al menos dos posiciones seleccionadas entre (a), (b) y (c).
3. Un polipéptido de Estefina A según la reivindicación 2 que comprende una inserción de péptido heterólogo adicional de modo que al menos un péptido heterólogo está insertado en la proteína en todas las tres posiciones (a), (b) y (c).
- 15 4. Un polipéptido de Estefina A según cualquiera de las reivindicaciones 1 a 3 en donde cuando un péptido heterólogo está insertado en la proteína en la posición 67 a 84, dicho péptido heterólogo está insertado en la proteína en la posición 82/83.
5. Un polipéptido de Estefina A que comprende la secuencia de aminoácidos que tiene al menos 80% de identidad con SEQ ID NO: 1; comprendiendo dicho polipéptido una mutación en la posición 4 en donde la Glicina de Estefina A está sustituida por Arginina;
- 20 caracterizado por que el polipéptido comprende además una inserción de péptido heterólogo, en donde
 - i) dicho péptido heterólogo está insertado en la proteína en la posición 46 a 54; o
 - ii) dicho péptido heterólogo está insertado en la proteína en la posición 67 a 84.
- 25 6. Un polipéptido de Estefina A según la reivindicación 5 que comprende una inserción de péptido heterólogo adicional de modo que al menos un péptido heterólogo está insertado en la proteína en cada una de las posiciones (i) y (ii).
7. Un polipéptido de Estefina A según la reivindicación 6 en donde cuando un péptido heterólogo está insertado en la proteína en la posición 67 a 84, dicho péptido heterólogo está insertado en la proteína en la posición 82/83.
- 30 8. Un polipéptido que comprende la secuencia de aminoácidos como se expone en una cualquiera de las SEQ ID NO: 10, 13, 16, 17, 22, 23, 24 y 25.
9. Un polipéptido de Estefina A que comprende la secuencia de aminoácidos que tiene al menos 80% de identidad con SEQ ID NO: 1; comprendiendo dicho polipéptido una mutación en la posición 4 en donde la Glicina de Estefina A está sustituida por Arginina (G4R);
- 35 caracterizado por que el polipéptido comprende además una inserción de péptido heterólogo, en donde
 - a) dicho péptido heterólogo está insertado en la proteína en el sitio G4R de Estefina A; o
 - b) dicho péptido heterólogo está insertado en la proteína en la posición 46 a 54; o
 - c) dicho péptido heterólogo está insertado en la proteína en la posición 67 a 84.
10. El uso de un polipéptido según cualquier reivindicación precedente como proteína de andamiaje.
- 40 11. Un método para identificar un péptido diana capaz de unirse a una estructura de interés que comprende:
 - (i) proporcionar una proteína de andamiaje de estefina A como se define en una cualquiera de las reivindicaciones 1 a 9 que comprende un péptido diana;
 - (ii) poner en contacto dicha proteína de andamiaje con dicha estructura de interés; y

- (iii) hacer un seguimiento de la asociación entre el andamio y la estructura de interés, en donde la asociación de la proteína de andamiaje con la estructura de interés identifica el péptido diana como péptido diana candidato capaz de unirse dicha estructura.
- 5 12. El uso de un polipéptido como se define en una cualquiera de las reivindicaciones 1 a 9 para presentar péptido(s) sobre micromatriz(ces).
13. El uso de un polipéptido como se define en una cualquiera de las reivindicaciones 1 a 9 como agente para unir y detectar específicamente un biomarcador.
14. El uso de un polipéptido como se define en una cualquiera de las reivindicaciones 1 a 9 en un ensayo de interacción, en donde dicho polipéptido está inmovilizado.
- 10 15. Un método para detectar una estructura de interés que comprende:
proporcionar una proteína de andamiaje de estefina A como se define en una cualquiera de las reivindicaciones 1 a 9 que comprende un péptido diana capaz de unirse a dicha estructura de interés;
poner en contacto dicha proteína de andamiaje con dicha estructura de interés; y
hacer un seguimiento de la asociación entre el andamio y la estructura de interés, en donde la asociación de la proteína de andamiaje con la estructura de interés detecta dicha estructura.
- 15 16. Un método según la reivindicación 15 en donde dicha estructura de interés comprende un biomarcador.
17. Un ácido nucleico aislado que comprende la secuencia de nucleótidos que codifica la secuencia de aminoácidos de un polipéptido según cualquiera de las reivindicaciones 1 a 7, o reivindicación 9.
- 20 18. Un ácido nucleico aislado que comprende la secuencia de nucleótidos que codifica un polipéptido que comprende la secuencia de aminoácidos como se expone en una cualquiera de las SEQ ID NO: 10, 16, 17, 22, 23, 24 y 25.

FIGURA 1A

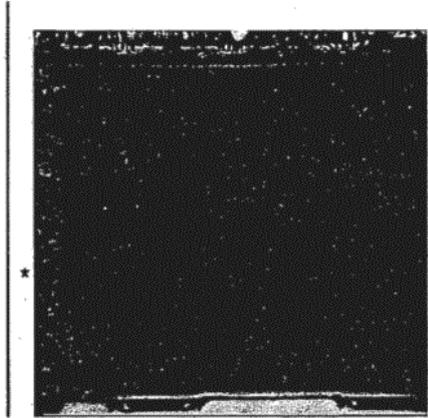


FIGURA 1B

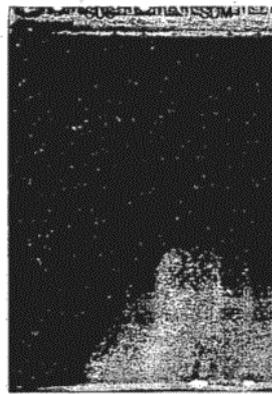


FIGURA 1C

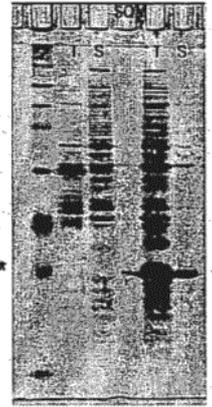


FIGURA 2

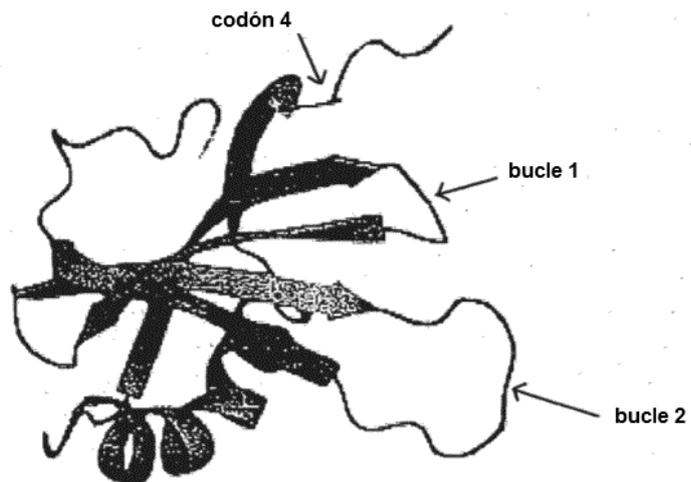


FIGURA 3

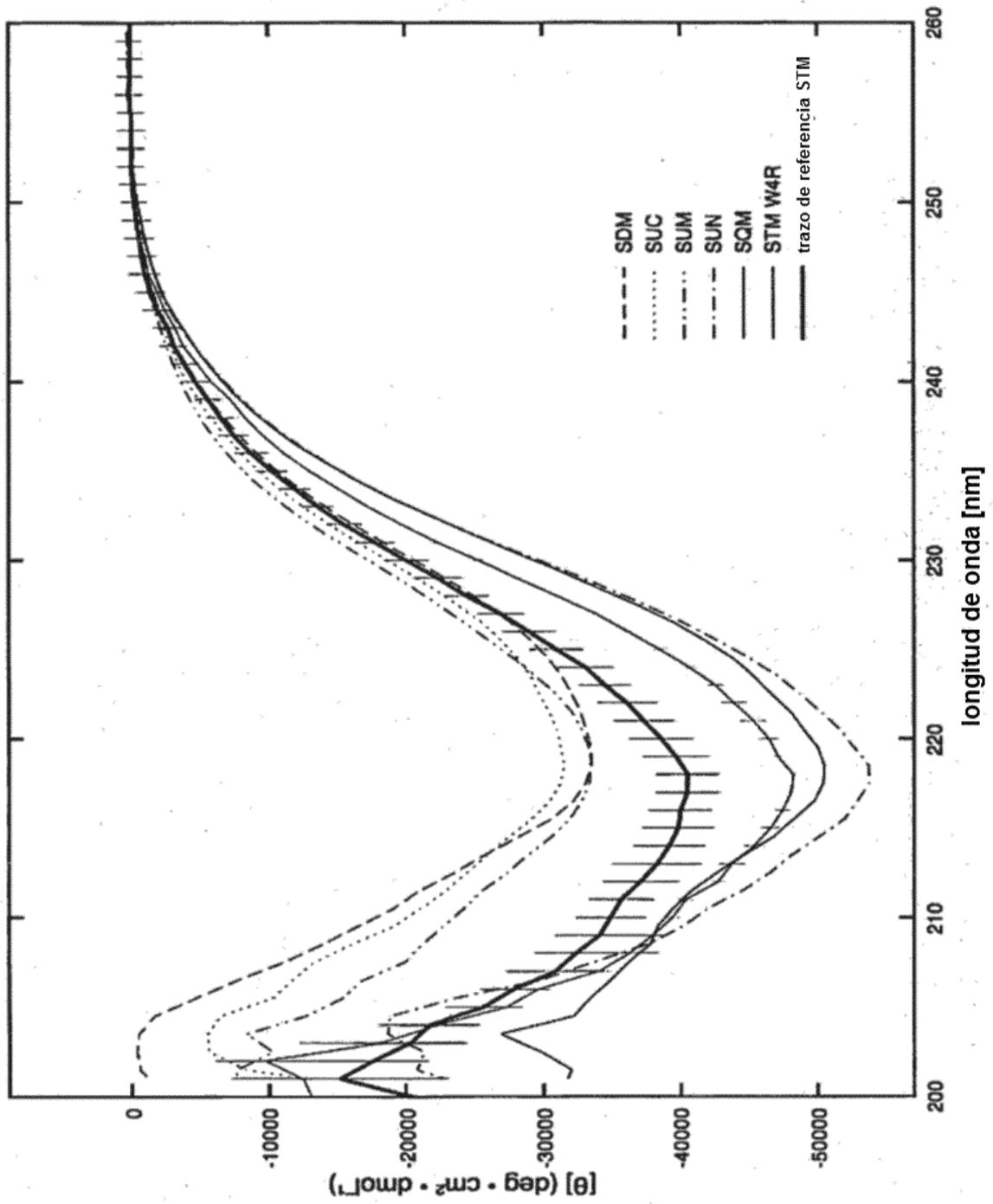


FIGURA 4

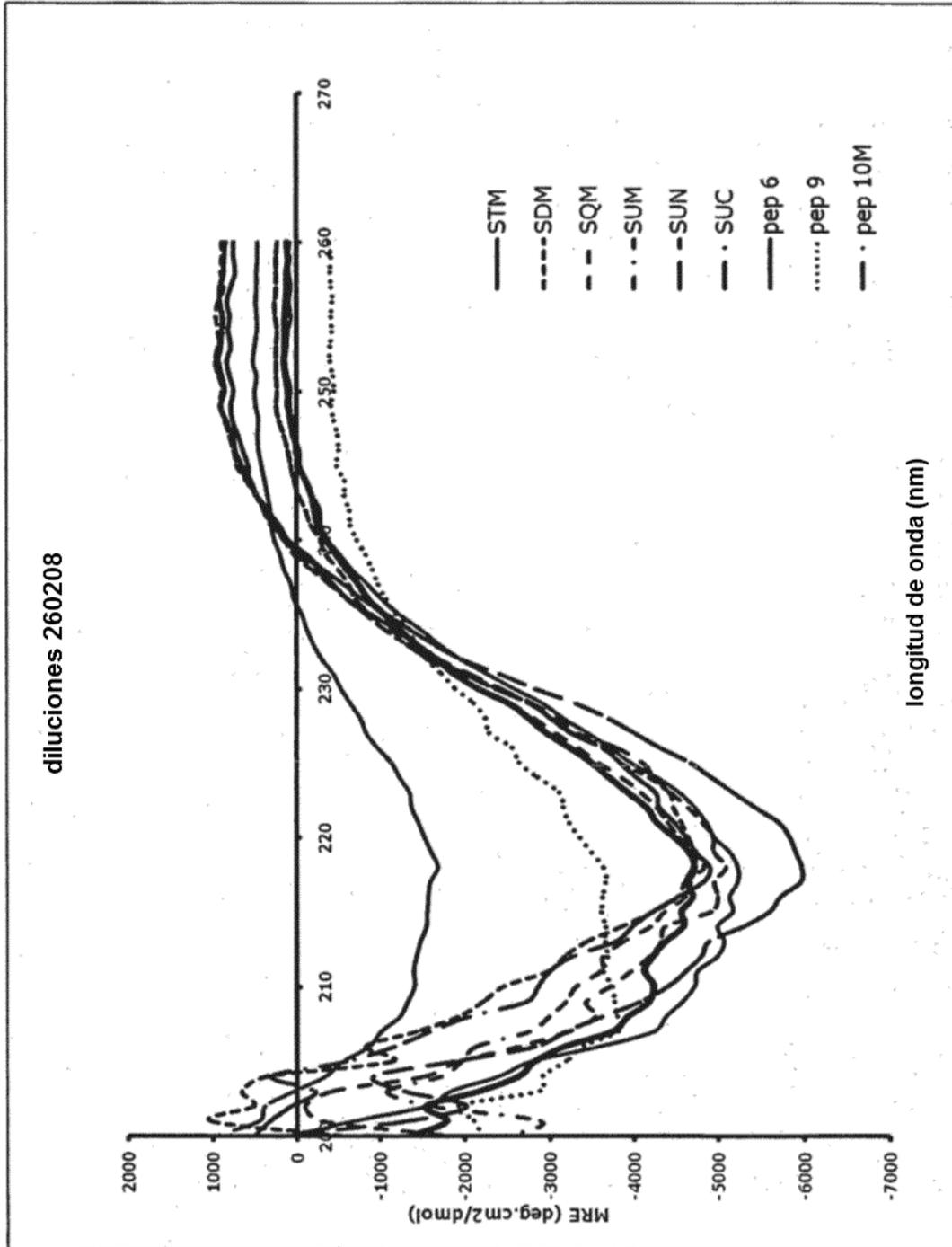


FIGURA 5

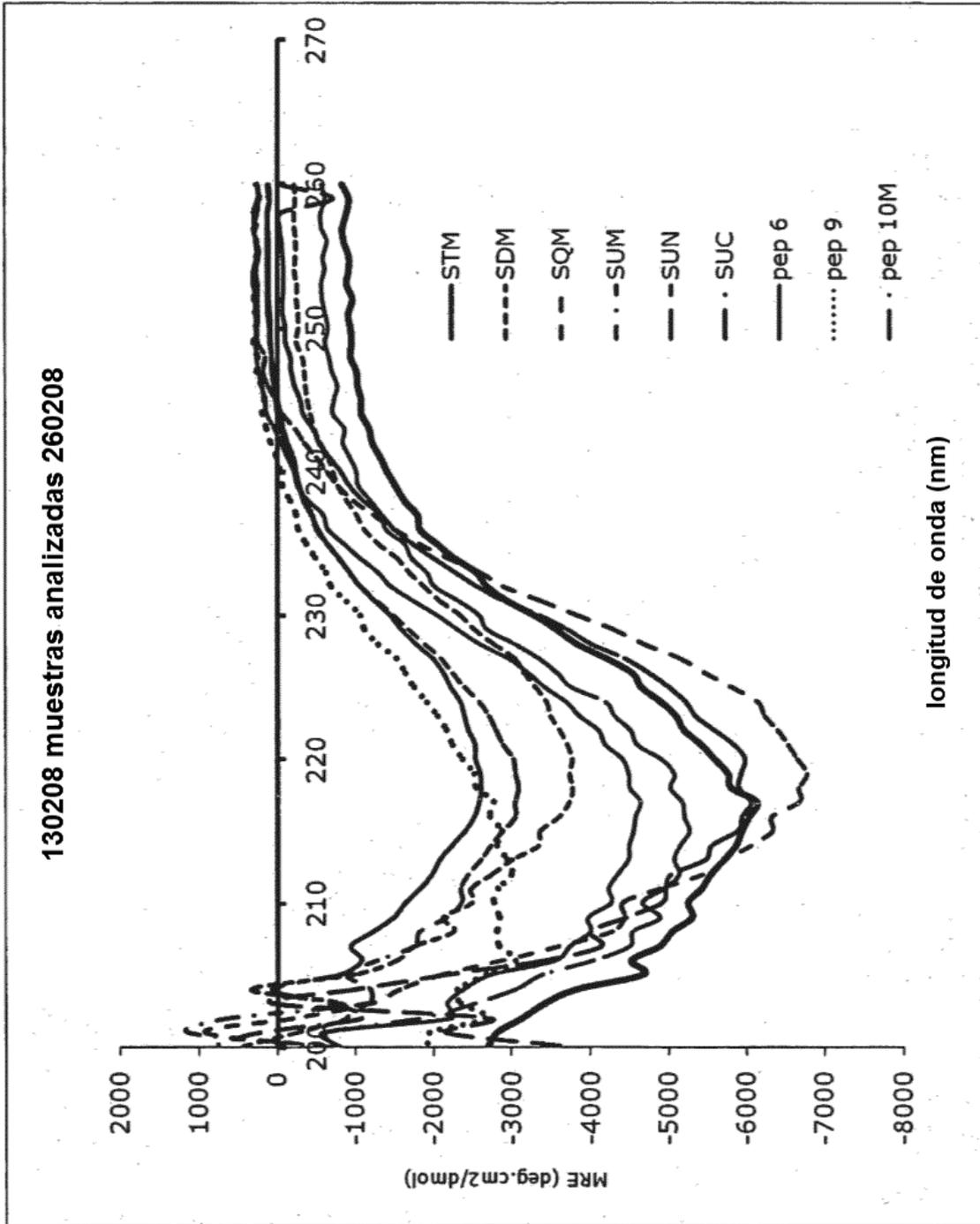


FIGURA 6

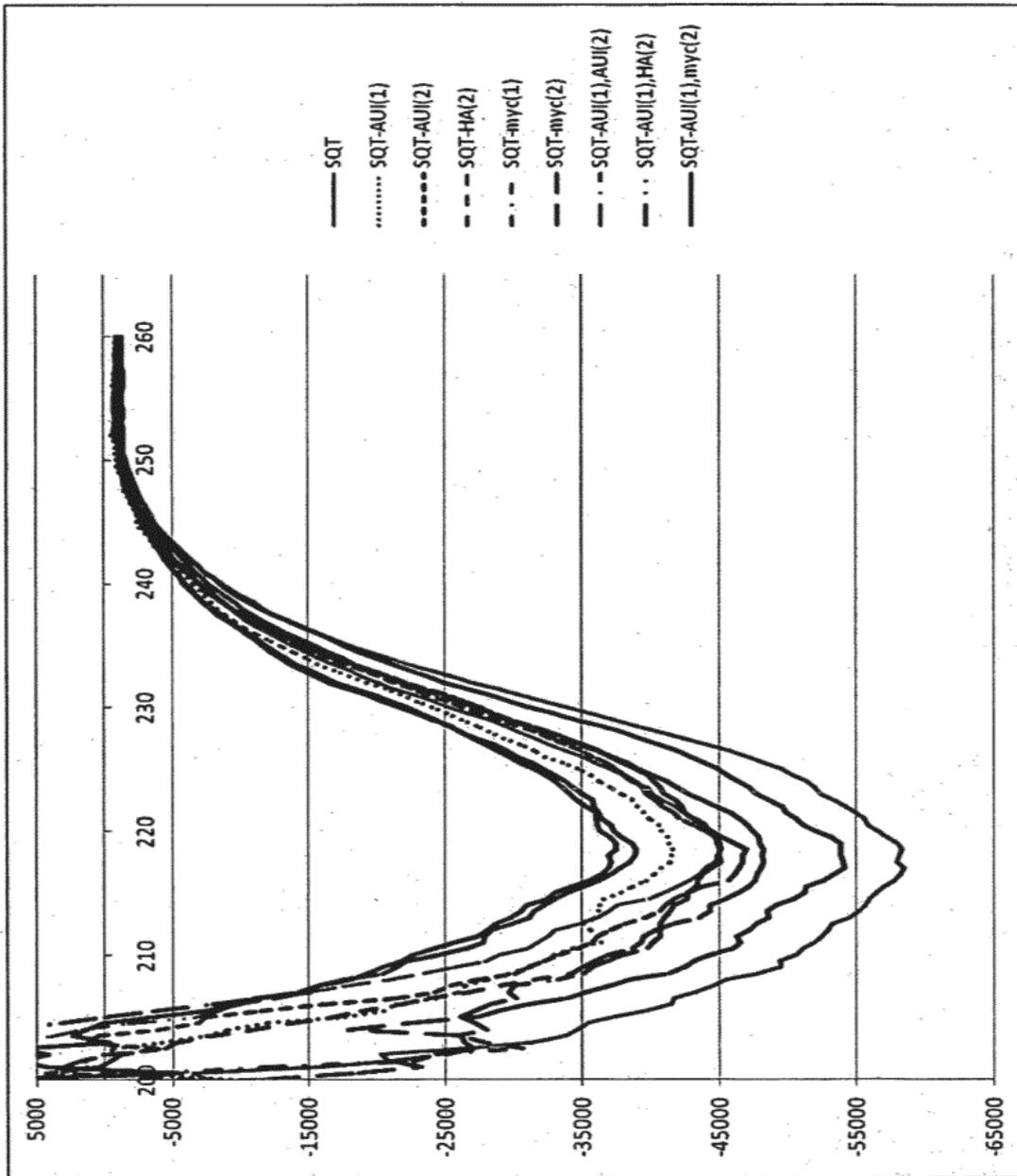


FIGURA 7

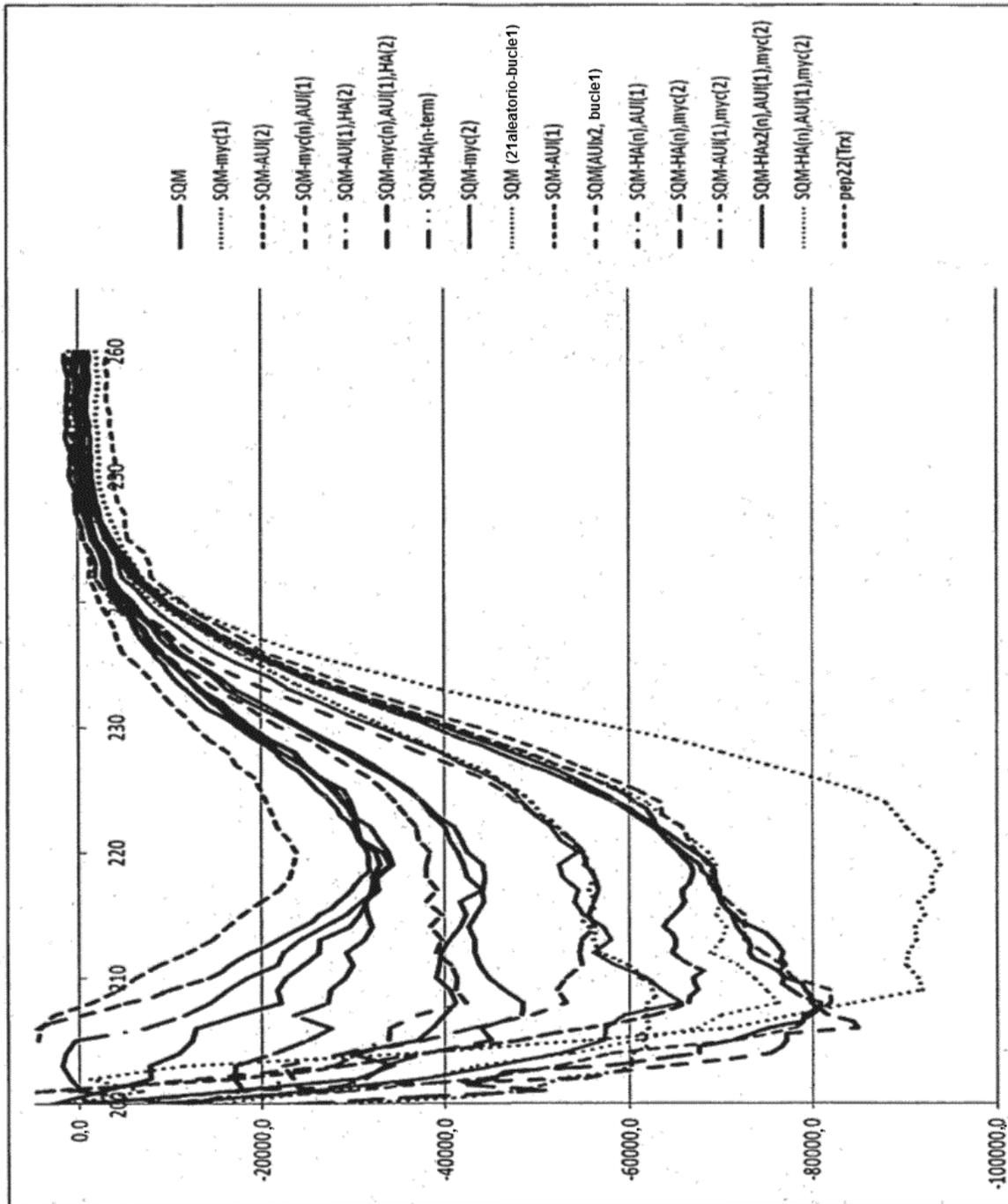


FIGURA 8A

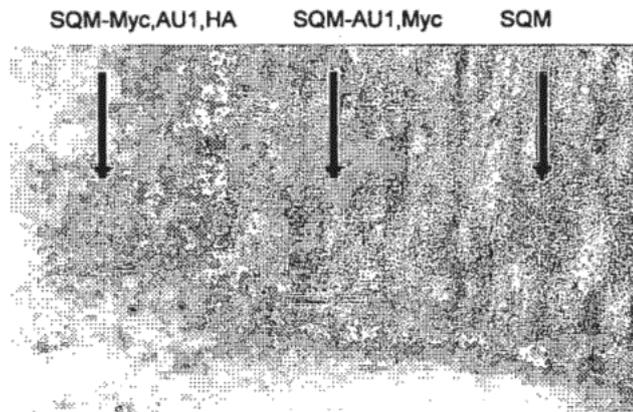


FIGURA 8B

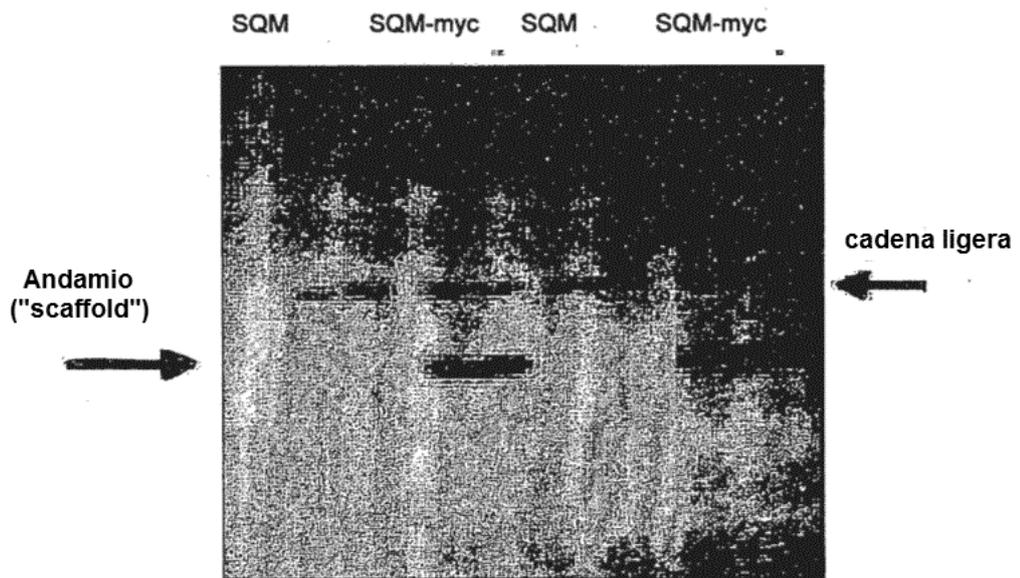
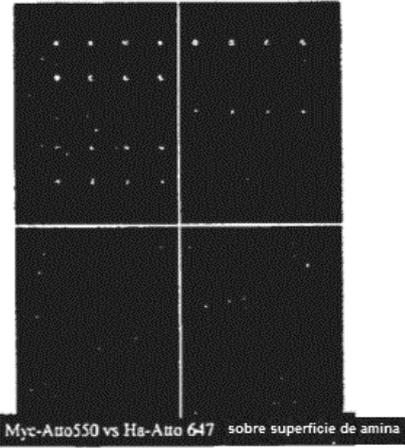
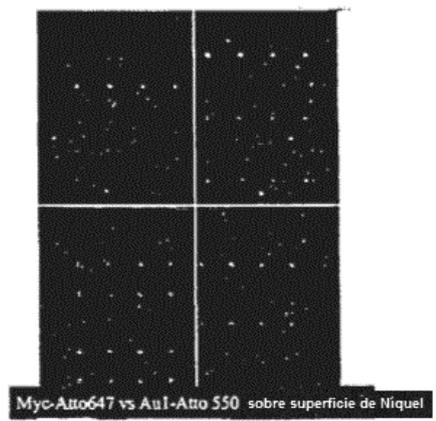
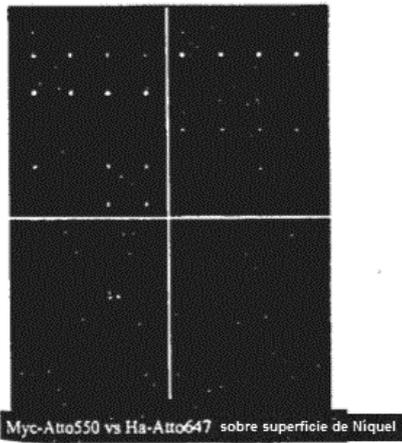
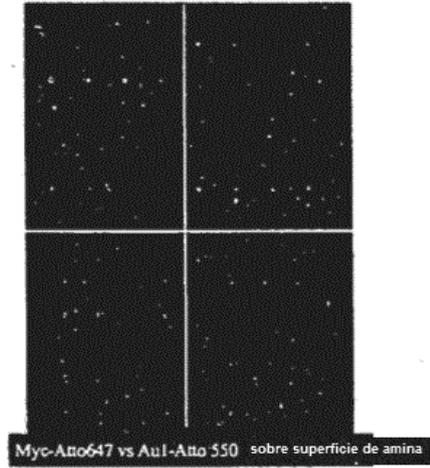


FIGURA 9

A



B



C

D

FIGURA 10A

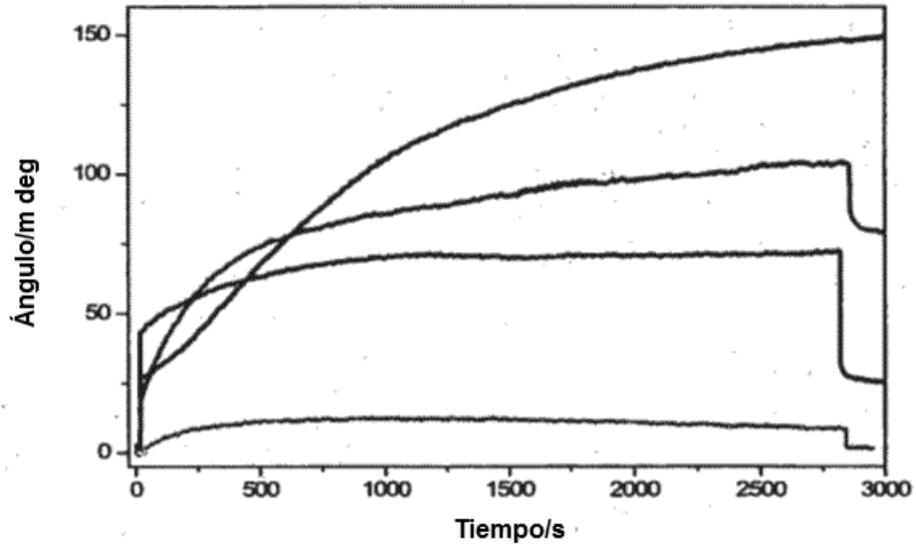


FIGURA 10B

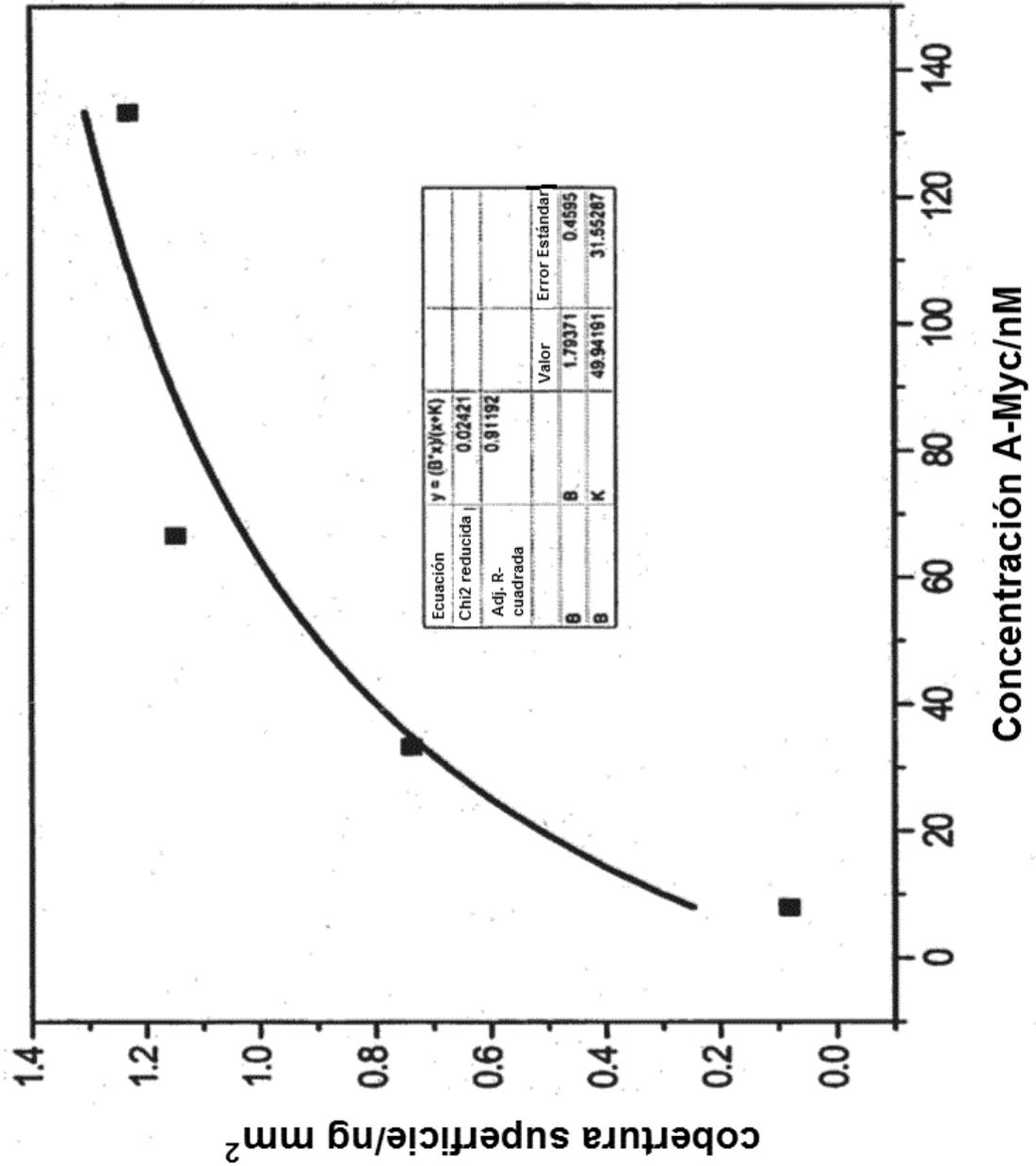


FIGURA 11

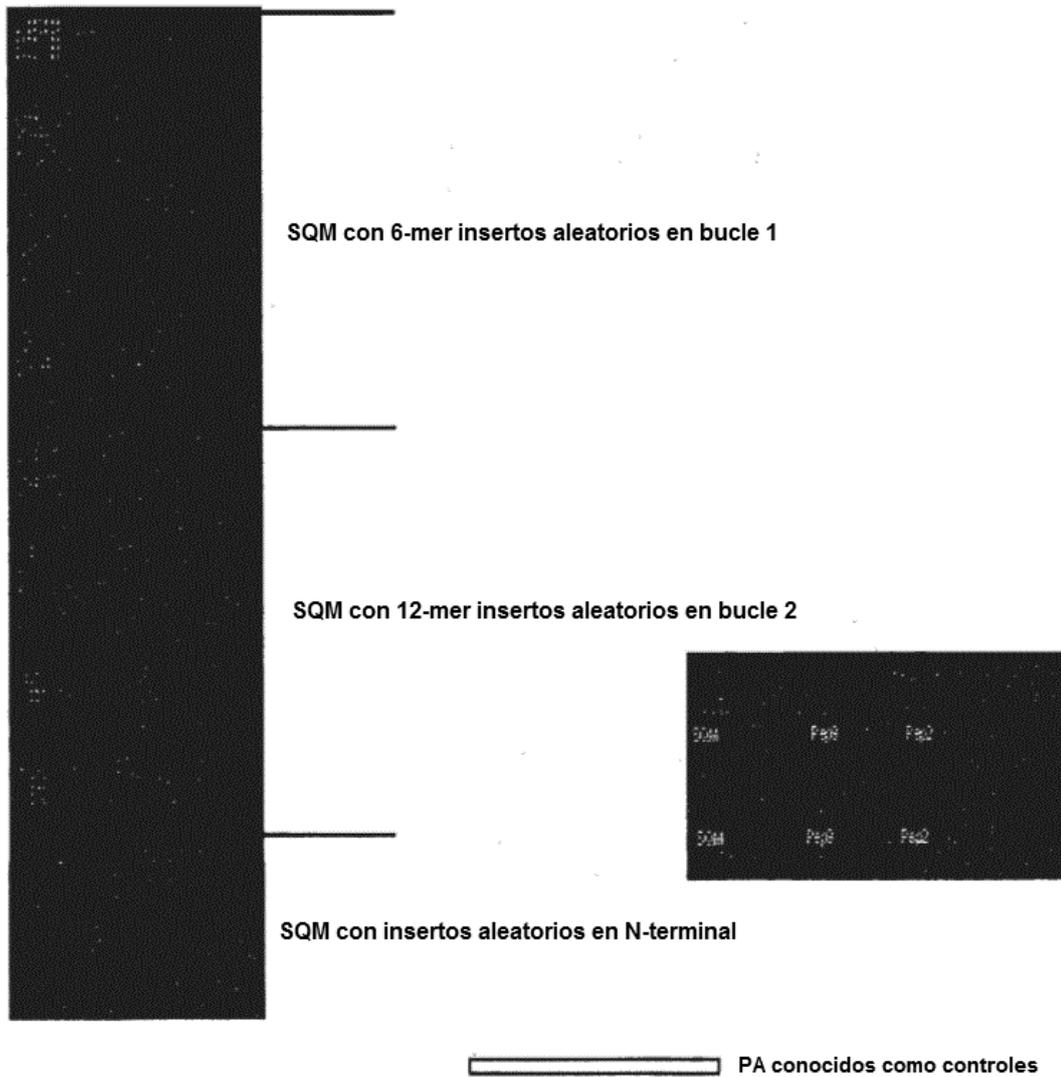
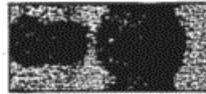


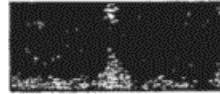
FIGURA 12



SQT



SQM-pep6



SQT-10M